ABSTRACT

KLUMP, ANDREW JAMES. Magnesium Doping and Surface Kinetics of III-Nitrides. (Under the direction of Drs. Ramón Collazo and Zlatko Sitar).

Despite the commercial success of the III-nitride based light-emitting and laser diodes over the past 30 years, two sources of inefficiency still limit the performance of these devices. The first is the reliance on non-native materials as the starting substrate for the epitaxial growth of the thin films. Dislocations are generated in the epi-layers as a result of the lattice mismatch to these substrates, and induce a spiral growth morphology – resulting in non-uniform layer thicknesses and doping content across the wafer. The second source is the low conductivity of the p-type layers based on the magnesium acceptor type dopant. The passivation and self-compensation of Mg during growth combine with a high ionization energy of the acceptor to limit the resistivity to a few tenths of an Ω cm.

To overcome these limitations, this dissertation approached the inefficiencies along two fronts. First, a series of gallium nitride (GaN) films was grown homoepitaxially on native GaN substrates to eliminate the spiral growth morphology occurring at high dislocation densities. A detailed introduction to the thermodynamics of the metal organic chemical vapor deposition is provided to understand the vapor supersaturation of the process, which guided the selection of growth conditions. A review of nucleation theory, and Burton, Cabrera, Frank theory of surface kinetics is given before extending the models to include asymmetry in adatom attachment rates to explain the films' morphology. Four growth conditions ranging in vapor supersaturation values from 70 to 36,000 were grown on offcuts of 0.3 and 0.4° to observe the growth mode and surface morphology. Guidelines for obtaining smooth step-flow surfaces are then established.

The second approach was to investigate the compensation of Mg in both GaN and aluminum gallium nitride ($Al_xGa_{1-x}N$, x=0.6). In GaN, two techniques were utilized during growth

to mitigate the effect of passivation and self-compensation. The first was to illuminate the samples with above bandgap light during the growth process, which can shift the quasi Fermi level of the unwanted defects. This technique was highly effective at reducing residual compensation before the onset of self-compensation (Mg~1.5×10^{19} cm^{-3}). The second technique was a chemical potential control scheme where a change in the vapor supersaturation of the growth affects the incorporation of defects. In this case, increasing the flow of ammonia during growth reduced the compensation occurring with Mg concentrations above the self-compensation onset. For the study of Mg doping in AlGaN, increased ammonia flow caused a pre-reaction in the gas phase and increased the compensation of the films. However, van der Pauw Hall effect was measured for the first time when the ammonia flow was restricted to a standard liter per minute. The results indicate that the passivation and self-compensation are much reduced for Mg:$Al_{0.6}Ga_{0.4}N$ in comparison to Mg:GaN films, and the limitation to the conductivity is due to poor mobility from alloy scattering.

Finally, in support of the study of Mg doping in the III-nitrides, secondary ion mass spectrometry (SIMS) was performed to determine total concentrations of Mg in the films. The ION TOF SIMS V equipment was previously unable to perform this measurement because of the low secondary ion intensity of Mg. Increasing the pulse width of the primary ion beam, and the number of analysis cycles per data cycle, enabled Mg concentrations as low as 10^{17} cm^{-3} to be reported for GaN. In AlGaN, the films become too insulating for accurate analysis. A method to reduce charging and allow unimpeded secondary ionization was developed. Additionally, correction factors to the composition, sputtering rate, and relative sensitivity factor were generated so that highly accurate multilayer depth profiles could be provided.

Magnesium Doping and Surface Kinetics of III-Nitrides

by
Andrew James Klump

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Materials Science and Engineering

Raleigh, North Carolina
2020

APPROVED BY:

_____	_____
Zlatko Sitar Committee Co-Chair	Ramón Collazo Committee Co-Chair
_____	_____
Elizabeth Dickey	Leda Lunardi

BIOGRAPHY

Andrew Klump was born on March 3, 1990 in St. Louis, Missouri. In 2012, he earned a BS in Physics from Seton Hall University in South Orange, New Jersey. He was then employed as SIMS technician at Evans Analytical Group, Inc. in East Windsor, New Jersey until 2015. In the Fall of 2015, he enrolled into the doctoral program in the Department of Materials Science and Engineering at North Carolina State University.

ACKNOWLEDGMENTS

There are many important people that have given me tremendous support throughout my doctoral studies that enabled me to complete this dissertation. First and foremost, to my family who has always had faith in me and have given me strength throughout the ups and downs of the last 12 years. I would like to acknowledge Professor Alper Sahiner, who advised me through the physics program at Seton Hall and gave me my first experiences in a laboratory setting. A heartfelt thanks to Temel Buyuklimanli, Charles Magee, Wei Ou, and Jeff Serfass at EAG in New Jersey. The professionalism of the lab, and their overall friendship and mentorship, gave me the model of how to conduct myself as a scientist. To Professors Ramón Collazo and Zlatko Sitar, for building a lab that generated the best AlGaN samples I had ever seen with SIMS, and for giving me this opportunity to learn about the growth side of this material system. Also, I would like thank the group members Alex Franke, Baxter Moody, Biplab Sarkar, Dennis Szymanksi, Dolar Khachariya, Dorian Alden, Felix Kaess, Hayden Breckenridge, James Tweedie, Ji Hyun Kim, Pegah Bagheri, Pramod Reddy, Qiang Guo, Ronny Kirste, Seiji Mita, Shun Washiyama, Will Mecouch, and Yan Guan for their friendship and support over the past 5 years. An additional thanks is given to Marc Hoffmann and Florian Hörich when I collaborated at OVGU. At AIF, I would like to thank Elaine Zhou and Fred Stevie for friendship and their vast SIMS knowledge they shared with me. I wish to thank Beth Dickey and Leda Lunardi for their advice and their participation as committee members. And lastly, for the tremendous support of Edna Deas and Berni Premachandra.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1 Introduction

1. Motivation and Overview

Over the past three decades, the III-nitrides have revolutionized the lighting industry with the emergence of the blue light-emitting diode (LED) based on gallium nitride (GaN) and indium gallium nitride (InGaN).[1,2] When paired with a yellow phosphor coating, this LED can generate white light and replace previous sources which are far less efficient, have shorter lifetimes, and are potentially toxic to the environment.[2,3] As the world is striving to reduce its reliance on fossil fuels for energy generation, this is a key technology to lower the demand placed on the power grid. Accordingly, the Royal Swedish Academy of Sciences awarded the 2014 Nobel Prize in Physics for this invention and validated the importance of the III-nitrides to the general public.[3]

Beyond visible lighting, the III-nitrides have promise to expand into the ultraviolet (UV) part of the electromagnetic spectrum with aluminum nitride (AlN) and the ternary alloy of aluminum gallium nitride (AlGaN).[4] One of the key characteristics of this high energy light is its ability to sterilize or disinfect surfaces.[4] As the recent pandemic has taught the world, having the capability to inexpensively and effectively clean surfaces or equipment that have been contaminated can help mitigate the spread of disease and save lives. Therefore, research in this field has even more importance than ever.

Although InGaN based commercial LEDs and laser diodes have been in production since the 2000s, there are still significant areas of improvement. Unlike its predecessors based on silicon or gallium arsenide (GaAs), the commercial III-nitride devices are produced primarily on the non-native sapphire substrates.[5] As a result, the devices suffer from a large density of dislocations which may act as a source of leakage and compensation in the films.[6-8] More significantly, the dislocations cause a spiral growth morphology across the surface of the epi-layer.[5] Such a

morphology can yield large differences in height and alter dopant/impurity incorporation across a film.[9,10] Device structures, such as the multi-quantum-well active regions of LEDs and laser diodes, require sharp layer interfaces and precise thicknesses which cannot be uniformly maintained on spiral growth surface.[9,10]

The other major source of inefficiency in the InGaN-based blue LEDs, and similarly limit the advancement of AlGaN-based UV-LEDs, is the poor p-type conductivity of the devices. One of the seminal achievements that enabled the creation of blue LEDs was the discovery of a post growth process to "activate" the magnesium dopants ($Mg_{(III)}$) in GaN used as the p-type layer of the devices. With these post-growth processes, the resistivity was reduced from several hundreds of Ω cm to single digit values.[11-13] In comparison, however, n-type resistivity obtained by Si dopants in GaN is in the hundredths to thousandths of Ω cm, i.e. three or more orders of magnitude lower.[14] When moving to the wider bandgaps of AlGaN or AlN, the conductivity of the p-type layers further degrades, with a minimum value of 47 Ω cm reported in literature.[15] As a result of these limitations, external quantum efficiencies of ~70% and <10% are reported for white and UV-LEDs, respectively.[1,3]

It is within the context of these two inefficiencies, and their effect on III-nitride LEDs, that this dissertation addresses. This work has been separated into three main chapters presenting the work on these problems of high dislocation densities and poor p-type conductivity. The present chapter includes this motivation, the material properties of the III-nitrides, and a brief introduction to the metal organic chemical vapor deposition (MOCVD) equipment where the growth of the films was performed. The fifth chapter will summarize main accomplishments before providing a list of future experiments. Chapters 2 through 4 present the main experimental results.

Chapter 2, titled “Crystal Growth and Homoepitaxy of GaN,” will introduce the theoretical and experimental results to understand the surface morphology of homoepitaxially grown GaN films. The background thermodynamics of MOCVD growth, nucleation of films, and the Burton, Cabrera, and Frank (BCF) theory of adatom kinetics will be reviewed. A extension to BCF will then be developed, which allows for asymmetry in the attachment of adatoms to the step edges across a surface. Such an asymmetry has been suggested for several surface morphologies to occur, especially step-bunching and step-meandering morphology. These morphologies, along with step-flow and a ‘faceting’ morphology were obtained for GaN homoepitaxy as a result of changing the vapor supersaturation during the growth.

The study of “Mg Doping of GaN and AlGaN” will be presented in Chapter 3. A brief introduction to defect energy of formation will be presented that highlights the importance of the chemical potential and the Fermi level on a defect’s incorporation. The experimental results for nitrogen diluent growth of Mg:GaN will be presented, with a focus on reduction in compensation and passivation through the use of above bandgap illumination during the growth. Then, Mg:GaN growths with increased ammonia (NH_3) flow under both nitrogen and hydrogen diluent conditions will be compared, and the influence of the chemical potential on the defects will be discussed. Lastly, the Mg:$Al_{0.6}Ga_{0.4}N$ study will be presented, with the key observations of reduced hydrogen concentration, and minimized resistivity at higher doping concentrations. Also, mobilities and free hole concentrations will be reported for the first time in Mg:$Al_{0.6}Ga_{0.4}N$.

The final experimental section is the “SIMS of the III-Nitrides” in Chapter 4, which allowed for the characterization of the Mg-doped films in Section III. An overview of the operation of the main secondary ion mass spectrometry (SIMS) instruments will be provided as well as the data acquisition procedure. Following this background, the improvement made to the detection

limit on the ION TOF SIMS V will be introduced along with charge compensation of insulating samples. Lastly, the composition, sputtering rate, and relative sensitivity factor (RSF) corrections for Mg-doping in $Al_xGa_{1-x}N$ ($0 \leq x \leq 1$) will be derived.

2. III-Nitride Material Properties

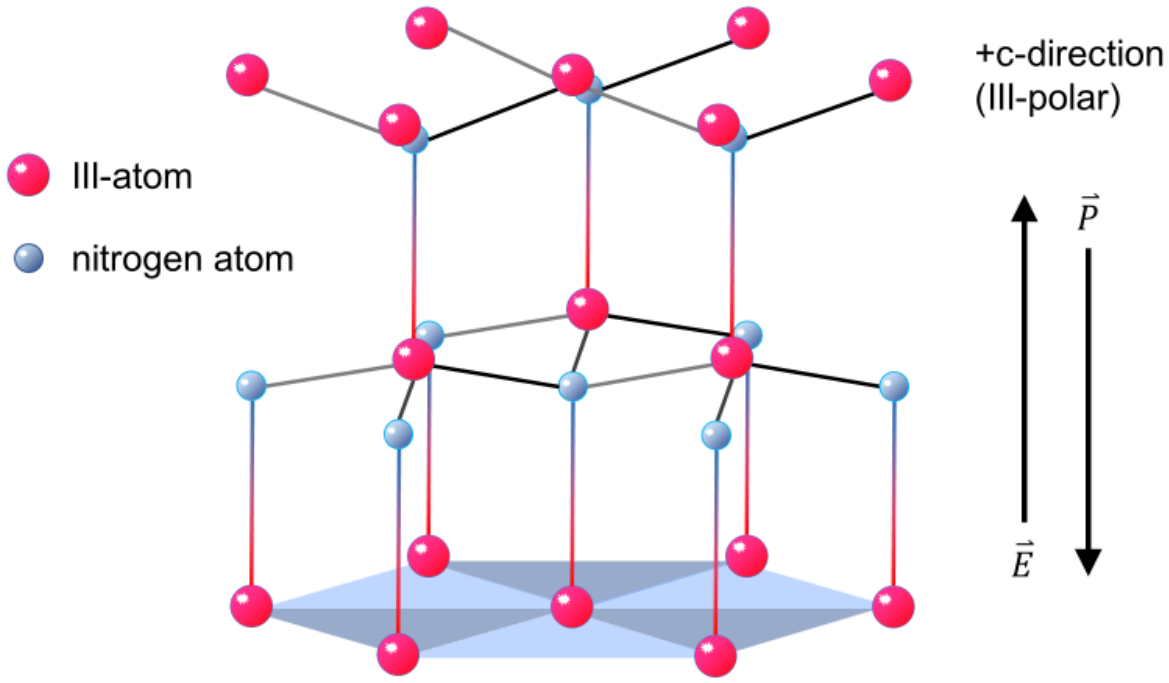


Figure 1.1 III-nitride wurtzite crystal. Direction from III-atom to nitrogen along the c-axis determines polar direction (in this case III-polar). Spontaneous polarization field, $\vec{P}$, points from nitrogen to III-atom. Internal electric field, $\vec{E}$, opposes.

The group III-nitride family of materials includes AlN, GaN, InN and their ternary and quaternary alloys. These semiconductors crystallize in either the zincblende or wurtzite crystal structure, however, wurtzite is the stable polymorph under ambient conditions.[16] The wurtzite crystal structure belongs to the $P6_3mc$ space group and has a hexagonal unit cell comprised of two interpenetrating hexagonal close-packed sublattices. The packing sequence for the structure is AaBbAa… and is shown in Fig. 1.1.[16,17] Each sublattice consists of one atom type, either group

III (metal) or nitrogen, displaced from the other type along the *c*-axis by a fractional coordinate value of 0.375. The basic unit cell for this structure has four atoms, and each is four-coordinated to atoms of the other type forming a tetrahedron. The unit cell can be described by the basal plane lattice parameter, *a*, and the unit cell height, *c*. An additional internal parameter, *u*, is defined as the cation-anion bond length divided by the unit cell height, *c*.[17] For the ideal wurtzite structure, the *c/a* ratio yields a value of 1.633 ($\sqrt{8/3}$) and *u* a value of 0.375.

The bonding between the III-cation and the nitrogen-anion is neither completely ionic nor covalent in nature.[16,17] The degree of ionicity is related to the electronegativities of the atoms and results in the III-nitrides deviating from the ideal wurtzite structure, breaking the centrosymmetry.[18] A useful parameter to understand this behavior is the fractional ionic character:[19]

$$FIC = \frac{|Q_A^* - Q_B^*|}{(Q_A^* + Q_B^*)} \qquad 1.1$$

where Q_A^* is the effective charge for the electron of atom *A*. For a perfectly covalent bond between atoms *A* and *B*, the *FIC* = 0 (*FIC* = 1 is perfectly ionic). The values obtained for AlN, GaN, and InN are 0.72, 0.51, and 0.54, respectively.[19] As shown in Table 1.1, the closer the *FIC* value is to unity, the greater the deviation from the ideal wurtzite lattice parameters. The deviation off the ideal values is derived either by a change in the *c/a* ratio or the *u* parameter.[20] However, the two parameters exhibit a strong correlation, such that if the *c*/*a* ratio decreases, *u* will increase to maintain close to equal tetrahedral distances, i.e. bond lengths, by changing the bond angles α and β (Fig. 1.1).[18] For the bond lengths to be equivalent, the following condition must be met:[20]

$$u = \left(\frac{1}{3}\right)\left(\frac{a^2}{c^2}\right) + \frac{1}{4} \qquad 1.2$$

Because of the degree of ionic behavior between the atoms, and deviation from the ideal wurtzite structure, a dipole moment pointing toward the III-atom forms. The dipole moment is

defined as the partial charge, δ, multiplied by the bond length, $\vec{r}$. A spontaneous polarization in the *c*-direction results, and the density is defined as:[21,22]

$$\overrightarrow{P_0} = n\delta\vec{r} \tag{1.3}$$

where *n* is the density of dipoles. The associated electric field in the opposite direction is:

$$\vec{E} = -\frac{\overrightarrow{P_0}}{\varepsilon} \tag{1.4}$$

where ε is the dielectric constant. With this polarization density there is an associated charge density given by:

$$\rho_s = \nabla P_0 \tag{1.5}$$

Within the bulk of the semiconductor, this net charge is equivalent to zero. However, a disruption in the continuity of the dipole moment, either at the free surface or a heterojunction interface, will create a non-zero value. Devices such as high electron mobility transistors (HEMTs) rely on the discontinuity between a GaN and AlN interface creating a 2-dimensional electron gas (2DEG) of high carrier density.

For the III-nitride wurtzite materials, there can be an additional polarization from the unrelaxed lattice, thermal, and defect strains (from the growth on non-native substrates). This is termed the piezoelectric polarization, P_{zp}, and has a relationship to the strain via multiplication of their respective tensors:[22]

$$\begin{bmatrix} P_{xp} \\ P_{yp} \\ P_{zp} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & e_{15} & 0 \\ 0 & 0 & 0 & e_{24} & 0 & 0 \\ e_{31} & e_{31} & e_{33} & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \varepsilon_{zz} \\ \varepsilon_{yz} \\ \varepsilon_{xz} \\ \varepsilon_{xy} \end{bmatrix} \tag{1.6}$$

From hexagonal symmetry, $e_{15} = e_{24}$ so the above can be reduced:

$$\begin{bmatrix} P_{xp} \\ P_{yp} \\ P_{zp} \end{bmatrix} = \begin{bmatrix} e_{15}\varepsilon_{xz} \\ e_{15}\varepsilon_{yz} \\ e_{31}(\varepsilon_{xx} + \varepsilon_{yy}) + e_{33}\varepsilon_{zz} \end{bmatrix} \quad 1.7$$

If the crystal is strained in the *c*-plane and allowed to expand or contract along the *c*-axis (in the absence of shear), the following relations hold $\varepsilon_{xx} = \varepsilon_{yy} = \frac{a-a_0}{a_0}$ and $\varepsilon_{yz} = \varepsilon_{xz} = \varepsilon_{xy} = 0$. The parameters a and a_0 represent the strained lattice constant of the epi-layer on the equilibrium lattice constant of the substrate or underlying layer. Taking advantage of Poisson's relation, $\frac{\varepsilon_{zz}}{(\varepsilon_{xx}+\varepsilon_{yy})} = -\frac{C_{13}}{C_{33}}$, where C_{13} and C_{33} are stiffness coefficients, eq. 1.7 is reduced to:[22]

$$P_{zp} = 2(\frac{a-a_0}{a_0})(e_{31} - \frac{C_{13}}{C_{33}}e_{33}) \quad 1.8$$

Table 1.1 List of III-nitride material properties.[16-20,23-33]

	AlN	GaN	InN
a (Å)	3.1106	3.1890	3.538
c (Å)	4.9795	5.1864	5.703
c/a	1.6008	1.6263	1.6119
u	0.3821	0.377	0.379
α (°)	108.19	109.17	108.69
β (°)	110.73	109.78	110.24
Q_A^*	1.12	1.98	1.83
Q_B^*	6.88	6.02	6.17
FIC	0.72	0.51	0.54
P_0 (C m^{-2})	-0.09	-0.034	-0.042
e_{31} (C m^{-2})	-0.53	-0.34	-0.41
e_{33} (C m^{-2})	1.5	0.67	0.81
C_{13} (GPa)	94	68	70
C_{33} (GPa)	377	354	205
$E_g(0)$ (eV)	6.1	3.4	0.7
α (meV K^{-1})	1.799	0.909	0.245
β (K)	1462	830	624

Finally, in the wurtzite configuration, the III-nitrides have a direct bandgap which ranges from 0.7 eV for InN, 3.4 eV for GaN, to 6.1 eV for AlN. As a function of temperature, these 0 K values will reduce according to the Varshni equation:[34]

$$E_g(T) = E_g(0) - \alpha T^2/(T+\beta) \quad (1.9)$$

where α and β are the Varshni constants. Thus far, the discussion has been limited to the binary materials within the III-nitrides, and the values for their many characteristics are listed in Table 1.1. Now, the focus will shift toward the ternary and quaternary forms. To understand the change to the material properties, Vegard's law is applied. The two lattice parameters, *a* and *c*, follow a linear extrapolation:[35]

$$a(x) = xa^A + (1-x)a^A \quad (1.10)$$

and

$$c(x) = xc^A + (1-x)c^B \quad (1.11)$$

where superscripts and A and B represent material A and B, and *x* represents the molar fraction. The spontaneous polarization also transforms with the addition of more than one binary material. Eq. 1.3 now takes the form:

$$\overrightarrow{P_0}(x,y) = x\overrightarrow{P_{0,Aln}} + y\overrightarrow{P_{0,InN}} + (1-x-y)\overrightarrow{P_{0,GaN}} \quad (1.12)$$

where the subscript denotes the spontaneous polarization of the constituent binary material. Combining the III-nitrides into ternary alloys leads to a direct bandgap which ranges from 0.7 eV to 6.1 eV, covering the entire spectrum from visible to ultraviolet (UV) wavelengths. The energy bandgap trends as a function of the ternary alloys' composition, *x*, and again follows Vegard's law. Here the bowing parameter, *b*, is included because the linear interpolation is insufficient to capture the curvature of the bandgap energy:[35]

$$E_g(x) = xE_g^A + (1-x)E_g^B - bx(1-x) \quad (1.13)$$

where $E_g{}^A$ and $E_g{}^B$ are the bandgaps of material A and B, respectively. Using bowing parameter values of 1.4 eV for InGaN, 0.7 eV for AlGaN, and 2.5 eV for AlInN, Fig. 1.2 is constructed.[17-20,23,24]

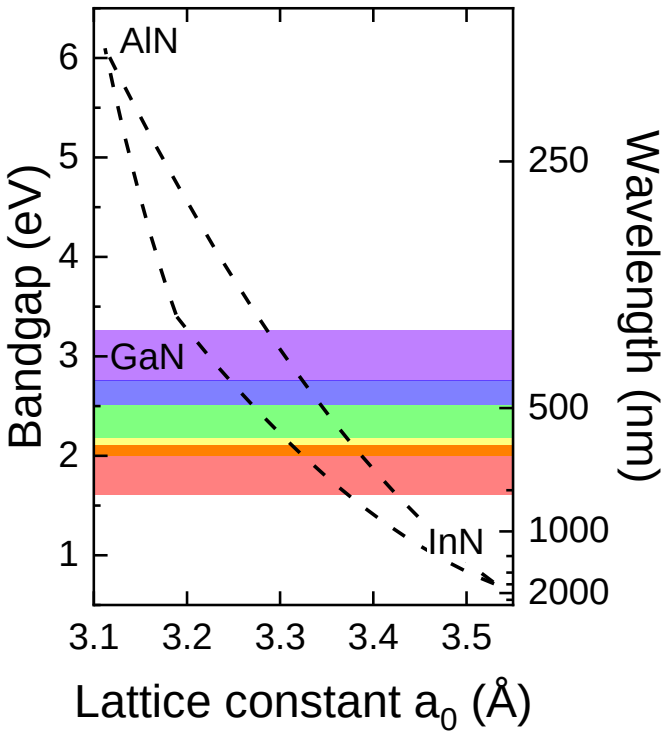


Figure 1.2 Bandgap of III-nitrides as a function of *a* lattice constant. Curves for ternary alloys are constructed using eqs. 1.10, 1.11, 1.13. Visible wavelength ranges are overlain. From[17-20,23,24].

3. Experimental Techniques: Growth

Many different techniques have been utilized to generate III-nitride thin films. Examples include molecular beam epitaxy (MBE), hydride vapor phase epitaxy (HVPE), and metal organic chemical vapor deposition (MOCVD). All of the thin films generated for this dissertation were grown using MOCVD. The following provides a brief introduction to the equipment used to grow the material, as the variables in growths experiments all stem from the growth conditions being supplied by the system.

3.1 MOCVD System

The most common technique used today for the growth of thin film III-nitride devices is via MOCVD. Its ability to achieve sharp layer interfaces for MQWs or growth rates up to several

μm per hour make it the ideal choice for the vast majority of commercialized devices today. All films in this dissertation were grown using this technique, in the reactor shown in Fig. 1.3. This reactor was designed and built in-lab, for the purposes of dedicated research on GaN, AlGaN, and AlN films. This reactor has the distinct advantage that all aspects of a growth run can be controlled by the operator, and are not restricted to pre-prescribed recipes given by the equipment manufacturer. As a result, a greater intuition about the influence of the growth parameters, including temperature, total pressure, and total flow of the various species is obtained. This flexibility played a major role in the investigations in this dissertation.

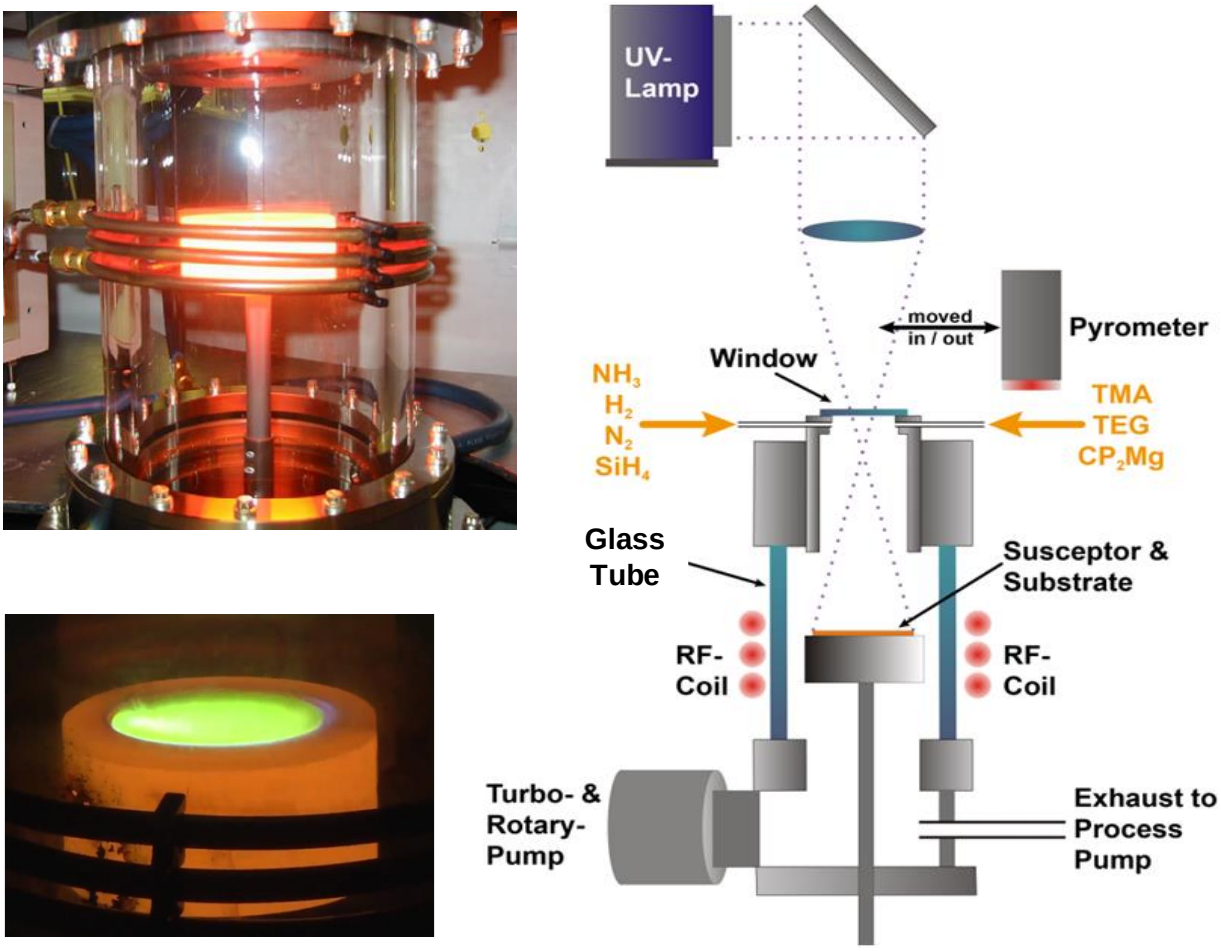


Figure 1.3 Top left: picture of the MOCVD reaction chamber at growth temperature. Bottom left: luminescence of a growing film under UV illumination. Right: schematic of the MOCVD system. From[37].

The schematic in Fig. 1.3 shows the major components in the MOCVD system: metalorganic sources and other gaseous species, the reaction chamber, the UV illumination system, and the exhaust system. Nitrogen (N_2) and hydrogen (H_2) gases serve as the carriers for the system. Ammonia (NH_3) is the source for nitrogen. For metal sources, the metalorganics triethylgallium (TEG), trimethylaluminium (TMA), and bis-cyclopentadienyl magnesium (Cp_2Mg) are used for Ga, Al, and Mg, respectively. These sources are kept at a specific temperature so that a given vapor pressure above is maintained. By flowing carrier gas through the bubbler, a molar flow of these species can be injected into reactor following the bubbler equation:

$$f_{MO} = f_{carrier} \frac{p_{MO}}{p_{bubbler} - p_{MO}} \qquad 1.14$$

where f are the flows in sccm, p the pressures in Torr, and MO representing metalorganic species (TEG, TMA, or Cp_2Mg). As mentioned above, the p_{MO} is vapor pressure of the metalorganic controlled by the temperature of the bubbler. Use of an upstream pressure controller determines the bubbler pressure, $p_{bubbler}$. That and the flow of the carrier gas, $f_{carrier}$, are the primary means for the operator to control the metalorganic flow into the reactor.

The gases are injected in the reaction chamber in a vertical showerhead configuration, which sits 5 cm above the substrate/susceptor in the reaction chamber. The chamber consists of a 13 cm long double quartz tube, with water cooling to prevent any reactions on the walls. The susceptor is a silicon carbide (SiC) coated graphite that is heated by radio frequency (RF) induction which enables temperatures up to ~1250° C. Wafers of 2 inch diameter or less can fit in the pocket of the susceptor, which is rotated by a motor underneath the chamber. The rotation ensures uniform distribution of the gases over the substrate. The pressure inside the growth chamber is controlled by a butterfly valve which is backed by a rotary vane pump to exhaust the gases from the system.

Also attached to the system is a turbomolecular pump, which can achieve base pressures between 10^{-6} to 10^{-7} Torr prior to growth.

The temperature of the substrate is monitored by a pyrometer located above the reactor chamber. A UV-transparent window made of sapphire is affixed atop of the showerhead and provides direct line-of-sight for the pyrometer to the wafer surface. When above bandgap illumination as a defect quasi Fermi level (dQFL) is desired, the pyrometer is removed and light directed onto the substrate. The UV illumination system consists of a mercury arc lamp sitting above the reactor, with a 45° mirror and UV lens to disperse the light uniformly across the substrate surface. For light with energies larger than the bandgap of GaN (~2.9 eV at 1040° C), a power density of ~1 W/cm^2 can be reached on the surface. This value was determined by subtracting the total power density collected at the substrate by the below bandgap power density. Greater details on the procedure are contained in Refs. 36 and 37.

The control of the growth parameters allotted by this instrument enabled the experimental investigations in the following chapters. In the subsequent chapter, the temperature range (<1500° C) and total flow rates for the gases up to 10 slm for the system allowed for a wide range of homoepitaxial growth conditions to be tested. In chapter 3, Mg concentrations ranging from mid-10^{18} cm^{-3} to 1x10^{20} cm^{-3} were easily obtained for GaN and AlGaN films. The overall ease with which to control the growth also allowed for vapor supersaturation calculations of the growth to be made. That key thermodynamic principle is the primary means of comparison within the studies and the derivation will be provided in the following chapter.

4. References

1. S. P. DenBaars, D. Feezell, K. Kelchner, S. Pimputkar, C-C. Pan, C-C. Yen, S. Tanaka, Y. Zhao, N. Pfaff, R. Farrell, M. Iza, S. Keller, U. Mishra, J. S. Speck, S. Nakamura. "Development of gallium-nitride-based light-emitting diodes (LEDs) and laser diodes for energy-efficient lighting and displays," *Acta Materialia* **61**, 945-951 (2013).

2. M. Kneissl. "Chapter 1 A Brief Review of III-Nitride UV Emitter Technologies and Their Applications." In: M. Kneissl, J. Rass, *III-Nitride Ultraviolet Emitters.* Springer Series in Materials Science, vol. 227. Springer, Cham.

3. I. Akasaki, H. Amano, and S. Nakamura. "Nobel Prize in Physics 2014: A new light to illuminate the world," *Press release of the Royal Swedish Academy of Sciences* (2014).

4. H. Hirayama, S. Fujikawa, and N. Kamata. "Recent progress in AlGaN-based deep-UV LEDs," *Electronics and Communications in Japan* **98**, no. 5, 1443-1448 (2015).

5. B. Heying, E. J. Tarsa, C. R. Elsass, P. Fini, S. P. DenBaars, and J. S. Speck. "Dislocation mediated surface morphology of GaN," *Journal of Applied Physics* **85**, 6470 (1999).

6. M. N. Gurusinghe and T. G. Andersson. "Mobility in epitaxial GaN: Limitations of free-electron concentration due to dislocations and compensation," *Physical Review B* **67**, 235208 (2003).

7. B. Sarkar, S. Mita, P. Reddy, A. Klump, F. Kaess, J. Tweedie, I. Bryan, Z. Bryan, R. Kirste, E. Kohn, R. Collazo, and Z. Sitar. "High free carrier concentration in p-GaN grown on AlN substrates," *Applied Physics Letters* **111**, 032109 (2017).

8. S. Usami, R. Miyagoshi, A. Tanaka, K. Nagamatsu, M. Kushimoto, M. Deki, S. Nitta, Y. Honda, and H. Amano. “Effect of dislocations on the growth of p-type GaN and on the characteristics of p-n diodes,” *Physics Status Solidi A* **214**, no. 8, 1600837 (2017).

9. T. Sugahara and S. Sakai. “Role of dislocation in InGaN/GaN quantum wells grown on bulk GaN and sapphire substrates,” *IEICE Transactions on Electronics* **E83C**, 4, 598-604 (2000).

10. J. H. Dycus, S. Washiyama, T. B. Eldred, Y. Guan, R. Kirste, S. Mita, Z. Sitar, R. Collazo, and J. M. LeBeau. “The role of transient surface morphology on composition control in AlGaN layers and wells,” *Applied Physics Letters* **114**, 031602 (2019).

11. H. Amano, M. Kito, K. Hiramatsu, and I. Akasaki. “P-type conduction in Mg-doped GaN treated with low-energy electron beam irradiation (LEEBI),” *Japanese Journal of Applied Physics Part 2-Letters* **28,** L2112 (1989).

12. S. Nakamura, T. Mukai, M. Senoh, and N. Iwasa. “Thermal annealing effects on p-type Mg-doped GaN films,” *Japanese Journal of Applied Physics Part 2-Letters* **31,** L139 (1992).

13. Y. Nakano, O. Fujishima, and T. Kachi. “Effect of p-type activation ambient on accepor levels in Mg-doped GaN,” *Journal of Applied Physics* **96,** 415 (2004).

14. F. Kaess, S. Mita, J. Xie, P. Reddy, A. Klump, L. H. Hernandez-Balderrama, S. Washiyama, A. Franke, R. Kirste, A. Hoffmann, R. Collazo, and Z. Sitar. “Correlation between mobility collapse and carbon impurities in Si-doped GaN grown by low

pressure metalorganic chemical vapor deposition," *Journal of Applied Physics* **120**, 105701 (2016).

15. T. Kinoshita, T. Obata, H. Yanagi, and S. Inoue, "High p-type conduction in high-Al content Mg-doped AlGaN," *Applied Physics Letters* **102**, 012105 (2013).

16. "14 Epitaxial Crystal Growth: Methods and Materials." and "32 Group III Nitrides," in *Springer Handbook of Electronic and Photonic Materials*. S. Kasap and P. Capper (eds.). Springer (2006).

17. H. Morkoç. "1 General properties of nitrides," and "2 Electronic band structure and polarization effects," in *Handbook of Nitride Semiconductors and Devices: Materials Properties, Physics and Growth.* John Wiley & Sons (2008).

18. G. A. Jeffrey, G. S. Parry, and R. L. Mozzi. "Study of the wurtzite-type binary compounds. I.* structures of aluminum nitride and beryllium oxide," *Journal of Chemical Physics* **25**, no. 5, 1024 (1956).

19. Y-N Xu. and W. Y. Ching. "Electronic, optical, and structural properties of some wurtzite crystals," *Physical Review B* **48**, no. 7, 4335 (1993).

20. H. Schulz and K. H. Thiemann. "Crystal structure refinement of AlN and GaN," *Solid State Communications* **23**, 815 (1977).

21. R. Resta. "Macroscopic polarization in crystalline dielectrics: the geometric phase approach," *Reviews in Modern Physics* **66**, no. 3, 899 (1994).

[22.] R. D. King-Smith and D. Vanderbilt. “Theory of polarization of crystalline solids,” *Physical Review B* **47**, no. 3, 1651 (1993).

[23.] Q. Guo and A. Yoshida. “Temperature dependence of band gap change in InN and AlN,” *Japan Journal of Applied Physics* **33**, 2453 (1994).

[24.] F. Yun, M. A. Reshchikov, L. He, T. King, H. Morkoc, S. W. Novak, and L. Wei. “Energy band bowing parameter in $Al_xGa_{1-x}N$ alloys,” *Journal of Applied Physics* **92**, 4837 (2002).

[25.] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan. “Band parameters for III-V compounds and their alloys,” *Journal of Applied Physics* **89**, 5815 (2001).

[26.] A. Zoroddu, F. Bernardini, P. Ruggerone, and V. Fiorentini. “First-principles prediction of structure, energetics, formation enthalpy, elastic constants, polarization, and piezoelectric constants of AlN, GaN, and InN: Comparison of local and gradient-corrected density functional theory,” *Physical Review B* **64**, 045208 (2001).

[27.] M. Leszczynski. “A1.1 Common crystal structures of the group III nitrides,” in *Properties, Processing and Applications of Gallium Nitride Related Semiconductors*. J. H. Edgar, S. Strite, I. Akasaki, H. Amano, and C. Wetzel (eds.). Institution of Engineering and Technology (1999).

[28.] M. Tanaka, S. Nakahata, K. Sogabe, and H. Nakata. “Morphology and x-ray diffraction peak widths of aluminum nitride single crystals prepared by the sublimation method,” *Japanese Journal of Applied Physics* **36**, L1062 (1997).

[29] M. Leszczynski, H. Teisseyre, T. Suski, I Grzegory, M. Bockowski, J. Jun, S. Porowski, K. Pakula, J. M. Baranowski, C. T. Foxon, and T. S. Cheng. "Lattice parameters of gallium nitride," *Applied Physics Letters* **69**, 73 (1996).

[30] W. Paszkowicz, J. Adamczyk, S. Krukowski, M. Leszczynski, S. Porowski, J. A. Sokolowski, M. Michalec, and W. Lasocha. "Lattice parameters, density and thermal expansion of InN microcrystals grown by the reaction of nitrogen plasma with liquid indium," *Philosophical Magazine A* **79**, 1145 (1999).

[31] W. Paszkowicz. "X-ray powder diffraction data for indium nitride," *Powder Diffraction* **14**, 4, 258-260 (1999).

[32] O. Ambacher, J. Majewski, C. Miskys, A. Link, M. Hermann, M. Eickhoff, M. Stutzmann, F. Bernardini, V. Fiorentini, V. Tilak, B. Schaff, and L. F. Eastman. "Pyroelectric properties of Al(In)GaN/GaN hetero- and quantum well structures," *Journal of Physics: Condensed Matter* **14**, 3399 (2002).

[33] V. Yu. Davydov, A. A. Klochikhin, R. P. Seisyan, V. V. Emtsev, S. V. Ivanov, F. Bechstedt, J. Furthmüller, H. Harima, A. V. Mudryi, J. Aderhold, O. Semchinova, and J. Graul. "Absorption and emission of hexagonal InN. evidence of narrow fundamental band gap," *Physica Status Solidi B* **229**, R1-R3 (2002).

[34] Y. P. Varshni. "Temperature dependence of the energy gap in semiconductors," *Physica* **34**, 149 (1967).

[35] L. Vegard. "Die constitution der mischkristalle unde die ramfüllung der atom," *Zeitschrift für Physik* **5**, 17-26 (1921).

[36] Z. Bryan, M. Hoffmann, J. Tweedie, R. Kirste, G. Callsen, I. Bryan, A. Rice, M. Bobea, S. Mita, J. Xie, Z. Sitar, and R. Collazo. “Fermi level control of point defects during growth of Mg-doped GaN,” *Journal of Electronic Materials* **42**, no. 5, 815 (2013).

[37] M. Hoffmann. “Polarity control and doping in aluminum gallium nitride,” *Technische Universität Berlin Dissertation* (2013).

Chapter 2 Crystal Growth and Homoepitaxy of GaN

1. Introduction

Commercially available native GaN substrates offer exciting new potential for the next generation of gallium nitride (GaN) based devices. Previously, GaN epitaxial layers suffered from a high density of dislocations due to growth on foreign substrates. These dislocations resulted in a rough surface morphology, limiting sharp layer transitions, and introducing doping non-uniformities. Growth on native substrates eliminates the lattice mismatch between film and substrate, and ultimately the presence of these dislocations.

However, simply transferring growth conditions to these new substrates does not ensure semiconductor grade surfaces. Typically, the substrates are processed from boules to achieve a misorientation (offcut) that yields a particular morphology of steps across the surface. These serve as sites for attachment of new adatoms to the substrate, and move the growth process forward. Depending on this initial offcut, and the growth parameters such as temperature, pressure, and flows of each of the gases, many different morphologies can evolve from the initial stepped surface. These include step-flow, step-bunching, step-meandering, and 2D nucleation. Of the four listed, only step-flow maintains the initial substrate conditions and ensures a smooth surface. It is vital to determine the window of growth conditions/offcuts that maintain this surface on GaN substrates. As will be discussed in Chapter 3, the growth conditions also influence the incorporation of impurities and dopants. So, a balance between these interests may be required.

To understand the morphological observations in this study, it is important to have a basic understanding of the processes involved in crystal growth. In the simplest sense, growth is a process where single atoms combine on the surface of the substrate. Yet, treating the entire system where 10^{22} atoms cm^{-2} or more are interacting is not feasible. Instead, the approach follows the

kinetic theory of gases, where averages are constructed for a vast number of atoms/molecules based on the behavior of a single atom.[1] The average behaviors correspond macroscopic variables, and these are described by thermodynamics and surface kinetics.[1-3]

First, the thermodynamics which describes the driving force for forming a solid crystal out of gases in the MOCVD process is introduced. Then nucleation theory will be covered, which gives the basis for new crystal phases forming on the surface. The last theoretical section will be utilizing Burton, Cabrera, Frank (BCF) theory of surface kinetics to understand how adatoms move across the steps on the crystal surface. Finally, the experimental results of the GaN homoepitaxy will be presented followed by analysis and conclusions.

2. Thermodynamics of MOCVD Growth

In its most simplified description, MOCVD is a process where a solid is deposited from a vapor phase.[4] To conceptualize this transition, the process is approached from a thermodynamic perspective comparing the distance from equilibrium for the system, i.e. the driving force. For example, the reaction forming GaN from triethylgallium (TEG) gas and NH_3 is:

$$Ga_{(g)} + NH_{3\,(g)} \leftrightarrow GaN_{(s)} + \frac{3}{2} H_{2\,(g)} \qquad 2.1$$

Since pyrolysis of the TEG providing the Ga species is considered 100% at MOCVD temperatures, the organic groups supplying $Ga_{(g)}$ are excluded.[2] N_2 can also be present in this system, however, it does not participate in the reaction because the temperature to break the nitrogen molecule far exceeds the decomposition temperature of GaN.[2] A simple way to understand the spontaneity of the reaction is to compare the standard free energy of formation, ΔG_f^o, for the various species. The free energy of formation is a measure of the free energy change to move from the standard state to a new state at the given temperature and 1 atm.[5] Taking the

difference of the free energies of formation between the products and reactants side of eq. 2.1 gives the free energy of the reaction defined:

$$\Delta G_{reaction} = \sum \Delta G^{o}_{f_{products}} - \sum \Delta G^{o}_{f_{reactants}} \quad 2.2$$

ΔG_f^0 for standard state species, such as $H_{2(g)}$, is zero. Others such as $Ga_{(g)}$ and $NH_{3(g)}$ can be found in the NIST-JANAF tables.[6] $GaN_{(s)}$ is found in the literature.[7] Table 2.1 lists the values and calculated free energy of reaction. It is important to note that the units for the ΔG_f^0 are energy per mole. As is clear from Table 2.1, the $\Delta G_{reaction}$ is negative for all T. One can therefore conclude that the reaction in eq. 2.1 can proceed to the right, and a solid is likely to form. However, this is an estimate of the spontaneity of the process just based on the free energy of formations (at standard pressure).[8] Instead, a formalism based on the experimental input parameters is desired.

Table 2.1 Values for free energy of formation and change in free energy of reaction as a function of temperature (Pressure = 1 atm).[6,7,9]

	ΔG_f^0 (kJ mol^{-1})			$\Delta G_{reaction}$
T (K)	$Ga_{(g)}$	$NH_{3(g)}$	$GaN_{(s)}$	(kJ mol^{-1})
298.15	233.736	-16.367	-98.741	-316.11
300	233.499	-16.183	-98.553	-315.869
400	222.497	-5.941	-86.432	-302.988
500	211.563	4.8	-74.199	-290.562
600	200.626	15.879	-62.002	-278.507
700	189.696	27.19	-49.87	-266.756
800	178.781	38.662	-37.823	-255.266
900	167.888	50.247	-25.864	-243.999
1000	157.023	61.91	-13.997	-232.93
1100	146.189	73.625	-2.217	-222.031
1200	135.385	85.373	9.475	-211.283
1300	124.613	97.141	21.094	-200.66
1400	113.872	108.918	32.624	-190.166

Highlighted qualitatively in Fig. 2.1, there are two methods to proceed from an equilibrium between two phases. The first is the method of supercooling, shown in Fig. 2.1 as ΔT. Under conditions of a constant pressure, the desired phase II, is obtained by decreasing the temperature from the equilibrium. A very obvious example of this process is the formation of ice from water.

Under conditions of atmospheric pressure, as the temperature of the system is reduced from the freezing point it enters a favorable condition for the stability of phase II (ice). As a result, the system can lower its enthalpy by transferring particles from one phase to another.[4]

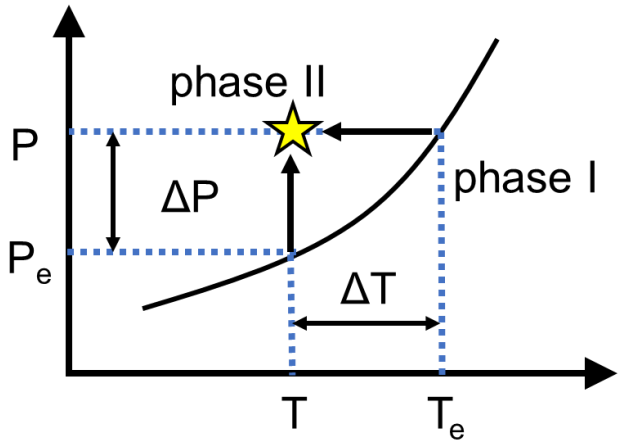


Figure 2.1 Transition from a less stable phase I to stable phase II. Occurs either by supercooling ΔT, at a constant pressure, or by supersaturation ΔP, at constant T. Adopted from[4]

The other method for inducing a phase transformation is by supersaturation, depicted by ΔP in Fig. 2.1.[4] In this process, the temperature is held constant while the pressure of the species is increased from the equilibrium. In a situation between a solid and vapor phase, if the solid is incompressible, the system is changed to a higher pressure by increasing the number of gas phase species (if temperature and volume of the system are constant). To alleviate the excess pressure above equilibrium, it becomes favorable for some of the excess species in the gas phase to form a solid, thereby decreasing the pressure until equilibrium is reestablished. This is the process utilized in MOCVD, except the non-equilibrium state is maintained by continually supplying excess gas species so that the transfer from gas into solid (growth) continues until the supply is removed.

To understand how the change in the process parameters results in a driving force for growth, the model will be derived in detail from a thermodynamic perspective. Thermodynamics

describes a system at equilibrium with its surroundings, and the derivation of the system's properties stems from its internal energy, U:[8]

$$U = TS - PV + \sum_{i=1}^{N_C} \mu_i n_i \quad 2.3$$

where T is the temperature, S entropy, P pressure, V volume, N_C is the total number of components, μ_i the chemical potential of the i^{th} component, and n_i is the number of moles for the i^{th} component. The thermodynamic state function for Gibb's free energy, G, is:[10]

$$G = U + PV - TS \quad 2.4$$

The chemical potential of the i^{th} component, is given by the partial derivative of the Gibb's free energy with respect to the number of moles of the i^{th} component:[4]

$$\mu_i = \left(\frac{\partial}{\partial n_i} G\right)_{T,P,n\neq 1} \quad 2.5$$

The G can be rewritten in terms of the chemical potentials for each component:[4]

$$G = \sum_{i=1}^{N_C} \mu_i n_i \quad 2.6$$

And the total chemical potential is given:[4]

$$\mu = \frac{G}{n} = \sum_{i=1}^{N_C} \mu_i x_i \quad 2.7$$

where n is the total number of moles, and x_i is the mole fraction defined as n_i/n. In addition to defining the system by the number of components, the Gibb's energy can also be defined for the sum of the amount of each component in their respective phases:[4]

$$G = \sum_{j=1}^{N_P} G(\varphi_j) \quad 2.8$$

with φ_j representing the j^{th} phase and N_P the total number of phases.[4]

$$G = \sum_{j=1}^{N_P} \sum_{i=1}^{N_C} \mu_i(\varphi_j) n_i(\varphi_j) \quad 2.9$$

By the definition of equilibrium for the system, where G is at minimum, then:[4]

$$\left(\frac{\partial}{\partial n_i(\varphi_j)} G\right)_{T,P,n(\varphi)\neq 1} = 0 \qquad 2.10$$

The meaning of the above statement is that under the given conditions of T, P, and n_i, the partial derivative of the Gibb's free energy with respect to the i^{th} component in the j^{th} phase is zero, i.e. a minimum. All components in all phases must follow this condition. Now consider a change in component 1 from phase I to phase II (but $n_1(\varphi_I) + n_1(\varphi_{II}) = constant$):[8]

$$\frac{\partial G}{\partial n_1(\varphi_I)} = \frac{\partial\left(G(\varphi_I)+G(\varphi_{II})+\cdots G(\varphi_{Np})\right)}{\partial n_1(\varphi_I)} = \frac{\partial G(\varphi_I)}{\partial n_1(\varphi_I)} + \frac{\partial G(\varphi_{II})}{\partial n_1(\varphi_I)} = \frac{\partial G(\varphi_I)}{\partial n_1(\varphi_I)} - \frac{\partial G(\varphi_{II})}{\partial n_1(\varphi_{II})} = 0 \qquad 2.11$$

where the only surviving terms are the Gibb's energy phase I and II of component 1, and the assumption $\partial n_1(\varphi_{II}) = -\partial n_1(\varphi_I)$. Substituting for the chemical potential on the right side:[8]

$$\mu_1(\varphi_I) - \mu_1(\varphi_{II}) = 0 \qquad 2.12$$

Again, by definition the total system change remains at 0, a minimum, despite a small transfer of component between two phases. To build on this concept, consider an arbitrary isothermal reaction ($\Delta T = 0$) between phases of solid species in their standard state:[8]

$$aA(s) + bB(s) = cC(s) + dD(s);\ \Delta G^\circ{}_1 \qquad 2.13$$

Where the lower case letter represents an integer or fraction of the upper case letter species. In such a situation, each of the solid species is considered at equilibrium with its vapor with standard partial pressure p°. For component A this is written:[8]

$$aA(s) = aA(g, p_A^\circ);\ \Delta G_A = 0 \qquad 2.14$$

Such an equation can be written for each component for a change from solid to gas phase. Eq. 2.13 becomes:[8]

$$aA(g, p_A^\circ) + bB(g, p_B^\circ) = cC(g, p_C^\circ) + dD(g, p_D^\circ);\ \Delta G_1 = \Delta G^\circ{}_1 \qquad 2.15$$

Now, consider changing the pressures of all the species to, p_e, an arbitrary pressure but maintains the equilibrium of the reaction, i.e. no free energy change of the reaction.[8]

$$aA(g, p_{Ae}) + bB(g, p_{Be}) = cC(g, p_{Ce}) + dD(g, p_{De});\ ;\ \Delta G_2 = 0 \quad 2.16$$

To convert from the reaction in 2.15 into 2.16, the partial pressure of each component is changed from $p°$ to p_e. Therefore, the free energy of each species has changed (the change for the total is still 0, eq. 2.12). To describe this change for a species, eq. 2.4 is differentiated:[8]

$$dG = dU + PdV + VdP - TdS - SdT \quad 2.17$$

Making use of the ideal gas law ($PV = nRT$),[8] the combined first and second laws of thermodynamics ($dU = TdS - PdV$),[10] and that temperature is constant; eq. 2.17 is reduced:[8]

$$dG = nRT\ d(\ln P) \quad 2.18$$

For each component n_i it is written:

$$\frac{\partial G}{\partial n_i(g)} = d\mu_i = RT\ d(\ln P)^i \quad 2.19$$

If Integrating from $p°$ to p_e:[8]

$$\mu_i(g, p_{ie}) - \mu_i(g, p_i°) = \Delta\mu_i = RT\ \ln\frac{(p_{ie})^i}{(p_i°)^i} \quad 2.20$$

Therefore, the free energy of reaction at p^e can be described by combining the change in free energy of standard pressures plus the sum of all components' Gibb's free energy change:[8]

$$\Delta G_2 = 0 = \Delta G°_1 + \sum_{i=1}^{N_C} n_i\,\Delta\mu_i = \Delta G°_1 + \sum_{i=1}^{N_C}\Delta G_i \quad 2.21$$

Then:

$$\Delta G_1° = -nRT\ln\left[\frac{(p_{Ce}/p_C°)^c(p_{De}/p_D°)^d}{(p_{Ae}/p_A°)^a(p_{Be}/p_B°)^b}\right] \quad 2.22$$

By definition, the fraction of pressures for each species is defined as the activity, *a*:[2,8]

$$a_i = \frac{p_i}{p_i°} \quad 2.23$$

Also, the equilibrium constant, *K*, is defined:[8]

$$K = \frac{\prod p_{ie}/p_i°(reactants)}{\prod p_{ie}/p_i°(products)} = \frac{\prod a_i(reactants)}{\prod a_i(products)} \quad 2.24$$

For the sake of developing the model, all the components in eq. 2.13 were treated as gases and taken from their standard equilibrium pressure to a new pressure, but one that maintains equilibrium in the reaction. The overall activities of each component depend on the standard state in the reaction. When a gas is the standard state, the standard state is 1 atm pressure. Thus, the activity of the gas is equal to the pressure (p^e) in atmospheres divided by 1 atm. If the standard state is solid or liquid, and it is in its pure form, then the activity is unity because it is in equilibrium with its vapor ($p_i^e / p_i° = 1$).

Now consider a reaction where the species have an arbitrary activity:[8]

$$aA(a_A) + bB(a_B) = cC(a_C) + dD(a_D);\ \Delta G_3 \qquad 2.25$$

To evaluate, begin by considering that each of the species are in their standard state ($a_j° = 1$) and change to a_j. The chemical potential change for each component is:[8]

$$\Delta\mu_i = (\mu_i(a_i) - \mu_i(a°)) = RT\ \ln(a_i)^i \qquad 2.26$$

Following eq. 2.21, the Gibb's free energy of reaction 3 can be written:[8]

$$\Delta G_3 = \Delta G° + \sum_{i=1}^{N_C} \Delta G_i \qquad 2.27$$

Substituting eq. 2.22 for $\Delta G°$ and eq. 2.26:

$$\Delta G_3 = nRT \ln\left[\frac{(a_C/a_C^e)^c (a_D/a_D^e)^d}{(a_A/a_A^e)^a (a_B/a_B^e)^b}\right] \qquad 2.28$$

Now, converting the activities into partial pressures:

$$\Delta G_3 = nRT \ln\left[\frac{(p_C/p_{Ce})^c (p_D/p_{De})^d}{(p_A/p_{Ae})^a (p_B/p_{Be})^b}\right] \qquad 2.29$$

where the standard pressures for each species in the natural logarithm cancel. Eq. 2.29 is a very powerful result. It explains that with knowledge of the equilibrium partial pressure of a species and partial pressure during the reaction, then Gibb's free energy change, i.e. the driving force for a reaction, can be calculated. In many reactions, the equilibrium constant, *K*, has been

calculated and a simplified form for GaN (eq. 2.1), InN, and AlN are provided in literature.[28] From this constant, the equilibrium partial pressure for each of the species can be calculated.

Now we apply eq. 2.29 to the reaction in eq. 2.1:

$$\Delta G_{GaN} = nRT \ln\left[\frac{(a_{GaN})(p_{H2}/p_{H2e})^{3/2}}{(p_{Ga}/p_{Gae})(p_{NH3}/p_{NH3e})}\right] \qquad 2.30$$

Notice that the partial pressures for GaN have be written as the activity on the right side of the equation. This was intentionally selected because in the MOCVD process, it is assumed that the vapor/solid interface is at equilibrium, i.e. $a_{GaN} = 1$.[2] Additionally, Koukitu et al. showed that like other III/V materials the solid phase is only formed when the V/III ratio of the input gas species is greater than unity.[11] In such a setting, it is expected that the III species is depleted, and as a result, the supply of the III species is what controls the growth rate.[2,11]

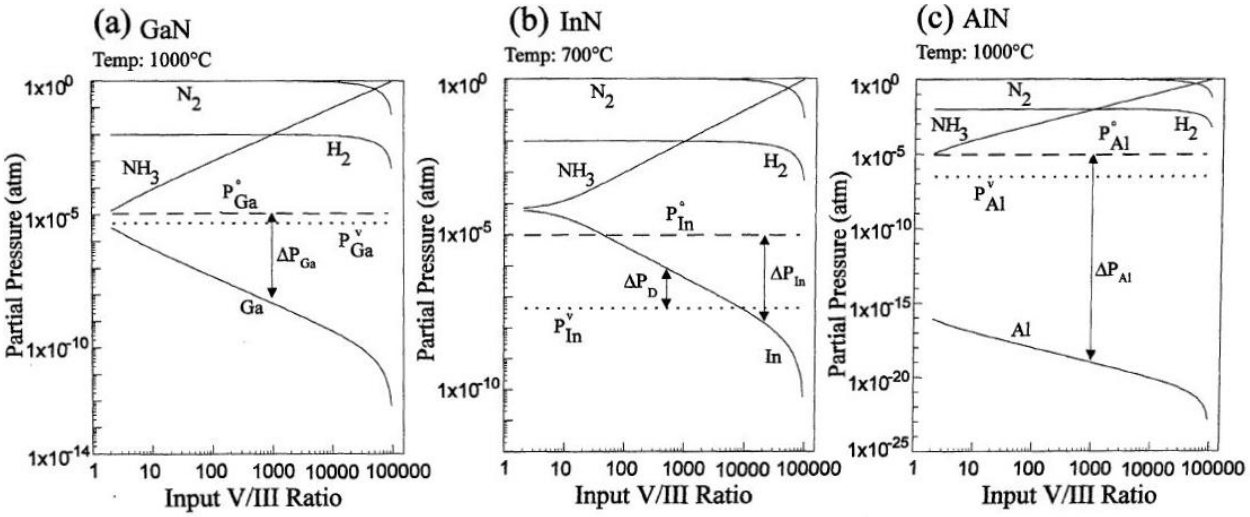


Figure 2.2 Equilibrium partial pressures over GaN, InN, and AlN. The input pressure of the III species is 10^{-5} atm, the total system pressure is 1 atm, the ratio of hydrogen gas to nitrogen gas is 0.01 and the ammonia decomposition parameter is 0 (complete). From[11].

From Fig. 2.2, and eq. 2.30, some major conclusions regarding the driving force for MOCVD growth can be made. First, is that equilibrium pressure for H_2 is invariant to the V/III ratio. The ratio p_{H2}/p_{H2}^{e} is going to be constant based on the condition selected. The equilibrium pressure for NH_3 increases with increasing V/III ratio, i.e. p_{NH3}/p_{NH3}^{e} also remains a constant.

The Gibb's free energy is then driven by the ratio between the III input gas pressure to the equilibrium. For AlN, the partial pressure can be 10 orders of magnitude higher than the equilibrium partial pressure, thus having the largest driving force of all the III-nitrides. For GaN, the deviation is much more modest, 4 orders of magnitude when the V/III ratio is at the typical 2000. InN requires V/III ratios above 10,000 to achieve growth.[11] Therefore, eq. 2.19 can be reduced:[9,12]

$$\Delta G \approx n_{Ga} RT \ln\left[\frac{p_{Gae}}{p_{Ga}}\right] \approx -n_{Ga} RT \ln[1+\sigma] \qquad 2.31$$

σ is the supersaturation, defined as: [9,12]

$$\sigma = \frac{p_{Ga} - p_{Gae}}{p_{Gae}} \qquad 2.32$$

The supersaturation is ultimately the driving force for growth of the solid and what was schematically depicted in Fig. 2.1. Once the pressure of the component in phase I is raised, it becomes favorable to form phase II. If the system were closed, this overpressure would be reduced as the excess in phase I forms more of phase II. Eventually, the pressure would equalize. This is in contrast to MOCVD, where the constant supply of new gas species keeps the system in a state of perpetual non-equilibrium and growth continues until the gas supplies are turned off.

The actual calculation of supersaturation from the determination of equilibrium partial pressure of the III-species is left to Refs. 9 and 12. To summarize these works, the Gibb's phase rule dictates 4 degrees of freedom for describing the system in eq. 2.3.[2,4,6,8-121] The degrees of freedom are the thermodynamic variables that determine the equilibrium state and they are the total pressure, input molar ratio of V to III species, temperature, and the ratio of inert gas such as N_2 to H_2. These are the so-called 'knobs' or parameters that the grower is able to control in the MOCVD process. Systems of equations were derived from thermodynamic constraints including law of mass action and stoichiometry that solved for the equilibrium vapor pressure of the III-

species.[9,11,12] From here the supersaturation is estimated, and link between experimental inputs and theoretical predictions is established. Such a tool cannot be underestimated; MOCVD growth does not need to be treated as a 'black box' where growth conditions are arbitrarily selected.[2]

3. Nucleation Theory

For now, a description for the driving force for growth of a crystal has been explained, but the mechanism for the incorporation of the atoms onto a native (homoepitaxy) or non-native (heteroepitaxy) substrate has not been addressed. The seminal understanding of how crystal growth occurs was put forward by Burton, Cabrera, and Frank in 1951, and for which BCF theory received its name.[13] In that work they described the process for growth to occur in the presence of a stepped surface, which explained the major discrepancy between the supersaturations predicted by nucleation theory and what were actually being employed for crystal growth. Since the study of GaN homoepitaxy and the resulting surface morphologies are a major topic in this thesis, a detailed outline of nucleation and BCF theory will be provided.

The beginnings of nucleation theory start with the formation of three dimensional nuclei from a parent phase. This is related to the previous discussion on how a pressure or temperature change can result in a solid phase being favorable (Fig. 2.1). However, if the initial starting conditions were homogenous in say the vapor phase, then a nucleus does not spontaneously form as soon as the system is forced from equilibrium.[4] Prior to the formation of this nucleus, an interface between the no-longer stable vapor phase and the new solid phase must form. The creation of which requires a certain amount of energy related to the surface energy, γ, of each phase. The surface energy can be thought of as the force acting opposite any force attempting to stretch the object and increase its area. Therefore, a certain level of supersaturation, or Gibb's free

energy change from the equilibrium via pressure in a system, must be reached before it is statistically favorable for a new solid nucleus to form.[4]

The formation of the nucleus within the parent phase has an associated Gibb's free energy change, ΔG_N. There are three contributions. The first is the reduction in free energy by some amount of material moving from the less stable to more stable phase, ΔG_V; which is proportional to the volume of the newly formed nucleus. The second is the energy to create the new interface, ΔG_S, proportional to the surface area multiplied by γ. Finally, there is an energy term related to the stress the nucleus is subjected to, ΔG_E. The total Gibb's free energy change is then:[4]

$$\Delta G_N = \Delta G_V + \Delta G_S + \Delta G_E \qquad 2.33$$

In the first case, consider the nucleus as a sphere suspended in the vapor with no strain. Eq. 2.33 becomes:

$$\Delta G_N = \frac{4}{3}\pi r^3 \frac{\Delta \mu}{\Omega} + 4\pi r^2 \gamma \qquad 2.34$$

where $\Delta\mu$ is the chemical potential, and Ω is the molar volume defined as the molecular weight divided by the density ($\Omega = M/\rho$). The first term represents the number of moles of the component formed by the new nucleus multiplied by the energy of the system removed from equilibrium. As already described in the previous section, for the solid to form from a reaction like eq. 2.3, the $\Delta\mu$ is negative. The second term is the surface area of the new sphere multiplied by the surface energy, which is always positive for a solid phase.[4] Eq. 2.34 is a competition between these two terms. As the nucleus size begins increasing, the Gibb's free energy overall increases. At this stage, the solid nucleus has not formed and can instead be thought of as a cluster of atoms of vapor. At a certain critical radius of this cluster, r^*, the total Gibb's free energy change is no longer increasing. With any additional atoms attaching to the cluster, the free energy change begins

decreasing, and the solid nucleus forms. This crossover point in the size of the cluster is described as the critical radius:

$$\frac{d}{dr}(\Delta G_N) = 0 \rightarrow r^* = -\frac{2\gamma\Omega}{\Delta\mu} \quad 2.35$$

Therefore, the critical Gibb's free energy change for nucleation:[4]

$$\Delta G_N(r^*) = \Delta G_N^* = \frac{16\pi\gamma^3\Omega^2}{3(\Delta\mu)^2} \quad 2.36$$

Eq. 2.36 describes the minimum change in ΔG for a nucleus of a solid phase to form. Since γ is property of the material, the only parameter that can be controlled by the grower in the above equation is the $\Delta\mu$ term. From eq. 2.32, it is the supersaturation or difference between the input and equilibrium partial pressures that determines how large the magnitude of ΔG becomes. Therein the grower, by careful selection of growth parameters in the MOCVD equipment, can promote a nucleus to form from the vapor.

However, this model is not necessarily an accurate depiction of the growth process because it does not account for the substrate on which the nucleus forms. Because of this solid crystal substrate, itself in a stable phase, then any nuclei forming on it has a smaller interface with the less stable vapor phase.[4] Again consider the system strain free, then eq. 2.33 balanced between the energy reduction from the volume of the nucleus created vs. the energy in creating three interfaces: nucleus-vapor (*nv*), substrate-vapor (*sv)*, and the nucleus-substrate (*ns*). Eq. 2.33 becomes:[4]

$$\Delta G_N = a_n r^3 \Delta G + a_1 r^2 \gamma_{nv} + a_2 r^2 \gamma_{ns} - a_2 r^2 \gamma_{sv} \quad 2.37$$

The 'a' variables in eq. 2.26 represent geometric factors because the nucleus is no longer spherical. Fig. 2.3 is representation of a nucleus forming on a substrate. Note, the image depicts the nucleus as a function of θ, the contact angle between the nucleus and substrate. The contact angle was first studied by Thomas Young, and has a relationship to line tension between the three interfaces given by:[4]

$$\gamma_{sv} = \gamma_{ns} + \gamma_{nv}\cos\theta \qquad 2.38$$

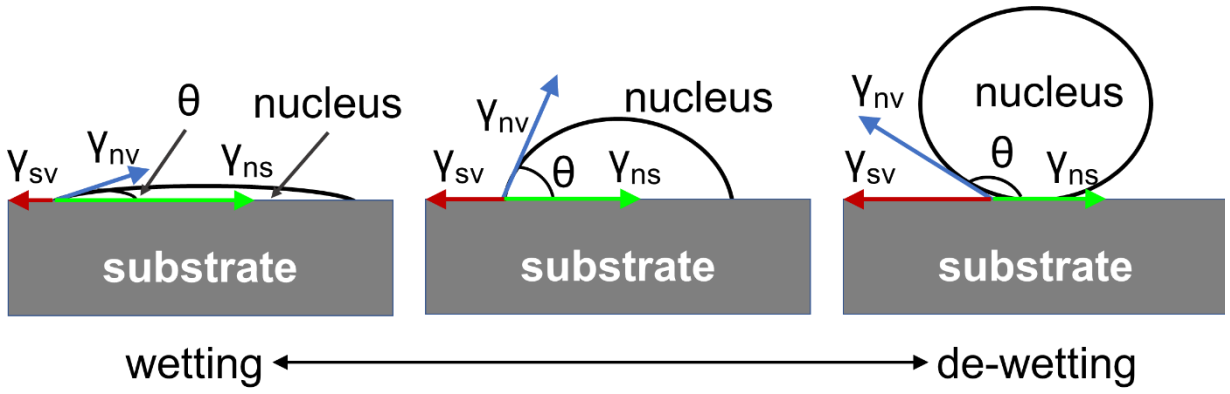


Figure 2.3 Wetting between solid nucleus and solid substrate. Left to right, increasing contact angle which decreases the wetting. Lines represent the direction of the surface energy but not the magnitude.

The three images in Fig. 2.3 represent three cases of the contact angle. Note the length of the arrow was drawn for clarity and does represent a magnitude satisfying eq. 2.38. Moving from a contact angle of 0° to 180° can also be thought of in terms of 'wetting' between the nucleus and substrate. At $\theta \sim 0°$ there is almost complete wetting between the nucleus and substrate, i.e. the nucleus does not 'bead' on the substrate. In the middle figure, the contact angle is approaching 90° and almost half a spherical nucleus forms on the substrate. Finally, the right image depicts when θ approaches 180° and almost complete de-wetting between the nucleus and substrate occurs. The mathematical representation of Fig. 2.3 is expressed by considering the volume of a spherical cap, the surface area of the spherical cap, and the area of the interface between the substrate and nucleus. In the absence of strain, eq. 2.37 becomes:

$$\Delta G_{N,hetero} = \frac{\pi}{3} r^3 (1-\cos\theta)^2 (2+\cos\theta)\frac{\Delta\mu}{\Omega} + 2\pi(1-\cos\theta) r^2 \gamma_{nv} + \pi r^2 \sin^2(\theta)\,(\gamma_{ns} - \gamma_{sv}) \qquad 2.39$$

Substituting in eq. 2.38 and using the identity of $\sin^2(\theta)$:

$$\Delta G_{N,hetero} = \frac{\pi}{3} r^3 (2 - 3\cos\theta + \cos^3\theta) \frac{\Delta\mu}{\Omega} + \pi r^2 \gamma_{nv} (2 - 3\cos\theta + \cos^3\theta)$$

$$\Delta G_{N,hetero} = \left(\frac{4\pi}{3} r^3 \frac{\Delta\mu}{\Omega} + 4\pi r^2 \gamma_{nv}\right) \frac{(2 - 3\cos\theta + \cos^3\theta)}{4}$$

$$\Delta G_{N,hetero} = \Delta G_{N,homo} \times f \qquad 2.40$$

Where f is the geometric factor equal to $(2 - 3\cos\theta + \cos^3\theta)/4$. To find r^*, again differentiate and solve at zero. From the above it is clear that the critical radius does not change between hetero- and homogenous nucleation, because the f factor is independent of *r*. Additionally, the critical free energy for heterogeneous nucleation will also be equal to the homogeneous case multiplied by f. It is important to discuss the implications of the f factor which is determined by the contact angle. When $\theta \sim 0°$ the f factor approaches 0. As a result, the barrier for heterogenous nucleation can be almost 0, significantly lower than the homogenous case. For the $\theta \sim 90°$ case, the heterogenous nucleation is half the barrier to homogeneous. Only when the $\theta \sim 180°$ does the f factor equal unity, and the barrier for heterogenous nucleation is equivalent to the homogenous case.

In both of the prior examples, the effect of strain was neglected to simplify the model. However, in reality strain plays a major role in the film growth, especially in heteroepitaxy where the nucleus and substrate are different materials. As a result, they often have differing lattices and lattice parameters. The nucleus will orient itself in such a way to minimize the energy of the system, where difference between the planes in lowest, but some residual mis-alignment can remain.[14] The nucleus is thus strained and has a strain energy ΔG_E (eq. 2.33). The derivation of this energy term is beyond the scope of this work, but it is shown that it is proportional to the square of the strain, in a form similar to Hooke's law.[15,16] It is always positive and therefore opposes the supersaturation $\Delta\mu$. With this inclusion, the critical barrier for heterogeneous nucleation is described:[4]

$$\Delta G^*_{N,hetero} = \frac{16\pi\gamma^3\Omega^2}{3(\Delta\mu+\Delta G_E/n)^2} \times f \quad 2.41$$

It is clear that in eq. 2.41, if the strain energy exceeds the supersaturation of the growth (through the change in chemical potential), then $\Delta G^*_{N,hetero}$ remains positive and there is no driving force for nucleation. Only when $\Delta\mu$ is extremely high can the strain be neglected and nucleation occur. In fact, these high supersaturation conditions are utilized to form low-temperature AlN buffer layer, which was a key technology in enabling GaN heteroepitaxy on sapphire.[17] Such a buffer layer was used extensively for heteroepitaxial growth in this dissertation.

Now, the formation of two-dimensional nuclei, such as in the case of homoepitaxial growth on flat native surfaces, will be presented. The earlier discussion proved that heterogeneous nucleation on a non-native surface could reduce the energy of formation depending on the wetting of the substrate. For growth on a native substrate, the wetting is zero, yet without a site for the adatom to attach to the bulk, there is an energy barrier to overcome before adatoms form new nuclei on the surface. The primary method of growth is the formation of 2D disks of a monolayer height. These provide sites on the sides of the disk for atoms to attach, and in the process of increasing their diameter, become a wide enough area for new disks to nucleate. Classically, this is called the Frank-van der Merwe or 2D layer-by-layer growth mode. Mathematically, the 2D disk is described by the volumetric decrease by forming the island and the formation of a new interface between the ledge of the disk and the vapor phase.[4] The Gibb's free energy change for the disk is:[18]

$$\Delta G_{2D} = \pi h r^2 \frac{\Delta\mu}{\Omega} + 2\pi r h \gamma_{step} \quad 2.42$$

where h is the height of the island, and γ_{step} is the free step energy.[4,18] The critical radius for the disk is again where the derivative of the change in Gibb's free energy is equal to 0:[18]

$$r_{2D}^{*} = -\frac{\Omega\gamma_{step}}{\Delta\mu} \quad 2.43$$

Then the critical barrier is defined:[18]

$$\Delta G_{2D}^{*} = -\frac{\pi h\Omega\gamma_{step}^{2}}{\Delta\mu} \quad 2.44$$

and the critical number of moles:

$$n_{2D}^{*} = \frac{\pi h(r_{2D}^{*})^{2}}{\Omega} = \frac{\pi h\Omega\gamma_{step}^{2}}{\Delta\mu^{2}} \quad 2.45$$

In homoepitaxial growth on flat surfaces devoid of kink sites, growth occurs by clustering of adatoms on the surface until a size is reached that when an additional adatom attaches, they create a solid nucleus. The source of the adatoms on the surface is derived from the vapor supersaturation of the overall MOCVD process. Yet, eq. 2.44 cannot describe the rate of this process, only when the probability is high enough for the disk to form. To describe the behavior of the adatoms on the surface, one has to think of the kinetics on the sample surface and that will be discussed in the following section. However, the derivation to eq. 2.44 is still important because at large enough supersaturations (and wide enough terraces), these 2D disks can form on the surface and simulate the perfectly flat crystal.

4. BCF Surface Kinetics Model

In the previous section, a model of growth was presented for the case of an infinitely flat surface with no attachment sites for adatoms on the surface. This would limit the formation of new layers on a crystal surface to formation of nuclei that would grow and expand across the surface. However, many researchers observed growth rates that far exceeded expectations based on the supersaturation values predicted from the growth parameters. In 1951, Burton, Cabrera, and Frank built on the work of Frenkel that described surfaces not as flat, but having many steps with kink sites which provide locations for the adatoms to attach along a step front (Fig. 2.4).[13] The source

of the steps could be either an inherent roughness of the sample surface or by screw dislocations having a burgers vector pointing in the growth direction. Steps also arise when substrates are intentionally cut with a misorientation, α, that is related to the average step heights h, and terrace widths d, by $\tan\alpha = h/d$. The kinks in the steps provide a lower energy position for the adatoms on the surface. Thus, adatoms have a propensity to attach at these sites well before forming nuclei.

Recent work on the homoepitaxy of AlN with MOCVD has demonstrated three distinct surface morphologies that result on a vicinal surface.[18,19] These are 2D nucleation on the terraces, step-flow growth, and step-bunching/macro step growth. In the first of these works,[18] a model was developed describing the transition between 2D nucleation and step-flow growth modes by linking the surface supersaturation, σ_s, which describes the relative difference between the adatom concentration on the surface and the equilibrium adatom concentration, to the critical supersaturation of 2D nucleation (eq. 2.44 in the previous section). In this model, the steps were considered to be perfect sinks of the adatoms and their incorporation at either step edge of the terrace was considered equivalent. At a combination of starting terrace width (from the vicinal offcut) and growth supersaturation condition, the authors were able to vary the growth mode between 2D nuclei on the terraces and step-flow. By either increasing the vapor supersaturation, σ, at a given terrace width, or increasing the terrace width at a given supersaturation, the surface supersaturation was able to exceed the critical value for 2D nucleation to occur.[18]

One of the limitations to this work was the lack of explanation for the transition between the step-flow growth mode to the step-bunching at narrower terraces or lower growth supersaturations.[18] To overcome this limitation, Ref. 38 invoked the idea of the Ehrlich-Schwoebel barrier (ESB), or an additional energy barrier for the adatom to attach to the step from an upper terrace to the lower step edge.[19,20] Qualitatively, the barrier is thought to arise from a simple bond-

breaking argument: an adatom on the upper surface will have an additional kinetic barrier energy to step-down vs. an adatom arriving along the surface. However, Ref. 19 did not invoke the normal application of the ESB and instead argued for an "inverse" ESB, where the barrier from the upper terrace is lower than the adatom arriving from the lower terrace. In addition, they also included an extra barrier to the adatom from the lower terrace, a so-called incorporation barrier.[21] As practically demonstrated in Ref. 21, step-bunching can occur in the model from either an inverse ESB or a positive barrier.

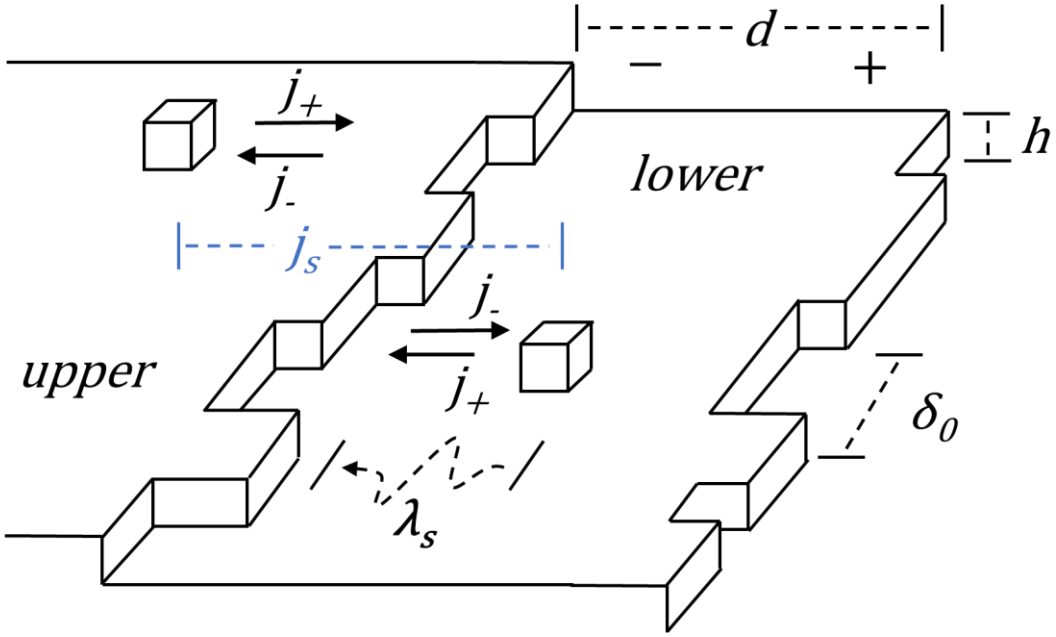


Figure 2.4 Schematic of the steps on a vicinal surface. Steps are considered to have many kinks for adatoms to attach, with an average kink spacing of δ_0. Adatoms are depicted as cubes on the surface. At a step edge, there is a flux of adatoms from the step edge (j_-) and toward the step edge (j_+) for both the upper and lower terraces. The total flux toward the step edge, j_s, is the sum of all these individual fluxes. The diffusion length, λ_s, is depicted as the mean length the adatom diffuses on the terrace before attaching to a kink or being desorbed. Since this is typically longer than the average kink spacing, δ_0, the problem is reduced to a one dimensional form (Fig. 2.5). The terrace width is d, the center of the terrace is $x = 0$, the upper step is $x = \frac{-d}{2}$, and the lower step is $x = \frac{d}{2}$.

In addition to the evolution of step-bunching on the surface, GaN homoepitaxial growth exhibits a fourth type of morphology – step meandering. This was not observed for the AlN

homoepitaxy,[18,19] qualitatively suggested because the growth supersaturation for Al is orders of magnitude greater than Ga. Thus, the formation of 2D nuclei is achievable for typical offcuts and growth conditions of AlN. GaN may need a truly flat surface devoid of any screw dislocations to observe these 2D islands. Instead of following AlN, it seems more suitable to follow modeling of GaAs, where both step-bunching and step meandering were observed.[22,23] From the work of Bales and Zangwill, it was predicted that any non-uniformity along the growth front could be made unstable in the presence of the positive ESB, i.e. the supply of adatoms from the lower terrace was greater than the upper terrace.[23] This would lead to the step-meandering described for both GaN and GaAs.

Here a BCF approach will be applied to create a model for the adatom behavior on the surface. There are several kinetic processes on the surface that need to be considered (Fig 2.4). As already discussed in Chapter 2, the driving force for growth is the input partial pressure difference with the equilibrium partial pressure. Under conditions of mass transport limited growth, the majority of the chemical potential drop between the gas phase and the surface takes place in the boundary layer, i.e. any local non-equilibrium at the solid interface is quickly restored to equilibrium.[2] That local non-equilibrium is a fluctuation of the adatom concentration above an equilibrium concentration derived from the excess vapor pressure. There are three methods of minimizing the excess adatoms: they (1) diffuse across the terrace and incorporate at the step edge, (2) conglomerate with other adatoms to form a 2D nucleus, or (3) desorb back into the gas phase. Note, (1) and (2) are growth processes while (3) is subtracted against the incoming flux. First, consider an infinitely flat surface under equilibrium conditions. The desorbing flux N_s/τ_s is equal to the impingement rate atoms adsorbing to the surface $p_e/\sqrt{2\pi m k_B T}$, where N_s is the adatom

concentration on the surface and τ_s is the mean residence time. In the case of equilibrium, N_s is the equilibrium concentration and is defined:[24]

$$N_{se} = \frac{p_e}{\sqrt{2\pi m k_B T}} \tau_s \quad 2.47$$

and for any pressure deviation above equilibrium:[24]

$$N_s = \frac{p}{\sqrt{2\pi m k_B T}} \tau_s = \frac{p}{p_e} N_{se} \quad 2.48$$

The mean residence time is the average time an adatom remains on the surface before it is desorbed back into the vapor. It is given by:[13]

$$\tau_s = \frac{1}{\vartheta} \exp\left(\frac{E_{des}}{k_B T}\right) \quad 2.49$$

where ϑ is the atomic vibration frequency and $E_{des}$is the energy for desorption. The mean residence time is linked to two other parameters that represent the movement of the adatoms on the surface. The first parameter is the diffusion coefficient, D_s:[13]

$$D_s = a^2 \vartheta \exp\left(-\frac{E_{sd}}{k_B T}\right) \quad 2.50$$

with a being the lattice constant and E_{sd} the diffusion barrier energy. This term represents per atomic vibration time, the lattice area that can be traversed. Multiplying the mean residence time by the diffusion coefficient gives Einstein's relation for the diffusion length, λ_s:[13]

$$\lambda_s^2 = D_s \tau_s \quad 2.51$$

The diffusion length is the average length an adatom travels during their lifetime on the surface. Substituting eqs. 2.49 and 2.50 into eq. 2.51:[13]

$$\lambda_s = \sqrt{D_s \tau_s} = a \exp\left(\frac{E_{des} - E_{sd}}{2 k_B T}\right) \quad 2.52$$

Depending on the sign of the numerator in the exponential in eq. 2.52, the diffusion length has very different properties with respect to temperature. In the case where $-E_{sd} > E_{des}$, the diffusion length will increase as the temperature is increased. An example was provided in Ref. 18

where increased temperature changed the morphology from 2D layer-by-layer into a step-flow growth mode because the diffusion length of the Al adatoms increased. In the other case when $-E_{sd} < E_{des}$, the diffusion length decreases when the temperature is increased. While the overall driving force is the concentration of the adatoms, if they do not have sufficient diffusion length in comparison to the terrace width and the capture lengths of the steps, then they cannot attach to the step ledge. In this situation, growth is reduced to the formation of nuclei on the terraces, unless the adatoms desorb from the surface. For AlN, desorption was minimal.[18] Thus, the real driver for the diffusion length was energy barrier for diffusion and by extension the diffusion coefficient.[19]

Now the model will be considered where a step is bounded by an upper (left) and lower (right) terrace (Fig. 2.4). This will be the equivalent to model described in Ref. 18 and assume a symmetric incorporation of the adatoms to the step edges.[18] One difference here is that the steps will not be treated as perfect sinks for the adatoms, so the adatom concentration in the near vicinity of the step is not considered as the equilibrium adatom concentration. Ultimately, eqs. 2.67 and 2.70 presented here match the model in Ref. 18 when the kinetic coefficient is treated as unity.

The step is not considered perfectly straight and is thought to have many kinks where adatoms can attach. An adatom approaching from the upper terrace is considered to be moving toward the 'descending' step. Meanwhile, an adatom approaching from the lower terrace is moving toward the 'ascending' step. Near the step, the adatom flux toward the kink site is made up of four separate fluxes: two fluxes, $j_{+,up}$ and $j_{+,low}$, of adatoms approaching the kink representing the upper and lower terraces and a two fluxes of adatoms, $j_{-,up}$ and $j_{-,low}$, off of the kink site to the respective terraces. They are given:[24]

$$j_{+,up/low} = a^2 \vartheta N_{st,up/low} \exp -\frac{E_{kin}}{k_B T} \qquad 2.52$$

$$j_{-,up/low} = a^2 \vartheta N_{se} \exp -\frac{E_{kin}}{k_B T} \quad 2.53$$

Where $N_{st,up/low}$ is the concentration of adatoms immediately near the kink site on either the upper or lower terrace E_{kin} is the kinetic energy barrier for the adatom to attach to the kink. The advance of a step is given by:[24]

$$v_{st} = a\frac{a}{\delta_0}\,(j_+ - j_-)\Big|_{upper} + a\frac{a}{\delta_0}(j_+ - j_-)\Big|_{lower} \quad 2.54$$

The ratio a/δ_0 represents the probability of finding a kink site, with δ_0 being the kink spacing. The kinetic coefficient of the step, β_{st}, can be defined:[24]

$$\beta_{st} = a\vartheta\frac{a}{\delta_0}\exp -\frac{E_{kin}}{k_B T} \quad 2.55$$

Combining eqs. 2.52, 2.53, and 2.55 into 2.54:

$$v_{st} = a^2\beta_{st}\,(N_{st} - N_{se})|_{up} + a^2\beta_{st}\,(N_{st} - N_{se})|_{low} \quad 2.56$$

To develop the model including the ESB and incorporation barrier the upper lower terraces are treated separately as written above. In addition, an extra barrier energy will be added to the exponential in eq. 2.55 to capture any asymmetry for these special cases. That will be dealt with later in the section. If the capture rates for the step of adatoms from the upper and lower terraces are equivalent, then eq. 2.56 reduces to.[24]

$$v_{st} = 2a^2\beta_{st}\,(N_{st} - N_{se}) \quad 2.57$$

An important note about the equation is that N_{st} is the adatom concentration in the near vicinity of the step. The great majority of the adatoms arriving from vapor phase due to the supersaturation during MOCVD lie elsewhere on the terraces, and must travel to the kinks on the step to be incorporated (or exceed the concentration to form a 2D nuclei). The rate of diffusion of the adatoms is given by the ratio of the diffusion length with the residence time:

$$\frac{\lambda_s}{\tau_s} = a\vartheta \exp(-\frac{E_{des}+E_{sd}}{2k_BT}) \quad 2.58$$

A direct comparison between eqs. 2.55 and 2.57 reveals the limitation for the advance of the step. When $\lambda_s/\tau_s \ll \beta_{step}$, the rate the step moves is considered diffusion limited, i.e. the adatoms moving across the terrace to the step are the limitation to step velocity. In the opposite case $\lambda_s/\tau_s \gg \beta_{step}$, sufficient adatoms are near the kink and instead it is the capture rate of the step limiting the step advance.[24] These cases will help define the surface morphology in the discussion going forward. Returning to the question of N_{st}, a driving force for the adatoms across the surface can be defined similar to the discussion of the thermodynamic driving force for growth. This driving force, or surface supersaturation σ_s, is related to the adatom concentration on the surface:[13,18,24]

$$\sigma_s = \frac{N_s - N_{se}}{N_{se}} \quad 2.59$$

Whereas the vapor supersaturation, σ, is constant over the entire surface, this surface supersaturation has a dependence on the x distance from the step edge. The flux of adatoms across the surface is given by Fick's first law:[13,24]

$$j_s = -D_s \frac{dN_s}{dx} = D_s N_{se} \frac{d\Psi}{dx} \quad 2.60$$

with D_s as the surface diffusion coefficient from Einstein's relation and $\Psi = \sigma - \sigma_s$.[13] The net flux of adatoms to the surface is given by the impingement rate of the gas subtracted from the desorption:[13,20]

$$j_V = \frac{p}{\sqrt{2\pi m k_B T}} - \frac{N_s}{\tau_s} = \left(\frac{p}{p_e} - \frac{N_s}{N_{se}}\right)\frac{N_{se}}{\tau_s} = \frac{N_{se}}{\tau_s}\Psi \quad 2.61$$

By continuity, the net flux of atoms in the vertical direction at a position x is equivalent to the change in horizontal flux in an infinitesimally small section of the surface dx. Therefore:

$$\frac{d}{dx} j_s = j_v \rightarrow \lambda_s^2 \frac{d^2}{dx^2} \Psi = \Psi \quad 2.62$$

The general solution to the above equation is:[18]

$$\Psi = A \exp\left(\frac{x}{\lambda_s}\right) + B \exp\left(\frac{-x}{\lambda_s}\right) \quad 2.63$$

The coefficients A and B in eq. 2.63 can be solved by the boundary conditions selected. For this solution two assumptions will be made. First, is that the step velocity is significantly lower than the adatom diffusion across the surface (the ratio of eq. 2.57 to 2.58 $\ll 1$). The second is that at the step edge $x = \pm d/2$, $\Psi = \sigma - \sigma_{st} = \chi\sigma$. where:[24]

$$\chi = \frac{\sigma - \sigma_{st}}{\sigma} \quad 2.64$$

and

$$\sigma_{st} = \sigma(1 - \chi) \quad 2.65$$

Therefore, eq. 2.57 can be rewritten:

$$v_{st} = 2a^2 \beta_{st} N_{se} \sigma(1 - \chi) \quad 2.66$$

Returning to eq. 2.63, the solution is given by the boundary conditions $\Psi = \chi\sigma$ at $x = \pm d/2$:[43]

$$\Psi = \chi\sigma \frac{\cosh\left(\frac{x}{\lambda_s}\right)}{\cosh\left(\frac{d}{2\lambda_s}\right)} \quad 2.67$$

To solve for χ, recall eq. 2.60 and that the step velocity is also given by $v_{st} = 2a^2 j_s$ at $x = \pm d/2$. The new v_{st} is given:

$$v_{st} = 2a^2 \frac{\lambda_s}{\tau_s} N_{se} \chi\sigma \tanh\left(\frac{d}{2\lambda_s}\right) \quad 2.68$$

Equating 2.66 and 2.68 results in:

$$\chi = \left(1 + \frac{\lambda_s}{\beta_{st}\tau_s} \tanh\left(\frac{d}{2\lambda_s}\right)\right)^{-1} \quad 2.69$$

Depending on the comparison between. λ_s/τ_s and β_{step}, the χ and ultimately v_{st} have different values. In general, the step velocity is given by combining eq. 2.68 and 2.69:[24]

$$v_{st} = 2a^2 \frac{\lambda_s}{\tau_s} N_{se}\sigma \tanh\left(\frac{d}{2\lambda_s}\right)\left(1 + \frac{\lambda_s}{\beta_{st}\tau_s}\tanh\left(\frac{d}{2\lambda_s}\right)\right)^{-1} \quad 2.70$$

When $\lambda_s/\tau_s \ll \beta_{step}$, $\chi \to 1$ and step velocity is based on 2.68:[24]

$$v_{st} = 2a^2 \frac{\lambda_s}{\tau_s} N_{se}\sigma \tanh\left(\frac{d}{2\lambda_s}\right) \quad 2.71$$

Conversely, $\lambda_s/\tau_s \gg \beta_{step}$, $\chi \to 0$ and step velocity is based on eq. 2.66:[24]

$$v_{st} = 2a^2 \beta_{st} N_{se}\sigma \quad 2.72$$

To recap, the model has been constructed with a surface having a vicinal offcut leading to parallel steps on the surface with many kink sites for attachment. Thus far the rates of capture from the upper and lower terraces are considered the same, and the concentration of adatoms is symmetric from either terrace. Three general solutions for the step velocity are found based on the relationship of λ_s/τ_s vs. β_{step}. Assuming a perfect capture rate of the steps ($\chi = 1$) like in Ref. 18, then eq. 2.71 describes the velocity of the steps. In this case, two surface morphologies could be accurately predicted depending on the values of σ and hyperbolic tangent. Recalling that 2D nucleation occurs at a critical concentration of adatoms (eq. 2.43 multiplied by Avogadro's number), if $N_{se}\sigma \tanh(d/2\lambda_s) > N_A n_{2D}^*$ then 2D nuclei formed in the center of the terraces. If $N_{se}\sigma \tanh(d/2\lambda_s) < N_A n_{2D}^*$, step-flow morphology was obtained. In that work, switching growth modes was performed by either changing the terrace width, d, at a set temperature or by keeping a set d and varying the temperature.[18] Changing the temperature affects N_{se}, σ, and λ_s, so it is difficult to conclude the ultimate driver for the reduction below the critical 2D nucleation concentration. Additionally, this paper could not account for the change from step-flow to step-bunching morphologies.[18,19]

Now the kinetic coefficients for the capture of an adatom from the upper and lower terraces will be treated as non-equivalent.[19,21] Eq. 2.55 becomes:

$$\beta_{st,up/low} = a\vartheta \frac{a}{\delta_0} \exp - \frac{E_{kin} + E_{ESB/IB}}{k_B T} \tag{2.73}$$

where E_{ESB} is the additional energy for adatom to attach from the upper terrace due to the Ehrlich Schwoebel barrier and E_{IB} is the incorporation barrier for adatoms to attach from the lower terrace. Either of these new energy terms can be positive or negative values, which can decrease or increase the capture coefficient for upper and lower terraces. As a result, eq. 2.56 now more accurately describes the step growth. A new solution for eq. 2.63 is required and will be constructed on new boundary conditions:

$$\begin{aligned} \mp D_s N_{se} \left.\frac{d\Psi}{dx}\right|_{\pm\frac{d}{2}} &= \beta_{st,\frac{up}{low}} \left(N_{st,\frac{up}{low}} - N_{se} \right) \\ &= \beta_{st,\frac{up}{low}} N_{se} \sigma \left(1 - \chi_{\frac{up}{low}} \right) \end{aligned} \tag{2.74}$$

The top line of eq. 2.74 qualitatively describes the rate of adatom attachment at the step edges. For a diffusion of adatoms to the kinks, the near step concentration must exceed the equilibrium concentration (which represents desorption of adatoms from the step front). Ref. 19 excludes this desorption term because AlN has an extremely low vapor pressure at typical growth conditions (Fig. 2.2). Some of the conditions for the GaN homoepitaxy were performed at very low supersaturations, so it is deemed appropriate to retain for this work. Additionally, χ in the form of eq. 2.64 is utilized, so a substitution to be performed in eq. 2.80 below. Deriving eq. 2.63 and setting it equal to the boundary condition at $x = +d/2$:

$$\frac{\beta_{st,up}\sigma\lambda_s}{D_s}\left(1 - \chi_{up}\right) = -A \exp\left(\frac{d}{2\lambda_s}\right) + B \exp\left(\frac{-d}{2\lambda_s}\right) \rightarrow \beta_u = -Ap + Bl \tag{2.75}$$

and for $x = -d/2$:

$$\frac{\beta_{st,low}\sigma\lambda_s}{D_s}(1-\chi_{low}) = A\exp\left(\frac{-d}{2\lambda_s}\right) - B\exp\left(\frac{d}{2\lambda_s}\right) \rightarrow \beta_l = Al - Bp \quad 2.76$$

where $p = \exp(d/2\lambda_s)$ and $l = \exp(-d/2\lambda_s)$. The coefficient A is found from eq. 2.76:

$$A = \frac{\beta_u + Bp}{l} \quad 2.77$$

by substituting eq. 2.77 back into 2.75, B is given:

$$B = \frac{p\beta_l + l\beta_u}{l^2 - p^2} \quad 2.78$$

eq. 2.75 is rewritten:

$$A = \frac{l\beta_l + p\beta_u}{l^2 - p^2} \quad 2.79$$

inserting eqs. 2.78 and 2.79 into eq. 2.63:

$$\Psi = \frac{l\beta_l + p\beta_u}{l^2 - p^2}\exp\left(\frac{x}{\lambda_s}\right) + \frac{p\beta_l + l\beta_u}{l^2 - p^2}\exp\left(\frac{-x}{\lambda_s}\right)$$

$$\frac{\sigma - \sigma_s}{\sigma} = -\frac{\lambda_{low}(1-\chi_{low})\cosh\left(\frac{2x-d}{2\lambda_s}\right) + \lambda_{up}(1-\chi_{up})\cosh\left(\frac{2x+d}{2\lambda_s}\right)}{\lambda_s\sinh\left(\frac{d}{\lambda_s}\right)} \quad 2.80$$

where $\lambda_{low/up} = \tau_s\beta_{st,low/up}$ is the capture length of the step for the lower and upper terraces. The vapor supersaturation was purposefully divided out because at the boundaries, $x = \pm d/2$, the left side of eq. 2.80 becomes $\chi_{up/low}$ (eq. 2.64). At $x = +d/2$:

$$\chi_{up} = -\frac{\lambda_{low}(1-\chi_{low}) + \lambda_{up}(1-\chi_{up})\cosh\left(\frac{d}{\lambda_s}\right)}{\lambda_s\sinh\left(\frac{d}{\lambda_s}\right)} = -SL(1-\chi_{low}) - CU\left(1-\chi_{up}\right) \quad 2.81$$

and for $x = -d/2$:

$$\chi_{low} = -\frac{\lambda_{low}(1-\chi_{low})\cosh\left(\frac{d}{\lambda_s}\right) + \lambda_{up}(1-\chi_{up})}{\lambda_s\sinh\left(\frac{d}{\lambda_s}\right)} = -SU\left(1-\chi_{up}\right) - CL(1-\chi_{low}) \quad 2.82$$

where the following substitutions were made: $SL = (\lambda_{low}/\lambda_s)\,\mathrm{csch}(d/\lambda_s)$, $SU = \left(\lambda_{up}/\lambda_s\right)\mathrm{csch}(d/\lambda_s)$, $CL = (\lambda_{low}/\lambda_s)\coth(d/\lambda_s)$, and $CU = \left(\lambda_{up}/\lambda_s\right)\coth(d/\lambda_s)$. By substituting eqs. 2.81 and 2.82 into one another, the solution for χ_{up} is:

$$\chi_{up} = \frac{-SL-CU+CLCU-SLSU}{1-CU-CL+CLCU-SLSU} \tag{2.83}$$

and for χ_{low}:

$$\chi_{low} = \frac{-SU-CL+CLCU-SLSU}{1-CU-CL+CLCU-SLSU} \tag{2.84}$$

so that eq. 2.80 becomes:

$$\frac{\sigma-\sigma_s}{\sigma} = \frac{SL(CU-SU-1)}{1-CU-CL+CLCU-SLSU}\cosh\left(\frac{2x-d}{2\lambda_s}\right) + \frac{SU(CL-SL-1)}{1-CU-CL+CLCU-SLSU}\cosh\left(\frac{2x+d}{2\lambda_s}\right)$$

$$\Psi = \sigma\left(\frac{SL(CU-SU-1)\cosh\left(\frac{2x-d}{2\lambda_s}\right)}{1-CU-CL+CLCU-SLSU} + \frac{SU(CL-SL-1)\cosh\left(\frac{2x+d}{2\lambda_s}\right)}{1-CU-CL+CLCU-SLSU}\right) \tag{2.85}$$

Eq. 2.85 represents the difference between the vapor supersaturation and the surface supersaturation at any position along the terrace. It can be reformulated to describe the surface supersaturation:

$$\sigma_s = \sigma\left[1-\left(\frac{SL(CU-SU-1)\cosh\left(\frac{2x-d}{2\lambda_s}\right)}{1-CU-CL+CLCU-SLSU} + \frac{SU(CL-SL-1)\cosh\left(\frac{2x+d}{2\lambda_s}\right)}{1-CU-CL+CLCU-SLSU}\right)\right] \tag{2.86}$$

or the adatom concentration:

$$N_s = N_{se} + N_{se}\sigma - N_{se}\sigma\left(\frac{SL(CU-SU-1)\cosh\left(\frac{2x-d}{2\lambda_s}\right)}{1-CU-CL+CLCU-SLSU} + \frac{SU(CL-SL-1)\cosh\left(\frac{2x+d}{2\lambda_s}\right)}{1-CU-CL+CLCU-SLSU}\right) \tag{2.87}$$

Eq. 2.87 has three main components making up the adatom concentration at any position along a terrace. The first is the equilibrium concentration of adatoms N_{se} and it is easy to imagine when the supersaturation term is small, then $N_s = N_{se}$. The second, $N_{se}\sigma$, is the supply of adatoms arriving to the surface from the vapor, and the third represents their diffusion away from a position based on attachment/detachment to the step edges or desorption back into the vapor.

This model was developed to gain an understanding of how the various kinetic parameters control the adatom density on the surface and show the propensity for their flux toward a step. Many of the parameters are difficult to find experimentally so it is hard to determine which values

should be inserted into the equations to represent the real behavior on the surface. Additionally, this model does not provide actual yield information on which morphologies evolve during growth, because it is a steady-state treatment of the surface which links to the starting offcut and growth parameters.

To briefly summarize, introducing an inverse ESB or positive IB, the adatom flux shifts to the lower step edge. When an adatom arrives from the vapor, there is a greater probability to attach at the step down position. If that happens in the simulation, then the upper terrace increases in size while the lower terrace shrinks. Because it has increased in size, the probability for the next adatom from the vapor to land on the upper terrace is increased. A feedback cycle is created where over time the larger terraces should expand faster than the smaller terraces and the characteristic bunched surface forms. Conversely, if a positive ESB or negative IB is utilized, then the adatom flux shifts to the upper step edge. This results in a terrace shrinking every time a new adatom arrives. This restores the equilibrium of the terrace width because the probability of the next adatom arriving decreases when the terrace shrinks. Finally, if the barriers are equivalent, then the parallel train step-flow is maintained because of the probabilities over a wide number of adatoms work to restore the equilibrium.

Coupling this kind of model with Monte Carlo simulations allows the evolution of the surface morphology. Many have had great success in replicating surfaces especially the transition from step-flow to step-bunched morphology.[19,21] These simulations consist of surfaces with several terraces and steps between them. At the starting point, the terraces may have equivalent height and width spacing, but in some cases, deviations are inserted. Regardless, the simulation begins by introducing an adatom from the vapor, and it has a probability to attach to any one of the terraces based on the width of the terrace. If the terraces begin with an equal spacing, then the

probability is equal for all terraces. Once the terraces grow based on adatom attachment, the probability increases for larger terraces, and shrinks for smaller terraces. When the adatom is on the surface, then the probability to approach the upper or lower step edge from the terrace is determined by the parameters in the model such as eqs. 2.85-87. The adatom then moves and attaches to one of the step edges, the terrace grows, and the next adatom is introduced from the vapor.

It is cautioned that in the literature, no single unifying model describing all growth conditions has been developed.[13,18-23] One such problem is the selective use of the ESB or IB, especially in the Monte Carlo simulations. These simulations consist of surfaces with several terraces and steps between them, and the surface evolves based on the probability for an adatom to attach at the step-down or step-up edge. Over many iterations, it is typical for step bunches to evolve from step-flow growth conditions when an inverse ESB or IB is utilized.[19,21] Yet, for the evolution from step-flow to step meandering, then a positive ESB is included.[22,23] None of the works have yet to describe why an ESB or IB exists, or how growth parameters such as the terrace width or vapor supersaturation influence their strength or appearance. The ESB and IB are simply included into the BCF models, so that the Monte Carlo simulations evolve into the appropriate surfaces.

In summary, the BCF modeling started by Burton, Cabrera, and Frank in the early 1950s was seminal in changing the perspective of the behavior of growing surfaces.[13] By treating the surface as a series of many terraces and steps, and the steps containing kinks where the adatoms could attach, they were able to show how crystals could grow without the need for nucleation. Including the ideas of diffusion length, residence time, and kinetic coefficients of the steps, allowed scientists to conceptualize the individual adatom behaviors on the terrace, and assume the

system behavior based on these individuals. As growth expanded to new materials and systems, this model has been expanded to describe the many different morphologies observed, but the field is still lacking a consensus on the source of these many changes.

5. Surface Morphology of GaN Homoepitaxy

The preceding sections introduced the background thermodynamic and surface kinetic processes that are used to understand the driving force for growth of a solid crystal from the vapor phase. This section will now introduce GaN growths performed on native GaN substrates and the resulting surface morphologies. As these new substrates become more commercially available, it is expected to bring a new generation of performance and devices because the substrates will minimize dislocations in the epitaxial layers. When growth is conducted on non-native substrates, the dislocations resulting from the lattice mismatch cause the surface morphology to be rough limiting the possibility of sharp layer transitions. It is also highly possible that the dislocations introduce non-uniformities for doping, composition (in the case of AlGaN), and are also a source of compensation themselves. So, the elimination of these deleterious effects could bring about a new generation in GaN devices.

With high dislocation densities, the step front becomes pinned at the dislocations and the resulting surfaces exhibit spiral growth morphology, virtually independent on the growth condition. Growing on the native vicinal substrates opens the possibility for many different surface morphologies compared to these non-native substrates. The purpose here was to observe the different morphologies, and link their appearance to specific combinations of substrate offcut and vapor supersaturation conditions. Using these trends, this work can advise how to obtain parallel step-flow morphology, or alter a given growth condition to transition out of the non-ideal morphology conditions.

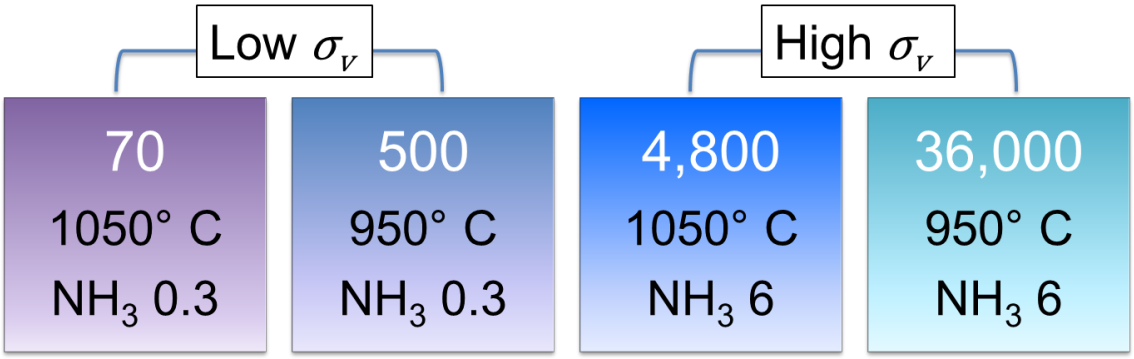


Figure 2.5 Growth condition variables for surface morphology study. Corresponding supersaturation value is listed in white. Supersaturation calculated following Refs. 29 and 30. Units for NH_3 flow are slm.

Four main growth conditions were utilized in this study that covered a wide range of supersaturation values (Fig. 2.5). The variables in the study were temperature and the ammonia flow. The growth process used only hydrogen diluent, and the H_2 flow rate raised or lowered so that the total flow was always 7.2 standard liters per minute (SLM). The total pressure in the MOCVD system was kept at 20 Torr. Finally, the total flow for triethylgallium (TEG) supplying the Ga species was always 280 standard cubic centimeters per minute (sccm) which equates to 134 μmol per minute. Fig. 2.7 gives the four conditions and corresponding vapor supersaturation values.

Before growth, the wafers were cleaved into quarters and cleaned with 0.1 M hydrofluoric acid (HF) for 1 minute, followed by 1 minute dip in 0.1 M hydrochloric acid (HCl). Samples were rinsed in flowing DI water to remove any residuals and dried using nitrogen air. Samples were immediately loaded into the MOCVD system. Sapphire wafers were cleaved to form a mask around the GaN substrate to prevent any turbulence at the edge of GaN wafers. Additionally, this allowed rotation of the susceptor to maintain a uniform boundary layer near the wafer surface and prevent large macroscopic defects that occur when rotation is not utilized. After pumping down main chamber to a pressure of $1x10^{-5}$ Torr growth was initiated. Nitrogen and ammonia were

introduced to the chamber at a flow rate of 1 slm each to stabilize the surface during temperature ramp up.[46] At 900° C, the nitrogen gas was replaced with hydrogen, and gas flows for H_2 and NH_3 were ramped to the final values. Once the temperature increased to 950° C, TEG was introduced to the chamber (and the temperature increased for the 1050° C conditions). Once the intended temperature was reached, approximately 100 nm of film were grown. After which the TEG was removed and the H_2 and NH_3 were quickly ramped down before changing to 2 slm N_2 gas for cooling.

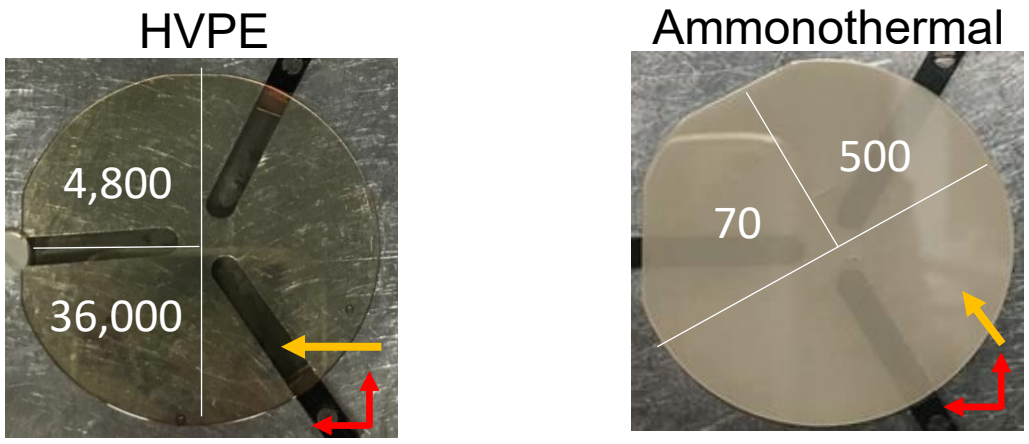


Figure 2.6 Wafers utilized in surface kinetics study with the corresponding supersaturation condition used for growth. Orange arrow indicates the direction of the offcut and red arrows are the directions where wafer curvature measured. Both measurements were acquired from the center position of the wafers.

Shown in Fig. 2.6 are the wafers utilized in the study of GaN homoepitaxy. Two different substrate growth methods were used and they provide for slightly different offcut behavior. The HVPE to the left yields slightly higher screw dislocation density ~10^6 cm^{-2} measured from 90x90 μm^2 atomic force microscopy (AFM) scans. Additionally, the lattice has significant bowing (~6 m radius of curvature) in both the $\varphi = 0°$ and at $\varphi = 90°$ directions measured from the center of the wafer. That corresponds to ~0.06° tilt in each of the red arrow directions. This would change the local offcut from the measured 0.3° in the center. It is expected that the HVPE wafer will be bowed

in additional areas of the wafer, or have other local variations, and not have a uniform tilt across the wafer. It was expected that the wafers would demonstrate quite different morphologies from center to edge. However, the high density of screw dislocations dominated the surface morphology for the low supersaturation growth conditions. Those results have been excluded from discussion. To observe the morphology at the low supersaturation growth conditions, growths were realized on the ammonothermal substrate (Fig. 2.6, right) which has $<10^4$ cm^{-2} screw dislocation density. From x-ray diffraction (XRD), the offcut was determined to be 0.4°, and from the large radius of curvature a small lattice bow is expected, i.e. constant offcut over the entire wafer.

With larger vicinal offcuts, the starting terrace width, d, is expected to reduce, while smaller vicinal offcuts yield larger terraces. However, quantifying the terrace width is difficult without the knowledge of the initial step height, h. For instance, with only half unit cell height steps at 0.3° offcut, the terrace width would approach ~47nm. If the height was full unit cell, then terrace width would be ~95nm. Recalling the discussion on the surface kinetics of the adatoms, the relationship between the diffusion length and terrace width determines if the adatoms reach the step edges and attach before desorption. Therefore, interpreting the evolution of a surface based on the diffusion length may be difficult if the step height is varying across the surface. Unfortunately, AFM could not determine the starting terrace width and height prior to growth. Instead the offcut was determined by XRD in the center of the wafers (Fig. 2.6), by 002 omega rocking-curves in four different phi angles.[25] Additionally, the wafer curvature was measured at 1 cm distance from the center in both the 0 and 90° phi rotations.

On the individual scans, the main offcut direction will be indicated by orange arrows outside the image. The teal arrows on the images represent the position and length of the linescans. The linescans are acquired in the offcut direction on the height images before flattening and the

data will be rotated according to the initial offcut of the substrate (example of non-rotated linescan is shown in Fig. 2.7). It must be noted that this method can yield errors by assuming the starting offcut, especially for the HVPE wafer where bowing may cause variations to the offcut. Finer adjustments to the rotation are made on the basis that smallest step heights must be either a full or half of the GaN unit cell in height. Height scans of the AFM surface will all be presented with the 'mud' color set and have been flattened for clarity. Amplitude images will be presented in grayscale. The amplitude image represents when the AFM tip is deflected off of the set amplitude. A completely flat surface would have an amplitude of 0. So, the peaks and troughs in the amplitude linescan show the step-up or step-down.

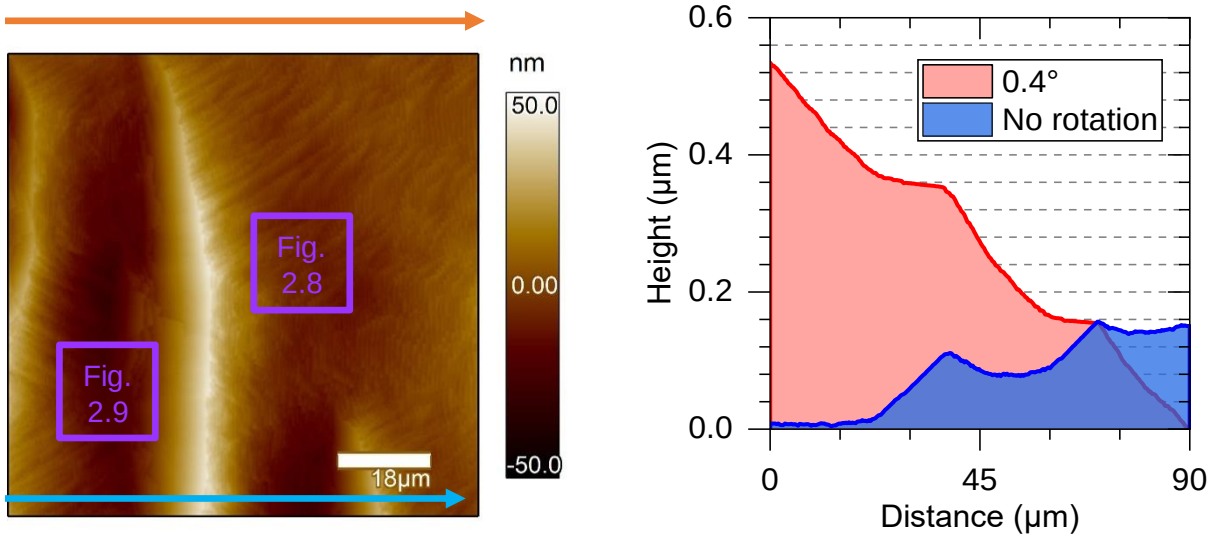


Figure 2.7 Left: 90x90 μm^2 AFM height image acquired in the center of the σ_v = 70 growth on the ammonothermal substrate. Orange arrow above points in offcut direction. Teal arrow on the image is the direction of linescan. Purple squares are areas for the Fig. 2.8 and 2.9. Right: linescan with and without the offcut angle rotation applied to the data.

Fig. 2.7 shows the 90x90 μm^2 image for the lowest vapor supersaturation condition, σ_v = 70, acquired at the center of the ammonothermal wafer (0.4° offcut). There are two prominent features of the surface: facetted regions surrounding wide macro terraces. The faceted region will be depicted in Fig. 2.8 which clarifies this 'faceted' designation. The intersection of the facets with

the macro terraces is investigated in Fig. 2.9 with a detailed image from the terrace itself. The linescan corresponds to the teal arrow on the height scan. The non-rotated linescan appears sawtooth before rotation is applied, and is not consistent when considering the direction of the offcut. When a rotation of the raw data by the angle of the offcut is performed, then width of the terraces is clearer and is approximated between 15 and 30 μm in diameter.

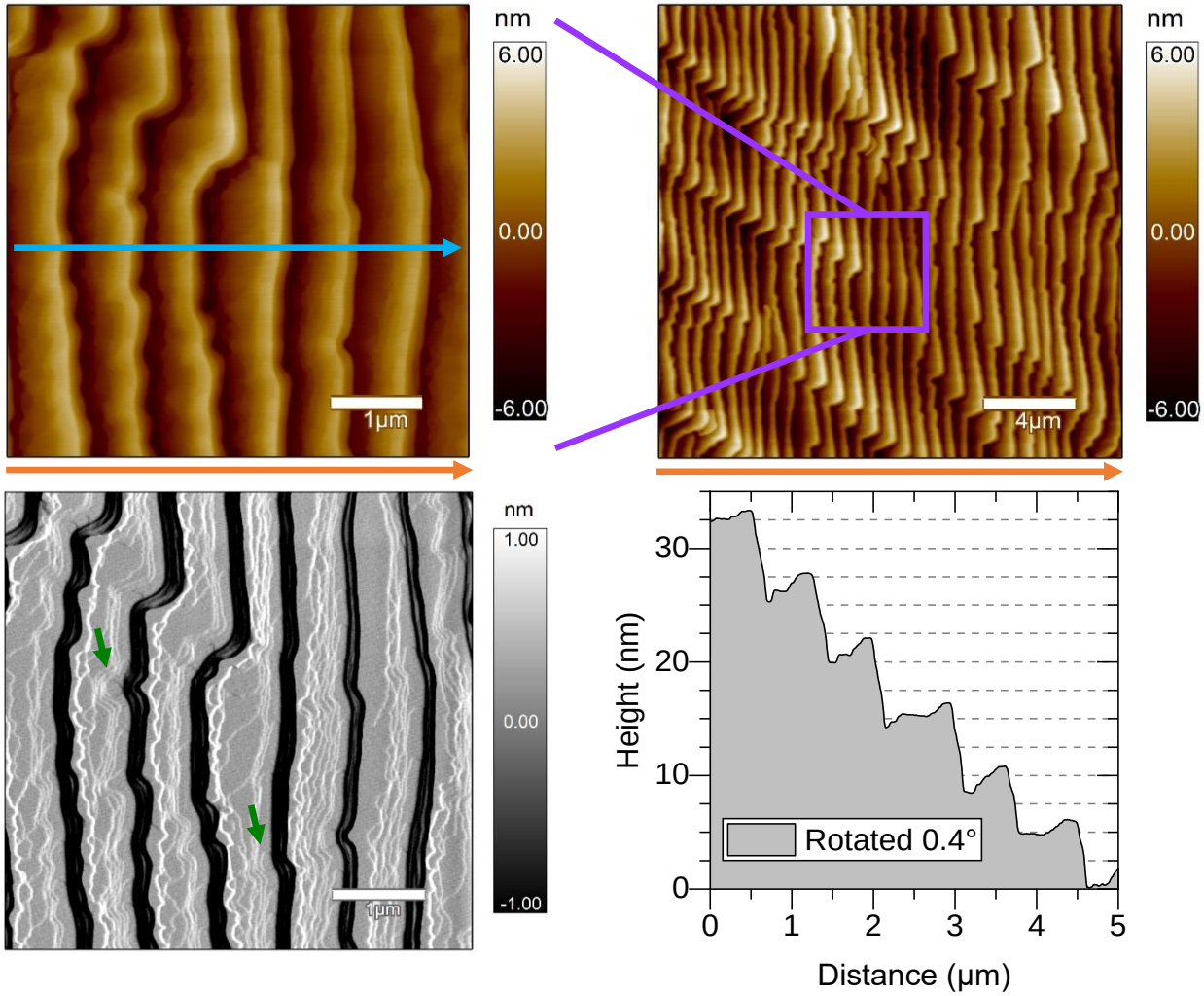
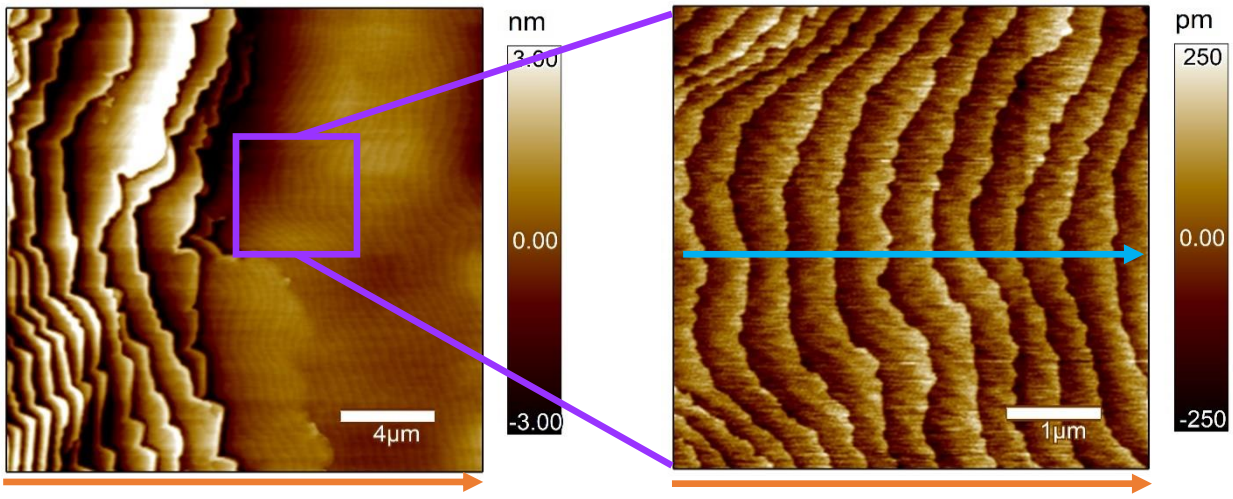


Figure 2.8 Top right: 20x20 μm^2 height image of the facetted region of (σ_v = 70 growth condition). Top left: 5x5 μm^2 height image. Above left: corresponding amplitude image from the top left image. 2D nucleation at the down step of the terraces is indicated by the white lines (when the tip is deflected vertically). Green arrows point to 2D discs. Above right: linescan for the height image. Nuclei are 1-2 nm and overall step height is ~3-6 nm.

Fig. 2.8 displays the shorter range scan ("higher magnification") view of the faceted region of Fig. 2.7. On the top right is the 20x20 μm^2 height image and shows the long range ordering of the steps with edges perpendicular to the offcut direction. However, many 90° jogs in the steps appear and seem to group in direction pointing 60° down from the offcut. At this time, it is not clear why they form but seem to be related to the crystal planes and directions of the GaN wafer. Considering the offcut direction is perpendicular to one of the m-planes, the 90° jog would represent the formation of a-plane wall while the 60° grouping would represent another m-plane direction for the crystal. The top left image of Fig. 2.8 is 5x5 μm^2 height image. Towards the step down edge in the offcut direction, the terraces appear to increase in height. This is clear in the corresponding amplitude image. Moving again towards the step down edge, the AFM tip was deflected vertically (white lines) in an increasing density. Surprisingly, many 2D discs appear near the step edges and appear to be related to the jogs in the step fronts. Finally, the right image of Fig. 2.8 is the linescan of the 5x5 μm^2 height image. Overall, the steps appear to be offset by ~5-8 nm in height and are almost 1 μm in diameter. Once the rotation has been applied, the resolution in the amplitude image is lost and only a couple of jumps in the height appear on the terrace. In general, the increase from the step up to step down edge is ~2.5 nm.

Fig. 2.9 shows the AFM images acquired near the intersection of the facets and the macro terrace region. The left image is the 20x20 μm^2 height image and the difference between the facets on the left of the image and the macro step on the right are very clear. In comparison to the facetted region, the terrace appears almost flat. Yet many steps appear on the macro terrace region, which is highlighted in the 5x5 μm2 height image on the right. Here many steps are formed and their edge shape is determined by the preceding facet (the left of the image). Beneath the image is the linescan across the height image. Whereas the height across the 5 μm distance in Fig. 2.8 was 32

nm, here it is only 3 nm. These steps do not have 2D nucleation island and instead are nearly uniform in their half unit cell height and 500 nm width. This result is surprising, because the same growth condition (and assuming offcut over the 90x90 μm^2) yielded such fundamentally different morphologies.



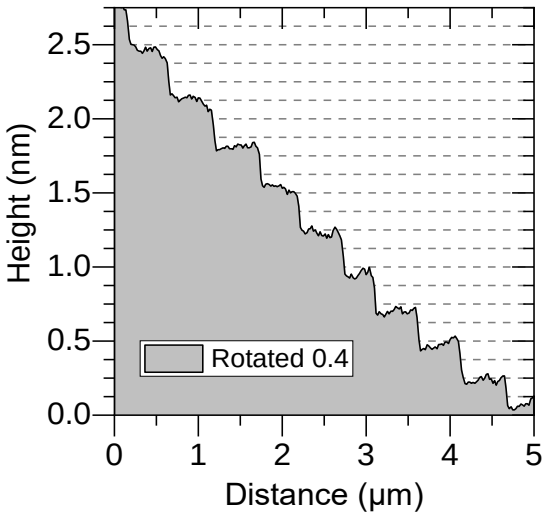


Figure 2.9 Above: 20x20 μm^2 AFM height image acquired at an intersection between the macro terrace and facetted region (σ_v = 70 growth condition).

Above right: 5x5 μm^2 AFM height image acquired on the terrace.

Right: Linescan on the 5x5 μm^2 image with the offcut angle rotation applied to the data. Step height is approximately half the unit cell height.

Fig. 2.10 displays the AFM images for the next growth condition, σ_v = 500, on the ammonothermal wafer. For the center of this quarter wafer, the density of terraces has increased dramatically as compared to the previous condition σ_v = 70. In comparison to Fig. 2.7, the macro steps are no longer separated by large regions of faceted morphology, and instead cover the entire surface. From the linescan after the rotation is applied, the large terraces have a width of ~10-15

μm and have step height of ~20-50 nm. The smaller AFM scan size acquired on the center terrace reveals the presence of the many smaller steps on the macro terraces as in Fig. 2.9. Examining the linescan of the 5x5 center image, over the 5 μm length there are 19 steps observed with a height change overall of ~5.8 nm. This roughly corresponds to the half unit-cell height steps but only ~200 nm in width (compared to the 500 nm in Fig. 2.7).

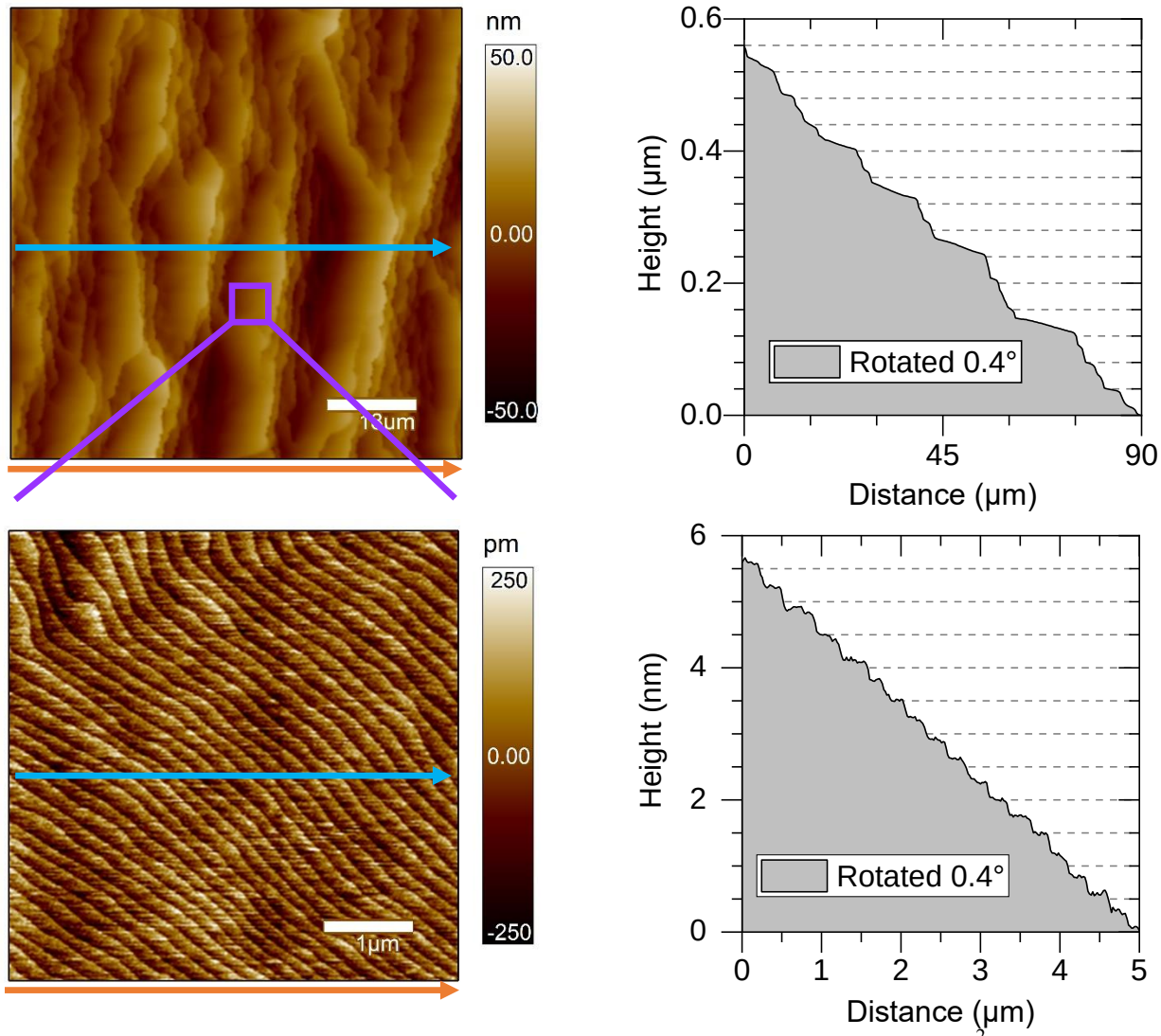


Figure 2.10 σ_v = 500 on ammonothermal substrate (0.4°). Top left: 90x90 μm^2 image reveals the surface is covered in the wide macro terraces and minimal facetted region. Top right: corresponding line can across the surface. Above left: 5x5 μm^2 height image taken on the center terrace. Above right: linescan revealing half unit cell height steps.

The results for the low supersaturation growth conditions provide an interesting discussion for the evolution of surface morphology based on the starting growth conditions alone. Because the ammonothermal substrate had small bowing, it is expected that the starting offcut of the substrate is uniform and the results of Figs. 2.7 through 2.10 are directly comparable on growth conditions. For the σ_v = 70 condition, two distinct morphologies were present on the surface: faceting and large macro terraces. The facetted regions consist of highly ordered steps in the m-direction but have jogs in the step front that represent new a-plane faces. At the step edges, some form of 2D nucleation is occurring at the step edges and many 2D discs are visible at the edges. This suggests that the local surface supersaturation exceeded the critical radius for the formation of nucleus.[18] However, this growth condition was the lowest in terms of vapor supersaturation, the exact opposite condition from Ref. 18, where 2D nucleation was observed at the highest supersaturation conditions. Besides these facetted regions, large macro terraces consisting of long range half unit-cell steps were observed in Figs. 2.9 and 2.10. The major difference here was that Fig. 2.10, where σ_v = 500, showed a much greater density of the macro terraces, and the width of the smaller steps decreased compared to Fig. 2.9.

A typical approach to describing step-bunching or macro step formation is to invoke the inverse ESB which makes it probable for adatoms on the surface to attach to lower step.[19,22] As the adatoms continue to attach at the lower step edge, the upper terrace extends in width until the majority of new adatom arrivals land on the large terrace. This would lead to a situation of wide terraces overtaking one another. However, the wide macro terraces in this study did not remain flat and instead consist of a wide region of step-flow morphology. This should not exist if the steps indeed have an inverse ESB. In the faceted regions, however, there appears a movement for the

adatoms towards the lower step edge. Instead of attaching preferably to the step-down edge, however, they build up and eventually form nuclei contrary to an inverse ESB.

To understand the behavior of these steps, the following model is proposed. Under the σ_v = 70 growth condition, the arrival flux of adatoms is low and a shift in surface supersaturation toward the lower step edge (due an inverse ESB) occurs. As a result, the steps increase in width and the number of impinging adatoms also increases. Eventually, the concentration of adatoms increases until they exceed the critical number for nucleation to occur. This is qualitatively how the faceting regions appear. In contrast, the σ_v = 500 growth condition has a sufficient supply of adatoms from the vapor that a shift in surface supersaturation does not appear, i.e. adatoms attach to both step edges equivalently. This results in the large regions of step-flow (the macro terrace). Due to some local inhomogeneity, however, the step-flow is not consistent across the entire surface. Thus, the large macro terraces outgrow regions of non-step flow, and lead to the bunched surface.

Since both growths were considered to have the same offcut, it would be expected that change in density of the macro terraces stems from the difference in the supersaturation used for each growth. For the σ_v = 70 condition, a greater desorption for the adatoms from the steps is expected compared to the σ_v = 500 condition because of the higher temperature. Additionally, the diffusion length (eq. 2.52) is expected to decrease with increased temperature. The 2D nucleation of the facets occurs because of the inverse ESB and wide steps. Conversely, the macro steps have step-flow regions where the adatoms were able to reach either step edge and attach.

The above discussion serves to highlight the difficulty in providing an explanation for the observed morphologies. Even when models are successful in demonstrating step-flow into step-bunched surfaces, they miss many of the nuances, such as the step-flow on the macro terraces.[19-23] These regions are especially interesting because the step height and terrace width do not match

the starting offcut of the wafer. Traditionally, step-flow maintains these features of the starting substrate, and only under an inverse ESB are the terraces able to increase in size. The purpose of this study was to observe the morphologies over the widest range of growth conditions possible for the MOCVD reactor utilized in this work, and find ideal conditions for growth of devices. It is clear that low supersaturation conditions yielded a macro-step morphology as previously described. Future experiments utilizing these conditions on new ammonothermal wafers with different starting offcut values is suggested to elucidate the connection between the terrace width and the growth condition. It is also advised to perform regrowths using these supersaturation conditions on wafers where step-flow morphology is confirmed. In this way, the starting terrace width and step height are confirmed. As will be shown later, the high supersaturation conditions yield steps that are mostly unit cell in height. It would be interesting to observe if macro terraces would evolve.

Next, the surface morphology for the higher supersaturation conditions, σ_v = 4,800 and 36,000 will be described. These significantly higher values arise from increasing the ammonia flow from 0.3 slm up to 6 slm, and decreasing hydrogen from 6.9 slm to 1.2 slm. Again, the growths were performed on the HVPE substrate which has an initial offcut of 0.3° which is expected to deviate from the large lattice bowing. Three locations will be shown for each growth within 2 mm of the edge separating the two pieces in Fig. 2.6 so that any wafer deviation can be minimized. 'Edge' regions are between 4 and 6 mm from the main flat, 'Middle' regions are between 1 and 2 cm from the main flat, and center are the regions further than 2 cm from the main flat. Due to the size of the features, only the 5x5 μm^2 images will be presented. Over the larger size images, the AFM was unable to detect the morphology due to the scanning speed utilized. Lacking the wide area image is compensated by acquiring many images in each area of the wafer.

The images presented are expected to give a good representation for the overall morphology in the region being discussed.

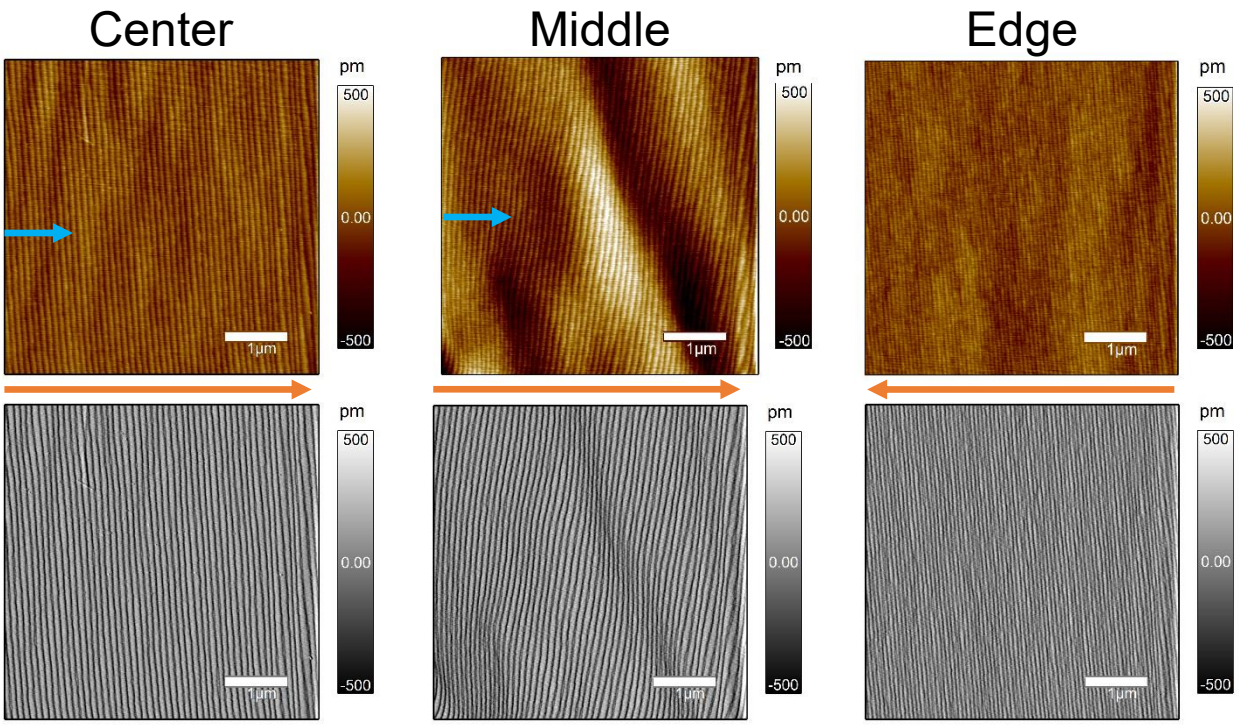


Figure 2.11 σ_v = 4,800 on HVPE (0.3°). Top row figures are 5x5 µm^2 AFM height images and below are the corresponding amplitude traces which make the morphology clearer. Left, middle and right columns reflect the center middle edge regions of the quarter wafer, respectively. Teal arrows are the linescan locations plotted in Fig. 2.12.

The results for the σ_v = 4,800 growth condition are displayed in Fig. 2.11. Starting with the center region of the wafer, the morphology of the surface is parallel step-flow. By counting the number of steps across the amplitude image, the width of the terraces is roughly ~90 nm. The corresponding linescan is plotted in Fig. 2.12. Since this scan was close to the center of the wafer, the measured offcut of 0.3° was used for the rotation. Many of the steps have a height of a full unit cell, while others are half a unit cell. Moving to the middle location on the quarter wafer, the parallel step flow is continued, but some ‘waves’ appear. These stem from some non-uniformity such as the lattice bow. Again, rough estimation of the number of steps yields a similar terrace

width of ~90 nm. Viewing the corresponding linescan in Fig. 2.12, again a mix of full unit cell and half unit cell height terraces are present. It must be noted that the same rotation angle for the offcut was used as the center of the wafer. This may not be valid due to the significant bowing of the wafer. However, the width of the terraces seems to suggest the morphology did not change between these two regions, so the utilizing the same rotation angle appears valid. At the edge of the wafer, the steps are again parallel steps, but have shrunk to a width of ~65 nm based on the number of steps counted in the amplitude image. Once the rotation is applied to the raw height data, no steps could be distinguished by the linescan. From the center and middle scans, it is expected that the steps would have mixed full and half unit cell heights.

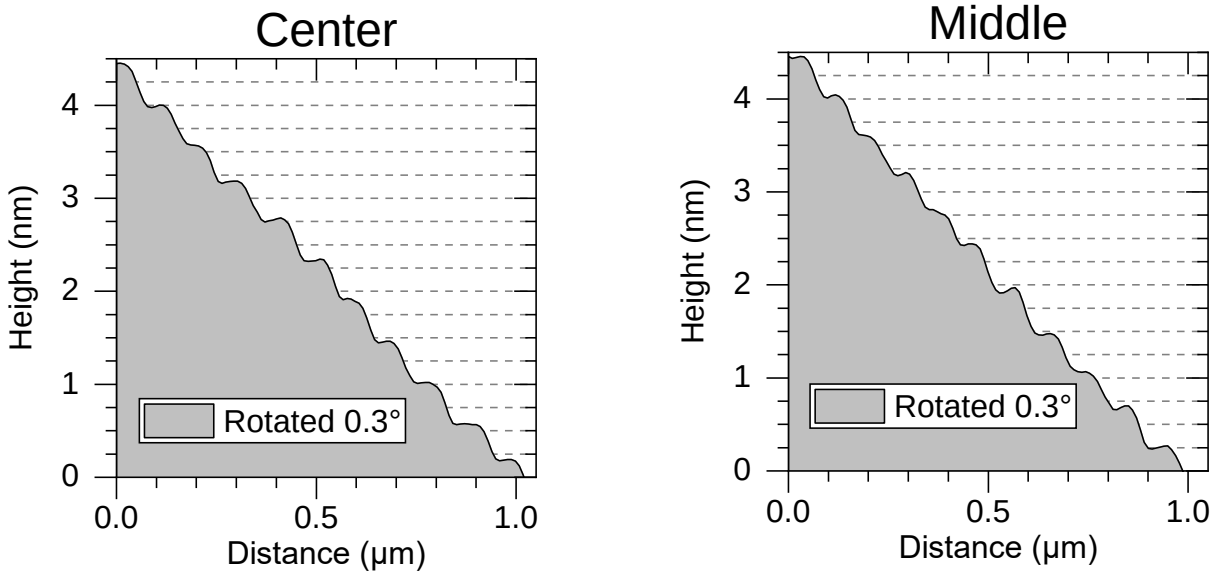


Figure 2.12 σ_v = 4,800 on HVPE (0.3°) linescans acquired in the teal arrows of Fig. 2.11.

Fig. 2.13 displays the results of the σ_v = 36,000 growth condition on HVPE quarter wafer. Starting with the center region of the wafer, the morphology of the surface is the step-meandering structure described previously in Si and GaAs.[22,23] For this structure, the long parallel steps with edges perpendicular to the offcut direction curve into 'waves' that pinch and create a curved surface. This was also observed for the middle region of the wafer, suggesting that like the σ_v =

4,800 growth, the HVPE wafer does not have a strong deviation in the offcut from center to the middle regions. Unfortunately, the size of the steps is at the limits of the detectability for AFM, so line traces in the offcut direction to find the height and width of the steps are not possible. However, the linescans were taken perpendicular to the offcut direction to see the variation of the height across the waves. In Fig. 2.14, both the center and middle region show that the between the crest and the trough of the wave there is a height change of roughly 1 nm or two unit cells. At the edge region in Fig. 2.13, the morphology returns to the straight step-flow morphology and similar to the σ_v = 4,800 condition, with steps having a width around ~65 nm. Additionally, across a pinch in the step the height appears to change by half of a unit cell.

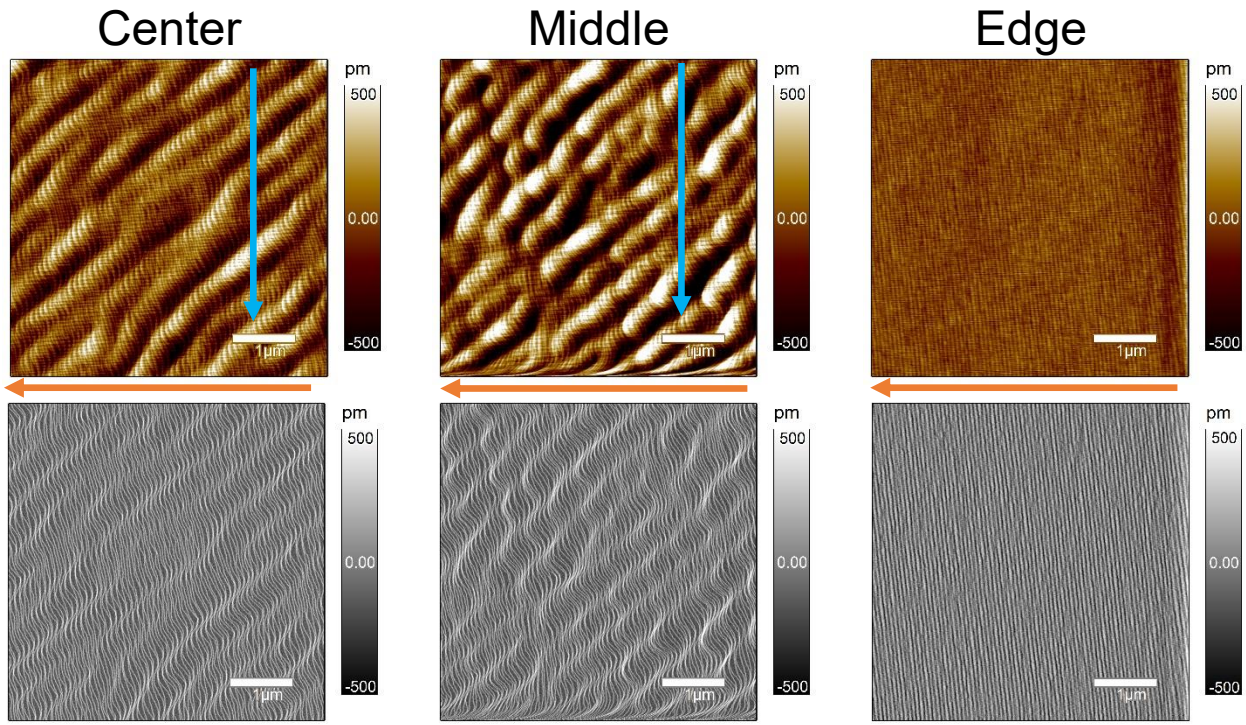


Figure 2.13 σ_v = 36,000 on HVPE (0.3°). Top row figures are 5x5 µm^2 AFM height images and below are the corresponding amplitude traces which make the morphology clearer. Left, middle and right columns reflect the center middle edge regions of the quarter wafer, respectively.

Comparing the results of Figs. 2.11 and 2.12, it is clear that the resulting surface morphologies are much more desirable for device applications than previously shown for the low

supersaturation conditions. The locations of these scans were within 2 mm of the centerline splitting these quarter wafers so the deviation perpendicular to the offcut is minimized, and the lattice bowing (if any) would be similar from the edge to center regions. For both growth conditions, the edge areas yielded the same step flow morphology, but at a step width smaller than the other regions of the σ_v = 4,800 growth. It is unclear exactly what is the offcut in these regions, but it cannot be the same as the center or middle regions, otherwise there should not be this deviation in morphology. Because the width of the terrace is smaller, it is assumed that the offcut is greater at the edge compared to center. By assuming roughly that the steps are a full unit cell in height, then offcut is approximately 0.4°. This is not unreasonable since the lattice bow 1 cm from the center of the wafer is ~0.06° based on the very large curvature of the wafer.

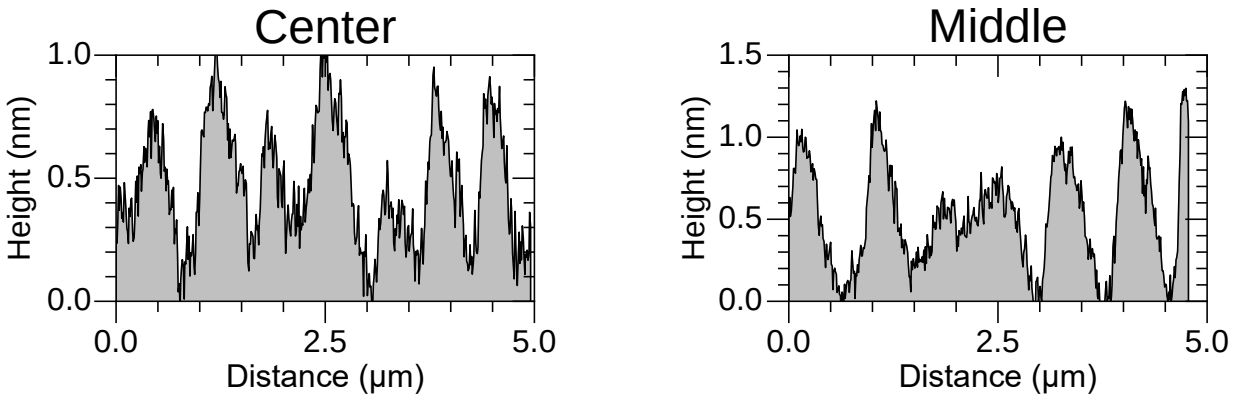


Figure 2.14 σ_v = 36,000 on HVPE (0.3°) linescans acquired across the teal arrows in Fig. 2.13.

These edge regions stand in very stark contrast to the low supersaturation growth condition where the surface was dominated by faceting (σ_v = 70) or by the macro terraces (σ_v = 500). Even considering the step-flow on the macro terraces, which had widths between 200 to 500 nm and step heights of half a unit cell, the edge regions of these high supersaturation growths yielded different heights and widths, i.e. the growth condition affects the kinetics. For step-flow growth to occur, it is understood that the diffusion length of adatoms is sufficient to reach the edges of the terrace and have a residence time that allows them to reach the edges.[18-22] It is also expected that

the adatoms have equal probability to attach to either the lower or upper step edge, or that a positive ESB exists.[23] Qualitatively, a positive ESB would result in the adatoms attaching to the upper step edge, causing the lower terrace to shrink in size. The subsequent adatoms would likely arrive at the upper terrace and move to its upper step edge, which would cause the upper terrace to shrink to the size of the lower terrace. In this way, the positive ESB restores the terraces to equivalent widths. For the high supersaturation growths, the greater flux of adatoms assumed that there is no apparent barrier or a positive ESB for the step-flow to prevail across wide areas of the wafer.

Considering the middle and center regions for the σ_v = 4,800 growth, the step-flow morphology is maintained except that the terrace width has changed (~90 nm). By making the same assumption that the steps are approximately a unit cell in height, then the offcut for these regions is 0.3° which matches the measured value from XRD taken before the growth. It is very difficult to determine the height of these terraces, since the starting offcut is unknown, and size of the steps is approaching the limits of resolution for the instrument. But it is assumed that due to the observed morphologies, the terrace width must have changed between center and edge on each of the quarter wafers. On the other hand, the 'middle' region for σ_v = 36,000 already displayed the step-meandering morphology. For the σ_v = 36,000 a small difference in the meandering morphology was observed between 'middle' and 'center' regions.

To describe the transition between a step-flow and a step-meandered surface, several different approaches will be presented. Bales and Zangwill offered the first explanation by treating the step edges as having an inherent curvature or waviness present.[23] Because of the waviness of the step, concave and convex regions are present with respect to the lower terrace. For adatoms on the lower terrace, it is assumed that they will approach the convex areas for attachment (while the upper terrace adatoms approach the concave regions). If the flux of adatoms and attachments rates

are equal, i.e. no ESB, then the convex and concave regions advance at the same rate and the step remains straight. However, in the presence of a positive ESB, where adatoms attach from the lower terrace more efficiently, then the convex regions of the step would advance faster than the concave regions, which would cause the step-meandering structure.[23]

Another approach to the step-meandering problem was put forward by Tersoff and Pehlke for the investigation of Si (001) surfaces.[26] Their work followed in the steps of Alerhand, et al. and who modeled elastic stress domains formed on the surface of strained Si crystals.[27] Under an externally applied strain, domains reflected 2x1 and 1x2 reconstructions of the surface would appear, and these different orientations lead to stresses on the surface.[27] Following this logic, Tersoff and Pehlke argued that steps would also alternate in these reconstructions, and the discontinuity in the stress at the step edges would generate a force on the step.[26] With increasing terrace width, it would be favorable for the step-edges to deviate parallel steps and instead form the step-meandering structure to alleviate this excess strain.[26] In their work, they numerically demonstrated that a sinusoidal edge would lead to a reduction in the strain for the structure, and the step-meandering morphology would appear when the supersaturation was too low for nucleation to occur.[26]

The final example that will be presented is a work by Zaluska-Kotur, et al. who discussed the possibility for GaN steps that are half of a unit cell in height.[28,29] Because of the AaBb stacking of the GaN crystal, the Ga atoms would be rotated by 60° and every other terrace would have an orientation away from the offcut direction. The key to this interpretation is that the energy barrier for an adatom to attach to the step edges varies because the GaN tetrahedra for each step would be more or less ‘complete’ based on the orientation. The fewer number of Ga nearest neighbor atoms in the tetrahedra, the lower the barrier for incorporation. Therefore, each terrace would have

different kinetic coefficients for the adatoms to be captured.[28] If the arrival flux of adatoms was low, then it is assumed that the terraces oriented in the offcut direction would outgrow the misaligned steps, and full unit cell high steps would propagate.[29] If the arrival flux was sufficient, the misaligned steps could grow in their preferred orientation (60° from the offcut direction) while the offcut-oriented terraces would grow in the offcut direction. Because of these competing directions, a step-meandering pattern would eventually emerge.[28,29]

While each of the above models provide some qualitative explanation for the appearance of the step-meandering morphology, each has their own flaws and/or assumptions which lack experimental backing. The Bales and Zangwill model is based on a 'waviness' in the step edges based on the starting growth conditions and positive ESB dictating the movement of adatoms on the surface. However, their model predicts with higher deposition fluxes the step wavelength is too short and step-meandering does not form.[23] This is opposite to the situation observed in this dissertation. In the Tersoff and Pehlke model, the step-meandering appears as a result of increasing terrace width alone,[26] which is also is at odds with the experimental results. Finally, the Zaluska-Kotur, et al. requires selective attachment to the steps based on the orientation of the Ga atoms at different positions in the unit cell.[28,29]

Qualitatively, some combination of all the above models may give some insight to the appearance of the step-meandering morphology. The experimental data suggests the following: 1) at small starting terrace widths, like the edge of the HVPE wafer, both growth conditions are suitable for step-flow morphology with mostly unit cell height terraces. 2) At wider terraces, the lower supersaturation (σ_v = 4,800) is stable for step-flow following the initial dimensions of the steps (full unit cell height, 90 nm widths). 3) For larger supersaturations, new steps offset by half of the unit cell form and the meandering pattern emerges. Based on these observations, some

suggestions can be made in terms of the growth conditions that can lead to for uniform step-flow morphology for semiconductor grade devices. It appears that the step-meandering surface is non-ideal for devices where very sharp layer interfaces are required. To avoid this morphology, the grower either can lower the supersaturation for instance by increasing the temperature when a low offcut substrate is utilized. However, making this change can have an effect on other impurities in the films. In that case, it is suggested to increase the offcut.

To confirm these suggestions, future experimental work is suggested to clarify how the models should be adapted. Ammonothermal substrates with offcuts values of 0.2°, 0.3°, and 0.4° will be acquired for growth. After confirming if lattice bowing is present by XRD, the wafers will be diced into many pieces for several possible growth regimes. First, both supersaturation conditions will be grown on the 0.4° wafer to confirm the 'edge' regions of the HVPE wafer in this work. Then the σ_v = 4,800 will be grown on both the 0.2° and 0.3° substrates, to confirm the center region of the HVPE wafer, and to see if even larger starting terraces yield step-flow or step-meandering morphology for this lower supersaturation. After these growths have been performed and AFM confirms the step morphology on the surface, the σ_v = 4,800 on the 0.3° offcut will be loaded back into the reactor, and the σ_v = 36,000 condition will be utilized. This allows for the starting terrace width and step height to be known prior to any growth condition and to confirm that new steps are being introduced when the step-meandering morphology is observed. The reverse of this growth is also suggested, begin with the σ_v = 36,000 on 0.3° offcut, then regrow the σ_v = 4,800 condition. This could confirm if the step-meandering morphology can be reversed.

6. Conclusion

The results from this first look at the morphology resulting from homoepitaxial growth of GaN by MOCVD introduces the wide range of surfaces that can be achieved for this material. Here

the supersaturation for GaN was changed four orders of magnitude, and four very unique morphologies were observed. Several models put forward in literature were introduced, however, there does not appear a universally applicable description that will yield all of the morphologies present in this work. It is the hope that the future studies suggested will help clarify many of the inconsistencies in the modeling, and complete the picture of how the starting terrace width and growth conditions dictate the morphology of the film. At the current status, high supersaturation growth conditions are suggested ($4{,}800 < \sigma_v < 36{,}000$) , with starting offcuts between 0.3 and 0.4°, to obtain smooth step-flow surfaces.

7. References

1. U. W. Pohl. “Introduction,” and “Chapter 4 Thermodynamics of Epitaxial Layer-Growth,” in *Epitaxy of Semiconductors*. Springer: Berlin, Heidelberg (2013).

2. G. B. Stringfellow. *Organometallic Vapor Phase Epitaxy: Theory and Practice*. Academic Press (1999).

3. H. M. Manasevit and W. I. Simpson. “The use of metal-organics in the preparation of semiconductor materials,” *Journal of the Electrochemical Society* **116**, 1725 (1969).

4. “14 Epitaxial Crystal Growth: Methods and Materials.” and “32 Group III Nitrides,” in *Springer Handbook of Electronic and Photonic Materials*. S. Kasap and P. Capper (eds.). Springer (2006).

5. J. W. Gibbs. *Scientific Papers*. Dover Publications: New York (1961).

6. *NIST-JANAF Thermochemical Tables*. fourth ed. The American Chemical Society and the American Institute of Physics for the National Institute of Standards and Technology: Woodbury, NY (1998) (janaf.nist.gov).

7. K. T. Jacob and G. Rajitha. “Discussion of enthalpy, entropy and free energy of formation of GaN,” *Journal of Crystal Growth* **311**, 3806 (2009).

8. R. A. Swalin. *Thermodynamics of Solids*. John Wiley & Sons: New York (1972).

9. S. Washiyama, P. Reddy, F. Kaess, R. Kirste, S. Mita, R. Collazo, and Z. Sitar. “A thermodynamic supersaturation model for the growth of aluminum gallium nitride by

metalorganic chemical vapor deposition," *Journal of Applied Physics* **124**, 115304 (2018).

10. R. DeHoff. *Thermodynamics in Materials Science*. Taylor & Francis Group: Boca Raton (2006).

11. A. Koukitu, N. Takahashi, and H. Seki. "Thermodynamic study on metalorganic vapor-phase epitaxial growth of group III nitrides," *Japan Journal of Applied Physics* **36**, L1136 (1997).

12. S. Mita, R. Collazo, A. Rice, R. F. Dalmau, and Z. Sitar. "Influence of gallium supersaturation on the properties of GaN grown by metalorganic chemical vapor deposition," *Journal of Applied Physics* **104**, 013521 (2008).

13. W. K. Burton, N. Cabrera, and F. C. Frank. "The growth of crystals and the equilibrium structure of their surfaces," *Philosophical Transactions of the Royal Society of London. Series A,* **243**, 299 (1951).

14. I. Markov and S. Stoyanov. "Mechanisms of epitaxial growth," *Contemporary Physics* **28**, no. 3, 267 (1987).

15. F. C. Frank and J. H. van der Merwe. "One-dimensional dislocations I. Static theory," *Proceedings of the Royal Society of London A* **198**, 205 (1949).

16. F. C. Frank and J. H. van der Merwe. "One-dimensional dislocations II. Misfitting monolayers and oriented overgrowth," *Proceedings of the Royal Society of London A* **198**, 216 (1949).

[17.] H. Amano, I. Akasaki, K. Hiramatsu, N. Koide, and N. Sawaki. “Effects of the buffer layer in metalorganic vapour phase epitaxy of GaN on sapphire substrate,” *Thin Solid Films* **163**, 415 (1988).

[18.] I. Bryan, Z. Bryan, S. Mita, A. Rice, J. Tweedie, R. Collazo, and Z. Sitar. “Surface kinetics in AlN growth: a universal model for the control of surface morphology in III-nitrides,” *Journal of Crystal Growth* **438**, 81 (2016).

[19.] K. Bellman, U. W. Pohl, C. Kuhn, T. Wernicke, and M. Kneissl. “Controlling the morphology transition between step-flow growth and step-bunching growth,” *Journal of Crystal Growth* **478**, 187 (2017).

[20.] R. L. Schwoebel and E. J. Shipsey. “Step motion on crystal surfaces. II,” *Journal of Applied Physics* **40**, 614 (1969).

[21.] M. H. Xie, S. Y. Leung, and S. Y. Tong. “What causes step bunching – negative Ehrlich-Schwoebel barrier versus positive incorporation barrier,” *Surface Sciences* **515**, L459 (2002).

[22.] M. Vladimorova, A. Pimpinelli, and A. Videcoq. “A new model of morphological instabilities during epitaxial growth: from step bunching to mounds formation,” *Journal of Crystal Growth* **220**, 631 (2000).

[23.] G. S. Bales and A. Zangwill. “Morphological instability of a terrace edge during step-flow growth,” *Physical Review B* **41**, no. 9, 5500 (1990).

[24.] I. V. Markov. *Crystal growth for beginners: fundamentals of nucleation, crystal growth and epitaxy*. World Scientific: River, Edge NJ (2003) online.

[25.] M. A. G. Halliwell and S. J. Chua. "Determining substrate orientation using a high resolution diffractometer," *Journal of Crystal Growth* **192**, 3-4, 456 (1998).

[26.] J. Tersoff and E. Pehlke. "Sinuous step instability on the Si (001) surface," *Physical Review Letters* **68**, no. 6, 816 (1992).

[27.] O. L. Alerhand, D. Vanderbilt, R. D. Meade, and J. D. Joannopoulos. "Spontaneous Formation of Stress Domains on Crystal Surfaces," *Physical Review Letters* **61**, no. 17, 1973 (1988).

[28.] M. A. Zaluska-Kotur, F. Kryzewski, and S. Krukowski. "Double step structure and meandering due to the many body interaction at GaN(0001) surface in N-rich conditions," *Journal of Applied Physics* **109**, 023515 (2011).

[29.] M. A. Zaluska-Kotur, F. Kryzewski, and S. Krukowski. "Emergence of regular meandered step structure in simulated growth of GaN (0001) surface," *Journal of Crystal Growth* **343**, 138-144 (2012).

Chapter 3 Mg Doping of GaN and AlGaN

1. Introduction

The key distinction between a semiconductor material and an insulator is that by introducing trace quantities of foreign atoms or compounds into the crystal lattice, the insulating material is modified to gain new optical and electrical characteristics. This strength has enabled semiconductors to drive the technological revolution of the past half century and become commonplace in everyday life. During growth of semiconductor crystals, foreign (external) point defects can be intentionally incorporated, by introducing the appropriate supply to the growth region, or unintentionally, from previous contamination or other ambients in the growth system.[1] Intrinsic defects, including vacancies and related complexes, can be formed in response to the incorporation of the external defects. These intrinsic defects often counteract the changes in performance due to the external defects in a process called self-compensation.[1,2]

Just like the overall growth of the crystal is dictated by thermodynamics, the probability that the external and intrinsic defects are incorporated is determined by the environment at the growth surface, the chemical potential described in Section 2.2.[1-5] Additionally, the different valency of the dopants to the host atoms causes a shift in the Fermi level at the surface, which also influences the probability for incorporation. Beyond the dopants themselves, quasi-Fermi levels can be established for both the valence and conduction band if the crystal is subjected to an external source of electrons and holes during the growth, such as by above bandgap illumination.[3,4] Overall, these effects can be summarized into an energy of formation for a given defect, and it is through this lens that magnesium dopant in GaN and $Al_{0.6}Ga_{04}N$ will be explored.

To date, Mg has been the only dopant shown to generate viable p-type conductivity in the III-nitrides.[6,7] However, this p-type characteristic is far from efficient.[8] Typical hole concentrations

in the $1x10^{17}$ to $1x10^{18}$ cm^{-3} are measured for total Mg concentrations around $1x10^{19}$ cm^{-3} in GaN. The situation deteriorates rapidly with the introduction of Al for the $Al_xGa_{1-x}N$ alloy, such that Mg:GaN is typically relied upon for p-type contact layers in devices. That limitation can hamper the efficiency, especially for deep-UV optoelectronics, because the GaN layer can absorb photons from the active region of the devices.[9]

The difficulty with $Mg_{(III)}$ acceptor is two-fold. First, Mg is a deep acceptor in GaN, with an ionization energy of ~170-180 meV.[10] At room temperature, this means a significant portion of the Mg atoms are not contributing to the conduction of the crystal. The second limitation, however, is the passivation of the Mg atoms by hydrogen, and the self-compensation by vacancies and related complexes.[11] This work will investigate how each of these are affected by varying process parameters and adding above bandgap illumination during growth.

Before describing the experimental results in this section, the energy of formation framework will be reviewed. This will highlight the main conclusions from Refs. 2-5, and for greater details the reader is directed to those investigations. After this introduction, a review of the defect quasi Fermi level control of Mg doping in GaN will be presented.[11] This work will serve as the basis of comparison for the other growths that are primarily focused on changing the chemical potential during growth via the process conditions. Finally, the experimental results for the study of Mg doping in $Al_{0.6}Ga_{0.4}N$ will be introduced before concluding the chapter.

2. Energy of Formation for Point Defects

Controlling the incorporation of point defects into a semiconductor is pivotal to obtaining the desired properties from the material. From thermodynamics, in the grand-canonical ensemble, the equilibrium concentration of point defects (external, intrinsic, or compounds) is given by:[2]

$$[X^q] = N_{sites} N_{config} \exp\left(\frac{-E^f}{k_B T}\right) \quad 3.1$$

where $[X^q]$ is the concentration of defect X with charge state q, N_{sites} are the number of lattice sites where the defect can incorporate, N_{config} are the number of configurations of the defect on the site included as a compound or different charges, and E^f is the energy of formation for the defect. The energy of formation does not include the entropy but is a useful approximation for modeling. It is important to mention, that this is based on a system at equilibrium and MOCVD growth is a non-equilibrium process. However, the equilibrium formalism can be justified on the basis that chemical potential change primarily takes place in the boundary layer separating the surface and the vapor during a mass transport limited regime.[3,4] On the surface, the adatoms have sufficient mobility to quickly incorporate and restore equilibrium to the system.[2] Equilibrium modeling is therefore justified.

The energy of formation term is the key to eq. 3.1 as it gives the probability for the defect to incorporate. It is clear, that a low energy of formation drives the exponential closer to unity, while a high energy of formation makes the exponential go to zero. In other words, employing techniques during growth to minimize the energy of formation for the desired dopants/impurities while maximizing the energy for the defects is the target.

The energy of formation is defined as:[2-5]

$$E^f(X^q) = E_{ref}(X^q) - \sum_j n_j \mu_j + q[E_F + E_V] \quad 3.2$$

E_{ref} is the free energy of the crystal containing a single defect in relative to an ideal crystal, n_j is the number of atoms exchanged with the reservoir to create the defect, μ_j is the chemical potential associated to this reservoir, and $E_F + E_V$ is the Fermi level referenced to the valence band. As already discussed in Section 2.2, the chemical potential of the growth process is

related to the growth conditions by the supersaturation. Here the choice of the chemical potential will depend on the defect being investigated.[4] For instance, the chemical potential of nitrogen is a more appropriate selection for the nitrogen vacancy. For a detailed exploration of the theoretical treatment of the reader is directed to Refs. 3-5. Here the purpose is to set the framework for the experimental data and not solve absolute values for the above equations. Instead, the focus will be looking at the changes to the defect concentration between two growth conditions, a reference level and a new growth based on changing one of the process variables. Therefore, a change in the defect concentration is related to a change in the energy of formation:

$$\Delta[X^q] \sim \exp(-\Delta E^f) \tag{3.3}$$

and the energy of formation change is related to a change in the chemical potential from a process condition change:

$$\Delta E^f_{ref \to 1}(X^q) \sim \Delta \mu_X \tag{3.4}$$

or by introducing above bandgap illumination:

$$\Delta E^f_{ss}(X^q) \sim (E_{FX} - E_F) \tag{3.5}$$

where E_{FX} is the quasi Fermi level of the defect subtracted from the Fermi level. This framework will be the lens to understand experimental observations of Mg:GaN and Mg:$Al_{0.6}Ga_{0.4}N$. Specifically, the different electrical resistivities and free hole concentrations obtained by changing the process parameters will be linked to experimental observations of the hydrogen concentrations in secondary ion mass spectrometry, and the nitrogen vacancy 2.8 eV blue luminescence in photoluminescence measurements.

While the total concentration of H can be determined, the concentration as a residual complex that compensates Mg is not clear. Similarly, the exact concentration of the nitrogen

vacancy (and complexes) is not measurable. Use of the charge balance equation can estimate the unknown compensator concentration, N_d:

$$p = N_A^- - N_D^+$$

$$p = \frac{[Mg]}{1 + 4\exp\left(\frac{\Delta E_a - (E_F - E_V)}{k_B T}\right)} - N_d$$

$$p = N_V(T)\exp\left(\frac{-(E_F - E_V)}{k_B T}\right)$$

$$-(E_F - E_V) = k_B \mathrm{T} \times \ln\left(\frac{p}{N_V(T)}\right)$$

$$N_d = \frac{[Mg]}{1 + 4\exp\left(\frac{\Delta E_a}{k_B T} + \ln\left(\frac{p}{N_V(T)}\right)\right)} - p \qquad 3.6$$

where p is the free hole concentration measured from Hall effect, N_A^- is the number of ionized acceptors, N_D^+ are the number of ionized donors, $[Mg]$ is the total Mg concentration measured from SIMS, ΔE_a is the ionization energy, and N_V is the vacancy concentration for a given temperature. This estimation will provide useful comparison of the effects of the above bandgap illumination and the chemical potential control on the electrical characteristics for the films.

3. Passivation and Compensation of Mg:GaN Grown by MOCVD

Obtaining p-type conductivity from Mg:GaN films was a pivotal breakthrough in the development of the III-nitrides, and was recognized as part of the 2015 Nobel Prize in Physics for the blue light emitting diode (LED). However, this conductivity is far from efficient. Typical Mg doping concentrations of ~2x10^{19} cm^{-3} yield free carrier concentrations in the low 10^{18} cm^{-3} range with resistivities below 1 Ω cm seldom demonstrated.[12] Despite thirty years of research, these

numbers have seen little improvement. Yet, any advances would translate into instant gains in efficiency for the commercial LEDs and laser diodes[8,9] which could reduce the cost of the devices, and may enable further adoption.

For a semiconductor, the method to increase the conductivity of a layer is to increase the mobility of the free carriers or their concentration. These often can compete because an increase in dopants/acceptors will cause the free carrier concentration to increase, but lower the mobility due to increased scattering. Conversely, lowering the dopant/acceptor concentration reduces the scattering but the concentration of carriers also decreases. Achieving maximum conductivity requires a compromise. In the case of III-nitrides, the holes have a high effective mass (low mobility)[13] and the ionization energy is greater than room temperature (incomplete ionization of the total acceptor atoms).[10] That eliminates the possibility for a low Mg concentration to yield high conductivity values.

However, the H-passivation and self-compensation of the $Mg_{(III)}$ further complicates this simplified picture. When growing Mg:GaN layers with MOCVD, the H concentration passivates the Mg until a total concentration of ~1-2x10^{19} cm^{-3}.[14,15] In comparison, unintentionally doped (UID) GaN or n-type doped GaN typically have a typical concentration of H (~1x10^{17} cm^{-3}) regardless of growth condition or doping. Only when magnesium is introduced as dopant during MOCVD growth, does the H increase above this level and match the total Mg. Fortunately, early pioneers discovered that a post growth low-energy electron beam irradiation (LEEBI)[6] or thermal activation (in a H-lacking ambient)[16,17] could lower the total H, and yield p-type conductivity for the Mg:GaN layers. That led researchers to conclude that below 1x10^{19} cm^{-3}, the majority of the Mg enters as the Mg-H complex, and its dissociation is the source of the p-type behavior.

Mentioned above, the H concentration does not increase past a Mg concentration of 1-2x10^{19} cm^{-3}. Yet, increasing the Mg concentration above this point actually yields a decrease in conductivity.[18-24] Photoluminescence (PL) measurements reveal a change from the typical donor-acceptor pair (DAP) dominant peak to a broad luminescence at an energy of 2.8 eV.[18-20] This blue luminescence was associated to some type of nitrogen vacancy complex such as Mg-V_N[19] or V_N-H.[21-22] Recent positron annihilation spectroscopy studies, however, reveal increases in the $V_{Ga}(V_N)_s$ (s = 1, 2, or 3) complex when the blue luminescence appears in the PL spectra.[22-23] Therefore, the decreased conductivity at high doping concentration is considered due to the self-compensation, the formation of vacancies and complexes during MOCVD growth. Unfortunately, no methods to date for decreasing the vacancy concentration during or after growth have been presented.

Here in this work, the goal was to understand the extent of the H-passivation and the self-compensation by framing the problem in terms of the defect energy of formation. As described in the previous section, this energy of formation can be influenced by the chemical potential of the growth environment as well as the defect's quasi Fermi level at the growth surface. During MOCVD growth, the former is accessed by altering the growth conditions like temperature, pressure, and the gas flows, while the latter is changed by photo generated current by above bandgap illumination. Both of these methods are employed in this study of the electrical properties of Mg:GaN layers.

The growth conditions for the study are as follows. All growths were performed at 1050° C, a total pressure of 20 Torr, and TEG flow rate of 67 μmol/min. The Mg:GaN layers were grown under N_2 or H_2 diluency (~7 slm), and the NH_3 flow was either 0.3 slm or 3 slm (a total flow rate change was present). The N_2 diluent growth with an NH_3 flow rate of 0.3 slm is considered the

reference sample and was presented in literature. The results will be reviewed here before presenting the results for the other conditions. Within each growth condition, multiple Mg total concentrations were generated by changing the Cp_2Mg flows from 0.15 to 0.75 μmol/min. For each doping level, an additional growth was performed with above bandgap light from a mercury-xenon arc lamp illuminating the sample surface. At a power density of ~1 W/cm^2 for the mercury-xenon arc lamp, the steady state carrier (holes and electrons) concentration of 10^{12} to 10^{14} cm^{-3} is expected. A post-growth thermal anneal was employed for the reference samples by annealing in a furnace at 650°C in flowing nitrogen for 2 hours. The other samples were annealed at 700°C in air ambient for 20 minutes. Repeat experiments confirm that both yielded similar performance for the same Mg doping levels.

SIMS analysis for the reference growth condition was realized using a CAMECA IMS-6f with magnetic sector analyzer. The error for all SIMS analyses is 20%, based on calibration against an ion-implanted standard for each species of interest. Analysis for H, C, O, and Si was achieved using Cs^+ primary beam and detection of negative secondary ions. The 15 nA primary beam was typically rastered over a 120 x 120 μm^2 area with ions detected from a 30 μm diameter region at the center of the raster. The analysis of Mg was made using O_2^+ primary beam with 50 nA current rastered over a 180 x 180 μm^2 area and positive secondary ions detected from a 60 μm diameter area at the center of the raster. The Mg concentrations were determined for the other growth conditions by measurement on an ION- TOF-SIMS with cesium source for sputtering and bismuth liquid metal ion gun (LMIG) source in non-interlaced sputtering mode for analysis. For the depth profiles, a Cs^+ beam with 10 keV energy and 25 nA current was rastered over a 120 x 120 μm area. The Bi^{3+} beam was 0.4 pA at 25 keV and rastered over a 50 x 50 μm area at the center of the sputtered crater. The angle of incidence was 45° from normal for both beams. The monitored

species were $CsGa^+$ and $CsMg^+$ in positive ion detection mode. Repeats of the reference sample were reanalyzed with the TOF system to ensure compatibility between the two instruments.

The photoluminescence measurements were performed at room temperature and 3 Kelvin using a 325 nm (56 mW) HeCd laser. The PL setup consisted of a Janis (SHI-RDK-415D) helium closed-cycle cryostat and a Princeton Instruments (SP2750 0.75 m) spectrometer with an attached PIXIS 2K charged-coupled-device (CCD) camera. Electrical characterization was performed on an Ecopia HMS-5500 AHT55T5 Hall effect measurement system. Room temperature resistivity measurements on the annealed films were performed using the van der Pauw method. Ni/Au contacts (20/40 nm thick, respectively) annealed in atmosphere for 10 min at 650°C were applied.

3.1 Mg:GaN Reference Conditions (N_2 Diluent, 0.3 slm NH_3)

To begin, the Mg:GaN films grown under the reference growth condition will be compared. To demonstrate the effectiveness of the light in suppressing the Mg-H, ladder doping samples were prepared for SIMS where the Mg concentration was introduced to a concentration of $3x10^{19}$ cm^{-3}, and lowered to a value of $6x10^{18}$ cm^{-3}. These represent Mg concentrations above and below the self-compensation onset of $1.5x10^{19}$ cm^{-3}, i.e. where blue luminescence in PL appears. Atomic concentrations of O, and Si were also measured by SIMS but they were below their corresponding background levels for the particular measurement: less than $1x10^{18}$ cm^{-3} and $1x10^{17}$ cm^{-3}, respectively. In addition, the atomic concentration of C was constant at ~$1x10^{18}$ cm^{-3} independent of UV-illumination. Consequently, C, O, and Si are not considered compensators in Mg:GaN.

Fig. 3.1 displays the results for the as-grown sample with and without illumination during growth (no post-growth thermal annealing) showing the Mg and H concentrations. The noise difference in the curves is a result of the sensitivity, the overall secondary ion counts during the SIMS measurement for each of the species (See Chapter 4 for details). The top plot displays the

sample grown without UV-illumination and has the highest concentration of H for both Mg doping levels of all the samples measured with SIMS. Moving right to left on the depth axis corresponds to the direction of growth. As Mg is introduced into the film, the H is seen rapidly increasing with the Mg until it saturates around ~1.5x10^{19} cm^{-3}. At approximately 550 nm in depth, the Cp_2Mg flow rate was decreased so the total Mg would decay. At that depth, the H starts to decrease and matches the total Mg as expected from H passivation.

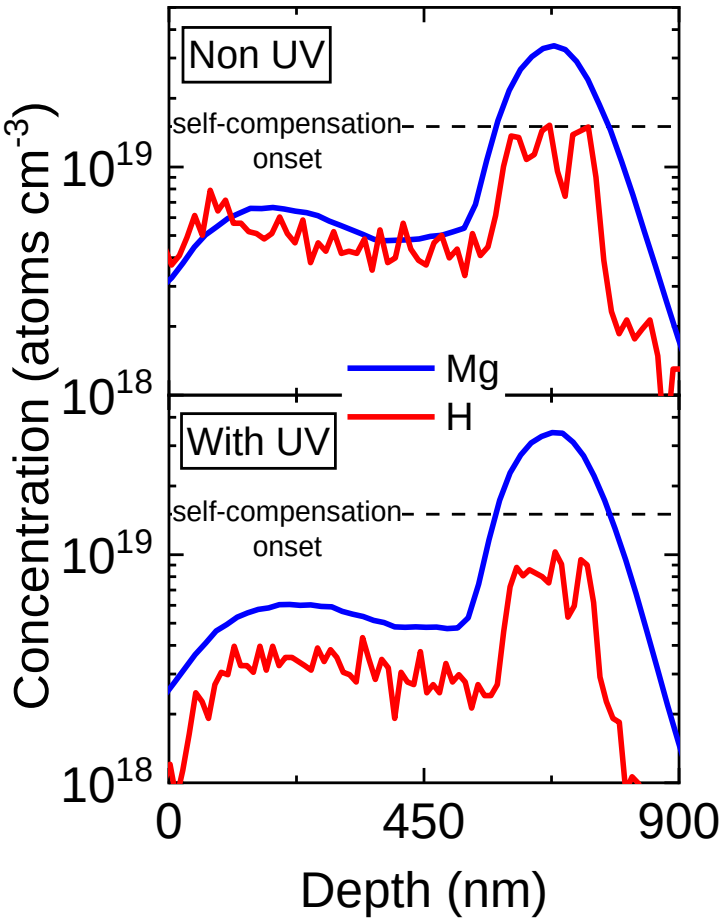


Figure 3.1 Cameca SIMS measurement of the Mg ladder structures grown using the reference conditions. Measurements acquired before a post growth thermal anneal (650° C 2 hours).

The lower plot in Fig. 3.1 displays the results for the sample grown under UV-illumination. It is clear that the overall Mg concentration does not change as a result of the above bandgap light, and is one of the tenets of defect quasi Fermi level approach to point defect reduction.[3,4] However,

the H concentration does change dramatically as result of the injection of the steady state carriers. Here, for the Mg concentration of $3x10^{19}$ cm^{-3}, the H leveled at a concentration of $8x10^{18}$ cm^{-3}. For the lower Mg concentration of $6x10^{18}$ cm^{-3}, the H level dropped to a concentration of $2\text{-}3x10^{18}$ cm^{-3} instead of matching the total Mg. Furthermore, for the Mg concentration above ~$2x10^{19}$ cm^{-3}, a reduction in the H level is observed down to a concentration of $8x10^{18}$ cm^{-3}.

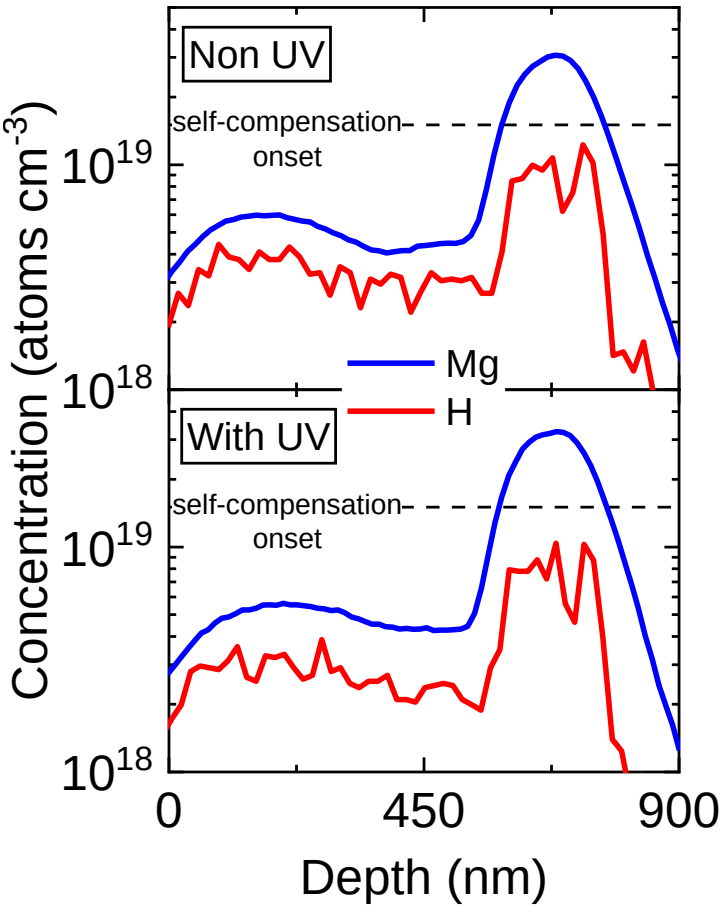


Figure 3.2 Cameca SIMS measurement of the Mg ladder structures grown using the reference conditions. Measurements acquired after a post growth thermal anneal (650° C 2 hours).

Fig. 3.2 shows the SIMS depth profiles for the samples presented in previous figure, after the samples were annealed for 2 hours at 650° C. After the annealing step, the Mg again remains at the same concentrations in both of the films. However, all H concentrations have been reduced to similar concentrations as presented in the lower plot of Fig. 3.1. This demonstrates that growth

under UV-illumination reduces the amount of H to a similar degree as the post growth annealing, but the combination yields no further reduction in total H. Electrically, it is not possible to test whether the UV-illumination is sufficient to provide conductive samples after growth because Ni/Au contacts require an anneal to provide conduction. Regardless, the H concentrations minimized to concentrations of $8x10^{18}$ and $2\text{-}3x10^{18}$ cm^{-3} for total Mg values of $3x10^{19}$ and $6x10^{18}$ cm^{-3}, respectively. This corresponds to a reduction of almost half compared to the concentrations in the top plot of Fig. 3.1. This remaining H could be present in a different charge state or form a different complex, such as bound to V_N (like the V_N-H complex),[21-22] making it stable against annealing or illumination during growth.

Individual Mg:GaN samples were then grown for the photoluminescence and electrical measurements using Hall effect under the van der Pauw geometry. The PL spectra were all acquired after the post-growth thermal activation, so that direct comparisons could be made to the electrical data. As mentioned in the opening of this section, appearance of 2.8 eV blue luminescence appears when the Mg concentration is increased above ~$1.5x10^{19}$ cm^{-3}.[18-20] This also corresponds to a decrease in the conductivity of the film and was associated to some form of V_N related complex. This was recently confirmed using positron annihilation spectroscopy, and the $V_{Ga}(V_N)_s$ (s = 1, 2, or 3) complexes were measurable at the higher Mg concentrations.[23,24] Therefore, the total Mg concentration of $1.5x10^{19}$ cm^{-3} is referred to as the self-compensation onset for Mg:GaN.

Fig. 3.3 displays the results for the room temperature photoluminescence of the Mg doping levels ranging from 1 to $5x10^{19}$ cm^{-3} for the reference growth condition. Two immediate observations can be made. The first, is that all samples have broad defect transition centered around a concentration of 2.2 eV. The origin is highly controversial and been contested as V_{Ga},[25] C_N,[26] or

C-related complexes such as C_N-O_N.[27] Since this peak remained invariant to the doping level and use of illumination, it is believed to be related to C since that concentration was the relatively high $1x10^{18}$ cm^{-3} regardless of the doping level. In the H_2 and N_2 diluent growths where the ammonia was increased to 3 slm, the total C in the films reduced to ~$1x10^{17}$ cm^{-3}, and this peak disappeared. It seems evident that the peak has a relationship to this total C in the film, but its presence did not correspond to the decreased conductivity in the films.

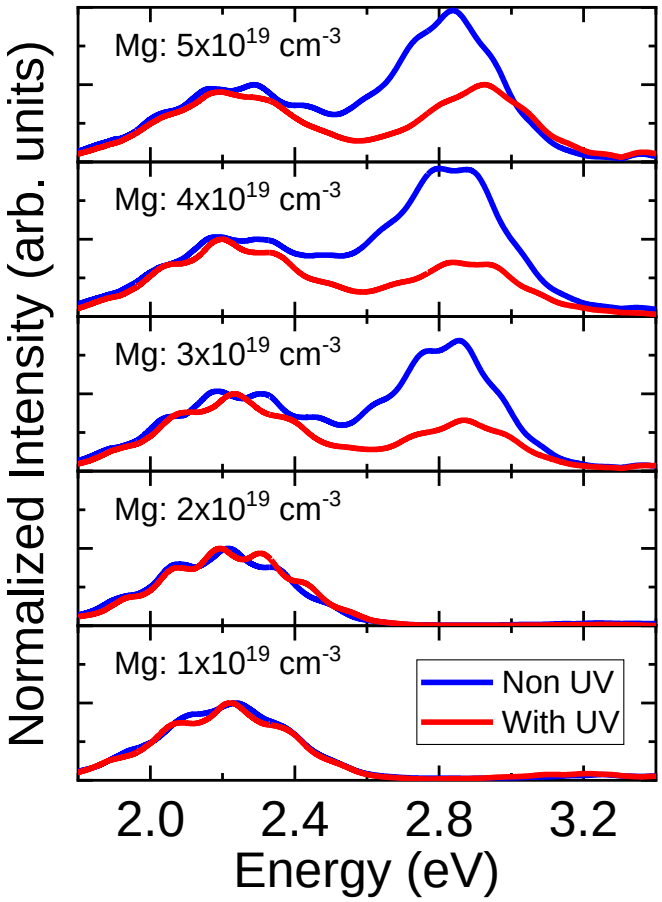


Figure 3.3 Room temperature photoluminescence of the Mg:GaN films grown using the reference conditions. Samples were normalized to the 2.2 eV luminescence. All acquired after post-growth thermal annealing step.

The second observation that is clear from Fig. 3.3 is that above a Mg doping concentration of $2x10^{19}$ cm^{-3}, a strong donor acceptor pair (DAP) luminescence at 2.8 eV appears.[28] The origin of the blue luminescence has been related to a V_N-Mg donor and Mg acceptor DAP transition,[29]

but more recent studies have shown a relationship to $V_{Ga}(V_N)_2$ or $V_{Ga}(V_N)_3$.[23,24] The corresponding decrease in conductivity as this peak increases is an indication for the onset of self-compensation in Mg:GaN. Under UV illumination, however, the samples doped above $2x10^{19}$ cm^{-3} show a diminished peak. Later electrical results will corroborate that the conductivity was superior for the UV-illuminated samples, and indicate the light leads to a reduction in self-compensation of Mg.

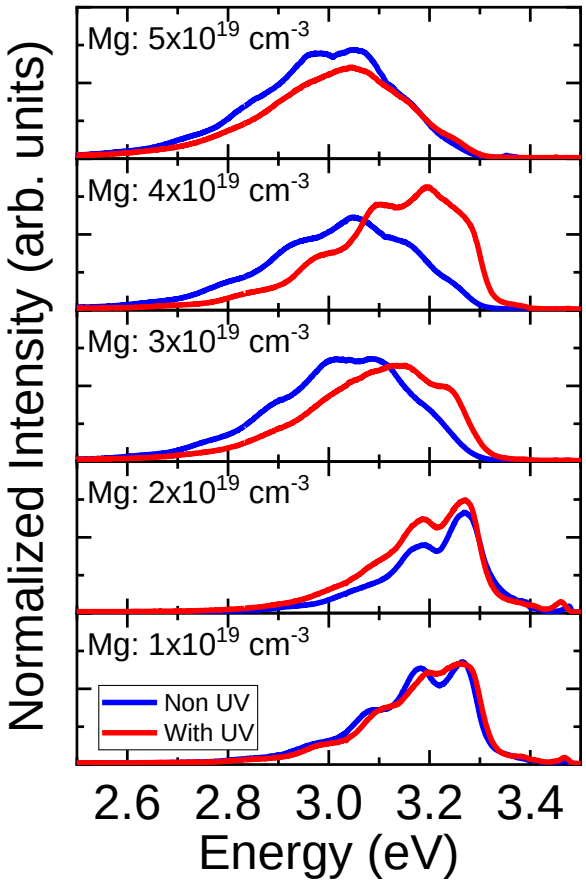


Figure 3.4 Low temperature (3K) photoluminescence of the Mg:GaN films grown using the reference conditions. Samples were normalized to the 2.2 eV luminescence (not shown). All acquired after post-growth thermal annealing step.

Similarly, Fig. 3.4 displays the spectra acquired at 3K for the samples presented in Fig. 3.3 and again the effect of the illumination is clear. Here the spectra have excluded the strong defect peak at 2.2 eV so that the prominent DAP transition at 3.27 eV related to Mg acceptors is observed.[29,30] Typically, this peak is observed because of the onset of self-compensation and

corresponds to the results in this study. Once the Mg doping exceeds the onset concentration, the non-illuminated samples show the transition from Mg acceptor DAP at 3.27 eV to the nitrogen vacancy related compensation peak at ~2.8 eV. However, the use of UV illumination delays the transition to this compensation luminescence. Illuminated samples with Mg concentrations of 3-4x10^{19} cm^{-3} do not have the strong emission at 2.8 eV and maintain the acceptor related transition at 3.27 eV. At the highest Mg doping level, the self-compensation related peak at ~2.8 eV dominates both spectra, but the UV-illumination reduces the intensity of this peak.

The final set of data that links the observations from SIMS and PL are the van der Pauw/ Hall effect results. Again, these measurements have been taken after a thermal annealing step as well as a post-contact deposition anneal for the Ni/Au contacts. Initially, the samples were tested as-grown with soldered In for contacts to avoid any potential "activation" by the post contact deposition anneal. The non-UV grown samples were resistive for measurement as expected from literature.[14-17] However, resistivity of the UV-grown samples was measured and values of 11, 3, and 16 Ω cm were obtained for Mg concentrations of 1, 2, and 3x10^{19} cm^{-3}, respectively. These values are higher than the samples measured with the Ni/Au contacts and treated with the post growth thermal activation in Table 3.1. It is impossible to determine if the worse conductivity is a result of incomplete "activation" or if the soldered In yielded poor ohmic contacts. Yet, it is still an important observation because it indicates that some of the Mg atoms are entering the film as acceptor dopants when the films are illuminated by above bandgap light.

Table 3.1 lists the resistivity, free hole concentration, and mobility for the samples with Mg concentrations ranging from 1 to 4x10^{19} cm^{-3}. Starting at a Mg doping level of 1x10^{19} cm^{-3}, the resistivity was reduced from 9 to 2 Ω cm, the free hole concentration increased by a factor of 2.5, and the mobility increased by 2.7 when the sample was illuminated during growth. Because

vacancy related defects are not expected until doping exceeds $1x10^{19}$ cm^{-3}, the improvement to the characteristics is suggested as an effect on the residual H present in the film. Consider the change in free hole concentration between the two films. In the non-UV growth, the free hole concentration is $8.4x10^{16}$ cm^{-3} and increases to $2.2x10^{17}$ cm^{-3} for UV-growth. Using these values in the charge balance eq. 3.6, the total compensation N_d was calculated to be $1x10^{18}$ cm^{-3} and $2.5x10^{18}$ cm^{-3}, respectively.

Table 3.1 Van der Pauw Hall effect of activated Mg:GaN samples grown with the reference conditions (N_2 diluent, NH_3 0.3 slm). Relative standard deviation was 8%, 10%, and 15% for the resistivity, hole concentration, and mobility, respectively.

	Mg Concentration (atoms cm^{-3})	$1x10^{19}$	$2x10^{19}$	$3x10^{19}$	$4x10^{19}$
Resistivity (Ω cm)	non-UV	9	1.5	12	36
	with UV	2	1.4	8	11
Hole Concentration ($x10^{16}$ cm^{-3})	non-UV	8.4	41	9.4	4.7
	with UV	22	47	10	7.5
Mobility (cm^2 V^{-1} s^{-1})	non-UV	8.3	11	6.7	3.7
	with UV	14	9.5	7.8	7.6

Above the compensation onset level (Mg ~$2x10^{19}$ cm^{-3}), samples also demonstrated a clear improvement to their electrical characteristics. At the Mg doping concentration of $3x10^{19}$ cm^{-3}, the measured resistivity decreased from 12 to 8 Ω cm., with slight increases in the hole concentration and mobility. The calculated compensation concentration was ~6-$7x10^{18}$ cm^{-3} for both UV and non-UV grown samples. A more pronounced effect was observed for the $4x10^{19}$ cm^{-3} Mg doping level, where the resistivity decreased to 11 Ω cm from 36 Ω cm, and an increase in free hole concentration by a factor of 1.5 and mobility by 2 was observed. Since the H concentration saturates well before these doping levels, the change in resistivity is interpreted as a decrease in the V_N-related compensation correlated to the diminished blue luminescence signals in Figs. 3.3

and 3.4. Via the charge balance equation, the number of compensators was reduced from ~$3x10^{19}$ cm^{-3} to $1x10^{19}$ cm^{-3}, respectively, when the illumination was applied during growth.

Surprisingly, no enhancement of the electrical performance was observed at the minimum resistivity value (Mg total concentration of $2x10^{19}$ cm^{-3}). From the charge balance equation, the estimated compensation concentration for both non-UV and UV-grown samples is ~$1.2x10^{18}$ cm^{-3}, only slightly larger than the compensation concentration for the UV-grown sample with an Mg concentration of $1x10^{19}$ cm^{-3}. Examining the low temperature PL data at the $2x10^{19}$ cm^{-3} Mg level, the with-UV sample appears to have a slightly broader, less defined DAP signal at ~3.2 eV. As already discussed, the broadening of the signature and shift towards the blue luminescence is a signal for increased V_N-related complexes and lower conductivity. Considering the SIMS results, the total H was suppressed during the growth to approximately $8x10^{18}$ cm^{-3}, which can also be attained for the non-UV grown film after thermal activation. The following hypothesis is proposed: the vacancy related complexes only form at growth temperatures and in response to the "free" Mg ionized acceptors that cause the film to be p-type conductive. At lower Mg doping levels, the above bandgap light suppresses H incorporation and allows free Mg ionized acceptors. However, the lower total Mg does not yield sufficient ionized acceptors (as a result of the ionization energy) to generate a high enough probability for vacancy complexes to form. At the self-compensation onset Mg doping level, however, the total number of Mg acceptors "free" from the H-passivation are sufficiently large to start generating the vacancy-complexes during the growth.

It is clear from the discussion of the low and high Mg-doped films, the residual compensators once self-compensation begins ($1x10^{19}$ cm^{-3}) are greater than the residuals from H ($1x10^{18}$ cm^{-3}). Delaying the onset of self-compensation could be the solution to achieve higher conductivity Mg:GaN. This discussion highlights the difficulties in obtaining p-type behavior in

GaN films, but the use of above bandgap illumination offers a promising glimpse in how changes to the growth process can influence the concentration of the unwanted defects by changing their formation energy during growth. Above it is proposed that the illumination changes the defect's quasi Fermi level for H (both as the total concentration that can be removed through thermal annealing, and whatever residual configurations) and the V_N-complexes characterized by the 2.8 eV luminescence in PL. In both instances, the reduction in these defects yielded higher hole concentrations and mobilities which lead to reduced resistivities for the films.

3.2 Mg:GaN N_2 Diluent, 3 slm NH_3

In the previous section, the reference growth conditions (N_2 diluent, NH_3 ~0.3 slm, TEG 67 μmol/min, total flow 7.2 slm, total pressure 20 Torr, and 1050° C) yield a Ga vapor supersaturation value of ~60. As mentioned in the beginning of this section, the vapor supersaturation is linked to the chemical potential of the growth environment, and this ultimately influences the probability for a defect to incorporate into the growing film. Here, that hypothesis is tested by changing the growth conditions first by increasing the NH_3 to 3 slm (total flow increased to 10 slm) and second by changing to H_2 diluent and maintaining the 3 slm of NH_3. These two conditions result in Ga supersaturation values of 55 and 225, respectively. Note, additional growths of Mg:GaN with 0.3 slm NH_3 under H_2 conditions were attempted (σ_v ~30), but were completely resistive and lacked luminescence in PL. They are considered of too poor quality for further consideration.

Similar to the reference conditions, several different Mg total concentrations were grown for each condition, along with repeats using above bandgap illumination. At this time, only three doping levels have been completed and these conditions do not have SIMS data for the H concentrations. The Mg concentration has been confirmed at this time so that direct comparisons

can be made to the reference samples. Fig. 3.5 displays the PL spectra acquired at room temperature for the N_2 diluent, 3 slm NH_3 condition. From the SIMS results, the Cp_2Mg flow of 0.3 μmol/min yielded a total Mg concentration of $5x10^{18}$ cm^{-3}, instead of the $2x10^{19}$ cm^{-3} condition measured for the reference (the other concentrations match expectation from molar flow). The difference appears as a result of the bubbler pressure and gas flow through the bubbler being halved in comparison to the reference growths (molar flow still is constant). This lower gas flow is closer to the limits of the mass flow controller on the line, and may lead to less Mg being introduced to the reactor. Regardless, the concentrations were consistent for both non- and UV-illuminated growths, and the doping concentrations were constant throughout the layer. They are therefore considered valid for the purposes of this study in investigating the H-passivation side of the Mg:GaN films. Future growths are required to achieve the low 10^{19} cm^{-3} concentrations in the other studies.

Beginning at this lower concentration of $5x10^{18}$ cm^{-3} Mg, the PL reveal the expected DAP centered around an energy of ~3.2 eV and the presence of near band edge emission at 3.4 eV. The UV-illuminated growth resulted in less signal intensity overall. Here, no broad C-related peak at 2.2 eV is present, likely a result of the concentration reduction to $1x10^{17}$ cm^{-3} for the growth condition. As a result, these extra emissions were not hidden by the strong signal present in Fig. 3.3. Since all spectra were acquired at the same condition, the luminescence intensity is displayed. The lowest Mg doped level in in Fig. 3.4 yielded a strong DAP around 3.2 eV with no band emission, and the corresponding UV-illuminated sample's spectra appeared slightly broader. However, this doping concentration $1x10^{19}$ cm^{-3} is much closer to the self-compensation onset of ~$1.5x10^{19}$ cm^{-3}. As mentioned in the reference growth section, that level of Mg may start allowing vacancy complexes to form, which could have broadened the peak. So, comparison to the $5x10^{18}$

cm^{-3} level in Fig. 3.5 may not be appropriate. At the higher Mg doping levels of 3 and $4x10^{19}$ cm^{-3} in Fig. 3.5, all the spectra are dominated by the blue luminescence at 2.8 eV related to the self-compensation by vacancy complexes. Once again, the UV-illuminated growths yielded lower peak intensities but they also do not display as strong of near band edge emission around 3.4 eV.

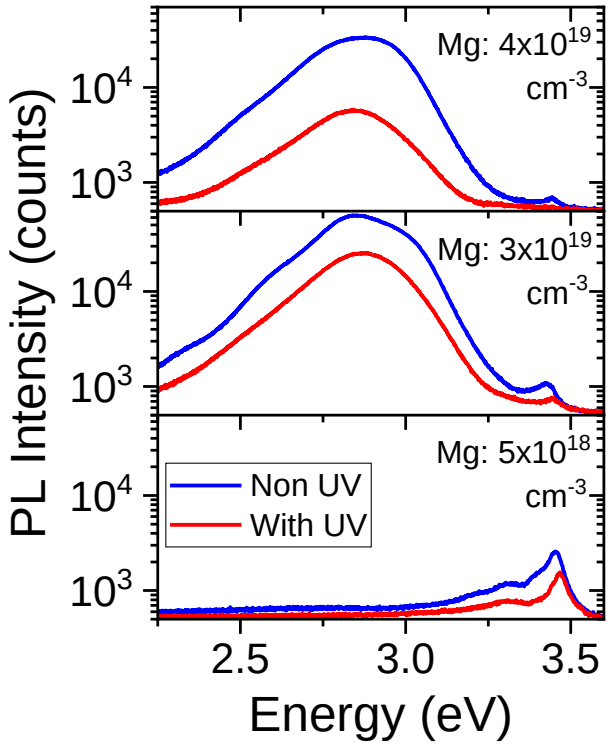


Figure 3.5 Room temperature photoluminescence of the Mg:GaN films grown using the N_2 diluent growth conditions (NH_3 3slm).

Table 3.2 presents the van der Pauw Hall effect results corresponding to the samples measured in Fig. 3.5. At the lowest Mg doping level of $5x10^{18}$ cm^{-3}, the non-UV sample was unable to be measured with the Hall effect system, despite the post growth thermal activation (700° C, 20 minutes in air) and contact anneal (650° C, 10 minutes in air). In contrast, the UV-illuminated growth for the Mg concentration of $5x10^{18}$ cm^{-3} was measurable in the Hall effect system and yielded a resistivity of 11 Ω cm, free holes concentration of $6x10^{16}$ cm^{-3}, and a mobility of 9.4 cm^2 V^{-1} s^{-1}. These values are somewhat similar to non-UV, $1x10^{19}$ cm^{-3} Mg concentration in Table 3.1,

and are not unreasonable considering the total Mg here is half. However, the calculated compensation level is ~$2x10^{18}$ cm^{-3}, which is 2x higher than the compensation for the UV-grown reference sample ($1x10^{18}$ cm^{-3}).

Table 3.2 Van der Pauw Hall effect of activated Mg:GaN samples grown using the N_2 diluent, NH_3 ~3 slm conditions. Relative standard deviation was 8%, 10%, and 15% for the resistivity, hole concentration, and mobility, respectively.

	Mg Concentration (atoms cm^{-3})	$5x10^{18}$	$3x10^{19}$	$4x10^{19}$
Resistivity (Ω cm)	non-UV	*	3.1	9.3
	with UV	11	2.3	9.6
Hole Concentration ($x10^{16}$ cm^{-3})	non-UV	*	15	7.8
	with UV	6	41	12
Mobility (cm^2 V^{-1} s^{-1})	non-UV	*	14	8.7
	with UV	9.4	6.7	5.4

*sample was too resistive for measurement

Moving to the higher doping concentrations of $3x10^{19}$ cm^{-3}, considered above the self-compensation onset because of the presence of the 2.8 eV luminescence in PL, the electrical performance corroborates the diminished 2.8 eV peak in the PL spectra for the UV-illuminated samples. The resistivity of the films decreased from 3.1 to 2.3 Ω cm, with a 2.5x increase in free hole concentration, but the mobility dropped by nearly double. Calculations from the charge balance equation indicate a reduction in the compensation concentration from ~$5x10^{18}$ cm^{-3} to ~$2.5x10^{18}$ cm^{-3} when the above bandgap light was employed. The $4x10^{19}$ cm^{-3} Mg concentration shows a slight increase in the resistivity, 9.3 to 9.6 Ω cm, when the illumination was utilized. This is somewhat surprising since the PL showed a reduction in the 2.8 eV luminescence, but this is within the error for the resistivity measurement. Yet the trend of increased holes (by 1.5x) but decreased mobility (by 1.5x) was observed. As a result, the calculated compensation reduced from $1x10^{19}$ cm^{-3} to $7x10^{18}$ cm^{-3}.

In general, these higher doping concentrations do not show a large improvement between the conventional and the UV-grown films amongst themselves. Yet, both the conventional and UV grown films with NH_3 flow of 3 slm outperform the electrical behavior of the reference conditions, which is reflected in the reduction to the calculated compensation concentration. This result is surprising because the supersaturation slightly decreased from 60 to 55 from the increase in NH_3 and total flow. As a result, the energy of formation for the defects would decrease via the decrease in the Ga chemical potential. However, the reservoir for the V_N-related defect may not be described by this chemical potential alone. Regardless, the increase in ammonia alone results in a greater reduction in the compensation than using the above bandgap light at the reference condition, and indicates the chemical potential has a stronger influence on these compensators.

3.3 Mg:GaN H_2 Diluent, 3 slm NH_3

The final set of data for Mg:GaN are the growths performed under H_2 diluent conditions, but maintain the 3 slm NH_3 flow. From the SIMS, the total Mg concentrations have been changed relative to the reference growth conditions. Here the Cp_2Mg flows of 0.3, 0.45, and 0.6 μmol/min yielded total Mg concentrations of $5x10^{18}$, $1.5x10^{19}$, and $4.5x10^{19}$ cm^{-3}, respectively. Fig. 3.6 displays the room temperature luminescence for these films measured after the post growth thermal activation. Similar to the higher ammonia, N_2 diluent growth, no C-related peak is present around ~2.2 eV because of the reduced C concentration in the layer (~$5x10^{17}$ cm^{-3}). Since all of the films were analyzed with the same conditions, the overall count intensity will be used to compare the films. At the lowest Mg concentration of $5x10^{18}$ cm^{-3}, both the conventional and illuminated samples show a strong near band emission ~3.4 eV and DAP around ~3.2 eV and no clear distinction is made between the samples, similar to the reference samples. For the $1.5x10^{19}$ cm^{-3} Mg doping level, the UV and non-UV grown samples have a stark difference in PL emission. The

non-UV sample has a strong DAP emission at 3.2 eV as well as the band edge near 3.4 eV. The UV-illuminated sample by contrast has a lower band edge and a strong vacancy related peak near 2.8 eV. From the later electrical measurements, vacancy-related compensators are expected for this growth. For the highest doping concentration of $4.5x10^{19}$ cm^{-3}, both samples display a strong 2.8 eV luminescence expected for the doping level. For the N_2 diluent samples, the UV-illumination reduced this 2.8 eV emission, and was correlated to improved resistivity. Here, no difference suggests the vacancy concentration is similar, which will be later shown in the electrical characteristics.

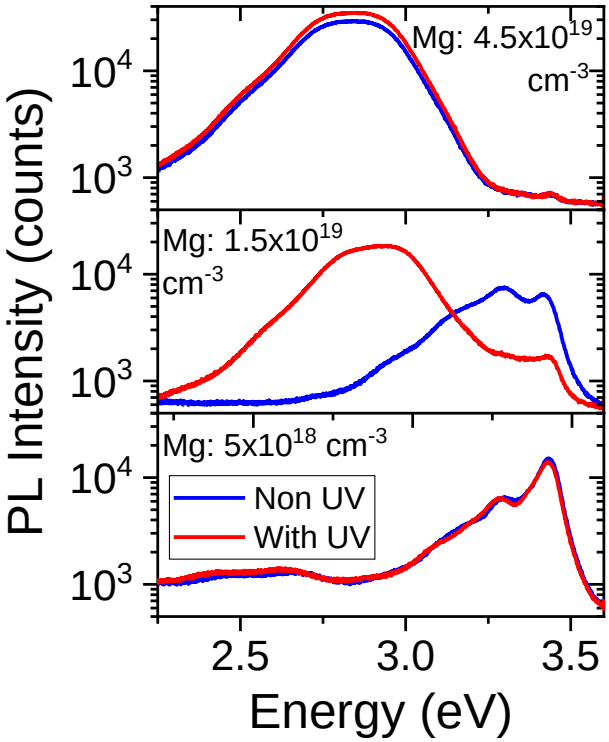


Figure 3.6 Room temperature PL of the Mg:GaN films grown using the H_2 diluent growth conditions (NH_3 3slm).

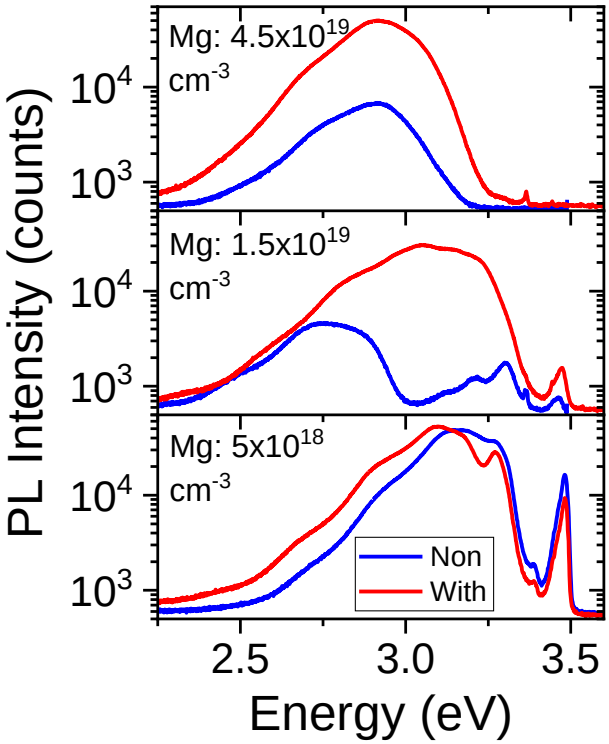


Figure 3.7 Low temperature (3K) photoluminescence of the Mg:GaN films grown using the H_2 diluent growth conditions (NH_3 3slm).

The samples analyzed in Fig. 3.6 also had low temperature photoluminescence measured, which is presented in Fig. 3.7. At the low Mg concentration of $5x10^{18}$ cm^{-3}, very sharp band edge emission is now separated from DAP. For the conventional growth, none of the individual peaks comprising the DAP around 3.2 eV are visible, but the UV-illuminated has the 3.2 and 3 eV peaks clearly separated. The later electrical results do show a greater difference in behavior, but it is not clear from these PL spectra. The Mg doping concentrations of $1.5x10^{19}$ cm^{-3}, on the other hand, showed a stark contrast in the luminescence of the film. Both films had a reduced near band edge emission, but the lower energies are quite different. The conventional growth had the DAP peaks at 3.2 and 3 eV drop significantly, while the peak at 2.8 eV rose compared to the lower Mg doping level. The H saturation limit was expected around this Mg level of $1.5x10^{19}$ cm^{-3}, and the 2.8 eV luminescence did not appear in the conventional doping level until the higher Mg doping level of

$3x10^{19}$ cm^{-3} in the reference (Fig. 3.4). Without SIMS for this H_2 diluent case, it is not certain that the H saturation limit changes, and therefore the onset of self-compensation, but the spectra indicate the presence of some vacancy-related defect. For the UV-illuminated sample, the DAP peaks at 3.2 and 3 eV are still strong but are not as clearly defined as the $5x10^{18}$ cm^{-3} sample, and it also has a strong luminescence around 2.8 eV adding to the overall broad peak. Combined with the room temperature PL for this sample, it is expected that some vacancy related compensation is present in the film. Similarly, the Mg doping level of $4.5x10^{19}$ cm^{-3} in Fig. 3.7 yield strong 2.8 eV signatures that confirm the observations from the room temperature measurements.

Table 3.3 presents the electrical characteristics for the H_2 diluent samples in Figs. 3.6 and 3.7. Beginning with the Mg concentration of $5x10^{18}$ cm^{-3}, the use of above bandgap light resulted in a resistivity reduction from 27.4 to 2.6 Ω cm. Unfortunately for the conventionally grown sample, Hall effect measurement was conflicted so the mobility and free carrier concentration were invalid. For the UV grown sample, the calculated compensation concentration was ~$8x10^{17}$ cm^{-3} which was the lowest compensation level in this study. Overall, these values match the reference sample (just with a slightly higher Mg concentration), but outperform the higher NH_3 flow.

Table 3.3 Van der Pauw Hall effect of activated Mg:GaN samples grown using the H_2 diluent, NH_3 ~3 slm conditions. Relative standard deviation was 8%, 10%, and 15% for the resistivity, hole concentration, and mobility, respectively.

	Mg Concentration (atoms cm^{-3})	$5x10^{18}$	$1.5x10^{19}$	$4.5x10^{19}$
Resistivity (Ω cm)	non-UV	27.4	1.3	12.4
	with UV	2.6	1.6	13.1
Hole Concentration ($x10^{16}$ cm^{-3})	non-UV	*	37	6.6
	with UV	11	28	6.7
Mobility (cm^2 V^{-1} s^{-1})	non-UV	*	13	7.6
	with UV	21.4	13.8	7.1

*Hall effect measurement was conflicted

Moving the $1.5x10^{19}$ cm^{-3} Mg doping level, the resistivity actually increased from 1.3 to 1.6 Ω cm. This difference is considered real because measurements across an individual 2 inch wafer (not shown) showed a < 10% relative error in the resistivity, and these values between these two growths are outside that range. Combined with the PL measurements, which showed stronger 2.8 eV luminescence for the with-UV sample, this resistivity change is considered a real difference. Similar to the reference growth condition, this indicates that near the self-compensation onset, the above bandgap light suppressed the H total concentration, and the H residuals, and resulted in V_N-related compensators associated with self-compensation. From the charge balance equation, the compensation concentration increased from 1 to ~$1.5x10^{18}$ cm^{-3}, and was the only instance where the UV caused higher compensation.

Lastly, the Mg doping concentration of $4.5x10^{19}$ cm^{-3} showed virtually no difference in resistivity, hole concentration, or mobility between the conventional growth and UV-illumination. Similarly, the calculated compensation concentration was ~$1x10^{19}$ cm^{-3} for each of the growths. This follows the observations from the N_2 diluent, NH_3 ~ 3 slm, growths, where the higher Mg doping levels showed a greater reduction vs. the reference growth condition, rather than by using above bandgap illumination. Since these H_2 diluent growths have a slightly higher concentration of Mg, they resulted in higher values overall vs. that N_2 diluent 3 slm NH_3 condition, and the illuminated reference sample. Additional growths matching the total concentrations are required for a true comparison, but the trend of the higher NH_3, and this case a higher Ga supersaturation, seem to affect the self-compensation greater than the above bandgap illumination alone.

In summary, the H-passivation and the self-compensation of Mg:GaN grown via MOCVD has been thoroughly investigated. By invoking the energy of formation arguments for the defects causing these effects, a framework is provided to understand the observed changes in SIMS, PL,

and van der Pauw Hall effect and their relationship to the growth conditions. Here, the diluency (N_2 vs H_2), flow of NH_3 (also total flow), Mg total concentration, and a steady state injection of holes and electrons by above bandgap light, were the variables utilized. When the Mg total concentration remained below a concentration of $1.5x10^{19}$ cm^{-3}, the diluency and NH_3 flow had little effect on the residual donors that remain after half the H concentration is removed via thermal annealing. Instead, growing films with above bandgap illumination resulted in the same reduction in total H and additionally removed some other residual donor concentration, resulting in a much lower resistivity and higher free hole concentration. SIMS for the higher NH_3 growths is required to confirm the total H reduction via the light, but it appears that the reference and H_2 diluent conditions yield the lowest resistivity and highest free hole concentrations.

For the Mg doping concentrations near the self-compensation onset of $1.5x10^{19}$ cm^{-3}, the use of above bandgap illumination either had no effect (the reference growth) or leads to a lower self-compensation onset (H_2 diluent). In samples without UV-illumination, the H concentration matches Mg until Mg is doped above a $1.5x10^{19}$ cm^{-3}. It is hypothesized that during growth, the H forms a complex with Mg that remains neutral and inhibits the formation of compensating nitrogen vacancy complexes. Above bandgap illumination was shown to suppress the total H to half of the Mg concentration during growth (again up to the Mg doping level of $1.5x10^{19}$ cm^{-3}). However, no vacancy complex signature is present in PL. When the sample is Mg doped to $1.5x10^{19}$ cm^{-3} and illumination is utilized, the steady state condition promoted the formation of V_N-related compensators that increased the resistivity. In conventional growths, this condition is met when the total Mg is greater than $2x10^{19}$ cm^{-3}.

For the higher Mg doping levels beyond a concentration of $3x10^{19}$ cm^{-3}, nitrogen vacancy related compensators increase in the film, and clear defect luminescence around 2.8 eV is present

in the PL spectra. For the reference growth condition, the intensity of this peak was reduced by the introduction of above bandgap light during growth. A corresponding decrease in the resistivity was also observed. However, the higher NH_3 flow growths all yielded resistivities below that reference, and above bandgap illumination provided no additional benefit. The diluency also had no influence on the resistivity. Therefore, the greater influence on the self-compensation related defects appears to be the growth conditions, i.e. the chemical potential. When discussing the chemical potential of the growth in Section 2.1, only the Ga supersaturation, and its associated chemical potential, was considered. Here the defect concentration appears to the follow this trend for the H_2 diluent growths, but not for N_2. At this time, it is clear that the increased NH_3 plays a role, but translating that into the chemical potential of the defect needs further investigation.

4. Mg doping of $Al_{0.6}Ga_{0.4}N$

Aluminum gallium nitride ($Al_xGa_{1-x}N$) has tremendous potential for use in deep-ultraviolet (UV) optoelectronic devices because the emission wavelength can be tuned across the entire UVA, UVB, and UVC spectrum. This spectral range has many applications like sterilization, UV curing, biomedical instrumentation, as well as communication.[31] To date many groups have demonstrated light-emitting diodes (LEDs) with emission wavelengths below 300 nm by using AlGaN, albeit with significantly lower external quantum efficiencies than their visible LED counterparts.[32-35] Similarly, optical pumping of AlGaN multiple-quantum-wells (MQW) below 300 nm has produced lasing,[36-38] and very recently electrically injected 271 nm wavelength lasers have been demonstrated.[39] However, significant improvements are required before these devices are viable enough to replace current UV sources.

The major limitation to the AlGaN-based LEDs and laser diodes is the conductivity of p-type layer.[31] This stems from the high ionization energy of the Mg_{III} acceptor and the compensating

point defects that are incorporated during the MOCVD growth. As mentioned in the introduction, a linear increase in the Mg_{III} acceptor ionization energy was demonstrated when the Al composition x in $Al_xGa_{1-x}N$ is increased.[40] As the ionization energy in GaN is greater than room temperature, and even further decrease in the number of free holes is expected for $Al_{0.6}Ga_{0.4}N$. In the worst case scenario, AlN, the Mg_{Al} acceptor ionization energy was calculated with a value of ~590 meV.[40] Consequently, a small fraction of the typical 10^{19} cm^{-3} total Mg atoms is ionized at room temperature. Also demonstrated in the previous section, compensation of the Mg ions plays a pivotal role in determining the electrical performance of the films. With increasing bandgap, the point defect energy of formation is expected to decrease.[41] The combined effects limit resistivity values in the hundreds of Ω cm for high Al content Mg-doped AlGaN.[40]

Alternative approaches to bulk epi-doping have been attempted to achieve a higher free carrier concentration (lower resistivity) for Al-rich AlGaN films. One approach utilized $Al_xGa_{1-x}N$/p-$Al_yGa_{1-y}N$ ($0\leq x\leq 1$, $0\leq y\leq 1$) superlattice structures.[42-46] These structures generate a large sheet charge due to the polarization discontinuity at the heterojunction of different composition $Al_xGa_{1-x}N$ alloys. However, conductivity perpendicular to the plane is poor because the potential barriers in the valence bands and the high effective mass of holes lead to a very low mobility through the superlattice.[42-43,45,47] Another approach utilized Mg doped films with graded Al-content AlGaN layers, termed distributed polarization doping (DPD).[39,47-49] Because of the graded Al composition, a three-dimensional, polarization-induced charge is spread over the entire layer, increasing the hole concentration. Recent work has demonstrated lasing with an undoped AlGaN grading layer.[39] At this time, the physical explanation is not clear. However, both superlattice and graded structures rely on a conducting p-GaN layer for p-contact formation. Since this layer will absorb any emission below 364 nm (the entire UV region), it compromises light extraction from the active region, or

introduces losses in waveguides. In short, obtaining a low resistivity p-AlGaN layer would simplify the design challenges facing the next generation devices.

The work of Kinoshita et al. reopened the possibility for highly conductive bulk doping of high Al-content AlGaN. Under their optimal growth conditions, they demonstrated a reduced ionization energy of ~47-80 meV at temperatures below 500 K for p-type $Al_{0.7}Ga_{0.3}N$.[50] This ionization energy is 5x lower than previous reports, but surprisingly 2x lower than the value obtained for GaN, which was previously considered to yield the lowest ionization energy for the Mg acceptor. In addition, they demonstrated room temperature resistivity of 47 Ω cm for the optimized sample, which is orders of magnitude lower than previously measured.[50] They were unable, however, to determine free hole concentrations or mobility. It is suggested that alloy scattering reduced the mobility and made Hall effect measurement impossible.[51] However, the combined low ionization value and resistivity suggest the compensation may not have as profound.

Following the work of Kinoshita, et al., the growth of Mg:$Al_{0.6}Ga_{0.4}N$ was investigated to determine the influence of the growth conditions on residual compensation in the films. Similar to the previous study in GaN, the supersaturation of the growth condition was varied to see the influence on the point defects using SIMS, photoluminescence, and van der Pauw/ Hall effect. Here only the ammonia flow was changed to vary the Ga supersaturation from 8 to 90.[51] All samples were grown on AlN/sapphire templates, at low pressure (20 Torr), and 1050° C temperature during growth of Mg:$Al_{0.6}Ga_{0.4}N$ films. For the $\sigma \sim 8$ condition, the H_2 flow was 10 slm, NH_3 was 1 slm, trimethylaluminum (TMA) was 6.5 μmol per minute, and triethylgallium (TEG) was 9 μmol per minute. For the $\sigma \sim 90$ growth conditions, the H_2 flow was 10 slm, NH_3 was 3 slm, TMA was 33 μmol per minute, and TEG was 36.5 μmol per minute. Note the change in total flow was accounted for in the calculation. TMA and TEG were increased to maintain a

similar growth rate between the conditions (500 nm per hour), and similar composition. Likely the higher NH_3 had significant pre reaction in the gas phase, but was not included in the supersaturation calculation. Further discussion will be provided later. Mg total concentrations ranging from $5x10^{18}$ to $1x10^{20}$ cm^{-3} were introduced into each of these conditions by flowing bis-(cyclopentadienyl)magnesium (Cp_2Mg) from 0.3 to 1.2 μmol/minute. Ladder structures of three different Cp_2Mg flows were used for secondary ion mass spectrometry (SIMS) evaluation, while the optical and electrical properties were performed on a uniformly doped 0.6 μm thick layer.

To determine the concentrations of H, C, O, Mg, Si as well as the Al to Ga ratio, ladder structures were analyzed using quadrupole-SIMS at EAG, Inc. comparing samples as-grown and a set that underwent an activation anneal. Activation annealing in Mg:GaN is a requirement to measure p-type behavior, and is linked to a reduction in the H concentration in the film. Following Kinoshita, et al. activation was performed in flowing N_2 at 20 Torr pressure for 20 minutes at a temperature of 850° C. After SIMS, and corresponding PL and resistivity measurements, it was clear that this step to remove H was unnecessary. For the uniformly doped films, an ION-TOF TOF-SIMS V instrument was used for ToF-SIMS analyses to determine Mg and the group III atom fraction. Re-analysis of the ladder structures and matching to the EAG data allowed for consistency in the reported concentrations. The photoluminescence (PL) measurements of as-grown samples were performed at room temperature and 3 K using a 325 nm (56 mW) HeCd laser. The PL setup consisted of a Janis (SHI-RDK-415D) helium closed-cycle cryostat and a Princeton Instruments (SP2750 0.75 m) spectrometer with an attached PIXIS 2K charged-coupled-device (CCD) camera.

Electrical characteristics were tested by van der Pauw/ Hall effect measurement. Prior to contact deposition, samples were cleaned using acetone, methanol, and deionized water, followed by dipping in 1% HF and hot 1:1 $HCl:H_2O$ solution. Van der Pauw geometry pads were masked

in the corners. Standard Ni (20 nm)/Au (40 nm) contacts were deposited using e-beam evaporation in UHV (base pressure $1x10^{-9}$ Torr). This step was followed by contact annealing at 600º C for 20 minutes in air ambient. I-V and Hall measurements were performed on an Lakeshore Cryogenics 8404 AC/DC Hall Effect Measurement System.

4.1 Mg:AlGaN SIMS

In the case of Mg-doped GaN, the H concentration in the film plays a pivotal role in determining the performance of the device. During MOCVD growth, it passivates Mg to form the Mg-H complex[14,15,20] and leads to as-grown resistivities greater than $1x10^6$ Ω cm. With a post-growth thermal activation,[16,17] the H concentration reduced by one-third,[11,15] and the resistivity drops to <1 Ω cm.[12] This Mg-H complex saturates around a total Mg concentration of ~$2x10^{19}$ cm^{-3}, however, further increases in doping lead to self-compensation by nitrogen vacancies and increased resistivity. To date no technique has been developed to reduce V_N concentration post-growth. Thus, the minimal resistivity in p-type GaN is restricted at an Mg total concentration of $2x10^{19}$ cm^{-3} – where the difference between Mg and H concentration is greatest, but before the onset of V_N-complex formation.

Determining the H concentration is key to interpreting the optical and electrical performance of the Mg-doped AlGaN. Ladder structures of different Mg doping concentrations were grown for SIMS analysis to determine the saturation level of H in the AlGaN films. Fig. 3.8 displays the results for the Mg and H concentrations for both σ conditions from EAG. It should be noted that the concentrations of C, O, Si, and the Al/Ga group III composition were also measured. However, no discernable difference between the structures was made. The composition was ~60% and the C, O, and Si concentrations were $8x10^{16}$, $1.5x10^{17}$, and $2x10^{16}$ cm^{-3}, respectively,

throughout both samples. These values are all orders of magnitude lower than the H and Mg and are not considered to have an effect on the electrical performance.

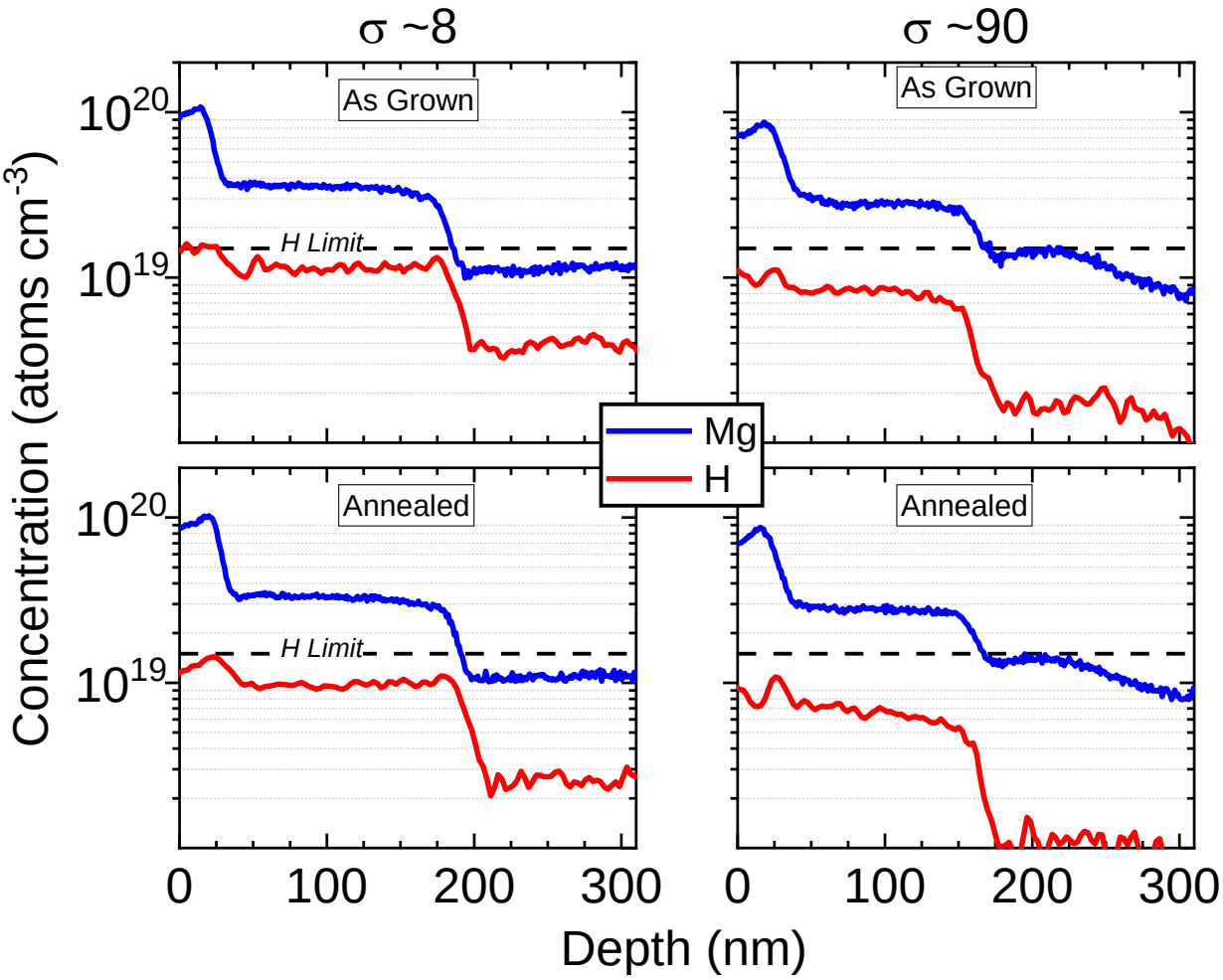


Figure 3.8 EAG SIMS of the ladder structures. Left and right columns represent the growth conditions. Top row are the samples analyzed before any activation anneal. Bottom row are samples that were annealed 850° C for 20 minutes (N_2 ambient).

In Fig. 3.8, two observations are immediately discerned. First, the total Mg concentration remains unaffected by a post-growth thermal annealing step. This is important confirmation that Mg atoms will be restricted to the intended regions of a given structure and not be reduced/migrate by the treatment. Second, and radically different than the established data for Mg-doped GaN, the H concentration never matches Mg. Regardless of doping level, growth condition, and post-growth

annealing, this remains true for these AlGaN films. This suggests that unlike GaN, a majority of Mg atoms are not incorporated into the film as a Mg-H complex.

For the σ ~ 8 condition, the Mg doping levels were $1x10^{19}$, $3.5x10^{19}$, and $1x10^{20}$ cm^{-3}. These levels represent three positions: below the H saturation limit, above the limit, and a typical p^+ doping level, respectively. For the lowest Mg doping level, the H concentration is $4x10^{19}$ cm^{-3} or 2.5x lower than Mg for the as-grown film. After annealing, this concentration was reduced further to $2.5x10^{19}$ cm^{-3}. At the next highest Mg doping level, the H concentration was ~$1x10^{19}$ cm^{-3} and had no appreciable decrease with the post-growth anneal. Similarly, the H concentration for the highest Mg doping reached its maximum at $1.5x10^{19}$ cm^{-3} but did not change with annealing.

For the σ ~ 90 growth condition, the Mg concentrations were slightly mismatched to the σ ~ 8 condition. Yet, the resulting total concentrations; $1.5x10^{19}$, $3x10^{19}$, and $8.5x10^{19}$; reasonably match the positions of below/above the H limit, and typical p^+ doping level, respectively. For the lowest Mg doping level, the highest H concentration as-grown was $2x10^{19}$ cm^{-3} and reduced to $1x10^{19}$ cm^{-3} after annealing. At the next highest Mg level, H was ~$8x10^{19}$ cm^{-3} and reduced to ~$6x10^{19}$ cm^{-3}. Finally, H concentration remained ~$1x10^{19}$ cm^{-3} for the as-grown and annealed case of the highest Mg doping level.

These observations lead to an interesting perspective on the analysis of Mg-doped AlGaN. Namely, the H concentration is significantly lower than the Mg level for all doping levels. This is in stark contrast to the observations made for Mg:GaN where the H concentration closely follows the Mg concentration until ~$2x10^{19}$ cm^{-3}, the minimum for resistivity, and is the main point defect responsible for the higher resistivity below this doping level. Doping above that limit sees the onset of self-compensation by V_N and related complexes. The H concentration is critical to understanding the increases in resistivity for Mg:GaN outside the doping level of $2x10^{19}$ cm^{-3}. For

the AlGaN films in this study, no such distinct transition between the passivation and self-compensation was observed as a function of incorporated Mg level. For the Mg levels <$3.5x10^{19}$ cm^{-3}, the activation annealing is able to reduce some of the H impurity in the samples. However, there was not an appreciable difference in PL between the as grown and activated films. Similarly, for the Mg doping levels above $3.5x10^{19}$ cm^{-3}, SIMS did not reveal a large change in the H concentration and no difference was measured in PL or in Hall effect. Therefore, the activation anneal was not performed in subsequent characterization techniques.

4.2 Mg:AlGaN PL

Following the results obtained from SIMS, a series of samples with a range of Mg doping concentrations reflecting the two growth conditions for AlGaN were prepared for optical and electrical testing. In the case of Mg:GaN, well defined signatures for the different compensation regimes (defined by the total Mg in the film) have been identified in PL. Typically, below a Mg concentration of $2x10^{19}$ cm^{-3}, a donor-acceptor pair (DAP) emission is present around 3.27 eV,[19] and acceptor bound exciton transitions from 3.4 to 3.5 eV are present only after post growth thermal activation.[30] Above this Mg doping level, an increase in blue luminescence at 2.8 eV is observed, which is described as a transition between the Mg acceptor and a donor complex of V_N – either V_N-Mg[28] or $V_{Ga}(V_N)s$ (s = 2 or 3).[23,24] The relationship of this signature in the spectra to V_N-complexes has been verified with positron annihilation spectroscopy and increased resistivity measured from Hall effect measurements confirms its compensation behavior.

For Mg:$Al_{0.6}Ga_{0.4}N$, limited studies have been published investigating PL, however, strong evidence linking mid-gap luminescence to the V_N-related defects has been presented.[53-56] Nakarmi et al. investigated the change in the peak position of the 2.8 eV emission with increasing Al content for heavily doped $Al_xGa_{1-x}N$:Mg doped films (~$1x10^{20}$ cm^{-3}).[56] Beginning with GaN and

increasing the Al-content until AlN, they traced the shift in the blue luminescence associated to V_N-related compensation from 2.8 eV to 4.7 eV, with the expected position for $Al_{0.6}Ga_{0.4}N$ at ~4.1 eV. Additionally, they investigated the effect of changing V/III ratio, i.e. changing the supersaturation, on the luminescence for $Al_{0.7}Ga_{0.3}N$:Mg. With an increase in V/III ratio to 5,000 from 2,000, the defect peak at 4.2 eV was significantly reduced, the band edge at 5.1 eV increased.[56] Electrically, the sample went from fully insulating to a measurable room temperature resistivity of ~10^5 Ω cm. Similarly, Kinoshita et al. presented a reduction in defect luminescence and increase in the band edge luminescence when they increased the V/III ratio to 1800 from 1300 in their $Al_{0.7}Ga_{0.3}N$:Mg ([Mg] = $3.6x10^{19}$ cm^{-3}) films.[50] The plausibility of the link to the V_N-related compensators is reasonable.

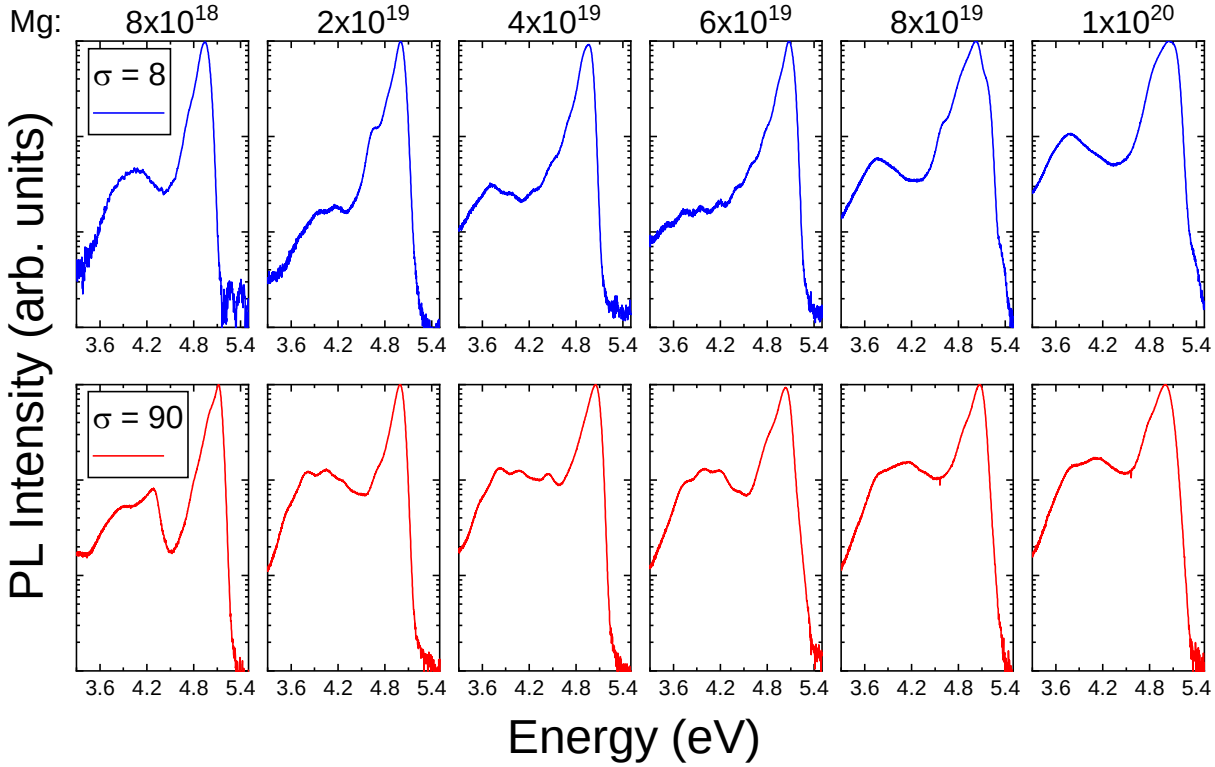


Figure 3.9 Low temperature (3K) PL of the Mg:$Al_{0.6}Ga_{0.4}N$ films doped with a single Mg concentration. Band edge was normalized to unity. Top row: σ = 8 growth condition. Above row: σ = 90 growth condition.

Fig. 3.9 displays the results of 3K photoluminescence measurements taken for the films with Mg concentrations ranging from $8x10^{18}$ to $1x10^{20}$ cm^{-3}. All spectra display the band edge at ~5 eV and this peak was used to normalize the spectra so the intensity and shape of midgap luminescences could be compared. Starting with the $\sigma \sim 8$ growth condition (blue), the Mg concentration of $8x10^{18}$ cm^{-3} yielded the peak ~4.1 eV which is expected to be V_N-related emission.[56] When the Mg concentration increased to $2x10^{19}$ cm^{-3}, this emission diminished, but a shoulder to the band edge appeared near ~4.6 eV. It is not clear what this emission is related to, but it may be similar to the DAP present around 3.2 eV for these doping levels in GaN.[30] However, GaN does not have the V_N-related emission present before the Mg exceeds the H saturation limit.[11] As AlGaN SIMS demonstrated, H never matches Mg and suggests H-passivation is not a dominant effect in AlGaN. It may be possible that both the V_N-related and DAP for low Mg doping can coexist in Mg:$Al_{0.6}Ga_{0.4}N$ because the formation of Mg-H complex is not prominent, leading to the more likely formation of V_N-related compensators. As the Mg concentration is increased, a stronger emission is present at ~3.7 eV vs. the expected 4.1 eV V_N-related emission. However, the later electrical characteristics show that stronger 3.7 eV vs. 4.1 eV luminescence corresponds with lower resistivity films.

The $\sigma \sim 90$ growth conditions (red) in Fig. 3.9 yield much stronger defect luminescence compared to the lower supersaturation condition. Starting with a Mg concentration of $8x10^{18}$ cm^{-3}, the major difference to the previous growth condition lies in the peak ~4.2 eV. Otherwise, similar peak difference between the band edge and 4.1 eV V_N-related emission is present. Moving to the higher Mg concentrations, the difference in the luminescence between the growth conditions is apparent. All of the films now have 4.1 eV and 3.7 eV emissions that are an order of magnitude lower than band edge. In comparison the previous growth conditions saw a 1.5x lower peak

intensity at 3.7 eV from the band edge, and a lower intensity peak at 4.1 eV. Additionally, the lower supersaturation conditions saw a rise in the midgap luminescence with increasing Mg concentration. For the σ ~ 90 condition, no difference was observed with doping concentration.

4.3 Mg:AlGaN Hall Effect

At this time, the peaks in the PL spectra for Mg:$Al_{0.6}Ga_{0.4}N$ are not well characterized and individual compensators not identified. The spectra presented above have a correlation with the resulting electrical measurements that will be presented below. The samples were measured by applying the traditional Ni/Au contacts in the van der Pauw geometry on the sample surface. A contact anneal of 650° C in air ambient for 10 minutes was performed to ensure working contacts. As mentioned previously, samples with and without the activation anneal were tested and yielded no difference in conductivity. This step was removed from the sample preparation process. The samples were measured using the Lakeshore Cryogenics 8404 AC/DC Hall Effect Measurement System, due to the low mobility of the samples. It is expected that alloy scattering is significant in AlGaN, and the limitation adds to the low mobility as a result of large effective mass of holes.

None of the samples grown at the σ ~ 90 displayed measurable electrical characteristics. From the SIMS data, minimal difference was measured for the H-passivation in the films, so V_N-related defects were expected to be the source of compensation. For this growth condition, PL showed a significantly larger midgap luminescence at both ~3.7 and 4.1 eV. Based on this experimental result, and the inability to measure the films electrically, it is concluded that the compensation of the σ ~ 90 growth condition is greater than for σ ~8. This is interesting because the previous section showed an overall reduction of V_N-related compensation in GaN when higher ammonia flows were utilized in the growth (and higher supersaturation for the H_2 growths). The expectation in this study was that the increased ammonia flow would result in lower compensation

of the films like for GaN since these were also H_2 diluent. However, the metalorganic flows were also increased to maintain constant composition and growth rate. AlGaN and AlN undergo pre-reaction between TMA and ammonia during the growth which consumes the precursors upstream of the substrate.[52] Therefore, the V/III was maintained between the two growth conditions. In GaN, the metalorganic flow was held constant and reduced compensation was observed at higher V/III ratios.

For the $\sigma \sim 8$ growth condition, all of the samples except for the Mg concentration of $8x10^{18}$ cm^{-3} growth were characterized using the Hall effect measurement at room temperature. Table 3.4 presents the results. With increasing Mg concentration, the films underwent a significant drop in the measured resistivity, decreasing from 310 Ω cm for a Mg concentration of $2x10^{19}$ cm^{-3} to a minimum of 35 Ω cm for the $8x10^{19}$ cm^{-3} Mg film. In contrast to GaN, the minimum resistivity in Mg:GaN occurs near a Mg concentration of $2x10^{19}$ cm^{-3} and increases with additional Mg. For the highest doping concentrations $6x10^{19}$, $8x10^{19}$, and $1x10^{20}$ cm^{-3}, the resistivities remained near this minimum value because the increase in free holes was accompanied by a drop in the mobility. However, none of the samples exceeded 0.5 cm^2 V^{-1} s^{-1} for the mobility regardless of doping level. It is suggested that alloy scattering plays a major role as well as the increased number of Mg atoms at the highest doping level.[47] Surprisingly, the free hole concentration increases with Mg doping. When charge balance equation was employed, using a ~47 meV activation energy,[50] the compensation concentration remained between 5 and $7x10^{18}$ cm^{-3}. In comparison, when the vacancy related self-compensation occurred in Mg:GaN, the hole concentration decreased when Mg was increased due to increased compensation values (~$1x10^{19}$ cm^{-3} when the Mg concentration exceeded $4x10^{19}$ cm^{-3}).

Table 3.4 Room temperature Van der Pauw Hall effect for Mg:$Al_{0.6}Ga_{0.4}N$ ($\sigma \sim 8$). Relative standard deviation was 8%, 10%, and 15% for the resistivity, hole concentration, and mobility, respectively.

Mg Concentration	$2x10^{19}$	$4x10^{19}$	$6x10^{19}$	$8x10^{19}$	$1x10^{20}$
Resistivity (Ω cm)	310	113	44	35	40
Hole Concentration ($x10^{16}$ cm^{-3})	13	12	60	50	100
Mobility (cm^2 V^{-1} s^{-1})	0.16	0.45	0.23	0.35	0.15

The measured free hole concentration and mobilities are the first to be reported for epitaxially doped Mg:$Al_{0.6}Ga_{0.4}N$. Historically, only the resistivity could be reported from the 4-point probe measurement with van der Pauw contact geometry.[40,50] The use of AC-Hall instrument allows for samples with low mobility to be characterized and is the strength of this research. Overall, the resistivity of the Mg:$Al_{0.6}Ga_{0.4}N$ is still orders of magnitude higher than can be obtained in Mg:GaN and limits the efficiency of electrical devices. However, the limited H-passivation and lower self-compensation at higher Mg doping levels are advantages not present in Mg:GaN.

Future work in the growth of Mg:$Al_{0.6}Ga_{0.4}N$ should investigate the influence of other process parameters on the compensation in the films. Specifically, changing the growth temperature should have the strongest effect on the growth supersaturation, and therefore the chemical potential of the defects. The temperature utilized in this study was 1050° C, and a comparison with 1000 and 1100° C would be useful in understanding if the chemical potential is driving the concentration of compensators in the film. Similar to the work on GaN,[12] it may be possible that a reduction in the dislocation density by moving to growth on AlN substrates can also yield lower resistivity values. Additionally, growth utilizing above bandgap illumination should be explored. Like in the case of Mg:GaN, an additional benefit past the overall reduction in H

concentration may occur. Since the post-growth anneal did not have an effect on the measured resistivity, utilizing the UV-light to suppress complex formation during growth could be a pathway for further compensation reduction.

Significant work is also required to characterize the defect luminescence below the band edge in these Mg:$Al_{0.6}Ga_{0.4}N$ films. There is a tentative link between the 3.7 and 4.1 eV peaks with the conductivity of the films. It is not clear what are the exact compensators. This will require a density functional theory and possibly positron annihilation spectroscopy experiments if the compensation source is a V_N-related complex. Since the films were grown on sapphire, the crystal quality may be lower than can be achieved on an AlN substrate. With the reduction in dislocation density, the surface morphology could improve dramatically and allow for greater uniformity across the films.

5. Conclusion

The $Mg_{(III)}$ acceptor type dopant has been extensively studied for the MOCVD growth of both GaN and $Al_{0.6}Ga_{0.4}N$. P-type doping has been one of the greatest limitations to the performance of III-nitride devices. However, the poor conductivity that has been achieved, in comparison to the easily attainable n-type behavior from Si or Ge, still was sufficient to enable LEDs to become commonplace in the current time. The problem with the $Mg_{(III)}$ acceptor is multifaceted: holes have limited mobility in the III-nitrides, the ionization energy is greater than room temperature, H passivates Mg in the low doping regime, and self-compensation occurs at high doping concentrations. Utilizing both the defect quasi Fermi level, through above bandgap illumination, or the defect chemical potential, by changing growth conditions, saw significant reductions in the compensation of the films.

From the work in Mg:GaN, it is clear that above bandgap illumination significantly reduces the H-passivation of Mg, and enables Mg to incorporate at concentrations previously where the majority were part of Mg-H complex. Even further, the work also showed that some residual passivation occurs even after the H is removed via a thermal activation process. The use of UV is highly suggested for obtaining the best conductivity in the films. Additionally, the self-compensation by V_N-related complexes was reduced by increasing the ammonia flow into the growth chamber (greater than the effect of the UV-illumination alone). The complete model of this defect chemical potential has not been determined, yet the Ga vapor supersaturation alone does not seem to capture the reduction.

From the Mg:$Al_{0.6}Ga_{0.4}N$ study, H-passivation seems to play no role in limiting the conductivity of the films as evidenced by its significantly reduced concentration below the Mg level. Virtually, no change was observed in the mobility regardless of the Mg doping level suggesting that the alloy scattering is limiting the conductivity. Again, the NH_3 flow was shown to have an influence on the PL spectra for the doped films as well as the conductivity. However, increasing the NH_3 made the compensation and conductivity worse in these films. Since significant pre-reaction occurs between the metalorganic precursors and NH_3, a more complex picture of the chemical potential is needed to understand the difference in compensation.

6. References

1. C. G. Van de Walle, C Stampfl, and J. Neugebauer. “Theory of doping and defects in III-V nitrides,” *Journal of Crystal Growth* **189/190**, 505-510 (1998).

2. C. G. Van de Walle and J. Neugebauer. “First-principles calculations for defects and impurities: Applications to III-nitrides,” *Journal of Applied Physics* **95**, 3851 (2004).

3. P. Reddy, M. P. Hoffmann, F. Kaess, Z. Bryan, M. Bobea, A. Klump, J. Tweedie, R. Kirste, S. Mita, M. Gerold, R. Collazo, and Z. Sitar. “Point defect reduction in wide bandgap semiconductors by defect quasi Fermi level control,” *Journal of Applied Physics* **120**, 185704 (2016).

4. P. Reddy, F. Kaess, J. Tweedie, R. Kirste, S. Mita, R. Collazo, and Z. Sitar. “Defect quasi Fermi level control-based C_N reduction in GaN: Evidence for the role of minority carriers,” *Applied Physics Letters* **111**, 152101 (2017).

5. P. Reddy, S. Washiyama, F. Kaess, R. Kirste, S. Mita, R. Collazo, and Z. Sitar. “Point defect reduction in MOCVD (Al)GaN by chemical potential control and a comprehensive model of C incorporation in GaN,” *Journal of Applied Physics* **122**, 245702 (2017).

6. H. Amano, M. Kito, K. Hiramatsu, and I. Akasaki. “P-type conduction in Mg-doped GaN treated with low-energy electron beam irradiation (LEEBI),” *Japanese Journal of Applied Physics Part 2-Letters* **28,** L2112 (1989).

7. I. Akasaki, H. Amano, M. Kito, and K. Hiramatsu, “Photoluminescence of Mg-doped p-type GaN and electroluminescence of GaN p-n junction LED,” *Journal of Luminescence* **48-9,** 666 (1991).

8. S. Nakamura, M. Senoh, S. Nagahama, N. Iwasa, T. Yamada, T. Matsushita, H. Kiyoku, and Y. Sugimoto, “InGaN-based multi-quantum-well-structure laser diodes,” *Japanese Journal of Applied Physics Part 2-Letters* **35,** L74 (1996).

9. M. Kneissl, T. Kolbe, C. Chua, V. Kueller, N. Lobo, J. Stellmach, A. Knauer, H. Rodriguez, S. Einfeldt, Z. Yang, N. M. Johnson, and M. Weyers. “Advances in group III-nitride-based deep UV light-emiting diode technology,” *Semiconductor Science and Technology* **26,** 014036 (2011).

10. W. Gotz, N. M. Johnson, J. Walker, D. P. Bour, and R. A. Street. “Activation of acceptors in Mg-doped GaN grown by metalorganic chemical vapor deposition,” *Applied Physics Letters* **68,** 667 (1996).

11. A. Klump, M. P. Hoffmann, F. Kaess, J. Tweedie, P. Reddy, R. Kirste, Z. Sitar, and R. Collazo, “Control of passivation and compensation in Mg-doped GaN by defect quasi Fermi level control,” *Journal of Applied Physics* **127**, 045702 (2020).

12. B. Sarkar, S. Mita, P. Reddy, A. Klump, F. Kaess, J. Tweedie, I. Bryan, Z. Bryan, R. Kirste, E. Kohn, R. Collazo, and Z. Sitar. “High free carrier concentration in p-GaN grown on AlN substrates,” *Applied Physics Letters* **111**, 032109 (2017).

13. M. Leszczynski, H. Teisseyre, T. Suski, I. Grzegory, M. Bockowski, J. Jun, S. Porowski, K. Pakula, J. M. Baranowski, C. T. Foxon, and T. S. Cheng. "Lattice parameters of gallium nitride," *Applied Physics Letters* **69**, 73 (1996).

14. W. Gotz, N. M. Johnson, J. Walker, D. P. Bour, M. D. McCluskey, and E.E. Haller. "Local vibrational modes of the Mg-H acceptor complex in GaN," *Applied Physics Letters* **69**, 3725 (1996).

15. A. Castiglia, J. F. Carlin, and N. Grandjean, "Role of stable and metastable Mg-H complexes in p-type GaN for cw blue laser diodes," *Applied Physics Letters* **98,** 213505 (2011).

16. S. Nakamura, T. Mukai, M. Senoh, and N. Iwasa. "Thermal annealing effects on p-type Mg-doped GaN films," *Japanese Journal of Applied Physics Part 2-Letters* **31,** L139 (1992).

17. Y. Nakano, O. Fujishima, and T. Kachi. "Effect of p-type activation ambient on accepor levels in Mg-doped GaN," *Journal of Applied Physics* **96,** 415 (2004).

18. H. Obloh, K. H. Bachem, U. Kaufmann, M. Kunzer, M. Maier, A. Ramakrishnan, and P. Schlotter. "Self-compensation in Mg doped p-type GaN grown by MOCVD," *Journal of Crystal Growth* **195,** 270 (1998).

19. U. Kaufmann, P. Schlotter, H. Obloh, K. Kohler, and M. Maier. "Hole conductivity and compenstion in epitaxial GaN:Mg layers," *Physical Review B* **62,** 10867 (2000).

[20.] R. Kirste, M. P. Hoffmann, J. Tweedie, Z. Bryan, G. Callsen, T. Kure, C. Nenstiel, M. R. Wagner, R. Collazo, A. Hoffmann, and Z. Sitar. "Compensation effects in GaN:Mg probed by Raman spectroscopy and photoluminescence measurements," *Journal of Applied Physics* **113,** 103504 (2013).

[21.] F. Shahedipour and B. W. Wessels. "Investigation of the formation of the 2.8 eV luminescence band in p-type GaN:Mg," *Applied Physics Letters* **76,** 3011 (2000).

[22.] O. Gelhausen, M. R. Phillips, E. M. Goldys, T. Paskova, B. Monemar, M. Strassburg, and A. Hoffmann. "Dissociation of H-related defect complexes in Mg-doped GaN," *Physical Review B* **69,** 125210 (2004).

[23.] A. Uedono, S. Takashima, M. Edo, K. Ueno, H. Matsuyama, H. Kudo, H. Naramoto, and S. Ishibashi. "Vacancy-type defects and their annealing behaviorsin Mg-implanted GaN studied by monoenergetic positron beam," *Physica Status Solidi B* **252**, No. 12, 2794 (2015).

[24.] S. F. Chichibu, K. Shima, K. Kojima, S. Takashima, M. Edo, K. Ueno, S. Ishibashi, and A. Uedono. "Large electron capture-cross-section of the major nonradiative recombination centers in Mg-doped GaN epilayers grown on a GaN substrate," *Applied Physics Letters* **112**, 211901 (2018).

[25.] J. Neugebauer and C. G. Van de Walle. "Gallium vacancies and the yellow luminescence in GaN," *Applied Physics Letters* **69,** 503 (1996).

[26.] J. L. Lyons, A. Janotti, and C. G. Van de Walle. "Carbon impurities and yellow luminescence in GaN," *Applied Physics Letters* **97**, 152108 (2010).

[27.] D. O. Demchenko, I. C. Diallo, and M. A. Reshchikov. “Yellow luminescence of gallium nitride generated by carbon defect complexes,” *Physical Review Letters* **110**, 087404 (2013).

[28.] M. A. Reshchikov, G. C. Yi, and B. W. Wessels. “Behavior of 2.8- and 3.2-eV photoluminescence bands in Mg-doped GaN at different temperatures and excitation densities,” *Physical Review B* **59**, 13176 (1999).

[29.] U. Kaufmann, M. Kunzer, M. Maier, H. Obloh, A. Ramakrishnan, B. Santic, and P. Schlotte. “Nature of 2.8 eV photoluminescence band in Mg doped GaN,” *Applied Physics Letters* **72**, 1326 (1998).

[30.] G. Callsen, M. Hoffmann, M. R. Wagner, J. Tweedie, T. Kure, Z. Bryan, J. S. Reparaz, S. Aygun, M. Bugler, R. Kirste, J. Brunnmeier, R. Collazo, C. Nenstiel, Z. Sitar, and A. Hoffmann. “Optical signature of Mg-doped GaN: Transfer process,” *Physical Review B* **86** (7), 075207 (2012).

[31.] M. Kneissl, “Chapter 1 A Brief Review of III-Nitride UV Emitter Technologies and Their Applications.” In: M. Kneissl, J. Rass, *III-Nitride Ultraviolet Emitters.* Springer Series in Materials Science, vol. 227. Springer, Cham.

[32.] Y. Taniyasu, M. Kasu, and T. Makimoto. “An aluminum nitride light-emitting diode with a wavelength of 210 nanometres,” *Nature* **441**, 325 (2006).

[33.] C. Pernot, M. Kim, S. Fukahori, T. Inazu, T. Fujita, Y. Nagasawa, A. Hirano, M. Ippommatsu, M. Iwaya, S. Kamiyama, I. Akasaki, and H. Amano. “Improved

efficiency of 255 -280 nm AlGaN-based light-emitting diodes," *Applied Physics Express* **3**, 061004 (2010).

34. A. Fujioka, T. Misaki, T. Murayama, Y. Narukawa, and T. Mukai. "Improvement in output power of 280-nm deep ultraviolet light-emitting diode by using AlGaN multi quantum wells," *Applied Physics Express* **3**, 041001 (2010).

35. T. Kinoshita, K. Hironaka, T. Obata, T. Nagashima, R. Dalmau, R. Schlesser, B. Moody, J. Xie, S-I. Inoue, Y. Kumagai, A. Koukitu, and Z. Sitar. "Deep-ultraviolet light-emitting diodes fabricated on AlN substrates prepared by hydride vapor phase epitaxy," *Applied Physics Express* **5**, 122101 (2012).

36. J. Xie, S. Mita, Z. Bryan, W. Guo, L. Hussey, B. Moody, R. Schlesser, R. Kirste, M. Gerhold, R. Collazo, and Z. Sitar. "Lasing and longitudinal cavity modes in photo-pumped deep ultraviolet AlGaN heterostructures," *Applied Physics Letters* **102**, 171102 (2013).

37. W. Guo, Z. Bryan, J. Xie, R. Kirste, S. Mita, I. Bryan, L. Hussey, M. Bobea, B. Haidet, M. Gerhold, R. Collazo, and Z. Sitar. "Stimulated emission and optical gain in AlGaN heterostructures grown on bulk AlN," *Journal of Applied Physics* **115**, 103108 (2014).

38. X-H. Li, T. Detchprohm, T-T. Kao, Md. M. Satter, S-C. Shen, P. D. Yoder, R. D. Dupuis, S. Want, Y. O. Wei, H. Xie, A. M. Fischer, F. A. Ponce, T. Wernicke, C. Reich, M. Martens, and M. Kneissl. "Low-threshold stimulated emission at 249 nm and 256 nm from AlGaN-based multiple-quantum-well lasers grown on sapphire substrates," *Applied Physics Letters* **105**, 141106 (2014).

[39] Z. Zhang, M. Kushimoto, T. Sakai, N. Sugiyama, L. J. Schowalter, C. Sasaoka, and H. Amano. “A 271.8 nm deep-ultraviolet laser diode for room temperature operation,” *Applied Physics Express* **12**, 124003 (2019).

[40] I. Bryan. “Al-rich AlGaN and AlN growth on bulk AlN single crystal substrates,” *NCSU Dissertation* (2015).

[41] J. S. Harris, J. N. Baker, B. E. Gaddy, I. Bryan, Z. Bryan, K. J. Mirrielees, P. Reddy, R. Collazo, Z. Sitar, and D. L. Irving. “On compensation in Si-doped AlN,” *Applied Physics Letters* **112**, 152101 (2018).

[42] I. D. Goepfert, E. F. Schubert, A. Osinsky, and P. E. Norris, “Demonstration of efficienct p-type doping $Al_xGa_{1-x}N$/GaN superlattice structures,” *Electronic Letters* **35**, 13 (1999).

[43] P. Kozodoy, M. Hansen, S. P. DenBaars, U. K. Mishra, “Enhanced Mg doping efficiency in $Al_{0.2}Ga_{0.8}N$/GaN superlattices,” *Applied Physics Letters* **74**, 3681 (1999).

[44] A. A. Allerman, M. H. Crawford, M. A. Miller, and S. R. Lee, “Growth and characterization of Mg-doped AlGaN-AlN short-period superlattices for deep-UV optoelectronic devices,” *Journal of Crystal Growth* **312**, 756 (2010).

[45] J. Simon, Y. Cao, and D. Jena, “Short-period AlN/GaN p-type superlattices: hole transport use in p-n junctions,” *Physica Status Solidi C* **7**, 2386 (2010).

[46] M. Martens, C. Kuhn, E. Ziffer, T. Simoneit, V. Kueller, A. Knauer, J. Rass, T. Wernicke, S. Einfeldt, M. Weyers, and M. Kneissl, “Low absorption loss p-AlGaN

superlattice cladding layer for current-injection deep ultraviolet laser diodes," *Applied Physics Letters* **108**, 151108 (2016).

47. J. Simon, V. Protasenko, C. Lian, H. Xing, and D. Jena, "Polarization-induced hole doping in wide-band-gap uniaxial semiconductor heterostructures," *Science* **327**, 60 (2010).

48. S. Li, T. Zhang, J. Wu, Y. Yang, Z. Wang, Z. Wu, Z. Chen, and Y. Jiang, "Polarization induced hole doping in graded $Al_xGa_{1-x}N$ (x=0.7~1) layer grown by molecular beam epitaxy," *Applied Physics Letters* **102**, 062108 (2013).

49. R. Dalmau and B. Moody, "Polarization-induced doping in graded AlGaN epilayers grown on AlN single crystal substrates," *ECS Transactions* **86**, 31 (2018).

50. T. Kinoshita, T. Obata, H. Yanagi, and S. Inoue, "High p-type conduction in high-Al content Mg-doped AlGaN," *Applied Physics Letters* **102**, 012105 (2013).

51. J. Simon, A. Wang, H. Xing, S. Rajan, and D. Jena. "Carrier transport and confinement in polarization-induced three-dimensional electron slabs: Importance of alloy scattering in AlGaN," *Applied Physics Letters* **88**, 042109 (2006).

52. S. Washiyama, P. Reddy, F. Kaess, R. Kirste, S. Mita, R. Collazo, and Z. Sitar. "A thermodynamic supersaturation model for the growth of aluminum gallium nitride by metalorganic chemical vapor deposition," *Journal of Applied Physics* **124**, 115304 (2018).

53. J. Li, T. N. Oder, M. L. Narkami, J. Y. Lin, and H. X. Jiang, “Optical and electrical properties of Mg-doped p-type $Al_xGa_{1-x}N$,” *Applied Physics Letters* **80**, 1210 (2002).

54. K. B. Nam, M. L. Nakarmi, J. Li, J. Y. Lin, and H. X. Jiang, “Mg acceptor level in AlN probed by deep ultraviolet photoluminescence,” *Applied Physics Letters* **83**, 878 (2003).

55. M. L. Nakarmi, K. H. Kim, M. Khizar, Z. Y. Fan, J. Y. Lin, and H. X. Jiang, “Electrical and optical properties of Mg-doped $Al_{0.7}Ga_{0.3}N$ alloys,” *Applied Physics Letters* **86**, 092108 (2005).

56. M. L. Nakarmi, N. Nepal, J. Y. Lin, and H. X. Jiang. “Photoluminescence studies of impurity transitions in Mg-doped AlGaN alloys,” *Applied Physics Letters* **94**, 091903 (2009).

Chapter 4 SIMS of the III-Nitrides

1. Introduction

For semiconducting materials, the inclusion of impurities and dopants as low as the 10^{14}-10^{15} cm^{-3} concentration range, and their location within the depth of a device or sample, can have a dramatic effect on the electrical and optical behavior. An analytical technique sensitive to these atomic fractions as a function of depth is required to understand and interpret the semiconducting material's electrical and optical performance. Many characterization techniques have been introduced over the years for determination of elemental information in a sample. Examples of these include energy dispersive spectroscopy (EDS), Auger electron spectroscopy (AES), x-ray photoelectron spectroscopy (XPS), and secondary ion mass spectrometry (SIMS).[1] However, only SIMS has the dynamic range suited for the detection of dopants or impurities, from 100 to parts per million atomic (ppma) and even parts per billion atomic (ppba) per cent.[1-4]

SIMS falls under the overall field of mass spectrometry, wherein a mass spectrometer is used to separate ions by their mass to charge ratio. The coupling of the mass spectrometer to some means of ion formation distinguishes the specific analytical techniques. Gases, liquids, and solids can all be analyzed, and the state of the sample dictates the method of ion generation.[1-3] The basic operating principle of SIMS is the bombardment of a surface with ions to generate emission of secondary ions from a sample. Other primary particles used for bombardment may include electrons, photons, or neutrals.[1-4] After this bombardment, the emitted secondary particles include electrons, neutral atoms and molecules, and atomic or cluster ions.[3] These secondary ions can then be extracted via electric fields into a mass spectrometer.

Continued bombardment can erode material, and detection of secondary ions during this process constitutes the dynamic SIMS process. The data is collected as secondary ion intensity as

a function of sputtering time and can be converted to concentration as a function of depth, establishing a depth profile for the sample. Static SIMS, on the other hand, uses very low current primary ions to generate secondary ions only from the surface of the material. Typically, the etching rate is on the order of only a few monolayers for tens of minutes of continued sputtering and is thus considered 'non-destructive'.[5] Etching rates for dynamic SIMS can exceed tens of nanometers per second in some cases.

The oft cited beginning of this technique is the work of J. J. Thomson, who first discovered secondary ion generation in a gas discharge tube in 1910,[6] and by 1913 developed the first mass spectrometer by passing these ions through a magnetic field. It would take another twenty years before Mattauch and Herzog introduced the first of the "modern" spectrometers used in SIMS, the double focusing magnetic sector.[7] Another twenty years would pass before the Time-of-Flight,[8] and the quadrupole[9] were introduced. By the middle of the 1950s, the trio of spectrometers still in use to this day were in place for analysis. Further discussion of the technical aspects of each will be covered later in the chapter.

As mentioned above, the combination of the mass spectrometer with a means of generating the secondary ions forms the analytical technique. For SIMS, these secondary ions are generated by rastering primary ions over a sample surface to achieve emission of ions from the sample.[1-5] Similar to the mass spectrometers, three main sources are used to generate the primary ions used in this bombarding process. The first of which, the cesium thermal ionization source, was discovered by Langmuir in 1923.[10] Manfred Von Ardenne patented the duoplasmatron in 1961,[11] and for SIMS applications typically oxygen gas is used as the gas source. The final source in major use today is the liquid metal ion source, which has its roots in the work of Zeleny[12] and Krohn,[13] who was studying its use in space propulsion.

It is clear that by the early 1960s, the major components comprising the SIMS systems had been invented. In those early years, the technique was still a novelty, and its original design was to study extraterrestrial material returning from the burgeoning space program.[3] As such it would be hard to predict widespread adoption, because material from the surface of the moon is hardly in large supply. Yet, early pioneers foresaw its potential in the field of semiconductor research.[3,4] The ability to gather information over a large dynamic range and microns of depth from the surface was a perfect match for the needs of the new field of microelectronics. To this day, this capability of depth profiling for these semiconductor materials is one of the primary applications for this SIMS technique.[1-5]

The following sections will serve as a general introduction to the primary instrumentation used in a SIMS analysis as well as basic operating considerations. Subsequently, the improvement made to the detection limit on the ION TOF SIMS V will be introduced. Finally, specific challenges pertaining to the difficulty of III-nitride SIMS analysis will be discussed and composition corrections for Mg-doping in $Al_xGa_{1-x}N$ ($0 \leq x \leq 1$) will be established. These last two points can be considered as the main contributions of this work as related to enhancing the SIMS analytical capabilities. The intention is not to create a user operating manual, for that several texts exist[1-4] and better serve an analyst. Instead, this is envisioned as brief introduction so that informed requests can be made for any future analytical needs.

2. Mass Spectrometers

As introduced above, there are three main types of mass analyzers used in SIMS to detect the ions ejected from a sputtered material. In general, the analyzer defines the type of SIMS instrument, as the vacuum system, transfer optics, primary ion sources are designed for the particular analyzer, and not the other way around. As such, the operational principles and

characteristic advantages and disadvantages of each instrument stem from the analyzer. Initially, the quadrupole and the magnetic sector analyzers were the equipment of choice for dynamic SIMS analysis.[1-5] Static SIMS typically utilized the quadrupole but has been phased out by the ToF analyzer.[5] With the addition of the second primary ion source used for sputtering operation, ToF systems could perform so-called dual beam depth profiles.[2,5] The bulk of the SIMS measurements in this dissertation were collected in this manner, and significant achievements were made to improve the sensitivity of this analysis. Those results will be discussed later in the chapter. Here, the main characteristics of each analyzer will be presented.

The first type of mass spectrometer is called the quadrupole analyzer (Fig. 4.1). It consists of four closely spaced circular rods that are grouped into two pairs. Both sets of pairs share a common DC voltage, but have an additional RF 180° out of phase from the other set.[1] This combination of the DC and RF voltages causes ions of a specific mass to charge ratio to undergo a stable oscillatory flight path through the quadrupole into the detector. Ions of other mass to charge ratios have unstable oscillations and will collide into one the rods. By increasing the voltage, but maintaining a constant frequency and ratio of the DC to RF voltage, the specific mass to charge ratio undergoing stable oscillation is increased. This allows for an entire mass spectrum to be collected, or sequential selection of several masses during a depth profile.[1,2,14] Typical geometry of the rods is 1 cm in diameter and ~15 to 20 cm in length.[2]

To obtain reasonable mass separation, the ions must spend sufficient time in the quadrupole to allow the unstable oscillations to occur. This puts a restriction on the energy of the secondary ions that can pass into the spectrometer, typically 10 eV and below.[1,2,14] As a result, only low extraction fields can be used to collect the secondary ions (<100 V cm^{-1}),[2] which limits the transmission of the system in two ways. The first is that the small field is unable to capture many

of the generated secondary ions, so only a fraction reach the transfer optics. Second, the field does not provide a massive change in acceleration, so the ions have a wide energy spread. To prevent overlap in the detector, the ions are passed through an energy filter which rejects another fraction of generated secondary ions. For these reasons, quadrupoles are characterized as having the lowest transmission and mass resolution of the three analyzers (Table 4.1). However, one benefit of the low extraction field is that much lower energy primary ion beams can be used. As a result, quadrupoles have great depth resolution and are ideally suited for shallow analyses.[1] Additionally, the more open chamber design allows for an electron gun to be incorporated, providing ample charge neutralization for insulating samples.[1,4]

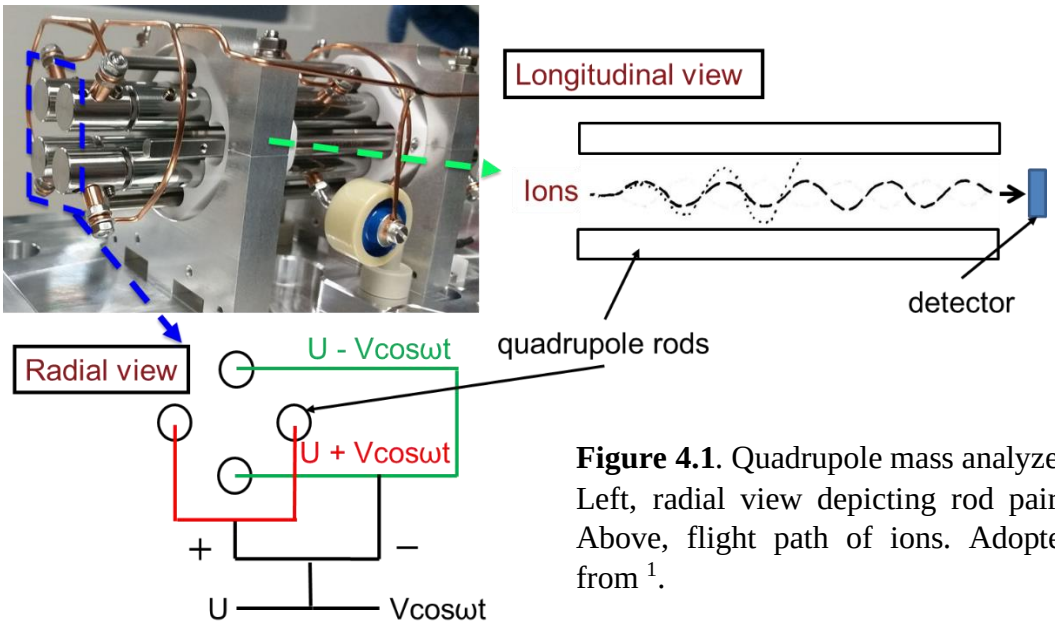


Figure 4.1. Quadrupole mass analyzer. Left, radial view depicting rod pairs. Above, flight path of ions. Adopted from [1].

Some improvements can be made for the mass resolution, for instance increasing the DC:RF ratio (entire mass range) or biasing the sample (lower masses). Yet, these reduce the transmission of high mass elements in the first case, and the low mass elements in the second.[2] Several 'hardware' updates have been employed, such as larger diameter quadrupoles and rods, increased RF-frequency components, or high-extraction field optics with retarder plates, but only

increased the transmission up to 1%.[2] While low in comparison to the magnetic sector and ToF, if mass interferences are minor, quadrupoles can still achieve sufficient sensitivity to measure low concentrations of impurities in semiconductors. Overall, the strengths of this the system, multiple mass detection, depth resolution, and charge neutralization, make quadrupole systems a very attractive and viable analysis tool for the III-nitrides.[1,2,4]

The second type of mass analyzer is called the magnetic sector. This style of analyzer consists of two sectors, an electrostatic and a magnetic, which provide double focusing capability in the following manner. The electrostatic sector generates a radial electrostatic field by a pair of cylindrical electrodes typically with a 90° geometry. As the secondary ions with slightly different energies pass into this field, they are dispersed by their energy and will have slightly different radii for their flight path. They are then directed towards the magnetic sector portion of the analyzer, which is also bent to an angle usually of 60°. Upon entering the magnetic field, the ions experience a force perpendicular to both the field and their initial direction of velocity. These ions are deflected according to their mass. With proper spacing between sample and electrostatic sector, and electrostatic sector to the magnetic sector, ions of same mass but different energy and angle of emission are focused at the detector.[1-3] Fig. 4.2 depicts the set-up and relevant equations.

Due to the length of the flight path, the magnetic sector instruments employ very high extraction fields near the sample surface, typically >10 keV cm^{-1}, so that the secondary ions have enough energy to reach the detector.[2] As a result, a much higher fraction of the generated secondary ions can be collected from the sample surface. Additionally, the field accelerates the ions to a similar high energy, which makes the initial energy spread much smaller. In practice all of the secondary ions created during the sputtering process are passed through the electrostatic portion of the detector, and only at the magnetic sector are the ions deflected, but typically some optical

gating is performed to avoid edge effects from the crater. The magnetic sector instruments are thus characterized as having high transmission (between 10 to 50%) and the highest mass resolution $m/\Delta m > 10^4$.[1,2]

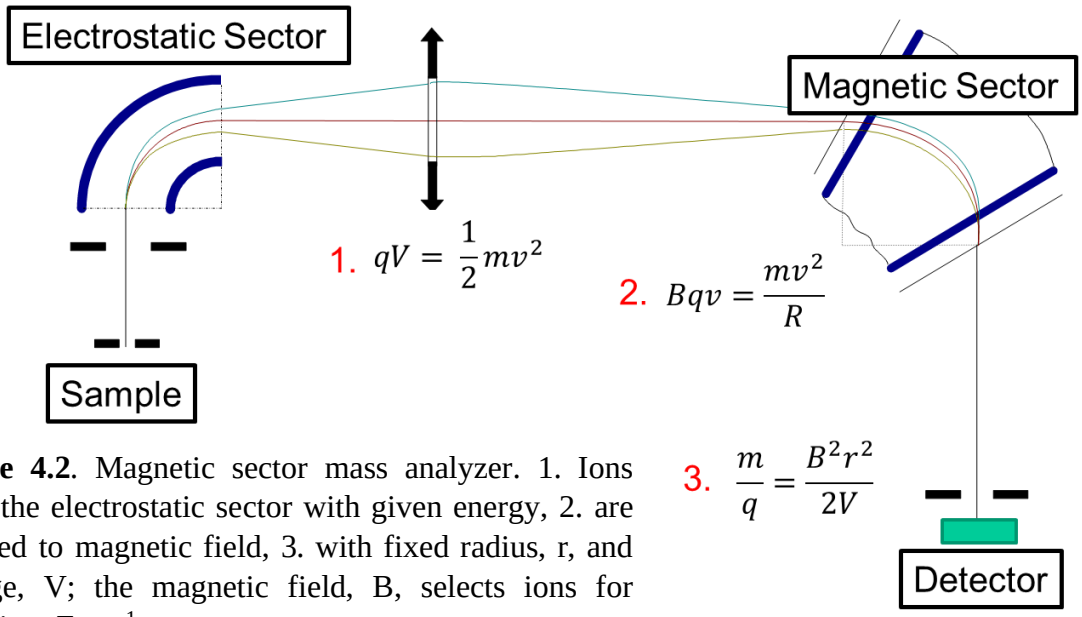


Figure 4.2. Magnetic sector mass analyzer. 1. Ions leave the electrostatic sector with given energy, 2. are exposed to magnetic field, 3. with fixed radius, r, and voltage, V; the magnetic field, B, selects ions for detection. From[1].

In theory, it would be possible to detect all masses if detectors were placed at the appropriate deflection points. However, such a detector would need to be impractically large.[1,2] Instead, the magnetic field is tuned so that a specific mass reaches the detector for each sputtered crater, i.e. only single mass detection. Another issue is that an electron gun cannot be easily placed in the chamber due to the tight geometry from the high extraction field.[1] So insulating samples can be difficult to analyze. As well, the primary ion beam is deflected by this field, so low energy beams for shallow analysis cannot be used unless the extraction voltage is also reduced.[1-3] In general, these are the most expensive instruments, but provide the highest sensitivity and mass resolving power. As such, they are ideally suited for low impurity detection and bulk analysis of the III-nitrides.

The final detector type is called the time-of-flight mass analyzer. Unlike the previous two analyzers, which use electric and magnetic fields to separate ions by their mass-to-charge ratios,

the ToF relates the time for an ion to strike a detector to its mass. In its most basic form, the analyzer consists of a drift tube typically greater than 1 m in length with a detector attached to the end.[2] Modern designs like the ION TOF V have the addition of electrostatic mirrors, which help mitigate initial energy spread of the generated secondary ions, and double the flight distance within the tube.[1] The ToF systems work by sending a short pulse of primary ions to the sample surface, typically < 30 ns in time.[5] Secondary ions are generated in a very discrete time, around < 10^{-12} s.[2,5] After the pulse strikes the surface, an extraction field is turned on to accelerate all of the secondary ions to a given potential, usually 5 kV, and inject them into the flight tube. Because all of the ions are at roughly the same energy, the time to reach the detector depends on their mass, i.e. the higher the mass, the longer the flight time. At this energy, typical time to reach the detector is approximately 100 μs.[1,25] After this period, the next primary ion pulse is directed toward the surface and the process repeats. Fig. 4.3 displays the typical design.

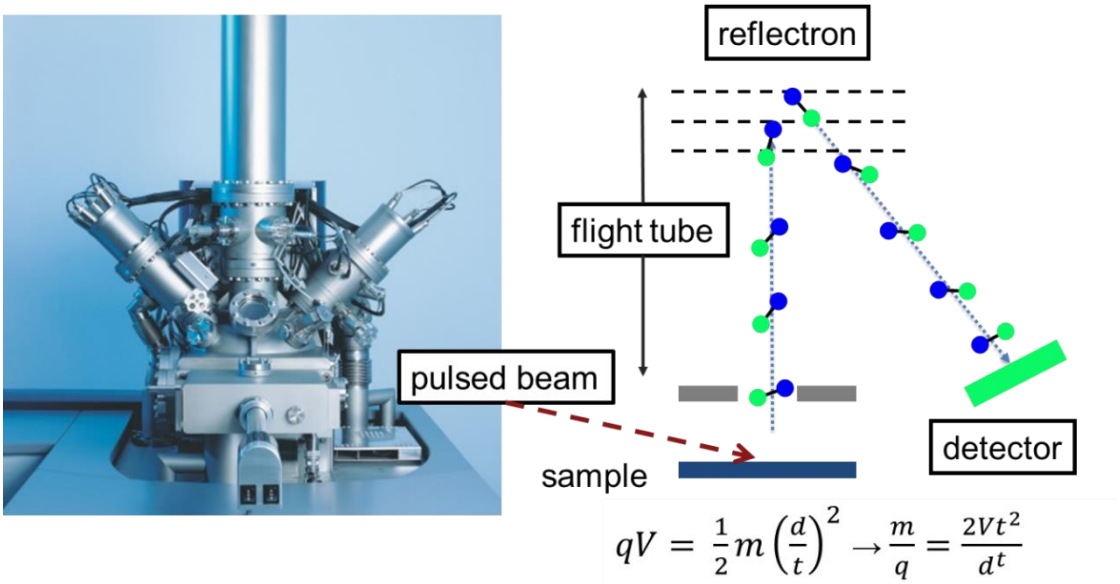


Figure 4.3. Time-of-flight mass spectrometer. Left, ION TOF SIMS V instrument. Right, schematic of the flight tube including reflectron mirror. Blue and green dots represent ions of the same mass but different starting energy. The mirror causes the green ion, with a lower starting energy, to have a shorter pathlength. Thus, both ions will arrive simultaneously to the detector. From ION TOF, gmbh.

Because this analyzer has a time-sensitive secondary ion generation and detection, there is a serious restriction in the primary ion system. Unlike the quadrupole or magnetic sector where ions are continuously collected during the sputtering process, the ToF requires that all ions be generated at a distinct point in time.[1,2,5] For example, if two carbon ions were formed but at separate times, the first would reach the detector and be registered as a carbon atom, but the second could be registered as nitrogen. That is why it is imperative to generate and collect all of the secondary ions in a very short window. This has two major consequences on the sensitivity of this analyzer. First, a pulse of primary ions has to be used. By definition, this pulse is a much lower current than continuous sputtering beams of the quadrupole and magnetic sector. Since the number of generated secondary ions depends on this primary ion current, the number of secondary ions generated is limited. Second, because the extraction window is very small, a majority of the secondary ions created are rejected. Therefore, the overall sensitivity of these systems is very low.

However, the ToF system does have some distinct advantages. First and foremost, it measures the entire mass spectrum, which can exceed m/z 1000, at every data point since it measures number of ion counts as a function of time.[1,2,5] This enables the user to go back and 'reconstruct' the data and include the signal from any element or molecule post-analysis. Second, the mass resolution of these systems can be very high when short pulses of the primary ion beam are selected. Third, with the dual beam operation for depth profiling, the ToF systems can have very robust charge neutralization. After the secondary ions are extracted into the flight tube, electrons can be directed to the sample surface to compensate the positive charge build up. With sufficient pause time in the data cycle, charging can be completely removed.[1] Lastly, because of low current of primary ions, the data spacing, i.e. the number of data points per unit depth, can be extremely high by appropriately tuning the sputtering primary ion source. Despite these many

attractive characteristics, the overall low sensitivity of the ToF precluded its use in depth profiling of semiconductors. Initially, this limitation prevented the use of the ION TOF SIMS V for the SIMS analysis of the samples produced in the Wide Bandgaps laboratory. Fortunately, a recent software update of the system allowed the analysis parameters to be modified, which was exploited to increase the intensity of the secondary ions. This will be discussed later in the chapter.

A final comparison will be made for the three analyzers by introducing the topic of useful yield. This is a figure of merit that relates the number of detected secondary ions per second divided by the sputtered number of atoms per second from the sampling volume.[3,5] It is the product of the ionization probability, α, the transmission of the spectrometer, *T*, and instrument specific term, *F*. For the quadrupole, the yield is modulated by the inverse of the number of elements for detection, *n*, because the masses are detected sequentially. The magnetic sector instruments have transmission terms that vary depending on the mass resolution selection, and are restricted to a single mass per analysis crater. Niehaus[5] has suggested an inverse 1.5 *n* relationship to account for multiple sputtering craters to detect additional elements. ToF, on the other hand, has an analyzed fraction term, S_A/S_{Sp} , which is the specific yield of the analysis beam divided by the specific yield from the sputtering beam. Specific yield is proportional to the beam current and total sputter yield and inversely related to the rastered area. Under the operating conditions utilized in this thesis, this S_A/S_{Sp} term has the 10^{-5} order of magnitude value.

The values for each analyzer can vary widely depending on a given application. For instance, bulk impurity profiling of the III-nitrides typical impurity species of interest are H, C, O, and Si with the matrix signals Al and Ga, for $n = 6$ value. For the quadrupole with transmission values ranging from 0.001 to 0.01, the useful yield is 0.001 to 0.1 times the ionization probability. Magnetic sector varies from 0.01 to 0.05. ToF reaches a maximum of 0.001. However, if a shallow

profile detecting these species is desired, then the ToF could reach 0.01 times the ionization probability or greater, because the sputtering beam current is reduced. When considering the static SIMS condition (all elements, no sputtering beam), then useful yield of ToF far exceeds the capabilities of the other two spectrometers.

In summary, three main types of analyzer are in use for SIMS depth profiling of semiconductors today. The design of each analyzer can provide many advantages and disadvantages that need to be considered. None of these systems are a 'one size fits all' that can be used for every type of analysis and material for that matter. For instance, if ultimate detection limit of impurities in a bulk layer is required, then a magnetic sector is the best choice. If doping in multiple thin layers near surface is required, then a quadrupole or ToF system is better suited. Table 4.1 is therefore constructed to inform the advantages and disadvantages of each system as it pertains to the III-nitrides. Included are some of the main properties such as mass resolution and transmission, as well as the capabilities for specific analysis requests.

Table 4.1 Comparison of mass analyzers.

	resolution M/ΔM	mass range	mass detection	transmission T
quadrupole	10^2-10^3	<1000	sequential	0.01-0.1
magnetic sector	10^4	>1000	single	0.1-0.5
ToF	>10^3	>1000	parallel	0.5-1

	useful yield	types of analysis
quadrupole	$U = \alpha\, T\, \frac{1}{n}$	shallow depths, multi-layer, insulating
magnetic sector	$U = \alpha\, T\, \frac{1}{1.5n}$	bulk, best detection limit, best resolution
ToF	$U = \alpha\, T\, \frac{S_A}{S_{Sp}}$	surface measurements, unknown contaminants

3. Sputtering and Secondary Ion Formation

Sputtering is the process in SIMS where the energetic beam of primary ions erodes atoms and molecules from a material. During the process, some of the atoms and molecules become

ionized and can then be extracted for mass analysis. As such, this process is the foundation for the entire SIMS measurement. The three main sources of the primary ions used in this process were briefly introduced above. However, their specific mechanisms for ion formation are less important to the SIMS process than the attributes of the primary ions themselves, i.e. type, mass, energy, and angle of incidence, because they dictate the key parameters in a SIMS analysis: sputter yield, secondary ion yield, and depth resolution. These, instead, will be the focus of this section.

In general, the sputtering process can be described by the linear collision cascade model, developed by Sigmund.[15] It explains the energy and momentum transfer from the primary ions to the near surface atoms via the concept of cascades, i.e. a series of binary elastic collisions. For this process to occur, the primary ions are typically high mass and accelerated to energies in the keV range.[16] In such a regime, the primary ion typically has a projected range of ~100 Å and the energy loss is due to nuclear collisions, not electronic effects like heating. Thus, as the incident primary ion enters the material, it strikes a stationary target atom within this shallow depth, transfers some energy, and the target atom recoils. If the energy transfer is sufficiently high, this 'recoiled' target atom can then strike an additional target atom and the process can be repeated creating the cascade. The primary species comes to rest in the substrate with most probable depth designated as projected range or R_p. Sputtering occurs when a cascade reaches the surface, and the energy transferred to the surface atom is greater than the binding energy of the lattice site allowing it to escape.[16] For context, if the primary ion energy is sub keV energy, no cascade occurs, and the only sputtering is due to single-knockon events.[15] Conversely, if the primary ion energy exceeds 100 keV, the range is too deep in the material for significant sputtering of the surface to occur.[16]

In terms of ionization, for most elements only a small fraction (~ 5%) of the total ejected atoms (or molecules) are charged.[16] Instead, the vast majority leave the sample as neutrals. To date,

there is not a strong consensus of how, or more specifically when, the ionization process occurs.[2] Suffice to say, however, (1) secondary ions do form and (2) they are enhanced/inhibited by the sputtering conditions.[1-4] Further discussion in this section will introduce the effects of the various sputtering parameters.

First and foremost is the selection of the primary beam species. This discussion will only focus on the following species O_2^+ (O^-), Cs^+, and Bi_3^+ (Bi^+), but many others have been tested to date.[1,4] The reason is the first two species are known to enhance the secondary ion yield, the number of secondary ions removed from the per incident ion, for positive and negative ions, respectively. The latter is the species for the primary analysis beam in the ION TOF SIMS V at NCSU and warrants comparison.

When searching for an element as an impurity, or even a matrix signal, the following general rules of thumb should be considered. For positive secondary ion formation, the key characteristic of the element is the ionization energy, the energy to remove an electron from an atom.[1-4] The lower the ionization potential the more likely the element will form positive secondary ions (electrons are easier to remove). The characteristic that signals an element's ability to form negative ions is the electron affinity, or energy change to take electrons outside an atom and place them in a vacant orbit.[1] The higher the electron affinity, the higher the probability that an element will form a negative secondary ion. Figure 4.4 depicts these tendencies across the periodic table.

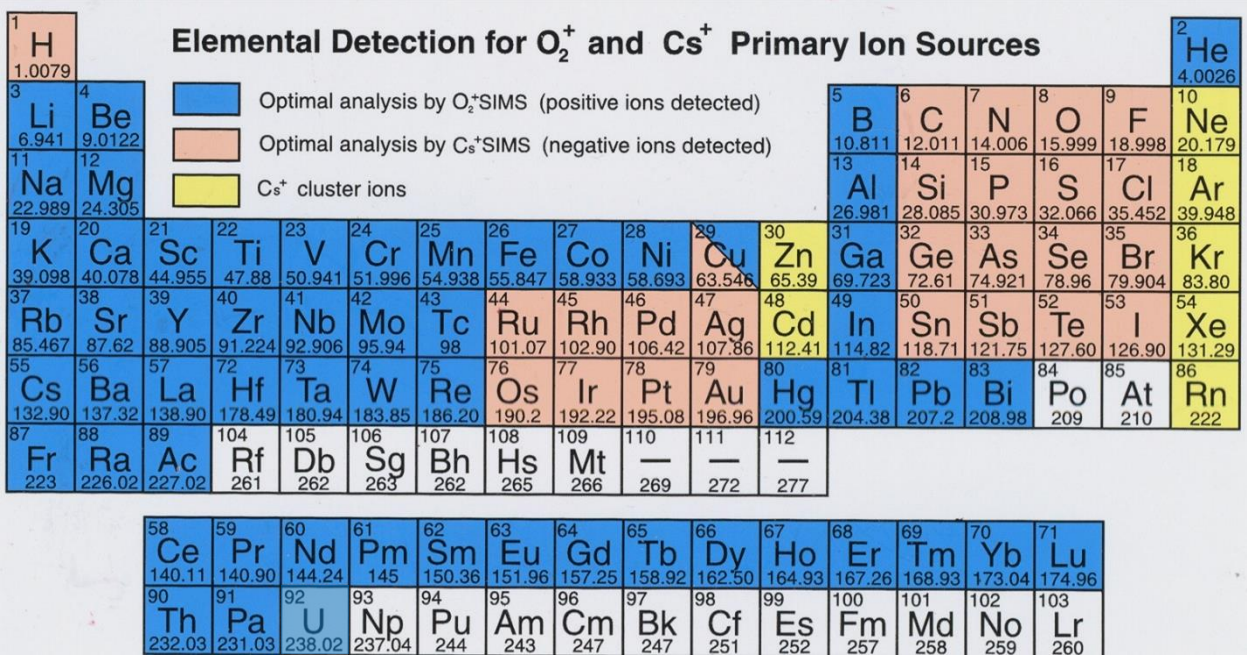


Figure 4.4. Periodic table of optimized polarity of secondary ion for each element, and the primary ion source generating the highest sensitivity. In blue are positive ions formed from O_2^+ primary source. Tan are negative secondary ions from a Cs^+ source. Yellow are cluster positive ions (CsM^+) also from Cs^+ source. Optimized detection has not been found for gray elements. From www.eag.com.

The selection of the primary ion species enhances the formation of different polarity species. O_2^+ is known to increase the secondary ion yield for positive ions via the bond-breaking method.[17] During the sputtering process, the surface becomes oxidized from the oxygen bombardment. As the sputtering process proceeds, the metal-oxygen bond is then broken, and electron transfers to the oxygen atom for a negative ion species. In turn, a greater number of positive metal ions are formed.[1,17] Cs^+, on the other hand, enhances the yield of the negative ion species. Negative ion formation relies on the difference between the work function at the surface of the material, and the individual atom's electron affinity. With sufficient coverage by cesium during the sputtering process, the work function at the surface is reduced, and the electron capture process is enhanced.[18] Therefore, a greater number of negative ion species are formed.

Fig. 4.5 demonstrates the effect of primary ion beam species on the resulting secondary ion yield. Plotted are the relative sensitivity factors (RSF) for various elements placed inside a Si sample. The RSF determination will be discussed later in the chapter, for now it is sufficient to mention that the RSF is inversely related to the secondary ion yield. Therefore, low RSF values represent high yield. Since the other conditions of the analysis are not reported, absolute comparison between the RSFs in the two figures cannot be made. For example, the reported RSF

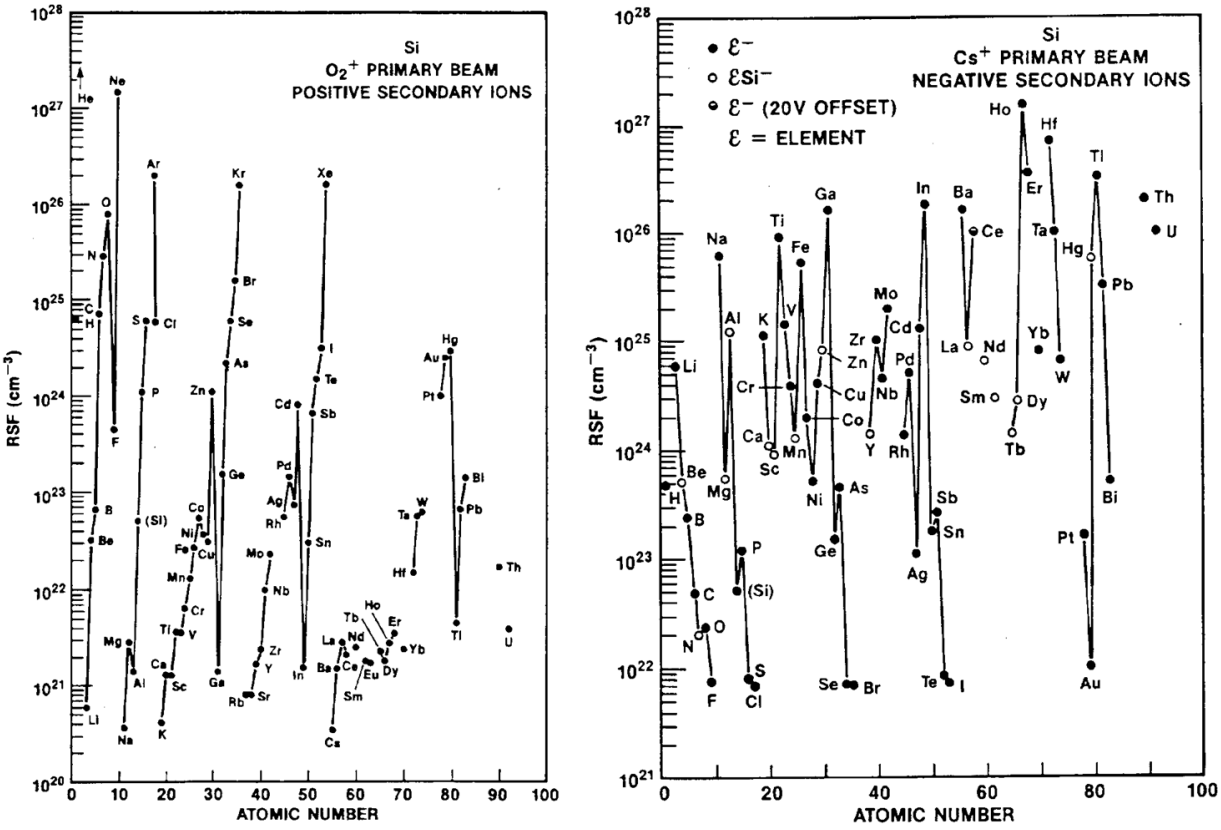


Figure 4.5. Left, RSFs for various elements implanted in Si. Detection of positive secondary ions from an O_2^+ source. Right, RSFs for elements implanted into Si. Detection of negative secondary ions under bombardment from Cs^+ source. From[4].

for silver under O_2^+ is lower than Cs^+, yet silver is primarily a negative ion species. It is more pertinent to relate a given element's placement on the figures, relative to the other elements on the periodic table. Continuing with the silver example, as a positive species silver has a RSF higher

than the majority of the periodic table. Conversely, as a negative species the only elements with lower RSF are gold and the nonmetals. In general, the nonmetals, semiconductors, and rare earth elements all tend to form negative secondary ions, and the alkali (except hydrogen), alkaline earth, and transition metals all form positive secondary ions. Several elements were noted in Figure 4 to form cluster positive ions, specifically zinc and cadmium. In this analysis, Cs^+ is used as the primary ion species and detected are positive ions of the form $CsM^{+,}$ where M represents the additional element.

Because bismuth does not have a high reactivity, no secondary ion yield enhancement is obtained using this as a primary ion source. However, the specific mass of the primary ion species has an effect on the measurement. Here, it comes as no surprise that for a given energy, the mass of the species increases the sputtering yield, the number of atoms or molecules removed per incident ion, since the related cross sections for energy transfer and nuclear stopping are directly proportional to mass.[9] Therefore, a greater number of collisions transferring energy are expected, and a faster rate of removal occurs. As a consequence of increased sputter yield, the secondary ion yield also will increase, since it also has a slight dependence on the sputtering rate.[2] To this end, bismuth is used for ToF-SIMS, because it can provide stable low current beams but increase the secondary ion yield over traditional liquid metal ion sources like Ga^+.[19] As well, the higher mass elements penetrate less into the material, providing better depth resolution, defined as the depth where signal intensity varies from 16% to 84% or vice versa when a sharp interface is crossed.[5,9] Oxygen, as the smallest mass species, has the lowest sputtering yield in terms of mass, and penetrates furthest into a material. However, use of O_2^+ as the primary species splits the energy over the two atoms, thereby reducing the penetration and generating two cascades.[1] This leads to no changes in the sputtering rate and depth resolution as a result.

The next beam parameter to consider is the energy of the primary ions. This primarily affects the sputtering yield and the depth resolution of the material. Over typical primary ion beam energies from 1 to 10 keV, the sputtering yield of the material increases with energy.[1-4] In terms of the linear cascade theory, one would expect greater energy transfer with increased primary ion energy. Therefore, one would expect larger cascades and greater probability of the cascades interacting with the surface atoms. At the same time, the higher energy could be thought of as transferring more energy to the surface atoms, overcoming the binding energy of the surface site. This increase is not indefinite and does reach a saturation point. Beyond 10 keV, a decrease in the sputtering yield is observed.[1,4] The likely cause is the increased penetration depth of the primary ion species, reducing the probability of a cascade reaching the surface atoms and having less interaction with the surface.

One drawback of using higher energy within the 1 to 10 keV window is a loss of depth resolution.[1,2,4] This is explained via the recoil and cascade mixing effects that occur during the sputtering process.[4,15] Recoil mixing occurs when the primary ion impacts a stationary atom and drives it deeper into the sample away from the surface. Not surprisingly, the higher the primary ion's energy, the greater depth this recoiled atom is driven. Cascade mixing is similar to the recoil mixing, except it is one of the recoiled atoms that drives a separate target atom to a different depth. When the target atoms are driven further into the sample, they will then be sputtered and detected at a deeper artificial depth than in the real crystal. As a result, the depth for signals coming from ion implantation, or multilayer structures, could be elongated. That is why lower primary ion beam energies are used for high depth resolution.

Closely related to the energy of the primary ion beam is the current of the primary ions. This is a measure of the number of ions bombarding the sample for a given length of time. It is no

surprise that with an increase in the number of primary ions at a given energy, sputter yield and secondary ion yield both increase due to increased cascades. As well, the depth resolution decreases, as the number of ions driven into the sample will undoubtedly increase. By moving to higher beam currents, however, the backgrounds and/or detection limits of impurities, especially atmospheric elements, can be lowered. Backgrounds can be caused by redeposition from residual gas in the chamber or come from material sputtered onto the secondary optics.[4] With increased number of primary ions, the removal rate of atoms greatly outpaces these redepositions. As with beam energy, however, the beam current cannot be indefinitely increased. The beam loses focus with increased current meaning that the area of beam itself being rastered on the sample spreads. Significant sputtering of the sidewalls of the raster area can occur, which distorts the true distribution of the elements in the layer.[1,4] Careful selection of the current, raster area, and acceptance area are needed to obtain best depth resolution and detection limits.[4]

The final parameter to consider for the primary beam of a SIMS instrument is the angle of incidence of the primary ions. For quadrupole and ToF systems, typically the primary ion source and secondary optics are fixed to the instrument and the sample stage is tilted to change in the incidence.[1] For a magnetic sector, the beam is deflected according the extraction voltage. Therefore, the angle of the beam is selected by the voltage between the sample and the optics.[1,4] The angle of incidence affects the sputtering yield, secondary ion yield, and the depth resolution of a profile. Fig. 4.6 depicts the effect of changing the incidence angle (measured from sample normal) on the sputtering yield and secondary ion yield. Starting from normal beam incidence, secondary ion yield decreases as a function of increasing angle of incidence.[1,4] Sputtering yield, on the other hand, doubles by moving from normal to 30 degrees. By 60 degrees from normal, the

sputtering yield is 10 times higher than normal.[6] Beyond this angle, however, the yield decreases. The last parameter, the depth resolution, also improves as the angle is made more oblique.[1,4]

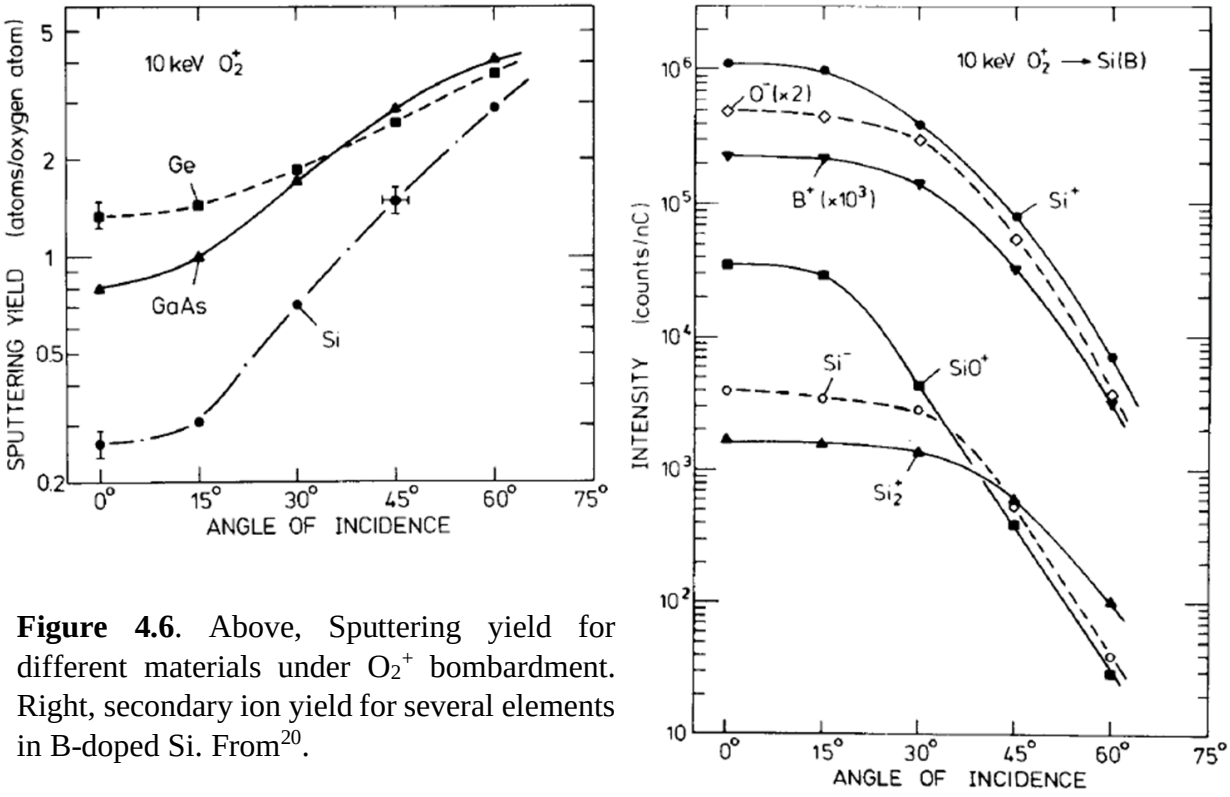


Figure 4.6. Above, Sputtering yield for different materials under O_2^+ bombardment. Right, secondary ion yield for several elements in B-doped Si. From[20].

The above discussion highlights the many parameters one must consider prior to a SIMS measurement. These are the very broad considerations – additional complications occur as a function of the material being analyzed. Most of the examples above were from silicon, and, generally speaking, the III-nitrides suffer worse depth resolution, slower removal rates, and higher background contamination than silicon. The cause is material related and not necessarily predictable. However, the overall trends based on the beam conditions discussed above still hold for a different material system.

The previous two sections were included to introduce the reader to the basics of the SIMS process and what information can be gleaned from the material based on the specific instrument in use. There is no such one-size-fits-all instrument that can be used to answer every type of

question that can be answered in SIMS. In a sense, these sections are intended to motivate the reader to understand what equipment will be used for analysis so that time and finances are not wasted. A company or facility that operates a magnetic sector should be used for ultimate detection limits and not shallow analysis. If a shallow analysis is requested, then a lower beam energy with higher incidence angle should be assumed. If data quality is poor, the above can guide the reader to ask the correct questions from an analyst, so that the best understanding can be reached.

4. Concentration Determination (Relative Sensitivity Factors)

One additional topic that will be introduced prior to the introduction of the various improvements made to the ToF system at NCSU, is conversion of the data from a SIMS analysis into concentration per volume values as a function of depth in the material. The central concept is the relative sensitivity factor (RSF), and the process to obtain this correction factor and apply it to the data will be presented.

In SIMS, the raw data is recorded as a number of counts for a given mass to charge ratio as a function of time. In the case of a bulk homogeneous material, the conversion from sputtering time into a depth is simple. The total depth of the profile crater, measured typically by profilometry, is divided by the total time of the analysis. Multiplying this rate by the time between data points establishes the depth scale. When samples contain multiple layers of different matrix material, sputtering rates vary depending on the matrix of the individual layer. Determining the depth scale requires a priori knowledge of the sputter rate change as a function of the matrix, or multiple measurements with profiles ending at each layer so that the total depth to each layer is known. While this second method provides the utmost accuracy, it is impractical from a cost and time standpoint. With a sufficient library of bulk films of varying matrix composition, sputtering rates

can be determined for a given analysis beam energy, current, and raster area. Such a technique was employed for $Al_xGa_{1-x}N$ ($0 \leq x \leq 1$), and will be presented later.

Determining the depth scale is a vital component for the conversion of the number of counts in the measurement to an actual concentration of the element in the material. The process involved for this conversion relies on the measurement of a standard sample, typically an ion implanted sample of a single matrix where the dose of the implant species is known. Cross calibration of this standard using other measurement techniques, such as x-ray diffraction for the composition and Rutherford backscattering spectroscopy for the dose, is required. However, if calibration has been performed historically in a different material, say silicon or gallium arsenide, a reference wafer from these materials can be co-implanted and analyzed with SIMS to determine the dose. It can then be inferred that the dose within the other co-implanted wafers is the same value.

Ultimately, the concentration of the implanted species is determined by the RSF calibration. For a robust derivation of the RSF equation, please refer to Benninghoven.[3] The derivation begins with the general SIMS equation describing the secondary ion current of element A (*I(A)*, counts s^{-1}) measured in the detector:

$$I(A) = I_{primary} Y_{total} c(A) \alpha(A) f(A) \qquad 4.1$$

where $I_{primary}$ is the primary ion beam current, Y_{total} is the sputter yield defined as the total number of secondary ions per primary current, $c(A)$ is the atomic fraction of element A (i.e. the concentration of A divided by the total concentration of atoms in the material), $\alpha(A)$ is the probability of element A forming an ion, and $f(A)$ is the transmission factor for the instrument, i.e. the fraction of the generated secondary ions of element A that reach the detector.

The above equation can be interpreted in the following manner. $I_{primary}$ multiplied by Y_{total} gives the total number of secondary atoms ejected. Of these total secondary atoms, only a

fraction is of the element A, given by the term $c(A)$. However, of the secondary atoms of A, only a small portion are ionized, represented by $\alpha(A)$, and of the ionized atoms A, again a fraction is only able to reach the detector $f(A)$. In summary, only a fraction of the real concentration within a material becomes ionized and is detected.

The derivation proceeds by taking the ratio of the secondary ion current for element, A, and secondary ion current for a reference element, R, usually a matrix element.

$$\frac{I(A)}{I(R)} = \frac{I_{primary} Y_{total} c(A) \alpha(A) f(A)}{I_{primary} Y_{total} c(R) \alpha(R) f(R)} = \frac{c(A)\alpha(A)f(A)}{c(R)\alpha(R)f(R)} = \frac{[A]\alpha(A)f(A)}{[R]\alpha(R)f(R)} \qquad 4.2$$

Where $[A]$ and $[R]$ represent the concentration (in atoms cm^{-3}) of element A and R, respectively. Rearranging this equation to solve for the concentration of element A:

$$[A] = \frac{[R]\alpha(R)f(R)}{\alpha(A)f(A)} \frac{I(A)}{I(R)} = RSF \frac{I(A)}{I(R)} \qquad 4.3$$

In this form, the concentration of element A can be determined by multiplying the ratio of secondary ion currents from element A to R and multiplying by the RSF. As mentioned in the beginning, the RSF is determined by a measurement reference sample implanted with a known dose of element A. Making use of the above equation for secondary ion current, the RSF can be rewritten:

$$RSF = \frac{[R]\alpha(R)f(R)}{\alpha(A)f(A)} = [R] \frac{\frac{I(R)}{I_{primary} Y_{total} c(R)}}{\frac{I(A)}{I_{primary} Y_{total} c(A)}} = \frac{[R]c(A)}{c(R)} \frac{1}{I(A)} I(R) = \frac{[A]}{I(A)} I(R) \qquad 4.4$$

Before the next substitution, it is important to consider the units in the above expression. $[A]$ is measured in atoms cm^{-3} and the secondary ion currents are measured in counts s^{-1} (or per data cycle in the case of ToF SIMS). What the RSF is trying to give is the concentration of element A per the ratio of the secondary ion intensity of A to the reference R, for a given data point in the

depth profile. For a bulk doped sample used as a standard, the above equation is sufficient for RSF determination since the intensities are measurable and the concentration is known. In an ion implanted sample, however, the concentration distribution is gaussian with a certain decay length into the bulk of the material. Therefore, concentration varies as a function of depth in the material. To capture this varying concentration, secondary ion intensities for the element are detected over the entire depth of the profile, instead of using singular data points. Combining with the other knowns; dose, depth of analysis, and number of data points to complete that depth; the left side of the equation be determined:

$$RSF = \frac{\Phi}{d} \Big/ \frac{\sum_{total} I(A)}{C} I(R) = \frac{\Phi C}{d \sum_{total} I(A)} I(R) \qquad 4.5$$

where Φ is the dose of element A measured in atoms cm-2, C is the number of data cycles, d is the depth of the analysis, and $\sum_{total} I(A)$ is the summation of the secondary ion intensity of A at each data cycle.

In the case of magnetic sector and quadrupole instruments, secondary ions are constantly being generated by the DC operation of the primary ion gun. However, element A is not always being detected during the data cycle. The above is modified by multiplying by an additional term in numerator, t, which represents the time dwelt on element A during a cycle to account for potential loss of ions due to off-time not measuring this species. ToF does not have such a correction, since the secondary ions are generated simultaneously at the specific depth corresponding to the data cycle.

In the following, the RSF determination and application to an ion implanted standard will be presented. A bulk GaN epitaxially grown film was doped with magnesium with a total dose of $1x10^{14}$ Mg cm^{-2}. This dose was determined via co-implantation in a GaAs standard wafer and calibration against in-house standards at NCSU. The standard was analyzed using the ION TOF

SIMS V instrument and the monitored species were $CsMg^+$ for Mg, and $CsGa^+$ for the reference Ga. Fig. 4.7 depicts the raw data acquired from the implant. For a total sputtering time of ~285 seconds (201 data cycles), the depth of the profile was 570 nm ($5.7x10^{-5}$ cm). The sum of secondary ion counts for the $CsMg^+$ signal over the 201 cycles was 3127. The ion intensity selected for the reference $CsGa^+$ was at cycle 100, corresponding to 76,900 counts, where it has reached an equilibrium count level. Substituting into eq. 4.5, the RSF was determined to be $8.67x10^{21}$ cm^{-3}. Multiplying this RSF by the ratio of $CsMg^+$ to $CsGa^+$ ion intensity, and applying the sputtering rate of 2 nm s^{-1}, the raw data is converted to atoms cm^{-3} vs. depth.

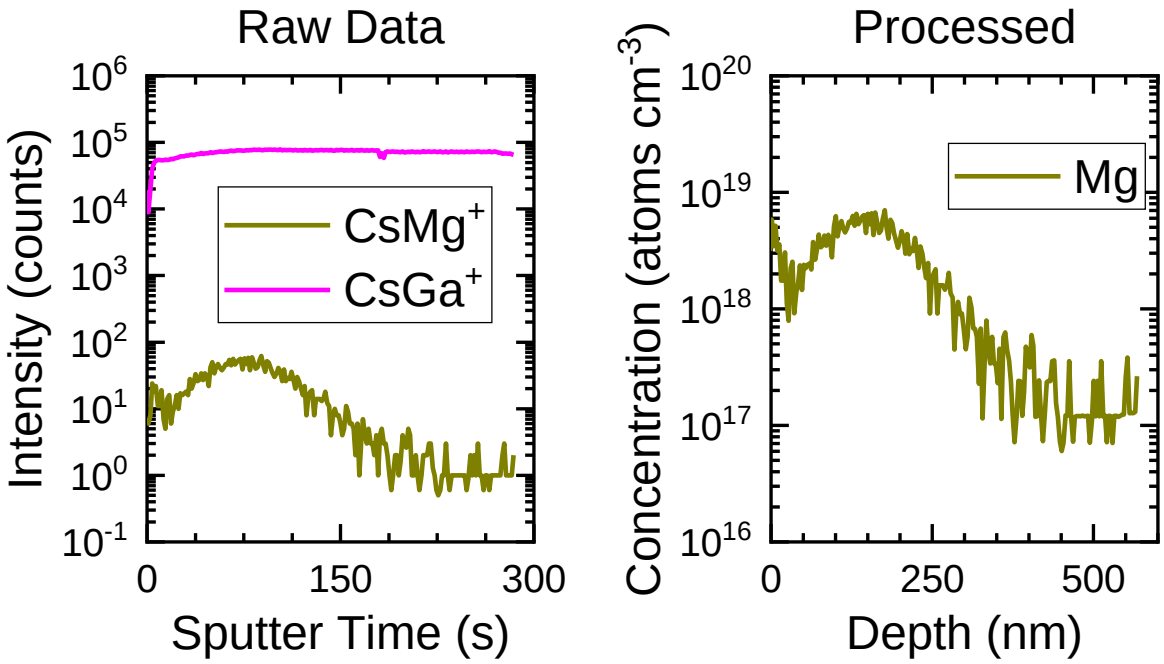


Figure 4.7. Mg-implanted GaN standard. Left, raw data acquired from measurement. Right, depth profile after sputter rate and RSF have been applied.

The above discussion encapsulates the data acquisition and conversion process for SIMS. This is the backbone of the depth profiling process, and provides SIMS with quantitative power that no other technique can provide. In the above example, the magnesium detected concentrations exceeded the mid-10^{17} atoms cm^{-3} concentrations. Considering GaN has a density of ~$9x10^{22}$

atoms cm^{-3}, the dynamic range is of the order of 10^5. This far exceeds other techniques like AES and EDS, but orders of magnitude from best values obtained for SIMS. It is worth noting, however, this analysis was only possible after significant improvement was made to the ION TOF SIMS V at NCSU. Before, the dynamic range was one order of magnitude reduced, and $CsMg^+$ counts were <10 for the entire analysis. The implant could not be captured, and no measurement of Mg was possible in any of the III-nitrides. The impact of not having this analytical capability cannot be stressed, so the process to overcome this limitation is presented in the following section.

5. Improvement to Sensitivity of ToF-SIMS Measurement[21]

For much of its history, ToF-SIMS instruments were primarily used in the chemical analysis of the top monolayers of a solid.[1,4,5,22] Despite having a sputtering secondary primary ion beam, this instrumentation was typically reserved for collecting mass spectra and never depth profiles for concentration determination. Because of its intended application – 'non-destructive' surface analysis – the primary ion source was designed to operate in the pA current range to bombard a surface with 10^{13} atoms cm^{-2} per pulse.[5] Considering eq. 4.1, this puts a significant limitation on the secondary ion yield that can be obtained.

To overcome the limited sensitivity in the ION TOF SIMS V system, the following investigation was performed. A GaN wafer was analyzed using the non-interlaced, dual beam depth profiling mode following the conditions provided in Table 4.2. The wafer was epitaxially doped during MOCVD growth with two Mg concentrations of $9x10^{18}$ and $1.5x10^{20}$ atoms cm^{-3}. For comparison, two analysis parameters were adjusted: the primary ion beam pulse width and the number of analysis frames collected per data cycle. Previously, these two parameters were locked to the user, and only after a recent software update was the user allowed access. The primary ion pulse width describes the amount of time the Bi_3^+ ions are directed down the primary optical

column before being bunched and pulsed toward the sample surface. The analysis frame is the sequence of events where the primary ion beam is pulsed on the sample surface, secondary ions from the surface are generated and extracted, and, after a certain flight time, the ions reach the detector for counting. The number of secondary ions detected in the analysis frame is displayed as the output for the given data cycle number. When the number of analysis frames is increased, the system repeats this sequence according to the number of frames before proceeding to the sputtering cycle. The secondary ions detected at the end of each sputtering frame are summed, and this summation is output as the total secondary ions detected for the data cycle.

Table 4.2. ION TOF SIMS V, experimental parameters for non-interlaced mode.

	Bi_3^+ (primary ions for analysis)	Cs^+ (sputtering ions)
energy	25 keV	10 keV
current	0.4 pA	30 nA
incidence angle	45°	45°
area	50 μm^2	120 μm^2
frames per cycle	1 or 10	1
pulse width	20 or 30 ns	
detected species	$CsMg^+$ and $CsGa^+$	

To understand the effect of varying the pulse width of the primary ion beam, the sample was measured under pulse width conditions of 20 and 30 ns while keeping the analysis frames constant. The 20 ns condition is considered the “normal” condition, because it provides the best combination of mass resolution and secondary ion intensity. Lowering the pulse width results in too few secondary ions being generated. Increasing the pulse width causes the mass resolution to degrade (Fig. 4.9, right). The comparison between the 20 and 30 ns pulse widths is plotted in Fig. 4.8. As expected, the secondary ion intensity increased for both $CsMg^+$ and $CsGa^+$ species due to the increase of primary ions striking the sample. For CsMg+, the first layer from the surface saw an increase in average concentration from 107 counts to 211 counts, or roughly 2x. The second

layer also saw an increase of ~2x, from 5 counts for the 20 ns pulse to 10 counts for 30 ns pulse. For CsGa+ the average of the secondary ion counts was selected for the entire profile. Between the two conditions this average increased from 8,101 to 14,814.

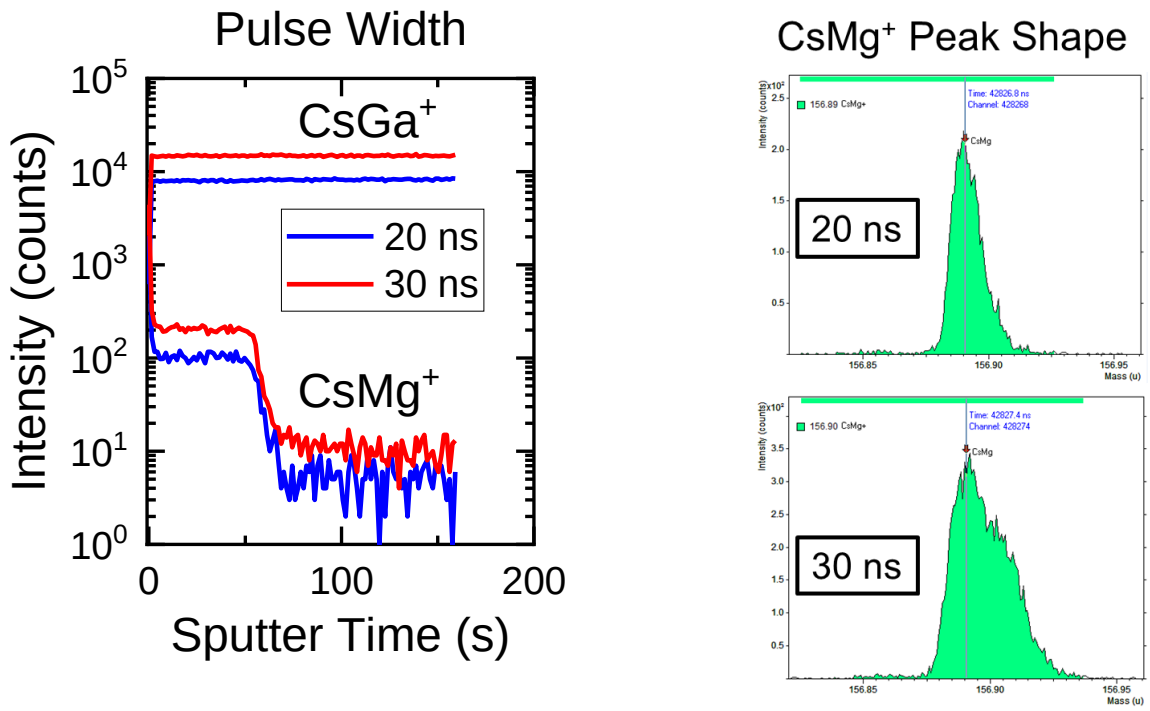


Figure 4.8. ToF-SIMS raw data for wafer 1 comparing pulse width conditions. Left, secondary ion intensities for $CsMg^+$ and $CsGa^+$ over the entire analysis. Right, peak shape of $CsMg^+$ displaying loss of mass resolution.

Next, the effect of an increased number of analysis frames on the measured secondary ion intensity was tested. The real time for a data cycle on the ION TOF SIMS V when a single analysis frame is selected is 3.6 s: 1.64 s for the analysis frame, 1.46 s for the sputtering frame, and 0.5 s of pause time. Increasing the number of analysis frames to 10 per data cycle results in a total time of 18.34 s. This had the most significant effect on secondary ion intensity. From a statistical standpoint, it is obvious that 9 additional acquisitions per depth should have a large increase in the detected ions, so long as the counts for a single frame were greater than one. The results comparing the increase of analysis frames for both 20 and 30 ns pulse widths are plotted in Fig. 4.9.

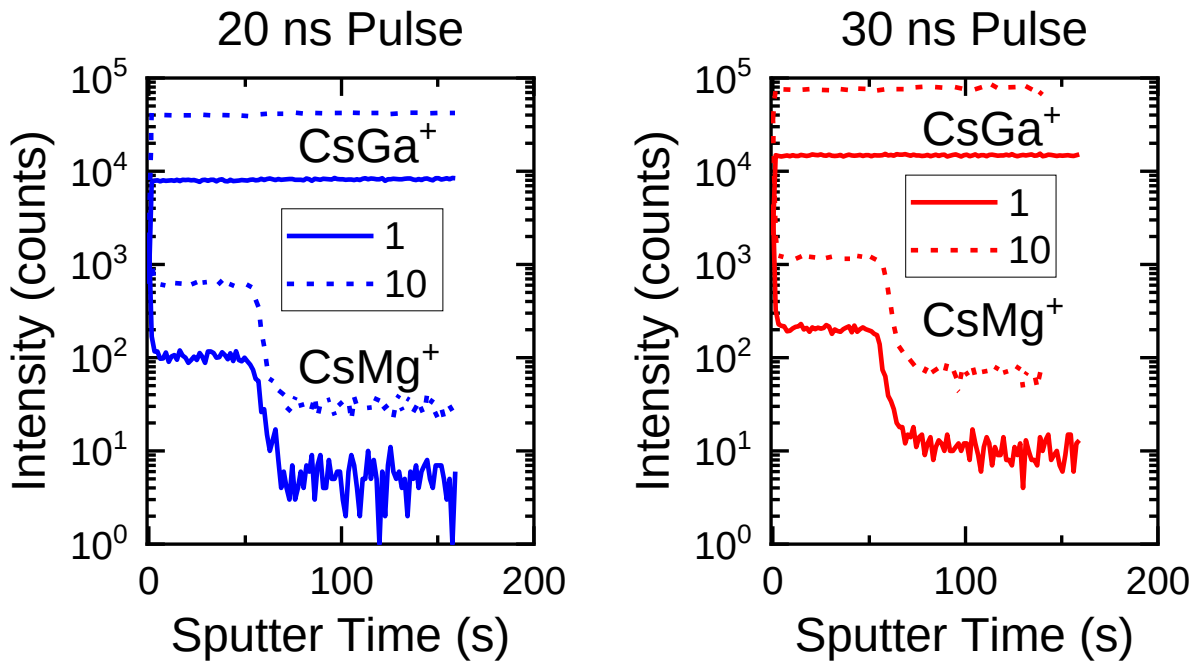


Figure 4.9. ToF-SIMS raw data for wafer 1 comparing number of analysis frames. Solid lines are the one frame condition. Dashed lines are the 10 frame condition. Left, comparing at 20 ns pulse width. Right, comparison at 30 ns pulse width.

In the case of the 20 ns pulse width condition in Fig. 4.9, left, $CsMg^+$ counts increased from 100 to 639 in the more highly Mg-doped layer near the surface. In the second doping layer, the counts increased from 5 to 31. Overall, a ~6x increase was observed for this secondary ion molecule. Meanwhile, the average $CsGa^+$, taken over the entire profile, increased by 5x to 40940 counts from 8,101. Moving to the 30 ns pulse width condition in Fig. 9, right, $CsMg^+$ jumped to 1213 and 71 counts, from 210 and 10, respectively. Roughly a 7x increase was observed by moving to the 10 frame condition. Similar to the 20 ns condition, the $CsGa^+$ concentration increased only 5x with the increase to 10 frames. A summary of all counts is listed in Table 4.3.

Both the pulse width and analysis frame changes employed in this study yielded increases in the secondary ions for both matrix and impurity species monitored in the sample. As alluded in the beginning, ToF-SIMS sensitivity limitations are due to a limited number of primary ions

striking the sample surface during an analysis cycle. The pulse width of the Bi_3^+ beam directly addresses this issue. By definition, the pulse width multiplied by the beam current gives the total ions through a given area. Thus, the increased time window boosts the number of primary ions striking the sample, and yields an increase in secondary ion yield. The main concern is the loss of mass resolution as shown in Fig. 4.8. However, a quick scan at 20 ns condition can confirm if any interferences exists. If not, then the pulse width could be increased without fear of interference.

Table 4.3. Average, standard deviation, and relative standard deviation results for $CsMg^+$ and $CsGa^+$ for wafer 1, all analysis conditions.

		1 Frame		10 Frames	
		20 ns Pulse	30 ns Pulse	20 ns Pulse	30 ns Pulse
$CsMg^+$	Average Counts	106.5	210.9	639.5	1,213.8
Layer 1	Standard Deviation	15.1	25.4	77.3	149.9
	RSD	14%	12%	12%	12%
$CsMg^+$	Average Counts	5.6	10.4	31.1	71.6
Layer 2	Standard Deviation	2.4	2.8	4.9	10.4
	RSD	43%	26%	16%	15%
$CsGa^+$	Average Counts	8,101.7	14,814.7	40,940.5	77,160.8
	Standard Deviation	169.9	215.6	1,239.0	2,931.7
	RSD	2%	1%	3%	4%

Increasing the analysis frames also increases the number of primary ions impacting the sample for a given cycle. Since the analysis frame has a given number of primary ions projected toward the sample, acquiring multiple analysis frames in each data cycle multiplies the number of ions by the number of frames. Thus, every data point was comprised of 10 analysis scans, or 10 cycles of primary ions striking the sample, that were summed together. The tradeoff for the increased primary ions is the total time for the depth profile. The data cycle time is increased by 6x with the additional 9 analysis frames in the cycle. One way to retain the time would be to increase the sputtering frames in a data cycle. This increases the data spacing, or the depth between

data points, so layer resolution is problematic. In the end, a trade-off occurs between sensitivity, sputtering rate, depth resolution, and analysis time.

To highlight the effect of this improved capability, two additional metrics will be presented. The first is the relative standard deviation (RSD) of the Mg doping layers taken from Fig. 4.9 (values listed in Table 4.3). The RSD is a ratio of the standard deviation to the mean value, and gives an idea of the variation a data point can have from the average value. For SIMS, it can represent a 'confidence' in a reported concentration because it gives the bounds for where 95% of the signal is captured. The lower the RSD, the lower the deviation from the average, and the more likely the reported concentration is the real value in the material.

Here, focus will be placed on the second layer where the Mg doping concentration was $9x10^{18}$ cm^{-3}. For the normal conditions (20 ns, 1 analysis frame), the Mg secondary ion intensity had a RSD value of 43%. Essentially, the actual concentration could vary $\pm 3.4x10^{18}$ cm^{-3} from the average $9x10^{18}$ cm^{-3} concentration reported. This stemmed from the $CsMg^+$ intensity never exceeding 10 counts for the spectrometer. Assuming the $CsGa^+$ average intensity and constant RSF value, that variation in concentration is the difference between the mass analyzer detecting 5.6 (the average) and 2 counts for $CsMg^+$. Considering only 5.6 counts on average were detected for 16 s of analysis time, and that other species like $CsGa^+$ have upwards of several thousands of counts for the same analysis window, is a difference in roughly 3 counts physically meaningful? That question captures the boundaries limiting the quantification of a material. However, with the optimized sensitivity conditions (30 ns pulse width and 10 analysis frames per cycle), the $CsMg^+$ counts now on average reach 71.6 counts for the layer. Despite the standard deviation increasing, the RSD in Mg layer 2 drops to 15%, or a variation of $\pm 1.2x10^{18}$ cm^{-3}. Therefore: (1) any changes in the measured counts for $CsMg^+$ can be considered meaningful and (2) there is confidence in the

reported concentration of $9x10^{18}$ cm^{-3} because it varies only between $8x10^{18}$ and $1x10^{19}$ cm^{-3} at the maximum.

The second feature of increased secondary ion concentrations is that a lower minimum concentration of Mg can be analyzed. In Figure 10, a comparison is made between the so-called normal and optimized conditions for the implant standard originally presented in Fig. 4.7. For consistency with Fig. 9, the same color and lines will be used according the analysis condition. Again, 'normal' represents 20 ns pulse width with a single analysis frame and 'optimized' is the 30 ns pulse width with 10 analysis frames. Note, the data presented in Fig. 4.8 were acquired with the optimized conditions, so they are re-plotted here in the dashed red lines. In contrast are the blue curves using the normal conditions. Viewing the raw data plot, it is clear that both the $CsMg^+$ and the $CsGa^+$ signals are roughly an order of magnitude lower for the normal conditions. For $CsGa^+$, the intensity is still sufficient for accurate data acquisition. For $CsMg^+$, however, the sensitivity is low. The peak signal is only 10 counts for the layer with a large noise level across the implant. Additionally, less than an order of magnitude of signal was detected for the implant, and after 150 s sputter time the signal bottoms at 0 counts detected for the analyzer. This acquisition would be insufficient to (1) determine an RSF from the implant and (2) report meaningful concentrations of Mg. For additional comparison, an artificial RSF was applied to the normal conditions, to match the peak concentration of the implant for the optimized conditions. These are plotted in Fig. 4.10, right. Demonstrated in this plot is the difference in the detection limit, or the minimum measurable concentration. Under the normal conditions, the zero secondary ion count line corresponded to a concentration of $1x10^{18}$ cm^{-3}, so the measurable concentrations for reporting are in the mid 10^{18} cm^{-3}. Under the optimized condition, the minimum measurable concentration is in the low 10^{17}

concentration, comparable to detection limits in an ION TOF IV system using an oxygen sputtering source.[23]

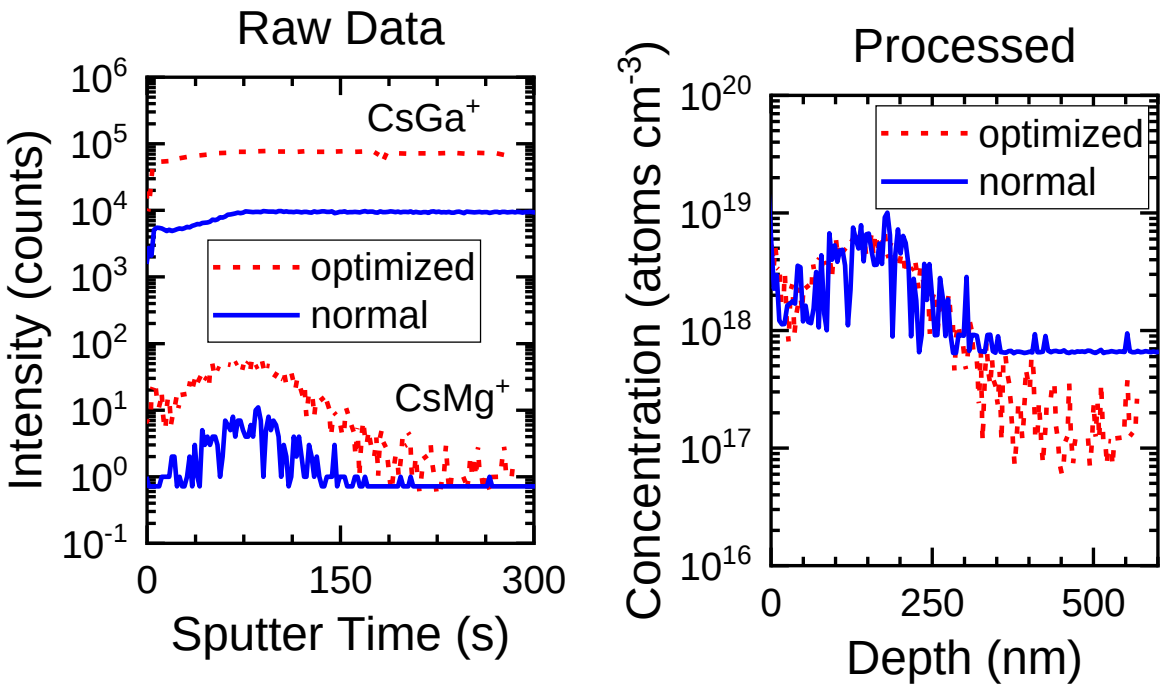


Figure 4.10. Comparison of the normal (20 ns pulse, 1 frame) to the optimized (30 ns pulse, 10 frames) analysis condition for the implant standard in Fig. 8.

In summary, the improvement in the secondary ion intensity for the ION TOF SIMS V paved the way for the study of Mg-doping within this thesis. Prior to this capability, the concentrations of Mg could not be measured at NCSU due to inability to produce an RSF from in-house implant standards (Fig. 4.10). Knowing the total concentration of Mg atoms was necessary to both correlate flow rates of Cp_2Mg to total concentrations, and relate the total concentrations to the electrical performance of the layers. In both cases, SIMS data was able to correct initial expectations for the optimum concentrations in the film.

The final section of this chapter detailing SIMS and the improvements made to ION TOF SIMS V at NCSU are the corrections made for the Mg-doping in AlGaN films. There were several

fundamental challenges once compositions in the materials exceed 60% Al, and these will be addressed in the next section.

6. SIMS of Mg in $Al_xGa_{1-x}N$ ($0 \leq x \leq 1$)

To obtain elemental and compositional data from a SIMS analysis, the data has to be corrected by co-analysis of a standard to establish an RSF. This process is straightforward when a single crystal matrix is analyzed. However, many device structures, including light emitting diodes (LEDs), laser diodes, or films containing polarization doping, contain multiple matrix compositions. Standards of each composition would need to be analyzed to accurately quantity the data over the entire structure. It is therefore easy to imagine a situation where a greater number of standards are required to calibrate the data than the actual number of devices needing characterization. Such a situation is wildly inefficient in both time and cost for an analysis.

In the III-nitrides, the different matrix compositions are obtained by changing the fraction of group III element, such as the Al composition x in $Al_{1-x}Ga_xN$, while nitrogen maintains 50% stoichiometry. For a given impurity, the RSF value changes as a function of this composition x. If possible, finding a trend to relate the change in RSF with the change in the composition x would solve the issue of running multiple standards for a multi-layer device. It is with this intention that the study of Mg in $Al_xGa_{1-x}N$ ($0 \leq x \leq 1$) was undertaken.

One of the major challenges facing the SIMS analysis of the III-nitrides is the tendency for the materials to charge under ion bombardment. Because of their increased bandgap, they behave more as insulators and build up charge instead of conducting. This excess charge can alter the path of the primary ion beam, resulting in reduced secondary ion generation in the region of interest, or it can affect the extraction of the secondary ions. As a result, the targeted peaks in the mass spectrum can be shifted. Fig. 4.11 is an example of such a peak shift.

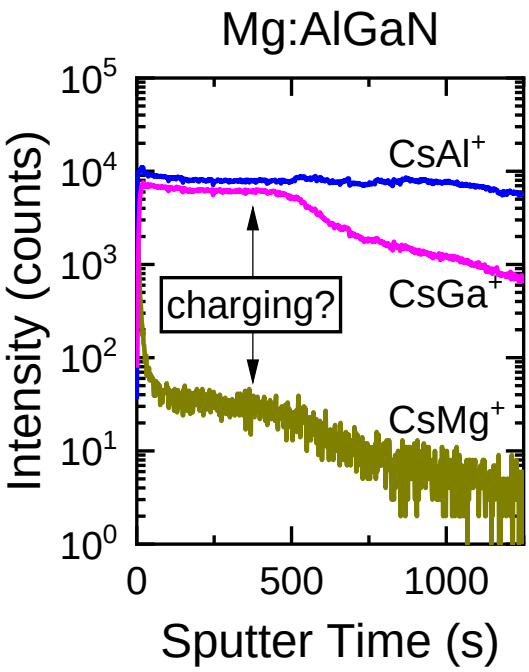


Figure 4.11. SIMS raw data (secondary ion intensity vs. sputter time) for an $Al_{0.75}Ga_{0.25}N$ sample with Mg doping in the top layer of the film.

Comparing the signal of $CsGa^+$ and $CsAl^+$ before and after 500 seconds of sputter time in Fig. 4.11, there is a clear change in the signal > 500 s. However, it is not clear if the $CsMg^+$ signal has been affected since the doping profile ends near this 500 second sputter point. There are two possibilities: all followed species underwent an equal peak shift, or each of the species were affected differently. SIMS quantification is performed by multiplying a relative sensitivity factor by the ratio of species of interest, in this case Mg, to the matrix elements. In the first charging possibility, all species were suppressed equally and would be factored out once the ratio is taken. If the species are affected differently by the charging, however, the ratio is artificially increased/decreased by the difference in the charging. As a result, inaccurate concentrations are reported.

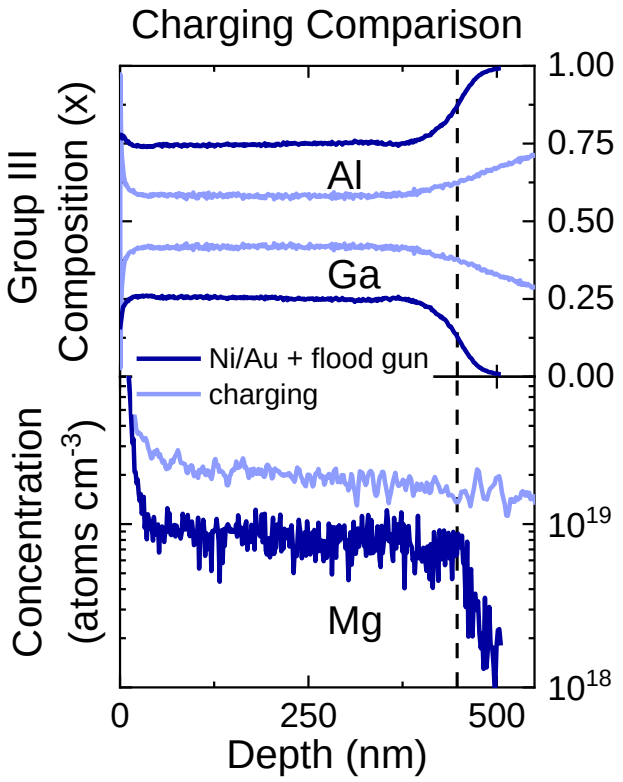


Figure 4.12. Comparison of the quantified SIMS data acquired from the sample (sample charging) with the same sample after the charging has been neutralized.

An additional analysis was performed with the instrument's electron flood gun in operation and the sample coated with the traditional Ni/Au p-type contact for the III-nitrides to remove the sample charging. The electron flood gun works by adding a negative charge of electrons over the sample surface. As the sample starts charging, these electrons can compensate the buildup of charge, which is typically positive. The metal contacts serve as a conduction pathway from the sputtering crater to the sample holder so that a circuit is created. The two methods were individually tested, but insufficient to completely remove the charging effects. The combination yielded the greatest reduction in the charging effects.

From the quantified results in Fig. 4.12, it is clear that the sample charging observed in Fig. 4.11 affected the species differently from one another. In the top plot is a comparison of the Al

and Ga group III atom fraction under the different analysis setups. Before charge neutralization, the resulting composition was ~60% Al for the doped region. This changed to ~75% Al after neutralization was performed. Similarly, the reported Mg concentration changed between the analysis conditions. Here the reported concentration changed from $1x10^{19}$ cm^{-3} to $2x10^{19}$ cm^{-3}. As mentioned in the beginning, if the species were affected similarly by the charging then no change should be observed because the ratio of species should have removed the sensitivity decrease. The ~15% change in reported Al is well outside the instrument error and clear evidence for a difference between $CsGa^+$ and $CsAl^+$. For Mg, the 2x increase is also beyond the 20% error expected from the analysis.

Table 4.4. Composition of wafers sent for ion-implantation.

	1	2	3	4	5	6
XRD composition	0.63	0.7	0.82	0.9	AlN	*GaN

With the ultimate goal of quantifying various dopants and impurities across the GaN to AlN composition space, a series of ion-implanted standards were generated to determine the change in relative sensitivity factor (RSF) as a function of Al composition. Six unintentionally doped samples were grown for ion implantation with varying Al composition. The compositions were determined by x-ray diffraction and are listed in Table 4.4. All 2-inch wafers were simultaneously implanted so the total dose would not vary between the samples. The implanted species include H, C, O, Si, and Mg. At this time only the Mg corrections have been determined. GaN was included because previously calibrated GaN standards could be used to determine the dose for all samples. The other compositions were >60% Al because correction factors <60% Al have been published in literature.[22]

As a requirement for the RSF corrections, the Al composition needs to be quantified with SIMS so that the proper RSF change is applied. Since $CsMg^+$ is the implant species of interest, the

matrix signals of $CsAl^+$ and $CsGa^+$ are followed. The results are displayed in Fig. 4.13. For the composition range $0.63 \leq x \leq 0.9$, a linear trend was determined between the ratio $CsAl^+/CsGa^+$ and x/(1-x). The function displayed in Fig. 13 was utilized at every data point for each of the implant standards, so a correction at every data point is applied.

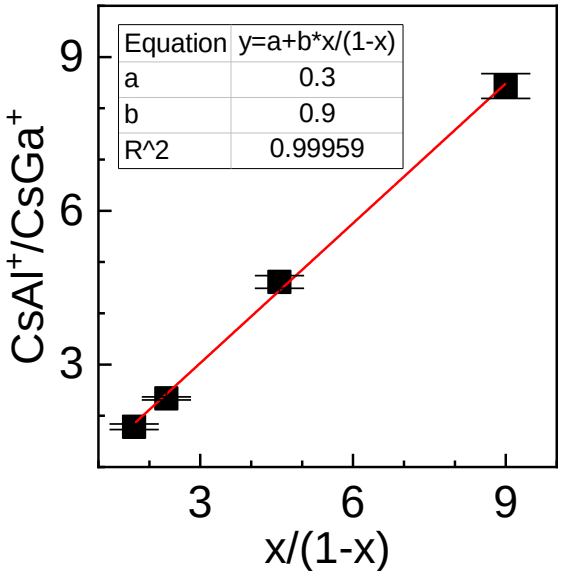


Figure 4.13. Matrix intensity ratio vs. group III composition ratio.

Additionally, corrections to the sputtering rate for the analysis were determined. Since the density of the samples changes with the composition, the etching rate will vary as well. Understanding a priori the change in etching rate as a function of composition will allow for accurate layer determination for an entire depth profile. Without such a correction, multiple depth profiles stopping at each layer are necessary to determine each layer thickness. Because each of the implant standards consists of a single composition, the sputtering corrections as a function of composition were determined. Analyzing each with the same sputtering conditions of 10 keV, 30 nA current, and 120 x 120 μm^2 raster, the total depth of the sputtered crater was converted into a sputtering rate. A polynomial trend was found across all Al composition (x), and the results are plotted in Fig. 4.14.

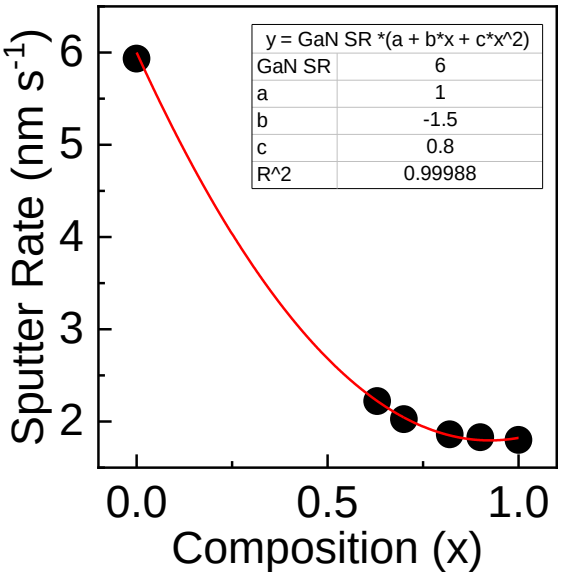


Figure 4.14. Sputtering rate as a function of Al composition.

Finally, the results for the composition corrections to the Mg RSF are presented in Fig. 4.15. To determine the RSF, the $CsMg^+$ signal was used for Mg and compared to the summation of the $CsGa^+$ and the $CsAl^+$ signals for the matrix. The combination was selected so at the extremes of composition, GaN and AlN, the function would not break down (if only $CsGa^+$ or $CsAl^+$ was used then the equation would fail at AlN and GaN, respectively). Two old standards were also included in the data: the previous GaN used to calibrate the doses and a 36% Al AlGaN sample. In Fig. 15, little change to the Mg RSF is observed below this 36% Al value. With only a single data point between GaN and $Al_{0.63}Ga_{0.37}N$ further investigation is needed before an absolute determination is made. However, the compositions greater than 36% Al show a linear increase in the Mg RSF with increasing composition.

It should be noted that the equation in Fig. 4.15 is only valid for the Bi primary gun conditions in this study: Bi_3^+ species, 25 keV, 0.4 pA, 50 ns pulse, and 50x50 μm^2 raster area. For different primary gun operation conditions, the RSF value will change according to an

increase/decrease in energy or number of bombarding species. However, the conditions selected in this study are the typical operating parameters for the instrument. Even after replacement of the Bi filament, the system is able to return to these settings, which is validated by the less than 20% deviation in RSF values for the old GaN implanted standard over a period of three years. Thus, there is some confidence that over time these corrections will remain true for the instrument, barring any charging differences.

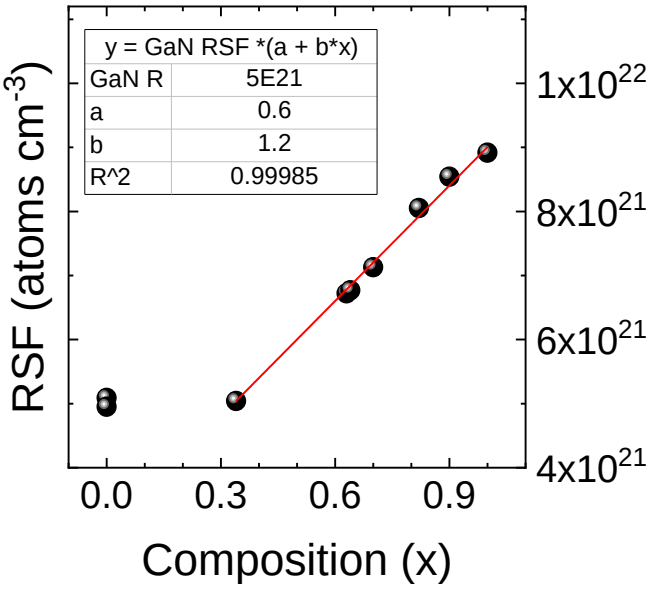


Figure 4.15. Mg RSF (determined from the comparison of the CsMg+ intensity to the combination of the CsGa+ and CsAl+) as a function of Al composition.

7. Conclusion

Secondary ion mass spectrometry is the only analytical measurement technique that can provide the type of sensitivity to measure sub-atomic percent fractions of dopants and impurities in a material. For semiconductors, these parts per million (or billion) fractions control the desired electrical and optical properties of the crystal, and quantifying the precise concentration is required. Within the context of this dissertation, the measured total concentration of Mg is vital to (1) qualify

a particular growth setting and (2) make sense of the electrical and optical performance because of the low ionization efficiency and high compensation of the films. In the case of Mg:AlGaN, the measurements were particularly intriguing because they broke with the previous conventions in the Mg:GaN, where magnesium and hydrogen were expected to match in concentration. This was one of the fundamental discoveries in the dissertation, and only possible as a result of the work presented in the last two parts of this section. These developed capabilities were instrumental to this work.

8. References

1. F. Stevie. Secondary Ion Mass Spectrometry: Applications for Depth Profiling and Surface Characterization. Momentum Press: New York, 2016.

2. J. C. Vickerman, A. Brown, and N. M. Reed. Secondary Ion Mass Spectrometry, Principles and Applications. Clarendon Press: Oxford, 1989.

3. A. Bennighoven, F. G. Rüdenauer, H. W. Werner. Secondary Ion Mass Spectrometry: Basic Concepts, Instrumental Aspects, Applications and Trends. John Wiley & Sons: New York, 1987.

4. R. G. Wilson, F. A. Stevie, and C. W. Magee. Secondary Ion Mass Spectrometry: A Practical Handbook for Depth Profiling and Bulk Impurity Analysis. John Wiley & Sons: New York, 1989.

5. E. Niehuis and T. Grehl. "28 Dual Beam Depth Profiling," in ToF-SIMS: Surface Analysis by Mass Spectrometry. J. C. Vickerman and D. Briggs (eds.), IM Publications: West Sussex, 2001.

6. J. J. Thomson. "Rays of positive electricity," The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 20, 752 (1910).

7. J. Mattauch and R. Herzog. "About a new mass spectrograph," Zeitschrift fur Physik 89, 786 (1934).

8. W. E. Stephens. "A pulsed mass spectrometer with time dispersion," Review of Scientific Instruments 24, no. 8, 616 (1953).

9. W. Paul and H. Steinwedel. “A new mass spectrometer without a magnetic field,” Z. Naturforschg. 8a, 448 (1953).

10. I. Langmuir and K. H. Kingdon, “Thermionic effects caused by vapours of alkali metals,” Proceedings of the Royal Society of London A 107, no. 741, 61 (1925).

11. M. Von Ardenne. “Ion Source,” US2975277A. United States Patent and Trademark Office, 14 March, 1961.

12. J. Zeleny. “The electrical discharge from liquid points, and a hydrostatic method of measuring the electric intensity at their surfaces,” Physical Review 3, 69 (1914)

13. V. E. Krohn, Jr. “Liquid metal droplets for heavy particle propulsion,” in Progress in Astronautics and Rocketry, 5, Electrostatic Propulsion, D. B. Langmuir, E. Stuhlinger, and J. M. Stellen, Jr. (eds.), Academic Press: New York, 1961.

14. J. C. Vickerman. “Static secondary ion mass spectrometry,” in Methods of Surface Analysis. J. M. Walls (ed.), University Press: Cambridge, 1989.

15. P. Sigmund, “Sputtering by ion bombardment: theoretical concepts,” in Topics in Applied Physics, 47, Sputtering by Particle Bombardment I. R. Behrisch (ed.), Springer-Verlag: Berlin, 1981.

16. R. Behrisch, “Introduction and Overview,” in Topics in Applied Physics, 47, Sputtering by Particle Bombardment I. R. Behrisch (ed.), Springer-Verlag: Berlin, 1981.

17. K. Wittmaack. “Oxygen-concentration dependence of secondary ion yield enhancement,” Surface Science 112, 168-180 (1981).

[18] M. L. Yu. “Work-function dependence of negative-ion production during sputtering,” Physical Review Letters 40, no. 9, 574 (1978).

[19] R. Hill. “4 Primary Ion Beam Systems,” in ToF-SIMS: Surface Analysis by Mass Spectrometry. J. C. Vickerman and D. Briggs (eds.), IM Publications: West Sussex, 2001.

[20] K. Wittmaack. “The effect of the angle of incidence on secondary ion yields of oxygen-bombarded solids,” Nuclear Instruments and Methods in Physics Research 218, 307-311 (1983).

[21] A. Klump, C. Zhou, F. A. Stevie, R. Collazo, and Z. Sitar. “Improvement in detection limit for time-of-flight SIMS analysis of dopants in GaN structures,” Journal of Vacuum Science and Technology B 36, 03F102 (2018).

[22] C. Gu, SIMS Quantification Matrix and Impurity Species in III-Nitride Alloys. Diss. North Carolina State University, 2005.

[23] X. Wei, L. Zhao, J. Wang, Y. Zeng, and J. Li. “Characterization of nitride-based LED materials and devices using TOF-SIMS.” Surface and Interface Analysis. 46, 299-302 (2014).

Chapter 5 Conclusion

1. Summary of Key Achievements

The previous three chapters conclude the research investigations to overcome some of the inefficiencies affecting III-nitride LEDs and related devices. As mentioned in the introduction, the high dislocation densities due to the non-native growth substrate and poor p-type conductivity of the Mg_{III} acceptor type dopant are some of the causes for the inefficiency. This dissertation, "Magnesium Doping and Surface Kinetics of III-Nitrides," demonstrates several original accomplishments to address these problems:

- **High supersaturation (4,800 < σ_v < 36,000) growth conditions at wafer offcuts between 0.3 and 0.4° yield smooth step-flow surfaces for GaN homoepitaxy**. A total of four supersaturation conditions ranging from σ_v = 70 to 36,000 were grown on HVPE or ammonothermal GaN substrates. The vicinal offcuts selected ranged from 0.3 to 0.4°. At the lower supersaturation conditions (σ_v = 70 and 500), the surface morphology was dominated by highly facetted regions with 2D nucleation at the lower step edge and very large macro terraces with step-flow morphology. It is hypothesized that the facetted morphologies result from asymmetric adatom fluxes along the surface (ESB) and short residence times on the surface (compared to the diffusion time). The large macro terraces result when the residence time is increased at the σ_v =500 condition. At the higher supersaturations, step-flow morphology is obtained at the higher offcut angle of 0.4°. When the offcut is reduced to 0.3°, step-flow morphology is maintained for σ_v = 4,800 growth. However, a step-meandering morphology appears for the σ_v = 36,000 condition.

- **Compensation and passivation of Mg acceptors in GaN have been controlled using above bandgap illumination and chemical potential control**. A series of GaN films with Mg

concentrations ranging from mid-10^{18} cm^{-3} to mid-10^{19} cm^{-3} were generated under varying diluent gas mixtures (N_2 and H_2) and flow of NH_3. Repeats were also grown with above bandgap illumination directed on the sample surface. Below a concentration of $1.5x10^{19}$ cm^{-3}, minimal nitrogen vacancy related defects appear and instead the film resistivity is dominated by the presence of H passivating the Mg atoms. At this doping level, the growth condition alone had a small effect on the compensation level. The UV-grown samples had significant reduction in the residual donor concentration. For Mg concentrations above $1.5x10^{19}$ cm^{-3}, nitrogen vacancy related defects appear and introduce large donor compensation. Increasing the NH_3 for either diluent gas mixture reduced the residual compensation and the UV-illumination did not change the compensation further. At Mg concentrations of $1.5x10^{19}$ cm^{-3}, the UV-illumination changed the steady-state condition and induced nitrogen vacancy formation as compared to non-UV grown samples.

- **Free hole concentrations have been measured for the first time in Mg:$Al_{0.6}Ga_{0.4}N$ and indicate reduced H-passivation and self-compensation compared to Mg:GaN**. Two growth conditions varying the NH_3 flow from 1 to 3 slm were utilized for Mg concentrations ranging from mid-10^{18} cm^{-3} to $1x10^{20}$ cm^{-3}. SIMS of the samples revealed that the hydrogen concentration remained approximately an order of magnitude lower than the Mg concentration for all doping levels (for Mg:GaN, hydrogen matches the Mg concentration until ~$1.5x10^{19}$ cm^{-3}). PL of the films revealed increased luminescence at 3.7 and 4.2 eV for the growths with NH_3 of 3 slm. It is believed that the emission at 4.2 eV corresponds to the nitrogen vacancy related defect at 2.8 eV in GaN, and plays a role in the highly resistive behavior of these films. For the 1 slm NH_3 growths, the 3.7 eV emission gradually increases with increasing Mg concentration. At the same time, the resistivity decreased with increasing Mg concentration for this growth condition, with a minimum

of 35 Ω cm obtained for an Mg concentration of $8x10^{19}$ cm^{-3}. Alloy scattering is considered the main limitation to the conductivity, as none of the films exceeded a mobility of 0.5 cm^2 V^{-1} s^{-1}. In contrast to GaN, the free hole concentration increased when the Mg doping increased above the saturation limit for hydrogen and the compensation level calculated from the charge balance equation remained near mid-10^{18} cm^{-3}.

- **Secondary ion intensities have been increased for ToF-SIMS by increasing the pulse width of the primary ion beam and by increasing the analysis frames collected for each data point**. ION TOF SIMS V instrument operates a primary ion gun with a pA current. This limited the total number of secondary ions generated because of the dependence on this primary ion fluence. Lower sensitivity elements, such as Mg in the III-nitrides, could not be measured as a result. Previously, the pulse width and the analysis frames conditions for a depth profile were not available to the user, but a recent software update granted access. Increasing the pulse width decreases the mass resolution, but increases sensitivity due to a greater number of primary ions striking the sample surface. Increasing the analysis frames per data cycle allows for multiple iterations of the primary ion beam bombarding the surface and collection of the generated secondary ions. Each data point becomes a summation of all the analysis frames collected before the sputtering ion gun is directed toward the surface, increasing the total secondary ions for a given element.

- **Charge compensation, composition, sputtering rate, and RSF corrections for Mg dopants have been determined for SIMS of $Al_xGa_{1-x}N$ ($0 \leq x \leq 1$) using ION TOF SIMS V**. With increasing Al content in AlGaN, III-nitride films become more insulating and induce charging problems during SIMS analysis. This effect was mitigated by applying a Ni/Au metal stack to the sample surface, and utilizing the flood gun during the sputtering cycle of the depth profile.

Additionally, the sputtering rate of the SIMS profile, and the secondary ion intensity of Mg, Al, and Ga, are altered by an Al content change in the films. To accurately characterize devices such as LEDs, correction factors when the composition changes are required. A series of six $Al_xGa_{1-x}N$ films ($0.63 \leq x \leq 1$), were implanted with Mg. Under constant SIMS analysis conditions depth profiles were acquired on the films. Measuring the profile crater with profilometry allowed for the sputtering rates to be established (total depth divided by sputtering time), and a polynomial correction to the sputtering rate in GaN as a function of Al content x was calculated. For the composition correction, the ratio of the secondary ion intensity of Al to Ga was measured and related to the composition x. Finally, the RSF of Mg was determined for each of the wafers, and a correction based on the Al composition x was made to the RSF in GaN.

The application of these achievements to future visible and UV-LED devices opens the possibility for exciting new performance capabilities. In terms of the visible LEDs, the next generation will be grown on the commercially available GaN bulk substrates produced by HVPE or the ammonothermal method. The work in this dissertation provides major guidelines for growth on these substrates, including the starting offcut and selection of the vapor supersaturation, to maintain smooth step-flow morphology. In terms of the p-type layers, higher ammonia flows to suppress the nitrogen vacancy type compensators in GaN are suggested. It may also enhance the widely utilized p^{++} layer directly on the surface of the thin film, which has Mg concentrations typically in the $\sim 1 \times 10^{20}$ cm^{-3} range. For the p-type layers in high Al-content AlGaN devices, moderate ammonia flows (1 slm) to avoid pre-reaction and higher compensation levels are suggested. The discovery that Mg concentrations in excess of 6×10^{19} cm^{-3} are possible because of the diminished H-passivation and self-compensation is the seminal finding in this dissertation, and should encourage further study.

2. Future Work

In addition to the studies already presented, the following future work is proposed.

- **Surface kinetics of GaN homoepitaxy**: Starting offcuts between 0.2 to 0.4° should be grown with the σ_v = 4,800 as a pretreatment before additional conditions are tested. One of the major limitations in this study was the inability to determine the starting terrace width and step height before the growths. The σ_v = 4,800 condition provided unit cell height terraces that varied in width according to the starting offcut. With this information it would be possible to know the starting conditions on the surface, and the evolution after a different supersaturation condition is utilized. The reversibility of the morphologies should also be tested. For the lower supersaturation conditions, longer growth times to see if the macro terraces stabilize or eventually overtake one another is proposed.

- **Mg doping in GaN**: The doping series for the N_2 diluent, 3 slm NH_3 flow should be extended to determine the minimum achievable resistivity. Then SIMS of the films before and after the activation anneal are necessary to determine the concentration of H relative to the Mg doping levels. For the H_2 diluent, 3 slm NH_3 flow, SIMS to determine the concentration of H are also required. Temperature dependent Hall effect measurements should be performed on all samples to confirm the ionization energy of the Mg acceptor, and determine compensation levels as a function of temperature. DFT calculations to determine the source of self-compensation are needed. Last, growths can be repeated at lower temperatures to see if the chemical potential control of the compensation can be further enhanced.

- **Mg doping in AlGaN**: Further exploration of the growth thermodynamic space is suggested to further reduce the compensation of the films. Since the mobility appears restricted by alloy scattering, the only pathway to increase the conductivity in the films is to drive the

compensation level down. Specifically, changing the temperature and utilizing above bandgap illumination should be explored. The dislocation density may also have an effect, and it is suggested that these growths be repeated on native AlN substrates. Characterization of the defect luminescence below the band edge in these Mg:$Al_{0.6}Ga_{0.4}N$ films is also necessary. There is a tentative link between the 3.7 and 4.1 eV peaks with the conductivity of the films. However, the exact compensators are not known at this time.

- **SIMS for $Al_xGa_{1-x}N$ ($0 \leq x \leq 1$)**: Wafers utilized for the Mg RSF calibration for high Al-content AlGaN have been co-implanted with additional impurities: H, C, O, and Si. These can be measured using the negative ion detection mode for SIMS and correction factors would allow the complete measurement of UV-LED and laser diode devices with SIMS. However, high chamber backgrounds of H, C, and from the organic species being analyzed in the shared ION TOF SIMS V system mean that separation of ^{28}Si from $^{27}Al+H$ or $^{12}C^{16}O$ needs consideration. Additional implantation samples are necessary for the complete extension of the correction factors for $Al_xGa_{1-x}N$ with $x<0.6$.