Strategy for reliable growth of thin GaN Caps on AlGaN HEMT structures

Alexander M. Hinz a, Saptarsi Ghosh a, Simon M. Fairclough a, James T. Griffiths a, Menno J. Kappers a, Rachel A. Oliver a, David J. Wallis a,b,c

a Department of Materials Science and Metallurgy, 27 Charles Babbage Road, University of Cambridge, Cambridge, CB3 0FS, United Kingdom
b Centre for High Frequency Engineering, University of Cardiff, 5 The Parade, Newport, Cardiff CF24 3AA, United Kingdom

1. Introduction

With the maturing of III-nitride semiconductor device technology there is not only a continued focus on improving device performance but also on improving device reliability [1–4]. This is especially true for AlGaN-based high electron mobility transistors (HEMTs) that have already entered commercial production. The key feature of a HEMT is a 2-dimensional electron gas (2DEG) that provides a conductive path between the source and the drain of the transistor. In AlGaN based HEMTs this 2DEG is formed between a thick GaN layer and a thin AlGaN barrier. Formation of the 2DEG is further aided by the intrinsic spontaneous polarisation of the nitride semiconductors results in the formation of the 2DEG at the interface between the GaN layer and the AlGaN barrier. The formation of the 2DEG is further aided by the formation of electric fields due to piezoelectric effects [5]. If the AlGaN barrier is thin enough, i.e. below the critical thickness for relaxation, it grows pseudomorphically on the underlying GaN layer and is thus under tensile stress. The resulting electric field enhances the sheet carrier density n_s of the 2DEG. Thus, the integrity of the AlGaN barrier is very important to maintain the performance of these HEMTs. Relaxation of the barrier has to be avoided. One approach to protect the AlGaN barrier is to grow a thin (≤ 5 nm) GaN cap on top of the barrier. The GaN cap prevents oxidation of the barrier and has also been found to prevent the formation of extended surface defects in the barrier that are associated with degradation in device performance [6–8]. Thin GaN caps have also been shown to prevent current collapse in HEMTs, reduce gate leakage and, in general, increase device robustness [9–11].

However, growing such thin GaN caps is challenging [12,13]. It is well known that GaN is unstable above about 850 °C and under the aggressive atmospheres found in organometallic vapour phase epitaxy (OMVPE) reactors [14,15]. Over the years numerous studies have been devoted to investigating the instability of GaN under relevant conditions [15–30]. The GaN dissociates and desorbs from the surface if the flux of precursors is too small to maintain net growth of the layer. This is especially true when cooling down the epilayers from the growth temperature during which the group-III precursor flux into the reactor is usually stopped. In the worst case, the entire GaN cap could desorb which would leave the AlGaN barrier exposed. As the GaN cap is rather thin, transmission electron microscopy (TEM) is one of the few techniques, which is also widely available, that can directly determine the cap thickness. However, as preparing the necessary high-quality TEM cross sections is notoriously difficult, it would be advantageous to use a growth strategy that is robust enough so that constant quality
control by TEM is not necessary. The growth strategy for the GaN cap must not only be able to prevent its complete desorption but should also be able to maintain the desired cap thickness. If the actual cap thickness differs from the target thickness, interpretation of electrical characterisation data becomes treacherous. This is especially true if one compares the measured data to models that rely on precise estimates of layer thicknesses. Ultimately, this might also lead to problems with quality control in industrial settings. If electrical performance deviations are not correctly identified as missing GaN caps but are attributed to other problems resulting in futile changes to process recipes and a decreased yield. Therefore, the aim of this paper is to present a growth strategy that addresses all of the concerns above.

2. Experimental details

All HEMT structures were grown in an AIXTRON 1×8′′ close coupled showerhead reactor on (111) oriented 6′′ Si wafers (Shin-Etsu Handotai Europe Ltd.). If not mentioned otherwise always temperatures refer to wafer surface temperatures as measured by the emissivity-corrected pyrometer of an EpsCurve® PT system (LayTec AG) installed on the reactor. Prior to the growth of the epilayers the wafers were heated to ca. 1080 °C in a hydrogen (H₂) atmosphere for 30 min to desorb the native oxide layer. The epilayer stack of the HEMT structure consists of a 250 nm AlN nucleation layer, a 1700 nm graded AlGaN buffer (from 85 at.% Al to 25 at.% Al), a 1000 nm GaN buffer, a nominally 1 nm thick AlN spacer, a 21 nm AlGaN barrier with nominally 25 at.% Al, and a GaN cap with a target thickness of 2 nm. All layers were grown using trimethylaluminium (TMAI), trimethylgallium (TMGa), and ammonia (NH₃) as the precursors and H₂ as the carrier gas. The GaN caps were grown at 1045 °C in an H₂/NH₃-atmosphere at 100 mbar. Injecting a TMGa flux of 86 μmol/min and an NH₃ flux of 6 slm into the reactor for the cap growth resulted in a V/III ratio of 3060 and a net growth rate of 0.18 nm/s as determined by analysing the 633 nm reflectance trace of a thick GaN layer grown under the same conditions. None of the GaN caps grown in this study were intentionally doped. A FEI Titan 80–300 transmission electron microscope (TEM) running at 300 kV with a convergence angle of 24 mrad and collection angle of 60–80–300 transmission electron microscope (TEM) were used to make a direct quantification of the composition in a certain region, used to make a direct quantification of the composition in a certain region, in the ternary Al₆Ga₄N₁₋ₓ system it is a good indicator for changes in the metallic site fraction x. Regions with higher x, i.e. higher Al content, should appear darker than regions with lower x. The Z-contrast is not the only contribution to the overall contrast in darkfield TEM images. However, in the epitaxial layer stacks that are relevant for this work, one expects rather sharp changes in the Z-contrast of layers with sufficiently different metallic site fractions so that the contribution from, for example, changes in the TEM foil thickness should only be a minor contribution to the overall contrast. Thus, for cross sections of the HEMT structure studied in this work one would expect the following contrast pattern at the top of the structure: a bright layer corresponding to the GaN buffer layer, followed by a darker layer corresponding to the AlGaN barrier, followed by another layer corresponding to the GaN cap with a similar brightness as the GaN buffer layer. But, we have found that if one tries to grow a GaN cap with a simple growth procedure consisting of simply turning off the Ga source at the end of the growth one observes the following contrast pattern: as expected, the bright GaN buffer followed by the darker AlGaN barrier followed by another thin layer that is even darker than the AlGaN barrier. This is a complete reversal of the expected contrast change between the AlGaN barrier and a GaN cap. An example of this contrast change is depicted in Fig. 1. It indicates not only that there is no GaN cap, but that the top surface of the AlGaN barrier has a higher Al content than the remaining AlGaN barrier underneath. This is consistent with energy-dispersive X-ray spectroscopy (EDX) maps, an example of which can be found in the Supporting Information to this article. Note that in conventional bright field TEM imaging this thin high Al content surface layer could be easily mistaken for the GaN cap, leading to the erroneous conclusion that the GaN cap growth has been successful. This unexpected contrast pattern can be explained by a complete desorption of the GaN cap followed by a preferential desorption of GaN from the surface of the exposed AlGaN barrier. Preferential desorption of GaN from the AlGaN barrier is not unexpected as GaN has a lower bond strength than AlN and is thus more prone to dissociation and subsequent desorption. The enrichment of the Al at the top of the barrier is not an artifact from the barrier growth itself. The growth is never stopped during the transition from AlGaN barrier to GaN cap. To grow the GaN cap on top of the AlGaN barrier injection of TMAI into the reactor is simply stopped slightly earlier than that of TMGa. Therefore, one would rather expect a slightly higher Ga content at the top of the barrier.

As the preparation of TEM cross sections for verifying the successful growth of GaN caps on a regular basis is usually unfeasible, one can resort to AFM scans of the epilayer stacks′ top surfaces. Structures with exposed AlGaN barriers have characteristic fissure-like surface features [6–8,31,32]. An example of these features is depicted in Fig. 2. In this particular structure a nominally 2 nm GaN cap was grown on top of the AlGaN barrier. The entire stack was grown using a recipe similar to the one that resulted in the structure depicted in Fig. 1. Thus, it appears plausible that the AFM micrograph shows the exposed AlGaN barrier. In structures without or with thinner caps the fissures become deeper beginning to merge into a network. The onset of this can, for example, be seen in the lower half of Fig. 2. Similar surface features have been observed in various studies of GaN capped and uncapped AlGaN layers. In previous work by Cheng et al. it was shown that the formation of the fissures progresses over time if an AlGaN barrier is exposed to the H₂/NH₃-atmosphere of an OMVPE reactor at high temperatures. Furthermore they showed that the formation of fissures can be suppressed by the deposition of a SiNx protective layer immediately following terminating the growth of the AlGaN barrier. They have argued that the SiNx protective layer prevents the preferential desorption of GaN from the barrier and thus suppresses the formation of fissures [33]. As an uncapped AlGaN barrier that sits fully strained on the GaN buffer should be under considerable tensile stresses it is not difficult to imagine that these stresses are the driving force for the formation of the fissures.

Desorption rate. To develop a growth strategy for reliably growing thin GaN caps it is useful to get a reasonable estimate of the actual GaN desorption rate. This estimate can be conveniently obtained from analysis of the 633 nm reflectance trace following the method proposed by Grandjean et al. [34]. The method will be briefly explained below. During certain intervals of the oscillation period the reflectance trace can be approximated by a straight line, cf. left hand side of Fig. 3. For a constant refractive index, i.e. constant composition, temperature, and roughness, the reflectance is only a function of the GaN epilayer thickness and it can be shown that the growth rate $g$ is

$$g = \text{const.} \times n_g$$

(1)
is then proportional to the slope $m_g$ of that straight line. The proportionality constant is only dependent on the material of the epilayer, i.e. GaN in this case. Thus, if the growth rate is known for one straight line section of the oscillation period one can easily deduce the growth rate in another straight line section by comparing their slopes. This works of course also for negative growth rates, i.e. desorption rates $d$:

$$d = g \frac{m_d}{m_g} \quad (2)$$

using the corresponding straight line slopes $m_g$ and $m_d$. Practically, estimating $d$ consists of two steps illustrated in Fig. 3. In the first step one grows a sufficient thickness of GaN, i.e. more than one full oscillation of the trace, so that one can extract the growth rate $g$ from modelling the reflectance trace. The slope $m_g$ is extracted from the same part of the trace. This is best done by using portions of the reflectance trace close to the inflection points, as the reflectance trace is well approximated by a straight line in those areas. In the second step one obtains $m_d$ from the trace. Therefore, one has to ensure that the desorption only takes place during an interval of the oscillation in which the straight line approximation is valid. As the desorption rate is, under normal growth conditions, much smaller than the growth rate one can conveniently fulfill this requirement by simply shutting off the TMGa injection shortly after having passed the midpoint of the straight line interval during layer growth. This is illustrated in Fig. 3. The TMGa injection was shut off just a few seconds before the dashed vertical line marking the transition between growth and desorption. Before changing the growth conditions, for example the temperature, for the next cycle it was found to be useful to inject additional TMGa briefly into the reactor. As the reflectance will change rapidly as film growth resumes, the desorption period is well defined, cf. far right side of Fig. 3. The entire procedure can then be repeated for another set of parameters so that one can obtain multiple data points within the same run. The only limitation for the number of data points that can be obtained per run is the quality of the reflectance trace. If, for example, the GaN layer begins to roughen so much that one cannot extract the growth rate reliably anymore, one should start on a fresh structure.

The approach presented above was used to determine the desorption rate of GaN in an $\text{H}_2/\text{NH}_3$-atmosphere for a variety of surface temperatures. To speed up the growth of the GaN layers a higher TMGa flux was used than for the cap growth resulting in a growth rate of ca. 0.8 nm/s depending on the growth temperature. As expected, the desorption process is thermally activated and follows an Arrhenius-type law with an activation energy on the order of $E_a \approx 2$ eV. The derived Arrhenius law is subsequently used to predict the cap desorption. One should keep in mind that the wafer surface temperature was obtained by a pyrometer. Thus, the accuracy of the absolute value of the surface temperature depends strongly on the calibration of the pyrometer. Thus, it is highly advisable to determine the temperature dependence of the desorption rate for every reactor individually.

**Cap growth.** One strategy to prevent the desorption of the GaN cap would be to simply grow the cap at low temperatures where desorption is not an issue. However, that would leave the AlGaN barrier exposed during the cool down to the lower GaN cap growth temperature, which
A.M. Hinz et al.

Fig. 3. Representative data set for illustrating the approach to determine the desorption rate $\delta$ of GaN from the 633 nm reflectance trace. The dashed vertical lines mark the transition between the growth and the desorption periods. The desorption rate can be related to the growth rate $g$, which is obtained from modelling the reflectance trace in the left-hand side of the figure, from the ratio of the linear slopes $m_g$ and $m_d$, cf. (2).

would probably also result in the formation of fissures and thus defeat the purpose of having the GaN cap in the first place. A more elaborate version of this approach would start the growth of the GaN cap directly after finishing the growth of the AlGaN barrier at high temperatures and then rapidly ramp down the wafer surface temperature to a value at which desorption is negligible, for example 850 °C, while continue to grow the GaN cap. Thus one would avoid exposing the AlGaN barrier and prevent desorption of the cap as one maintains a finite growth rate during the cool down. This approach does indeed give the desired results (cf. Fig. 4), however there are two potential issues with this approach. Firstly, if one aims for a thin cap thickness it is difficult to maintain control of the cap thickness. As even cooling down a wafer in a small OMVPE reactor, like the one used in this work, to 850 °C takes several tens of seconds, one would either need to use a very small growth rate to begin with or ramp down the growth rate simultaneously to ramping down the wafer surface temperature. The former requires to grow the barrier at a low growth rate as well and being able to maintain control of a rather low growth rate which is in general more difficult. The latter makes predicting the final cap thickness difficult. One would not only need to know how the growth rate changes with temperature, but also at what rate the growth rate is being ramped down. In an industrial production system one could of course simply calibrate the recipe to the desired thickness and just maintain the stability of the process, however for a research reactor this can be challenging. If it is required to grow caps with varying thicknesses one would probably have to resort to verifying the cap thickness by TEM and adjust the recipes accordingly. Secondly, another potential issue might be the high carbon doping levels one would expect to build up in the GaN cap at lower growth temperatures [35,36].

Another strategy for preventing the desorption of the cap is using controlled desorption of the GaN cap. If the growth rate of the GaN cap is known, one can easily predict the thickness of the GaN cap before desorption. One then just needs to control the cool down of the wafer in such a way that the final GaN cap thickness coincides with the target value. If the temperature dependence of the desorption rate is known, the desorption of the cap can be predicted if the time evolution of the wafer surface temperature is known. An example for predicting the cap desorption is given in Fig. 5. The inset shows the time evolution of the surface temperature, i.e. the surface temperature of the Si wafer, after the growth of the GaN cap has been stopped, i.e. the TMGa flow into the reactor has been shut off and the heater starts to ramp down.

As the pyrometer only records the Si wafer surface temperature ca. every 4 s, additional data points have been interpolated between the actual data points. By integrating the time evolution of the desorption rate up to a given point after stopping the cap growth one obtains the cumulative thickness that has desorbed from the cap. As expected, most of the GaN desorbs from the cap while the wafer surface temperature is still high. Using plots like the one in Fig. 5 one can easily predict the desorption of the cap and adjust either the cool down rate or the starting thickness of the cap as required. But, the reader should be aware that the prediction of the cumulative desorbed thickness is based on a continuum approach, i.e. the GaN cap is treated as a continuous layer. However, on the relevant thickness scale this is actually a crude assumption. With a bulk $c$-lattice constant of 5.19 Å for wurtzite GaN, a desorbed thickness of 2 nm corresponds to only about 8 layers of Ga atoms. Thus, even if one has detailed knowledge about the desorption
process, controlling the cap thickness is a formidable challenge. At most one could expect a controllability down to 1 atomic layer, i.e. ca. 0.2 nm. At a total cap thickness of 2 nm this is already 10%. However, taking into account that controlling the cap thickness on this level does not only require to maintain a constant and stable growth rate across the entire wafer, but also a very high homogeneity of the wafer surface temperature across the wafer during the cool down, one should expect the controllability of the cap thickness to be considerably less.

Cap characterisation. Nevertheless, the approach outlined above has been successfully used to prevent the desorption of thin GaN caps and predict their final thickness. An AFM micrograph of a thin GaN cap grown by this approach is displayed in Fig. 6. There is no evidence of fissure formation or substantial roughening of the cap surface by the desorption. The root mean square roughness of \(5 \mu \text{m} \times 5 \mu \text{m}\) scans is lower than 0.6 nm. The small pits in the surface are common surface pits associated with threading dislocations that penetrate the top surface of the GaN cap. A notable difference between this intact cap and the exposed AlGaN barriers is the apparent absence of surface pits associated with threading dislocations in the AFM micrographs (cf. Fig. 2). The most likely explanation is that the surface fissures actually start at the threading dislocations. In that case the small pits would be indistinguishable from the surface fissures. This would be consistent with the findings by Cheng et al. although they identified only edge dislocations as the origin of the surface fissures [37]. It is notable that the small pits associated with threading dislocations are not readily apparent in samples with exposed AlGaN barriers (cf. Fig. 2). The most likely explanation is that the surface fissures actually start at the threading dislocations. In that case the etch pits would be indistinguishable from the surface fissures. This would be consistent with the findings by Cheng et al. although they identified only edge dislocations as the origin of the surface fissures [37].

Measuring the cap thickness in STEM cross-sections confirmed that the final cap thickness was predicted successfully. The initial cap thickness was predicted to be 3.8 nm. For the cool down ramp used in this particular growth run a desorption of 2.2 nm was predicted based on the data in Fig. 5. Thus, the final cap thickness was predicted to be 1.6 nm. The actual cap thickness was determined to be slightly thinner at about 1 nm corresponding to 5 rows of Ga atoms or 2 wurtzite GaN unit cells. (cf. Fig. 7). Thus, the difference between the predicted and the actual cap thickness is on the order of one wurtzite GaN unit cell. Given that the contrast between the cap and barrier in Fig. 7 is not particularly sharp the cap might be slightly thicker. As mentioned earlier, given that the desorption prediction is continuum in nature, but the GaN cap is fairly discrete one cannot expect the prediction to be more accurate than about half a unit cell. Furthermore, the initial thickness of the cap might also have been overestimated. It would also not be surprising that the amount of desorbed GaN is being underestimated. The precise onset of cap desorption is not known, so the prediction uses the shut-off of the precursor flow as a reference point. By including a suitable delay it should be possible to refine the accuracy of the prediction. Nevertheless, on an absolute scale, i.e. in terms of atomic layers, the accuracy of the prediction is promising and serves its purpose in ensuring that the GaN cap remains intact.

Stability of capped and uncapped barriers. An important aspect of using capped barriers is the long-term stability of the barrier. Fig. 8 shows AFM micrographs of HEMT structures with intact and desorbed GaN caps. The HEMT with intact GaN cap shows no apparent signs of degradation even 15 month after it had been grown. However, 17 month after its growth the uncapped barrier shows not only the typical
surface fissures but also long elevated streaks that appear to follow the symmetry of the GaN crystal structure. Such degradation, over long time periods, is expected to impact device reliability.

Another aspect that affects the performance of HEMT transistors is the thickness homogeneity of the GaN cap. Comparing the AFM micrographs depicted in Fig. 8(a) and (c), the rms roughness of the GaN cap and the AlGaN surface from which the GaN cap has desorbed are 1.0 nm and 0.78 nm respectively. One might infer based on this small difference that the GaN cap thickness varies over lateral distances of a few microns. However, it is more likely that the GaN cap grows conformally on the AlGaN barrier and the difference in RMS roughness reflects local inhomogeneities of the total film thickness. Deviations in the GaN cap thickness would affect transistor characteristics, like the pinch-off voltage, as the distance between the gate electrode and the 2DEG is changed. Because the GaN cap is in general much thinner than the AlGaN barrier, deviations in the transistor characteristics are expected to be minor, but for high reliability devices or applications requiring tight transistor specifications they could be relevant. Thus,
4. Conclusions

In this work it has been shown that the reliable growth of thin GaN caps on AlGaN HEMT structures is challenging as GaN is unstable at conventional OMVPE growth conditions and tends to desorb during cool down of the wafer. Complete desorption of the GaN results in exposure of the AlGaN barrier in these structures which leads to the formation of fissures and an AI enriched layer at the top of the barrier. Both the fissures and the Al enriched surface can be used as indicators for unsuccessful growth of the cap. Based on measurements of the temperature dependence of the GaN desorption rate a growth strategy has been developed to reliably grow thin GaN caps. It relies on intentionally growing a thicker cap and controlled desorption of the excess thickness to prevent exposure of the AlGaN barrier and maintain the target cap thickness. The GaN caps grown in this way show no signs of fissure formation or significant roughening by the controlled desorption. GaN caps with thicknesses as small as 1 nm can be reliably grown with this approach. These thin GaN caps also appear to prevent the degradation of AlGaN barriers over time.

CRediT authorship contribution statement

Alexander M. Hinz: Conceptualization, Formal analysis, Funding acquisition, Investigation, Writing – original draft. Saptarsi Ghosh: Investigation, Writing, review & editing. Simon M. Fairclough: Investigation, Writing – review & editing. James T. Griffiths: Investigation, Writing – review & editing. Menno J. Kappers: Methodology, Writing – review & editing. Rachel A. Oliver: Funding acquisition, Supervision, Writing – review & editing. David J. Wallis: Conceptualization, Funding acquisition, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

Dr. Alexander Hinz would like to acknowledge funding for his position at the University of Cambridge in the framework of a Research Fellowship by the Deutsche Forschungsgemeinschaft (DFG), Germany under the grant number HZ141/1-1. Prof. D J Wallis would like to acknowledge support of EPSRC, United Kingdom fellowship EP/N01202X/2. Funding from EP/P00945X/1 and the EPSRC National Epitaxy Facility, United Kingdom is also acknowledged.

References


A.M. Hinz et al.


