Title
Influence of plasma-based in-situ surface cleaning procedures on HfO2/In0.53Ga0.47As gate stack properties

Permalink
https://escholarship.org/uc/item/14z7c6b7

Journal
Journal of Applied Physics, 114(15)

ISSN
0021-8979

Authors
Chobpattana, Varistha
Mates, Thomas E.
Mitchell, William J.
et al.

Publication Date
2013-10-16

DOI
10.1063/1.4825259

Peer reviewed
Influence of plasma-based in-situ surface cleaning procedures on HfO$_2$/In$_{0.53}$Ga$_{0.47}$As gate stack properties

Varishta Chobpattana, Thomas E. Mates, William J. Mitchell, Jack Y. Zhang, and Susanne Stemmer

Materials Department, University of California, Santa Barbara, California 93106-5050, USA

(Received 25 September 2013; accepted 30 September 2013; published online 16 October 2013)

We report on the influence of variations in the process parameters of an in-situ surface cleaning procedure, consisting of alternating cycles of nitrogen plasma and trimethylaluminum dosing, on the interface trap density of highly scaled HfO$_2$ gate dielectrics deposited on n-In$_{0.53}$Ga$_{0.47}$As by atomic layer deposition. We discuss the interface chemistry of stacks resulting from the pre-deposition exposure to nitrogen plasma/trimethylaluminum cycles. Measurements of interface trap densities, interface chemistry, and surface morphology show that variations in the cleaning process have a large effect on nucleation and surface coverage, which in turn are crucial for achieving low interface state densities. © 2013 AIP Publishing LLC.

INTRODUCTION

In$_{0.53}$Ga$_{0.47}$As is currently being investigated as a channel material for III-V metal-oxide-semiconductor field effect transistors for post-Si complementary metal-oxide-semiconductor (CMOS) applications. To reduce the large trap densities at interfaces between In$_{0.53}$Ga$_{0.47}$As and the gate dielectric (typically Al$_2$O$_3$ or HfO$_2$), various surface preparation and passivation methods are currently being investigated. Interface trap densities ($D_{it}$), at least around midgap, are relatively straight-forward to detect for In$_{0.53}$Ga$_{0.47}$As, because of its narrow band gap (0.75 eV (Ref. 10)); they cause, for example, a frequency-dependent hump in the depletion region of capacitance-voltage (CV) characteristics, and peaks in normalized conductance maps. Recently, we have reported on an in-situ, cyclic nitrogen plasma/trimethylaluminum (TMA) surface treatment prior to atomic layer deposition (ALD) of gate dielectrics on In$_{0.53}$Ga$_{0.47}$As. We showed that this process results in significantly reduced $D_{it}$ of metal-oxide-semiconductor capacitors (MOSCAPs), while at the same time allowing for scaling of accumulation capacitance densities to greater than 2 $\mu$F/cm$^2$ for both HfO$_2$ and HfO$_2$/Al$_2$O$_3$ gate stacks. For further scaling and $D_{it}$ reduction, it is important to understand the mechanisms by which surface treatments control the gate stack properties. Towards this goal, we report here on comparisons of different implementations of this surface treatment process, and how these correlate with $D_{it}$, interface chemistry and surface morphology. We show that the details of the surface treatment critically influence uniformity and surface coverage.

EXPERIMENTAL

MOSCAPs were fabricated on 300-nm-thick, n-type In$_{0.53}$Ga$_{0.47}$As (Si: $1 \times 10^{17}$ cm$^{-3}$) grown by molecular beam epitaxy on (001) n$^+$-InP (IntelliEpi, Richardson, Texas). After a 3 min clean in buffered HF, samples were transferred to an ALD reactor (Oxford Instruments FlexAL ALD), where they were exposed to the in-situ nitrogen plasma/TMA surface clean immediately prior to HfO$_2$ growth. Two slightly different surface cleaning procedures are compared here. The first (recipe A) consisted of 7 cycles of nitrogen plasma/TMA pulses/nitrogen plasma, as described in detail elsewhere. The second (recipe B) is a modified version of recipe A. For recipe B, each cycle consisted of a nitrogen gas set-up step (5 s at 20 mTorr), a nitrogen plasma pulse (inductively coupled plasma (ICP) power of 100 W at 20 mTorr for 2 s), a short TMA pulse (40 ms), followed by an Ar gas draw/purge step (7 s). After 9 of these cycles, a final nitrogen plasma pulse was added (100 W ICP power at 20 mTorr for 2 s), followed by 4 s of a N$_2$ stabilization step. The valve between the plasma and main reaction chambers was opened only during the nitrogen plasma pulse to avoid precursor contamination to the plasma chamber. The main differences between the two recipes are the additional pump and Ar purging steps after the nitrogen plasma and TMA pulse steps, respectively, in recipe B, as well as the shorter total time that the sample was exposed to nitrogen plasma in recipe B.

The substrate temperature was 300°C for both cleaning and HfO$_2$ deposition. The chamber reactor was held at 200 mTorr during oxide deposition. The deposition cycle used was as follows: a TEMAH (tetrakis[ethylmethylamino]hafnium) pulse for 1 s (flowing 250 sccm Ar gas through a bubbler held at 60°C) followed by an Ar gas purge step (7 s) then a short pulse of deionized water (500 ms) followed by pump (7 s) and Ar gas purge (5 s) steps. The oxide thicknesses were determined ex-situ (immediately after growth) using variable angle spectroscopic ellipsometry. The growth rate was ~1.3 Å/cycle. After dielectric deposition, the samples were annealed in a tube furnace at 400°C for 15 min in forming gas (95% of N$_2$ and 5% of H$_2$) at atmospheric pressure. 85-nm-thick Ni gate electrodes were deposited by thermal evaporation through a shadow mask. The back Ohmic contact, consisting of Cr (20 nm)/Au (100 nm), was deposited using thermal evaporation. Frequency-dependent CV and conductance-voltage measurements were carried out in the dark from 1 kHz to 1 MHz at room temperature using an impedance analyzer (Agilent...
RESULTS AND DISCUSSION

Figure 1 compares the CV curves and conductance maps measured for the ~4 nm HfO2/InGaAs MOSCAPs that were cleaned with recipe A [Figs. 1(a) and 1(c)] and with recipe B [Figs. 1(b) and 1(d)], respectively. In both cases, the CV characteristics are steep, and accumulation capacitance densities of greater than 2.5 μF/cm² at 1 MHz are achieved, which corresponds to sub-nm equivalent oxide thickness (EOT). The large frequency dispersion in accumulation is typical for highly scaled MOSCAPs, for reasons that currently not fully understood.15–17 The measured accumulation capacitances at low frequencies suffer from leakage artifacts; however, leakage is not expected to be an issue for transistors with gate dimensions much smaller than the MOSCAPs investigated here. In particular, the current density was measured to be less than 2 mA/cm² at 2 V. At frequencies not affected by leakage, the measured accumulation capacitance dispersion is slightly smaller in the sample cleaned using recipe B than that cleaned using recipe A, indicating a reduced (interface) trap density.17 More importantly, the much smaller frequency dispersion (hump) at negative biases measured for recipe B [Fig. 1(b)] is a clear evidence of a significantly reduced $D_{it}$.11 These qualitative observations are confirmed in the measured conductance maps [Figs. 1(c) and 1(d)], which show the normalized conductance peaks, $G_{p}/Aωq$ at $D_{it}$, where $G_{p}$ is the parallel conductance, $A$ is the capacitor area, $ω$ is the modulation frequency, and $q$ is the elemental charge, as a function of gate voltage and $ω$. The $G_{p}/Aωq$ values (see vertical axes and scale bar on the right of each graph) can be used to estimate the $D_{it}$ by multiplication with a factor of ~2.5.18 Recipe B results in a $D_{it}$ that is ~4 times lower than for recipe A (values are in the mid-$10^{12}$ cm⁻²eV⁻¹ range). The maps also indicate how efficient the Fermi level moves around midgap as a function of gate bias.19 For both samples, the peaks shift more than two orders of magnitude in frequency as the gate bias is changed between 0 and $−1$ V. For the recipe B sample, the narrower trace additionally suggests larger band bending in response to a change in gate bias.

We note that there are great inconsistencies (amounting to several orders of magnitude) in the reported $D_{it}$ in the dielectric/III-V MOSCAP literature. Several sources of error exist for all $D_{it}$ quantification methods, in particular, near the conduction band/valence band edges, as discussed in the recent literature (see Refs. 11 and 20, and references cited therein). One source of error in conductance maps are incorrect estimates of the oxide capacitance, which cannot be directly obtained from the CV of MOSCAPs.11,12 To reduce this error, the dielectric capacitance used here (3.8 μF/cm²) was estimated from a thickness series (measurement of $C_{acc}^{-1}(t)$), where $C_{acc}^{-1}$ is the inverse of the accumulation capacitance density and $t$ is the thickness measured by transmission electron microscopy). This value includes the interfacial layer capacitance, $C_{IL}$ (see below). In particular, the intercept of $C_{acc}^{-1}(t)$ gives $(C_{IL})^{-1} + (C_{S})^{-1}$, where $(C_{S})^{-1}$ is the semiconductor capacitance at 2 V, which was calculated as described previously.12 Because of the low density of conduction band states typical for III-V semiconductors, the $C_{acc}^{-1}$ is lower than the dielectric capacitance. The slope of $C_{acc}^{-1}(t)$ can be used to extract the dielectric constant of HfO2, which was determined to be 17 ± 2, which is close to the expected value. Thus, the conductance maps reported here are expected to give realistic values for the $D_{it}$, at least around midgap.

To investigate the origins of the different electrical properties resulting from the different cleaning recipes, Fig. 2 shows AFM and SEM images of the HfO2 film surfaces after the postdeposition anneal (images taken of as-deposited films show qualitatively similar features). The surface of the recipe A sample [Fig. 2(a)] is rougher, and shows features that are up to 10 nm in height. The sample also exhibits regions of poor coverage, as seen in SEM [darker areas in Fig. 2(c)]. In contrast, the dielectric film of the recipe B sample is completely coalesced and exhibits a much more uniform surface, with lower roughness. The latter may be caused by shorter total plasma time, because long plasma exposures are known to result in surface roughening.23 Occasional surface features still exist but are only about 2 nm in height. Since a main difference between recipes A and B is the additional pumping/purging steps in B, we may conclude that these steps result in improved nucleation. Improved nucleation and growth of HfO2 on III-V surfaces using small amounts of TMA either before or during
deposition has been shown previously;\textsuperscript{24–26} hence the additional pumping/purging steps likely result in a more uniform distribution of the nucleation centers generated by TMA-exposure of the surface, due to the additional time for surface diffusion.

The improved nucleation also allows for further scaling. Specifically, if improved nucleation allows for complete coverage of the III-V surface at lower physical thicknesses, HfO\textsubscript{2} films with higher capacitance densities and good electrical properties should be achievable. Recipe B allows for MOSCAPs with capacitance densities of 3 \( \mu \)F/cm\textsuperscript{2} at 1 MHz for HfO\textsubscript{2} films with a 3 nm thickness, as shown in the inset of Fig. 1(b). This is capacitance density is higher than values in the literature,\textsuperscript{27–29} while a very low \( D_{it} \) is maintained. Attempts to scale films grown on recipe A-cleaned surfaces to comparable thicknesses resulted in higher \( D_{it} \), as exhibited by large frequency dispersion at negative biases.

If the nucleation is indeed facilitated by the TMA step, residual Al may be present at the interface. To investigate this, XPS was performed on thinner (\( \sim 1.5 \) nm) HfO\textsubscript{2} films to allow signals from the interface to be detected. Figure 3 shows the measured XPS spectra from these samples. The binding energy scales were calibrated by setting the surface aliphatic hydrocarbon peak to 285.0 eV. Quantification of the Al 2p binding energy scales were calibrated by setting the surface binding energy at 136.7 eV, and the Al 2p peak, deconvolved into two components, namely, Al 2p\textsuperscript{3/2} and Al 2p\textsuperscript{1/2}, which are separated by 0.6 eV.\textsuperscript{30} The peaks have a full width at half maximum (FWHM) of \( \sim 1.2 \) eV. A very weak nitrogen signal was also detected. Figures 3(c) and 3(d) show the As 3d and Hf 5p peaks. The As 3d\textsuperscript{3/2} and 3d\textsuperscript{5/2} peaks have a FWHM of 0.65 eV. Because \( \Delta_{As,3d} = 0.7 \) eV,\textsuperscript{30} we conclude that no detectable arsenic oxides are present in the interfacial layer. The Hf 5p\textsuperscript{3/2} and Hf 5p\textsuperscript{1/2} peaks have a FWHM of 2 eV and 1.6 eV, respectively, with \( \Delta_{Hf,5p} = 7.2 \) eV.\textsuperscript{31} SIMS (not shown) with a 3 kV Cs beam detected a negatively charged fragment at mass 41 at the oxide-arsenide interface, identified as AlN\textsuperscript{−}. This mass was not ascribed to AlCH\textsubscript{2}− because other masses (at 12, 13, and 14 Da) clearly associated with hydrocarbon followed a different depth profile than mass 41, decreasing monotonically from the top surface. Additional SIMS analysis with a 2 kV oxygen beam revealed a positively charged mass at 27 peaks at the HfO\textsubscript{2}/InGaAs interface region on both samples, which we identify as Al. The small N signal in XPS and the Al 2p\textsuperscript{3/2} peak position at 73.8 eV are consistent with an oxygen-rich AlO\textsubscript{x}N\textsubscript{y} interfacial layer.

The notable absence of As-(sub)-oxides at these interfaces provides evidence of the efficiency of the cleaning procedures (both recipe A and B) in native oxide removal, good coverage, and passivation of the In\textsubscript{0.53}Ga\textsubscript{0.47}As surface early in the ALD process, and that even \( \sim 1.5 \) nm thin HfO\textsubscript{2} provides protection from subcutaneous oxidation. The results support theoretical studies that indicate that avoiding As-(sub)-oxides is critical for achieving low \( D_{it} \).\textsuperscript{32–34} In particular, even for recipe A, the \( D_{it} \) is already significantly lower than for other cleaning methods,\textsuperscript{14} and this is likely due to combination of the absence of As-(sub)-oxides and the passivation properties of the AlO\textsubscript{x}N\textsubscript{y} interfacial layer.

Within the detection limit of XPS and SIMS, we find no difference in the interfacial chemistry between recipe A and B. This shows that these methods are not sufficiently sensitive to the atomic scale origins that are responsible for the differences in surface morphology, coverage and \( D_{it} \) between the two recipes. While this may simply be a detection limit issue, it is more likely that a reduction of defects, such as dangling bonds and subtle atomic rearrangements, which have been implicated as giving rise to \( D_{it} \) for III-V semiconductors,\textsuperscript{34–36} are responsible for the improvement in \( D_{it} \) with recipe B. This is in keeping with prior results that show that avoiding or reducing damage to the III-V surface results in improvements in the \( D_{it} \).\textsuperscript{7,27} Such defects cannot be detected by XPS or SIMS and will require the development of experimental methods, such as electron spin resonance,
which is, however, substantially more difficult for III-V semiconductors than for Si.36 However, the results reported here provide indirect evidence that complete surface coverage as early as possible in the ALD process, as facilitated by optimizing the cleaning process, reduces these defects through early passivation of the surface. Conversely, nucleation issues that may result in prolonged exposure of the surface after cleaning increase the $D_{nr}$.

CONCLUSIONS

In summary, we have shown that modifications to a pre-deposition in-situ cleaning using alternating cycles of remote nitrogen plasma and TMA can further substantially reduce the $D_{nr}$ of HfO$_2$/In$_{0.53}$Ga$_{0.47}$As gate stacks into the mid-10$^{12}$ cm$^{-2}$ eV$^{-1}$ for sub-nm-EOT gate stacks. Comparison of different implementations of the cleaning procedure showed that improving surface coverage is critical for significant $D_{nr}$ reduction and EOT scaling. Both TMA and N are essential for providing nucleation sites that facilitate the growth of HfO$_2$ on In$_{0.53}$Ga$_{0.47}$As surfaces,14,24–26 and both species were detected at the interface in XPS and SIMS. The present study shows that optimization of the pre-deposition cleaning procedure, such as sequences and exposure times, are needed for a high density of these nucleation sites, and can enable very low-$D_{nr}$ and highly scaled gate stacks.

ACKNOWLEDGMENTS

The authors gratefully acknowledge support for this work by the Semiconductor Research Corporation through the Nonclassical CMOS Research Center (Task ID 1437.008). V.C. thanks Adam Kajdos and Mark Cornish for help with plotting conductance maps. A portion of this work was performed in UCSB’s nanofabrication facility, which is part of the NSF-funded NNIN network, and MRL facilities. The work also made use of the MRL Shared Experimental Facilities, which are supported by the MRSEC Program of the NSF under Award No. DMR 1121053.