Low-Temperature Microwave Annealing for MOSFETs With High-k/Metal Gate Stacks

Yao-Jen Lee, Bo-An Tsai, Chiung-Hui Lai, Zheng-Yao Chen, Fu-Kuo Hsueh, Po-Jung Sung, Michael I. Current, and Chih-Wei Luo

***Abstract*— In this letter, low-temperature (480 °C) microwave annealing (MWA) for MOS devices with high-k/metal gate-stacks is demonstrated. The capacitance-voltage (*C–V*) characteristics of the MOS gate-stacks, TiN/HfO$_2$, and TaN/HfO$_2$, after different annealing methods are discussed. The increases in equivalent oxide thickness (EOT) of the MOS devices after dopant activation processing can be eliminated using low temperature MWA. In addition, the short channel effects in nMOSFETs annealed by MWA can be also improved because of the suppression of dopant diffusion and stabilization of EOT.**

***Index Terms*— High-*k*, low temperature, metal gate, microwave annealing (MWA).**

I. Introduction

To IMPROVE the performance of CMOS devices, it is advantageous to use a pair of metals with work functions that are near the conduction-band and valence-band edges of silicon to replace the conventional n$^+$/p$^+$ poly-Si gate materials [1]. Gate-last processing has been adopted to eliminate work function shifts and high-*k* dielectric degradation after high temperature dopant activation of the source and drain (S/D) regions [2], [3]. This complex process however, leads to restrictions on design rules and narrow processing windows [4].

If gate-first processing enables suppression of the work function shifts and the equivalent oxide thickness (EOT) increases after the dopant activation process, it will provide a viable and promising option to reduce costs and simplify current design rules [5], [6]. Microwave annealing (MWA) is an alternative to other rapid thermal processing methods [7]–[10]. Therefore, in this letter gate-first processing is evaluated by investigating and comparing the characteristics of MOS devices with high-*k*/metal gate-stacks annealed by either MWA or rapid thermal annealing (RTA).

Manuscript received August 12, 2013; accepted August 17, 2013. Date of current version September 23, 2013. The review of this letter was arranged by Editor K.-S. Chang-Liao.

Y.-J. Lee is with the National Nano Device Laboratories, Hsinchu 30010, Taiwan, and also with the Department of Physics, National Chung Hsing University, Taichung 402, Taiwan (e-mail: yjlee@narlabs.org.tw).

B.-A. Tsai and C.-W. Luo are with the Department of Electrophysics, National Chiao Tung University, Hsinchu 30050, Taiwan.

C.-H. Lai and Z.-Y. Chen are with the Department of Electronics Engineering, Chung Hua University, Hsinchu 30012, Taiwan.

F.-K. Hsueh and P.-J. Sung are with the National Nano Device Laboratories, Hsinchu 30010, Taiwan.

M. I. Current is with Current Scientific, San Jose, CA 95124 USA.

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LED.2013.2279396

II. Experiments

The starting substrates were 150-mm p-type Si (100) wafers with resistivity of 10–30 Ω-cm. For MOS capacitor fabrication, a chemical oxide was formed by immersion of Si into a 31% H$_2$O$_2$ solution at 100 °C for 300 s. The thickness of the chemical oxide is ∼0.6 nm, measured by an ellipsometer.

A 3-nm-thick HfO$_2$ film was then deposited by tetrakis-ethylmethylamino hafnium and O$_3$ via the atomic layer deposition process at 320 °C. Next, wafers were treated by postdeposition annealing. A 50-nm TiN or TaN layer was deposited using a physical vapor deposition system as the metal gate electrodes, followed by the lithography and etching steps. For the nMOSFETs, e-beam lithography was used to define the active region, followed by the high-*k* dielectric and metal gate electrode deposition process (TiN/HfO$_2$/chemical oxide/Si). After gate patterning, source and drain junctions were implanted with ^{31}P (10 keV at 1×10^{15} cm^{-2}), followed by different dopant activation conditions for comparisons, including low-temperature MWA and high-temperature RTA. The microwave frequency was 5.8 GHz and the MWA peak anneal temperatures on the wafer were ∼480 °C, measured by a line-of-sight pyrometer on the wafer backside. The MWA processing time was defined as the duration for which the microwave magnetron was turned ON. A N$_2$ purge was performed before the MWA was started and the N$_2$ flow was maintained until the process was completed. The operating frequency for the *C–V* analysis was 100 kHz. Work functions were extracted by CVC analysis software [11]. I_D–V_G curves were also measured and compared.

III. Results and Discussion

Fig. 1(a) shows the *C–V* curves of the MOS devices with TiN gate electrodes, including control split (w/o annealing). The annealing conditions include MWA (2.4 kW for 150 s and maximum temperature: 480 °C) and RTA (480 °C for 150 s, 800 °C for 60 s, and 1000 °C for 1 s).

The EOT of the MOS devices after RTA at 800 °C increased to 1.62 nm, whereas after RTA at 1000 °C, it increased to 2.24 nm. With MWA, a marginal increase of 0.1 nm occurred in the MOS devices with TiN gate-stacks. For the splits annealed by RTA at 480 °C, 800 °C, and 1000 °C, the work function shifts were 40, 384, and 420 mV, respectively. For the low-temperature MWA splits, the work function shift was 155 mV.

0741-3106 © 2013 IEEE

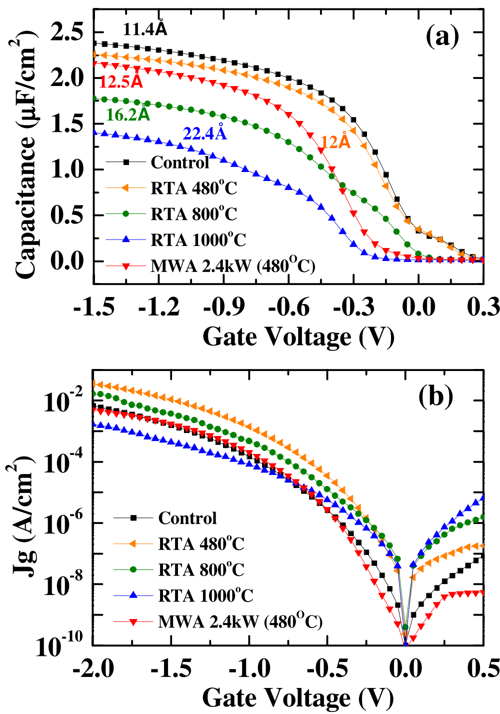

Fig. 1. MOS capacitance with TiN metal gate electrodes. (a) C–V with EOT values shown. (b) J_g–V_g curves. MWA can suppress both work function shift and EOT increases because of its lower temperature processing.

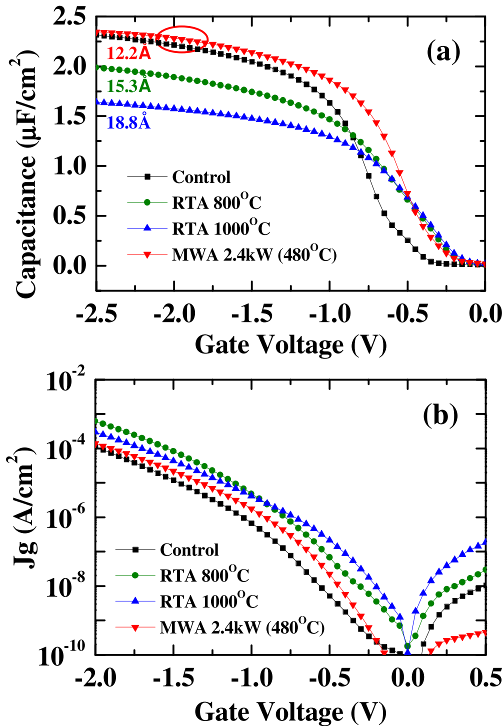

Fig. 2. MOS capacitance with TaN metal gate electrodes. (a) C–V curves. (b) J_g–V_g curves. MWA suppresses EOT increases because of its low temperature processing.

Gate leakage current of the MOS devices with TiN gate electrodes was measured, as shown in Fig. 1(b). Compared with the control split, the gate leakage current of the split annealed by RTA at 1000 °C is slightly lower than the MWA case because the EOT increased substantially. The gate leakage current of the split annealed by RTA at 800 °C is, however, slightly higher even with the EOT increase. The suspected cause is crystallization of the dielectric after high-temperature annealing [12].

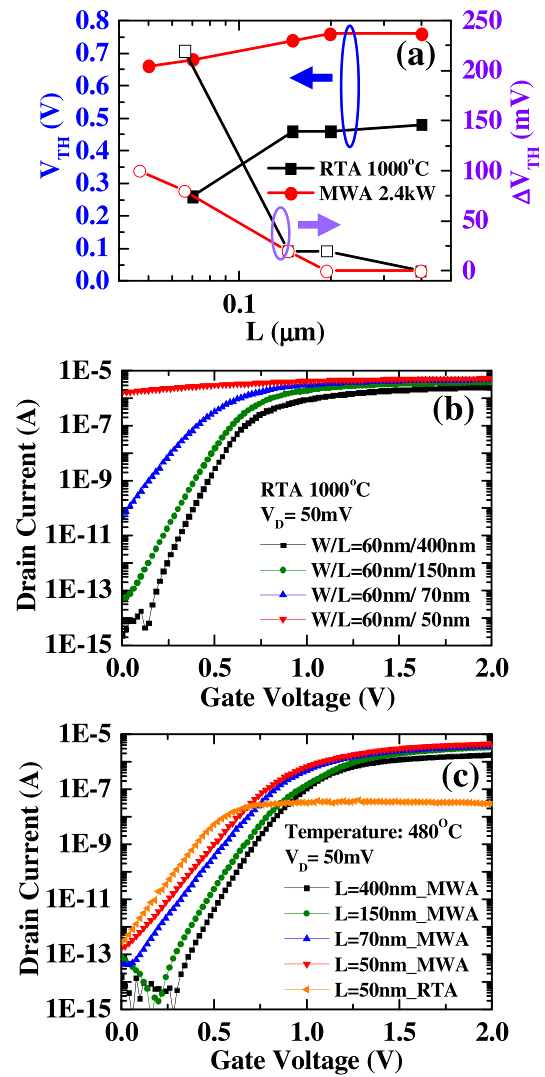

Fig. 3. MOSFET with TiN metal gate electrodes (a) characteristics of V_{TH} rolloff. I_D–V_G characteristics by (b) RTA at 1000 °C and (c) MWA and RTA annealing temperatures at 480 °C under the same S/D implant conditions.

Fig. 2(a) shows the C–V curves of the MOS devices with TaN gate electrodes. The EOT of the MOS devices increases to 1.53 and 1.88 nm after RTA at 800 °C and 1000 °C, respectively. For the MWA split, interestingly, there are no EOT increases for the MOS devices with TaN gate electrodes. In addition, a difference in the work function shift for all the splits after RTA at 800 °C and 1000 °C or MWA is not apparent.

Gate leakage current of the MOS devices with TaN gate electrodes was also measured, as shown in Fig. 2(b). The gate leakage current of the control and MWA splits is lower than that of RTA. The gate leakage current of the splits by RTA at 800 °C and 1000 °C, however, increases slightly, with EOT increasing to 1.53 and 1.88 nm, respectively.

Fig. 3(a) shows the threshold voltage (V_{TH}) rolloff characteristics of nMOSFET versus gate length (L) by RTA at

TABLE I
SUMMARY OF CHARACTERISTICS OF MOS AND MOSFETs WITH TiN GATE ELECTRODES BY DIFFERENT ANNEALING METHODS

		RTA 800°C	RTA 1000°C	MWA 2.4 kW (480 °C)
MOS				
ΔEOT (Å)		4.8	11	1.1
ΔWF (mV)		384	420	155
MOSFETs				
V_{TH}(V) @W=60nm	L=400nm	N.A.	0.48	0.76.
	L=200nm		0.46	0.76
	L=150nm		0.46	0.74
	L=70nm		0.26	0.68
	L=50nm		Punch-through	0.66

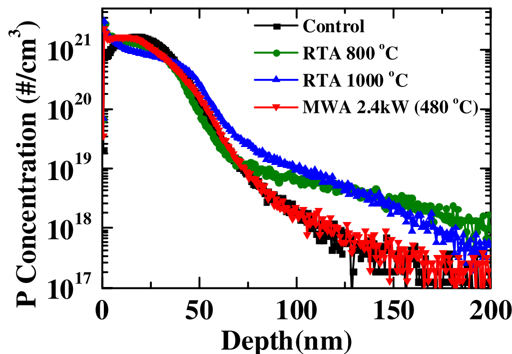

Fig. 4. SIMS profiles of ^{31}P concentration at a dose of 5 × 10^{15} ions/cm^2. MWA process resulted in no significant dopant diffusion.

1000 °C and MWA. ΔV_{TH} is defined by V_{TH} (L = 400 nm) and $V_{TH}(L < 400$ nm). As the gate length shrinks from 400 to 50 nm, short channel effects (SCEs) can be more effectively suppressed for the splits annealed by MWA because of its low temperature processing than for splits annealed by RTA. Table I summarizes the characteristics of MOS and MOSFETs with TiN gate electrodes by different annealing methods.

Fig. 3(b) shows the I_D–V_G characteristics of nMOSFETs with different gate lengths annealed by RTA at 1000 °C. In Fig. 3(b), for the splits annealed by RTA, the I_{ON}/I_{OFF} ratio is ∼10^5 (I_{ON} at V_G = 2 V and I_{OFF} at V_G = 0 V) with W/L = 60/70 nm. However, when the gate length falls below 70 nm, OFF-current increases, and I_{ON}/I_{OFF} ratio is 3.4 × 10^0 with W/L = 60/50 nm.

Fig. 3(c) shows the I_D–V_G characteristics of nMOSFETs with different gate lengths annealed at 480 °C and splits include MWA and RTA at W/L = 60/50 nm. For the splits annealed by MWA, SCEs are greatly improved. The I_{ON}/I_{OFF} ratio is ∼10^8 (I_{ON} at V_G = 2 V and I_{OFF} at V_G = 0 V) for nMOSFETs annealed by MWA for W/L = 60/50 nm. For the splits annealed by RTA at 480 °C with L = 50 nm, I_{ON} is limited at 3 × 10^8 A due to higher S/D series resistance, and the I_{ON}/I_{OFF} ratio is only 1.1 × 10^5.

Fig. 4 shows the SIMS profiles of the ^{31}P concentrations. The dopant profile after the low temperature MWA was indistinguishable from the as-implanted profiles, indicating no dopant diffusion motion for these low temperature anneals.

The junction depths (at a concentration 10^{19}/cm^3) for the anneals by RTA at 1000 °C and MWA were 110 and 72 nm, respectively, a difference of 38 nm. The EOT values were 1.25 and 2.24 nm for the splits by MWA and RTA at 1000 °C, respectively. The Rs of the splits by RTA at 800 °C and 1000 °C and by MWA are 78, 61, and 85 Ω/square, respectively. Although Rs of the split by RTA is the lowest among all splits, because of the substantial diffusion and EOT increases, SCEs appear to be severe. Therefore, considering entire MOSFETs performance, such as I_{ON}/I_{OFF} ratio and SCEs, the use of MWA can improve device performance because of lower dopant diffusion and better gate control ability.

IV. CONCLUSION

In this letter, different annealing methods for nMOS devices with high-k/metal gate-stacks were investigated. Both EOT increases and SCEs after dopant activation process after RTA appear to be better controlled using low-temperature MWA. In addition, work function shifts for TiN gate electrodes seen with RTA can be suppressed by MWA processing. Therefore, MWA processing appears to be a useful option for gate-first processing of MOSFETs with high-k/metal gate-stacks.

REFERENCES

[1] I. De, D. Johri, A. Srivastava, *et al.*, "Impact of gate work function on device performance at the 50 nm technology node," *Solid State Electron.*, vol. 44, no. 6, pp. 1077–1080, Jun. 2000.

[2] H. Y. Yu, C. Ren, Y. C. Yeo, *et al.*, "Fermi pinning-induced thermal instability of metal-gate work functions," *IEEE Electron Device Lett.*, vol. 25, no. 5, pp. 337–339, May 2004.

[3] J. Lee, H. Zhong, Y. Suh, *et al.*, "Tunable work function dual metal gate technology for bulk and non-bulk CMOS," in *Proc. IEDM*, Dec. 2002, pp. 359–362.

[4] L.-Å. Ragnarsson, Z. Li, J. Tseng, *et al.*, "Ultra low-EOT (5 Å) gate-first and gate-last high performance CMOS achieved by gate-electrode optimization," in *IEDM Tech. Dig.*, 2009, pp. 663–666.

[5] F. Arnaud, J. Liu, Y. M. Lee, *et al.*, "32 nm general purpose bulk CMOS technology for high performance applications at low voltage," in *Proc. IEEE IEDM*, Dec. 2008, pp. 633–636.

[6] T. Tomimatsu, Y. Goto, H. Kato, *et al.*, "Cost-effective 28-nm LSTP CMOS using gate-first metal gate/high-k technology," in *Proc. Symp. VLSI Technol.*, Jun. 2009, pp. 36–37.

[7] T. L. Alford, D. C. Thompson, J. W. Mayer, *et al.*, "Dopant activation in ion implanted silicon by microwave annealing," *J. Appl. Phys.*, vol. 106, no. 11, pp. 114902-1–114902-8, Dec. 2009.

[8] H. Zohm, E. Kasper, P. Mehringer, *et al.*, "Thermal processing of silicon wafers with microwave co-heating," *Microelectron. Eng.*, vol. 54, nos. 3–4, pp. 247–253, Dec. 2000.

[9] Y. J. Lee, F. K. Hsueh, S. C. Huang, *et al.*, "A low-temperature microwave anneal process for boron-doped ultrathin Ge epilayer on Si substrate," *IEEE Electron Device Lett.*, vol. 30, no. 2, pp. 123–125, Feb. 2009.

[10] F. K. Hsueh, Y. J. Lee, K. L. Lin, *et al.*, "Amorphous-layer regrowth and activation of P and As implanted Si by low-temperature microwave annealing," *IEEE Electron Device Lett.*, vol. 32, no. 2, pp. 194–196, Feb. 2011.

[11] J. Hauser and K. Ahmed, "Characterization of ultra-thin oxides using electrical C–V and I–V measurements," in *Proc. Characterizat. Metrol. ULSI Technol. Conf.*, 1998, pp. 235–239.

[12] W. J. Zhu, T. P. Ma, T. Tamagawa, *et al.*, "HfO_2 and HfAlO for CMOS: Thermal stability and current transport," in *IEDM Tech. Dig.*, Dec. 2001, pp. 463–466.