(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
5 July 2012 (05.07.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/090973 A1

(51) International Patent Classification:
H01L 29/786 (2006.01) *H01L 21/336* (2006.01)

(21) International Application Number:
PCT/JP2011/080141

(22) International Filing Date:
20 December 2011 (20.12.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2010-293047 28 December 2010 (28.12.2010) JP

(71) Applicant *(for all designated States except US)*: **SEMICONDUCTOR ENERGY LABORATORY CO., LTD.** [JP/JP]; 398, Hase, Atsugi-shi, Kanagawa, 2430036 (JP).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **YAMAZAKI, Shunpei** [JP/JP]; c/o SEMICONDUCTOR ENERGY LABORATORY CO., LTD., 398, Hase, Atsugi-shi, Kanagawa, 2430036 (JP).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

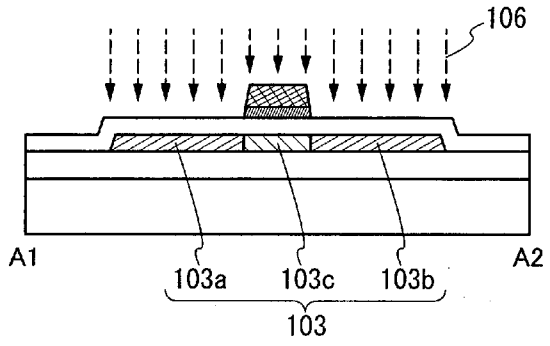
Published:

— *with international search report (Art. 21(3))*

WO 2012/090973 A1

(54) Title: SEMICONDUCTOR DEVICE AND MANUFACTURING METHOD THEREOF

FIG. 3C



(57) Abstract: A semiconductor device capable of high speed operation is provided. Further, a semiconductor device in which change in electric characteristics due to a short channel effect is hardly caused is provided. An oxide semiconductor having crystallinity is used for a semiconductor layer of a transistor. A channel formation region, a source region, and a drain region are formed in the semiconductor layer. The source region and the drain region are formed by self-aligned process in which one or more elements selected from Group 15 elements are added to the semiconductor layer with the use of a gate electrode as a mask. The source region and the drain region can have a wurtzite crystal structure.

DESCRIPTION

SEMICONDUCTOR DEVICE AND MANUFACTURING METHOD THEREOF

TECHNICAL FIELD

[0001]

The present invention relates to a semiconductor device which includes a circuit including a semiconductor element such as a transistor, and a method for manufacturing the semiconductor device. For example, the present invention relates to an electronic device which includes, as a component, a power device mounted on a power supply circuit; a semiconductor integrated circuit including a memory, a thyristor, a converter, an image sensor, or the like; an electro-optical device typified by a liquid crystal display panel; a light-emitting display device including a light-emitting element; or the like.

[0002]

Note that in this specification, a semiconductor device means any device that can function by utilizing semiconductor characteristics. An electro-optical device, a light-emitting display device, a semiconductor circuit, and an electronic device are all semiconductor devices.

BACKGROUND ART

[0003]

A transistor formed over a glass substrate or the like is manufactured using amorphous silicon, polycrystalline silicon, or the like, as typically seen in a liquid crystal display device. Although a transistor including amorphous silicon has low field-effect mobility, it can be formed over a larger glass substrate. On the other hand, although a transistor including polycrystalline silicon has high field-effect mobility, it is not suitable for being formed over a larger glass substrate.

[0004]

In view of the foregoing, attention has been drawn to a technique by which a transistor is manufactured using an oxide semiconductor, and such a transistor is applied to an electronic device an optical device. For example, Patent Document 1 and Patent

Document 2 disclose a technique in which a transistor is manufactured using zinc oxide or an In-Ga-Zn-based oxide as an oxide semiconductor and such a transistor is used as a switching element or the like of a pixel of a display device.

[0005]

Patent Document 3 discloses a technique in which in a staggered transistor including an oxide semiconductor, a highly conductive oxide semiconductor including nitrogen is provided as buffer layers between a source region and a source electrode and between a drain region and a drain electrode, and thereby the contact resistance between the oxide semiconductor and the source electrode and between the oxide semiconductor and the drain electrode is reduced.

[0006]

Further, as a method for forming a source region and a drain region of a transistor including an oxide semiconductor by a self-aligned process, Non-Patent Document 1 discloses a method in which a surface of an oxide semiconductor is exposed and argon plasma treatment is performed thereon for reducing resistivity of the exposed portion in the oxide semiconductor.

[Reference]

[Patent Document]

[0007]

[Patent Document 1] Japanese Published Patent Application No. 2007-123861

[Patent Document 2] Japanese Published Patent Application No. 2007-096055

[Patent Document 3] Japanese Published Patent Application No. 2010-135774

[Non-Patent Document]

[0008]

[Non-Patent Document 1] S. Jeon et al., "180nm Gate Length Amorphous InGaZnO Thin Film Transistor for High Density Image Sensor Application", IEDM Tech. Dig., p.504, 2010.

DISCLOSURE OF INVENTION

[0009]

An object is to provide a semiconductor device capable of high speed operation.

[0010]

An object is to provide a semiconductor device in which variation in electric characteristics due to a short channel effect is less likely to be caused.

[0011]

Further, an object is to provide a semiconductor device which can be easily miniaturized by formation of a source region and a drain region by a self-aligned process.

[0012]

Further, an object is to provide a semiconductor device in which contact resistance between a source region and a source electrode and between a drain region and a drain electrode is reduced and on-state current is improved, by formation of the source region and the drain region each having resistance lower than that of a channel.

[0013]

Further, an object is to provide a highly reliable semiconductor device.

[0014]

An embodiment of the present invention is a semiconductor device including an oxide semiconductor layer having crystallinity, a gate insulating layer, and a gate electrode. The oxide semiconductor layer includes a first oxide semiconductor region and a pair of second oxide semiconductor regions. The first oxide semiconductor region is sandwiched between the pair of second oxide semiconductor regions. The first oxide semiconductor region overlaps with the gate electrode with the gate insulating layer interposed therebetween.

[0015]

Further, an embodiment of the present invention is a semiconductor device including an oxide semiconductor layer having crystallinity, a gate insulating layer, and a gate electrode. The oxide semiconductor layer includes a first oxide semiconductor region, a pair of second oxide semiconductor regions, and a pair of third oxide semiconductor regions. The first oxide semiconductor region is sandwiched between the pair of third oxide semiconductor regions. The pair of third oxide semiconductor regions is sandwiched between the pair of second oxide semiconductor regions. The first oxide semiconductor region overlaps with the gate electrode with the gate insulating layer interposed therebetween.

[0016]

A non-single-crystal semiconductor may be used for the oxide semiconductor layer.

[0017]

The first oxide semiconductor region includes a c-axis aligned crystalline oxide semiconductor (CAAC-OS). The CAAC-OS includes crystal parts, in each of which a c-axis is aligned in a direction parallel to a normal vector of a surface where the CAAC-OS is formed or a normal vector of a surface of the CAAC-OS, triangular or hexagonal atomic arrangement which is seen from the direction perpendicular to the a-b plane is formed, and metal atoms are arranged in a layered manner or metal atoms and oxygen atoms are arranged in a layered manner when seen from the direction perpendicular to the c-axis.

[0018]

Each of the second oxide semiconductor regions may contain at least one of Group 15 elements at a concentration of higher than or equal to 5×10^{19} atoms/cm^3 and lower than or equal to 1×10^{22} atoms/cm^3. The second oxide semiconductor regions may have a wurtzite crystal structure.

[0019]

Each of the third oxide semiconductor regions may contain at least one of Group 15 elements at a concentration of higher than or equal to 5×10^{18} atoms/cm^3 and lower than or equal to 5×10^{19} atoms/cm^3. The third oxide semiconductor regions may have a wurtzite crystal structure.

[0020]

The second oxide semiconductor regions and the third oxide semiconductor regions can each have a crystal structure different from the first oxide semiconductor region. In this case, the oxide semiconductor layer included in the semiconductor device has a heterojunction. An oxide semiconductor having a heterojunction is used for a semiconductor layer of a transistor, whereby on-state current is expected to be increased. Further, off-state current is expected to be reduced.

[0021]

The oxide semiconductor can contain two or more elements selected from In,

Ga, Sn, and Zn.

[0022]

The first oxide semiconductor region serves as a channel formation region of a transistor. The pair of second oxide semiconductor regions serves as a source region and a drain region of the transistor. The pair of third oxide semiconductor regions serves as low-concentration regions of the transistor.

[0023]

In a top-gate transistor, a source region and a drain region can be formed by adding a dopant to an oxide semiconductor layer with the use of a gate electrode as a mask. The source region and the drain region are formed using the gate electrode as a mask, whereby the source region and the drain region do not overlap with the gate electrode and thus parasitic capacitance can be reduced. Since the parasitic capacitance can be reduced, the transistor can be operated at high speed.

[0024]

In the case where low-concentration regions are formed between a channel formation region and a source region and between the channel formation region and a drain region in a top-gate transistor, first, a dopant for forming the low-concentration regions is added to an oxide semiconductor layer with the use of a gate electrode as a mask; then, sidewalls are formed on side surfaces of the gate electrode; finally, a dopant for forming the source region and the drain region is added to the oxide semiconductor layer with the use of the gate electrode and the sidewalls as masks.

[0025]

In a bottom-gate transistor, a source region and a drain region can be formed by addition of a dopant to an oxide semiconductor layer with the use of a channel protective layer as a mask. The channel protective layer is formed to protect a back channel portion of an active layer, and is preferably formed using a single layer or a stacked layer using one or more of materials selected from silicon oxide, silicon nitride, aluminum oxide, aluminum nitride, and the like.

[0026]

A dopant used for forming a source region, a drain region, and low-concentration regions of a transistor can be added by an ion doping method, an ion implantation method, or the like. As the dopant, one or more elements selected from

Group 15 (Group 5B) elements such as nitrogen (N) or phosphorus (P) can be used. Further, the dopant is added to an oxide semiconductor layer through an insulating layer by an ion doping method or an ion implantation method, so that excessive damage to the oxide semiconductor layer in addition of the dopant can be reduced. Furthermore, the interface between the oxide semiconductor layer and the insulating layer is kept clean, so that characteristics and reliability of the transistor are improved. Moreover, the depth to which a dopant is added (addition region) is easily controlled, so that a dopant can be accurately added to an oxide semiconductor layer.

[0027]

The carrier density of an oxide semiconductor region can be increased as the concentration of a dopant to be added is increased; however, carrier transfer is inhibited and the conductivity is decreased if the concentration of the dopant to be added is too high.

[0028]

An oxide semiconductor to which a dopant is added is used for a source region and a drain region, whereby a curve of a band edge of a channel formation region to which the dopant is not added can be small. On the other hand, in the case where the source region and the drain region are formed using a metal material, a curve of the band edge of the channel which is the oxide semiconductor region is not negligible, so that the effective channel length is decreased in some cases. This tendency becomes more remarkable as the channel length of a transistor is reduced.

[0029]

Oxide semiconductor regions to which a dopant is added are used for forming a source region and a drain region of a transistor, whereby on-state current of the transistor can be increased without an increase in off-state current of the transistor.

[0030]

Further, the resistivity of the third oxide semiconductor regions is set higher than that of the second oxide semiconductor regions. The third oxide semiconductor regions are provided, whereby an electric field generated between the first oxide semiconductor region and the second oxide semiconductor regions can be relieved and degradation of transistor characteristics can be reduced. Further, negative shift of the threshold voltage due to a short channel effect can be reduced.

[0031]

An oxide semiconductor which is purified (purified OS) by reduction of an impurity such as moisture or hydrogen which serves as an electron donor (donor) can be made to be an i-type (intrinsic) oxide semiconductor or an oxide semiconductor extremely close to an i-type semiconductor (a substantially i-type oxide semiconductor) by supplying oxygen to the oxide semiconductor to reduce oxygen deficiency in the oxide semiconductor. Accordingly, a transistor including the i-type or substantially i-type oxide semiconductor in a semiconductor layer where a channel is formed has characteristics of very small off-state current. Specifically, the hydrogen concentration of the purified oxide semiconductor, which is measured by secondary ion mass spectrometry (SIMS), is lower than 5×10^{18} /cm^3, preferably lower than or equal to 1×10^{18} /cm^3, further preferably lower than or equal to 5×10^{17} /cm^3, still further preferably lower than or equal to 1×10^{16} /cm^3. In addition, the carrier density of the i-type or substantially i-type oxide semiconductor, which is measured by Hall effect measurement, is less than 1×10^{14} /cm^3, preferably less than 1×10^{12} /cm^3, further preferably less than 1×10^{11} /cm^3. Furthermore, the band gap of the oxide semiconductor is 2 eV or more, preferably 2.5 eV or more, more preferably 3 eV or more. With the use of the i-type or substantially i-type oxide semiconductor for a semiconductor layer where a channel is formed, off-state current of the transistor can be reduced.

[0032]

The SIMS analysis of the hydrogen concentration in the oxide semiconductor is described here. It is known that it is difficult to obtain accurate data in the proximity of a surface of a sample or in the proximity of an interface between stacked films formed using different materials by the SIMS analysis in principle. Thus, in the case where distributions of the hydrogen concentrations of the films in thickness directions are analyzed by SIMS, an average value in a region where the films are provided, the value is not greatly changed, and almost the same value can be obtained are employed as the hydrogen concentration. Further, in the case where the thickness of the film is small, a region where almost the same value can be obtained cannot be found in some cases due to the influence of the hydrogen concentration of the films adjacent to each

other. In this case, the maximum value or the minimum value of the hydrogen concentration of a region where the films are provided is employed as the hydrogen concentration of the film. Furthermore, in the case where a mountain-shaped peak having the maximum value and a valley-shaped peak having the minimum value do not exist in the region where the films are provided, the value of the inflection point is employed as the hydrogen concentration.

[0033]

According to an embodiment of the present invention, a semiconductor device including an oxide semiconductor, which has favorable electric characteristics and is easily miniaturized, can be provided.

[0034]

Further, a semiconductor device is provided in which variation in electric characteristics due to a short channel effect is not easily caused.

[0035]

When a dopant is added to an oxide semiconductor through an insulating layer, the oxide semiconductor is prevented from being thinned and the interface between the oxide semiconductor and the insulating layer is kept clean, so that characteristics and reliability of a semiconductor device can be increased.

BRIEF DESCRIPTION OF DRAWINGS

[0036]

In the accompanying drawings:

FIGS. 1A and 1B are a top view and a cross-sectional view illustrating an embodiment of the present invention;

FIGS. 2A and 2B are a top view and a cross-sectional view illustrating an embodiment of the present invention;

FIGS. 3A to 3D are cross-sectional views illustrating an embodiment of the present invention;

FIGS. 4A and 4B are cross-sectional views illustrating an embodiment of the present invention;

FIGS. 5A and 5B are a top view and a cross-sectional view illustrating an embodiment of the present invention;

FIGS. 6A and 6B are a top view and a cross-sectional view illustrating an embodiment of the present invention;

FIGS. 7A and 7B are cross-sectional views each illustrating an embodiment of the present invention;

FIG. 8 is a cross-sectional view illustrating an embodiment of the present invention;

FIGS. 9A and 9B illustrate band structures of an oxide semiconductor and a metal material;

FIGS. 10A and 10B are circuit diagrams each illustrating an embodiment of the present invention;

FIG. 11 is a circuit diagram illustrating an embodiment of the present invention;

FIGS. 12A and 12B are circuit diagrams each illustrating an embodiment of the present invention;

FIGS. 13A and 13B are circuit diagrams each illustrating an embodiment of the present invention;

FIG. 14A is a block diagram illustrating a specific example of a CPU and FIGS. 14B and 14C are circuit diagrams each illustrating part of the CPU;

FIGS. 15A to 15E show crystal structures of oxide materials;

FIGS. 16A to 16C show a crystal structure of an oxide material;

FIGS. 17A to 17C show a crystal structure of an oxide material; and

FIGS. 18A and 18B show crystal structures of oxide materials.

BEST MODE FOR CARRYING OUT THE INVENTION

[0037]

Embodiments of the present invention will be described below with reference to the accompanying drawings. Note that the present invention is not limited to the description below, and it is easily understood by those skilled in the art that various changes and modifications can be made without departing from the spirit and scope of the present invention. Therefore, the present invention should not be construed as being limited to the description in the following embodiments. Note that in the structures of the present invention described hereinafter, the same portions or portions having similar functions are denoted by the same reference numerals in different

drawings, and description thereof is not repeated.

[0038]

Note that the position, size, range, or the like of each structure illustrated in drawings and the like is not accurately represented in some cases for easy understanding. Therefore, the disclosed invention is not necessarily limited to the position, size, range, or the like as disclosed in the drawings and the like.

[0039]

Note that terms such as "first", "second", and "third" in this specification are used in order to avoid confusion among components, and the terms do not limit the components numerically. Therefore, for example, the term "first" can be replaced with the term "second", "third", or the like as appropriate.

[0040]

A transistor is one mode of a semiconductor device and can achieve amplification of current or voltage, a switching operation for controlling conduction or non-conduction, or the like. A transistor in this specification includes an insulated-gate field effect transistor (IGFET) and a thin film transistor (TFT).

[0041]

Functions of a "source" and a "drain" of a transistor are sometimes replaced with each other when a transistor of opposite polarity is used or when the direction of current flowing is changed in circuit operation, for example. Therefore, the terms "source" and "drain" can be used to denote the drain and the source, respectively, in this specification.

[0042]

In addition, in this specification and the like, the term such as "electrode" or "wiring" does not limit a function of a component. For example, an "electrode" is sometimes used as part of a "wiring", and vice versa. Furthermore, the term "electrode" or "wiring" can include the case where a plurality of "electrodes" or "wirings" is formed in an integrated manner.

[0043]

(Embodiment 1)

In this embodiment, a transistor in which an oxide semiconductor is used for a channel and a manufacturing method thereof will be described with reference to FIGS.

1A and 1B, FIGS. 2A and 2B, FIGS. 3A to 3D, and FIGS. 4A and 4B.

[0044]

FIG. 1A is a top view illustrating a structure of a transistor 100 which is one mode of a structure of a semiconductor device, and FIG. 1B is a cross-sectional view illustrating a cross-sectional structure of a portion indicated by a chain line A1-A2 in FIG. 1A. In FIG. 1A, a substrate and an insulating layer are omitted.

[0045]

In the transistor 100 illustrated in FIGS. 1A and 1B, a base layer 102 is formed over a substrate 101, and an oxide semiconductor layer 103 is formed over the base layer 102. A gate insulating layer 104 is formed over the oxide semiconductor layer 103, and a gate electrode 105 is formed over the gate insulating layer 104. An insulating layer 107 and an insulating layer 108 are formed over the gate electrode 105, and a source electrode 110a and a drain electrode 110b are formed over the insulating layer 108. The source electrode 110a and the drain electrode 110b are electrically connected to the oxide semiconductor layer 103 through contact holes 109 provided in the gate insulating layer 104, the insulating layer 107, and the insulating layer 108.

[0046]

The oxide semiconductor layer 103 includes a channel formation region 103c which overlaps with the gate electrode 105 with the gate insulating layer 104 interposed therebetween, a source region 103a which is electrically connected to the source electrode 110a, and a drain region 103b which is electrically connected to the drain electrode 110b.

[0047]

Further, the gate electrode 105 includes a gate electrode 105a which is in contact with the gate insulating layer 104 and a gate electrode 105b which is stacked over the gate electrode 105a.

[0048]

Although FIG. 1A illustrates an example in which a plurality of the contact holes 109 is provided over each of the source region 103a and the drain region 103b, only one contact hole 109 may be provided over each of the source region 103a and the drain region 103b. Further, it is preferable that the size of the contact hole 109 be as large as possible and the number of the contact holes 109 be large in order to reduce

contact resistance between the source electrode 110a and the source region 103a and contact resistance between the drain electrode 110b and the drain region 103b.

[0049]

A transistor 140 illustrated in FIGS. 2A and 2B includes, in addition to the structure of the transistor 100, sidewalls 111 on side surfaces of the gate electrode 105 and a low-concentration region 103d and a low-concentration region 103e in regions of the oxide semiconductor layer 103, which overlap with the sidewalls 111. The low-concentration region 103d is formed between the channel formation region 103c and the source region 103a, and the low-concentration region 103e is formed between the channel formation region 103c and the drain region 103b. FIG. 2A is a top view illustrating the structure of the transistor 140 and FIG. 2B is a cross-sectional view illustrating a stacked structure of a portion indicated by a chain line B1-B2 in FIG. 2A.

[0050]

The low-concentration region 103d and the low-concentration region 103e are provided, whereby deterioration of transistor characteristics and the negative shift in threshold voltage due to a short channel effect can be reduced.

[0051]

Each of the transistor 100 and the transistor 140 is one mode of a top-gate transistor.

[0052]

Next, a method for manufacturing the transistor 100 illustrated in FIGS. 1A and 1B will be described with reference to FIGS. 3A to 3D and FIGS. 4A and 4B. Note that FIGS. 3A to 3D and FIGS. 4A and 4B are cross-sectional views of the portion indicated by the chain line A1-A2 in FIG. 1A.

[0053]

First, the base layer 102 is formed with a thickness of greater than or equal to 50 nm and less than or equal to 300 nm, preferably greater than or equal to 100 nm and less than or equal to 200 nm over the substrate 101. As the substrate 101, a glass substrate, a ceramic substrate, a plastic substrate that has high heat resistance enough to withstand a process temperature of this manufacturing process, or the like can be used. In the case where a substrate does not need a light-transmitting property, a metal substrate such as a stainless alloy, whose surface is provided with an insulating layer,

may be used. As the glass substrate, for example, an alkali-free glass substrate of barium borosilicate glass, aluminoborosilicate glass, aluminosilicate glass, or the like may be used. Alternatively, a quartz substrate, a sapphire substrate, or the like can be used. Further alternatively, a single crystal semiconductor substrate or a polycrystalline semiconductor substrate made of silicon, silicon carbide, or the like, a compound semiconductor substrate made of silicon germanium or the like, an SOI substrate, or the like may be used as the substrate 101. Furthermore, any of these substrates further provided with a semiconductor element may be used as the substrate 101.

[0054]

The base layer 102 can be formed using a signal layer or a stacked layer using one or more of materials selected from aluminum nitride, aluminum oxide, aluminum nitride oxide, aluminum oxynitride, silicon nitride, silicon oxide, silicon nitride oxide, and silicon oxynitride. The base layer 102 has a function of preventing diffusion of an impurity element from the substrate 101. Note that in this specification, a nitride oxide is a substance which includes more nitrogen than oxygen, and an oxynitride is a substance which includes more oxygen than nitrogen. Note that content of each element can be measured by Rutherford backscattering spectrometry (RBS) or the like, for example.

[0055]

The base layer 102 can be formed by a sputtering method, a CVD method, a coating method, a printing method, or the like as appropriate. In this embodiment, a stack of layers of silicon nitride and silicon oxide is used as the base layer 102. Specifically, a 50-nm-thick silicon nitride layer is formed over the substrate 101, and a 150-nm-thick silicon oxide layer is formed over the silicon nitride layer. Note that the base layer 102 may be doped with phosphorus (P) or boron (B).

[0056]

When a halogen element such as chlorine or fluorine is contained in the base layer 102, a function of preventing diffusion of an impurity element from the substrate 101 can be further improved. The concentration of a halogen element to be contained in the base layer 102 is measured by secondary ion mass spectrometry (SIMS) and its peak is preferably greater than or equal to 1×10^{15} /cm^3 and less than or equal to $1 \times$

10^{20} /cm^3.

[0057]

The base layer 102 may be formed using a material from which oxygen is released by heating. "Oxygen is released by heating" means that the amount of released oxygen which is converted into oxygen atoms is greater than or equal to 1.0×10^{18} atoms/cm^3, preferably greater than or equal to 3.0×10^{20} atoms/cm^3 in thermal desorption spectroscopy (TDS).

[0058]

Here, a method in which the amount of released oxygen is measured by being converted into oxygen atoms using TDS analysis will be described below.

[0059]

The amount of released gas in TDS analysis is proportional to the integral value of a spectrum. Therefore, the amount of released gas can be calculated from the ratio between the integral value of spectrum of the insulating layer and the reference value of a standard sample. The reference value of a standard sample refers to the ratio of the density of a predetermined atom contained in a sample to the integral value of a spectrum.

[0060]

For example, the number of the released oxygen molecules (N_{O2}) from an insulating layer can be found according to Equation 1 with the TDS analysis results of a silicon wafer containing hydrogen at a predetermined density which is the standard sample and the TDS analysis results of the insulating layer. Here, all spectra having a mass number of 32 which are obtained by the TDS analysis are assumed to originate from an oxygen molecule. CH_3OH, which is given as a gas having a mass number of 32, is not taken into consideration on the assumption that it is unlikely to be present. Further, an oxygen molecule including an oxygen atom having a mass number of 17 or 18 which is an isotope of an oxygen atom is also not taken into consideration because the proportion of such a molecule in the natural world is minimal.

[0061]

$$N_{O2} = N_{H2}/S_{H2} \times S_{O2} \times \alpha \quad \text{(Equation 1)}$$

[0062]

N_{H2} is the value obtained by conversion of the number of hydrogen molecules desorbed from the standard sample into density. S_{H2} is the integral value of a spectrum when the standard sample is subjected to TDS analysis. Here, the reference value of the standard sample is set to N_{H2}/S_{H2}. S_{O2} is the integral value of a spectrum when the insulating layer is subjected to TDS analysis. α is a coefficient which influences spectrum intensity in TDS analysis. Refer to Japanese Published Patent Application No. H6-275697 for details of Equation 1. Note that the amount of released oxygen from the above insulating layer is measured with a thermal desorption spectroscopy apparatus produced by ESCO Ltd., EMD-WA1000S/W using a silicon wafer containing a hydrogen atom at 1×10^{16} atoms/cm^3 as the standard sample.

[0063]

Further, in the TDS analysis, oxygen is partly detected as an oxygen atom. The ratio between oxygen molecules and oxygen atoms can be calculated from the ionization rate of the oxygen molecules. Note that, since the above α includes the ionization rate of the oxygen molecules, the number of the released oxygen atoms can also be estimated through the evaluation of the number of the released oxygen molecules.

[0064]

Note that N_{O2} is the number of the released oxygen molecules. For the oxide insulating layer, the amount of released oxygen when converted into oxygen atoms is twice the number of the released oxygen molecules.

[0065]

In the above structure, the insulating layer from which oxygen is released by heating may be oxygen-excess silicon oxide (SiO_X ($X > 2$)). In the oxygen-excess silicon oxide (SiO_X ($X > 2$)), the number of oxygen atoms per unit volume is more than twice the number of silicon atoms per unit volume. The number of silicon atoms and the number of oxygen atoms per unit volume are measured by Rutherford backscattering spectrometry.

[0066]

Oxygen is supplied to the semiconductor from the base layer, so that the interface state between the base layer and the oxide semiconductor can be reduced. As

a result, charge or the like, which is generated due to the operation of the transistor or the like, can be prevented from being trapped at the interface between the base layer and the oxide semiconductor, so that the transistor with little degradation of electric characteristics can be obtained.

[0067]

Further, in some cases, charge is generated due to oxygen deficiency in the oxide semiconductor. In general, part of oxygen deficiency in an oxide semiconductor serves as a donor to generate an electron which is a carrier. As a result, the threshold voltage of a transistor shifts in the negative direction. This tendency occurs remarkably in an oxygen deficiency caused on the back channel side. Note that a back channel in this specification refers to the vicinity of an interface of the base layer in the oxide semiconductor. Sufficient release of oxygen from the base layer to the oxide semiconductor can compensate oxygen deficiency in the oxide semiconductor which causes negative shift of the threshold voltage.

[0068]

In other words, when oxygen deficiency is caused in the oxide semiconductor, it is difficult to suppress trapping of a charge at an interface between the base layer and the oxide semiconductor. However, by providing an insulating layer from which oxygen is released by heating for the base layer, the interface state between the oxide semiconductor and the base layer and the oxygen deficiency in the oxide semiconductor can be reduced and the adverse effect of the trapping of a charge at the interface between the oxide semiconductor and the base layer can be made small.

[0069]

The base layer 102 may be formed using an insulating material containing the same kind of component as the oxide semiconductor to be formed later. In the case where the base layer 102 is a stack of different layers, a layer in contact with the oxide semiconductor is formed using an insulating material containing the same kind of component as the oxide semiconductor. This is because such a material is compatible with the oxide semiconductor, and therefore, the use of such a material for the base layer 102 enables a state of the interface between the oxide semiconductor and the base layer 102 to be kept well. Here, "the same kind of component as the oxide semiconductor" means one or more of elements selected from constituent elements of

the oxide semiconductor. For example, in the case where the oxide semiconductor is formed using an In-Ga-Zn-based oxide semiconductor material, gallium oxide is given as an insulating material containing the same kind of component as the oxide semiconductor.

[0070]

Next, an oxide semiconductor is formed over the base layer 102. Before the oxide semiconductor is formed, in order that hydrogen, a hydroxyl group, and moisture are contained in the oxide semiconductor as little as possible, it is preferable to preheat the substrate 101 in a preheating chamber of a deposition apparatus so that an impurity such as hydrogen or moisture adsorbed on the substrate 101 or the base layer 102 is removed and exhausted. As an exhaustion unit provided in the preheating chamber, a cryopump is preferable. Note that this preheating treatment can be omitted. Further, this preheating treatment may be performed on the substrate 101 in a similar manner before formation of the base layer 102.

[0071]

The oxide semiconductor preferably contains at least indium (In) or zinc (Zn). It is particularly preferable that In and Zn be contained. As a stabilizer for reducing change in electrical characteristics of a transistor including the oxide semiconductor, gallium (Ga) is preferably additionally contained. Tin (Sn) is preferably contained as a stabilizer. Hafnium (Hf) is preferably contained as a stabilizer. Aluminum (Al) is preferably contained as a stabilizer.

[0072]

As another stabilizer, one or more lanthanoids which include lanthanum (La), cerium (Ce), praseodymium (Pr), neodymium (Nd), samarium (Sm), europium (Eu), gadolinium (Gd), terbium (Tb), dysprosium (Dy), holmium (Ho), erbium (Er), thulium (Tm), ytterbium (Yb), and lutetium (Lu) may be contained.

[0073]

As the oxide semiconductor, for example, indium oxide, tin oxide, zinc oxide, a two-component metal oxide such as an In-Zn-based oxide, a Sn-Zn-based oxide, an Al-Zn-based oxide, a Zn-Mg-based oxide, a Sn-Mg-based oxide, an In-Mg-based oxide, or an In-Ga-based oxide, a three-component metal oxide such as an In-Ga-Zn-based oxide (also referred to as IGZO), an In-Al-Zn-based oxide, an In-Sn-Zn-based oxide, a

Sn-Ga-Zn-based oxide, an Al-Ga-Zn-based oxide, a Sn-Al-Zn-based oxide, an In-Hf-Zn-based oxide, an In-La-Zn-based oxide, an In-Ce-Zn-based oxide, an In-Pr-Zn-based oxide, an In-Nd-Zn-based oxide, an In-Sm-Zn-based oxide, an In-Eu-Zn-based oxide, an In-Gd-Zn-based oxide, an In-Tb-Zn-based oxide, an In-Dy-Zn-based oxide, an In-Ho-Zn-based oxide, an In-Er-Zn-based oxide, an In-Tm-Zn-based oxide, an In-Yb-Zn-based oxide, or an In-Lu-Zn-based oxide, or a four-component metal oxide such as an In-Sn-Ga-Zn-based oxide, an In-Hf-Ga-Zn-based oxide, an In-Al-Ga-Zn-based oxide, an In-Sn-Al-Zn-based oxide, an In-Sn-Hf-Zn-based oxide, or an In-Hf-Al-Zn-based oxide can be used.

[0074]

The oxide semiconductor layer preferably includes In, more preferably In and Ga.

[0075]

Note that here, for example, an "In-Ga-Zn-based oxide" means an oxide containing indium (In), gallium (Ga), and zinc (Zn) and there is no particular limitation on the ratio of In:Ga:Zn. Further, a metal element in addition to In, Ga, and Zn may be contained.

[0076]

For the oxide semiconductor layer, a thin film expressed by the chemical formula, $InMO_3(ZnO)_m$ ($m > 0$), can be used. Note that M represents one or more metal elements selected from Sn, Zn, Ga, Al, Mn, and Co. Alternatively, a material represented by $In_3SnO_5(ZnO)_n$ ($n > 0$) may be used as the oxide semiconductor.

[0077]

For example, an In-Ga-Zn-based oxide with an atomic ratio of In:Ga:Zn = 1:1:1 (= 1/3:1/3:1/3) or In:Ga:Zn = 2:2:1 (= 2/5:2/5:1/5), or an oxide with an atomic ratio close to the above atomic ratios can be used. Alternatively, an In-Sn-Zn-based oxide with an atomic ratio of In:Sn:Zn = 1:1:1 (= 1/3:1/3:1/3), In:Sn:Zn = 2:1:3 (= 1/3:1/6:1/2), or In:Sn:Zn = 2:1:5 (= 1/4:1/8:5/8), or an oxide with an atomic ratio close to the above atomic ratios may be used.

[0078]

However, the composition is not limited to those described above, and a

material having an appropriate composition may be used in accordance with necessary semiconductor characteristics (such as mobility, threshold voltage, and variation). In order to obtain necessary semiconductor characteristics, it is preferable that the carrier density, the impurity concentration, the defect density, the atomic ratio of a metal element to oxygen, the interatomic distance, the density, and the like be set as appropriate.

[0079]

For example, with the In-Sn-Zn-based oxide, a high mobility can be relatively easily obtained. However, the mobility can be increased by reducing the defect density in the bulk also in the case of using the In-Ga-Zn-based oxide.

[0080]

Note that for example, the expression "the composition of an oxide including In, Ga, and Zn at the atomic ratio, In:Ga:Zn = a:b:c (a+b+c = 1), is in the neighborhood of the composition of an oxide including In, Ga, and Zn at the atomic ratio, In:Ga:Zn = A:B:C (A+B+C = 1)" means that a, b, and c satisfy the following relation: $(a-A)^2+(b-B)^2+(c-C)^2 \leq r^2$, and r may be 0.05, for example. The same applies to other oxides.

[0081]

The oxide semiconductor may be either single crystal or non-single-crystal. In the latter case, the oxide semiconductor may be either amorphous or polycrystalline. Further, the oxide semiconductor may have either an amorphous structure including a crystalline portion or a non-amorphous structure.

[0082]

An amorphous oxide semiconductor can have a flat surface with relative ease; therefore, when a transistor is manufactured with the use of the oxide semiconductor, interface scattering can be reduced, and relatively high mobility can be obtained with relative ease.

[0083]

In a crystalline oxide semiconductor, defects in the bulk can be further reduced and when a surface flatness is improved, mobility higher than that of an amorphous oxide semiconductor can be obtained. In order to improve the surface flatness, the

oxide semiconductor is preferably formed over a flat surface. Specifically, the oxide semiconductor is preferably formed over a surface with an average surface roughness (R_a) of less than or equal to 1 nm, preferably less than or equal to 0.3 nm, more preferably less than or equal to 0.1 nm. Note that R_a can be measured using an atomic force microscope (AFM).

[0084]

As the oxide semiconductor having crystallinity, a CAAC-OS (c-axis aligned crystalline oxide semiconductor) is preferable. The CAAC-OS is not completely single crystal nor completely amorphous. The CAAC-OS is an oxide semiconductor with a crystal-amorphous mixed phase structure where crystal parts are included in an amorphous phase. Note that in most cases, the crystal part fits inside a cube whose one side is less than 100 nm. From an observation image obtained with a transmission electron microscope (TEM), a boundary between an amorphous part and a crystal part in the CAAC-OS is not clear. Further, with the TEM, a grain boundary in the CAAC-OS is not found. Thus, in the CAAC-OS, a reduction in electron mobility, due to the grain boundary, is suppressed.

[0085]

In each of the crystal parts included in the CAAC-OS, a c-axis is aligned in a direction parallel to a normal vector of a surface where the CAAC-OS is formed or a normal vector of a surface of the CAAC-OS, triangular or hexagonal atomic arrangement which is seen from the direction perpendicular to the a-b plane is formed, and metal atoms are arranged in a layered manner or metal atoms and oxygen atoms are arranged in a layered manner when seen from the direction perpendicular to the c-axis. Note that, among crystal parts, the directions of the a-axis and the b-axis of one crystal part may be different from those of another crystal part. In this specification, a simple term "perpendicular" includes a range from 85° to 95°. In addition, a simple term "parallel" includes a range from −5° to 5°.

[0086]

In the CAAC-OS, distribution of crystal parts is not necessarily uniform. For example, in the formation process of the CAAC-OS; in the case where crystal growth occurs from a surface side of the oxide semiconductor, the proportion of crystal parts in

the vicinity of the surface of the oxide semiconductor film is higher than that in the vicinity of the surface where the oxide semiconductor film is formed in some cases. Further, when an impurity is added to the CAAC-OS, the crystal part in a region to which the impurity is added becomes amorphous in some cases.

[0087]

Since the c-axes of the crystal parts included in the CAAC-OS are aligned in the direction parallel to a normal vector of a surface where the CAAC-OS is formed or a normal vector of a surface of the CAAC-OS, the directions of the c-axes may be different from each other depending on the shape of the CAAC-OS (the cross-sectional shape of the surface where the CAAC-OS is formed or the cross-sectional shape of the surface of the CAAC-OS). Note that when the CAAC-OS is formed, the direction of c-axis of the crystal part is the direction parallel to a normal vector of the surface where the CAAC-OS is formed or a normal vector of the surface of the CAAC-OS. The crystal part is formed by film formation or by performing treatment for crystallization such as heat treatment after film formation.

[0088]

The CAAC-OS becomes a conductor, a semiconductor, or an insulator depending on its composition or the like. The CAAC-OS transmits or does not transmit visible light depending on its composition or the like. Note that nitrogen may be substituted for part of oxygen included in the CAAC-OS.

[0089]

With use of the CAAC-OS in a transistor, change in electric characteristics of the transistor due to irradiation with visible light or ultraviolet light can be reduced. Thus, the transistor has high reliability.

[0090]

An example of a crystal structure of the CAAC-OS will be described in detail with reference to FIGS. 15A to 15E, FIGS. 16A to 16C, and FIGS. 17A to 17C. In FIGS. 15A to 15E, FIGS. 16A to 16C, and FIGS. 17A to 17C, the vertical direction corresponds to the c-axis direction and a plane perpendicular to the c-axis direction corresponds to the a-b plane, unless otherwise specified. When the expressions "an upper half" and "a lower half" are simply used, they refer to an upper half above the a-b plane and a lower half below the a-b plane (an upper half and a lower half with respect

to the a-b plane). Furthermore, in FIGS. 15A to 15E, O surrounded by a circle represents tetracoordinate O and O surrounded by a double circle represents tricoordinate O.

[0091]

FIG. 15A illustrates a structure including one hexacoordinate In atom and six tetracoordinate oxygen (hereinafter referred to as tetracoordinate O) atoms proximate to the In atom. Here, a structure including one metal atom and oxygen atoms proximate thereto is referred to as a small group. The structure in FIG. 15A is actually an octahedral structure, but is illustrated as a planar structure for simplicity. Note that three tetracoordinate O atoms exist in each of an upper half and a lower half in FIG. 15A. In the small group illustrated in FIG. 15A, electric charge is 0.

[0092]

FIG. 15B illustrates a structure including one pentacoordinate Ga atom, three tricoordinate oxygen (hereinafter referred to as tricoordinate O) atoms proximate to the Ga atom, and two tetracoordinate O atoms proximate to the Ga atom. All the tricoordinate O atoms exist on the a-b plane. One tetracoordinate O atom exists in each of an upper half and a lower half in FIG. 15B. An In atom can also have the structure illustrated in FIG. 15B because an In atom can have five ligands. In the small group illustrated in FIG. 15B, electric charge is 0.

[0093]

FIG. 15C illustrates a structure including one tetracoordinate Zn atom and four tetracoordinate O atoms proximate to the Zn atom. In FIG. 15C, one tetracoordinate O atom exists in an upper half and three tetracoordinate O atoms exist in a lower half. Alternatively, three tetracoordinate O atoms may exist in the upper half and one tetracoordinate O atom may exist in the lower half in FIG. 15C. In the small group illustrated in FIG. 15C, electric charge is 0.

[0094]

FIG. 15D illustrates a structure including one hexacoordinate Sn atom and six tetracoordinate O atoms proximate to the Sn atom. In FIG. 15D, three tetracoordinate O atoms exist in each of an upper half and a lower half. In the small group illustrated in FIG. 15D, electric charge is +1.

[0095]

FIG. 15E illustrates a small group including two Zn atoms. In FIG. 15E, one tetracoordinate O atom exists in each of an upper half and a lower half. In the small group illustrated in FIG. 15E, electric charge is -1.

[0096]

Here, a plurality of small groups forms a medium group, and a plurality of medium groups forms a large group (also referred to as a unit cell).

[0097]

Now, a rule of bonding between the small groups will be described. The three O atoms in the upper half with respect to the hexacoordinate In atom in FIG. 15A each have three proximate In atoms in the downward direction, and the three O atoms in the lower half each have three proximate In atoms in the upward direction. The one O atom in the upper half with respect to the pentacoordinate Ga atom in FIG. 15B has one proximate Ga atom in the downward direction, and the one O atom in the lower half has one proximate Ga atom in the upward direction. The one O atom in the upper half with respect to the tetracoordinate Zn atom in FIG. 15C has one proximate Zn atom in the downward direction, and the three O atoms in the lower half each have three proximate Zn atoms in the upward direction. In this manner, the number of the tetracoordinate O atoms above the metal atom is equal to the number of the metal atoms proximate to and below each of the tetracoordinate O atoms. Similarly, the number of the tetracoordinate O atoms below the metal atom is equal to the number of the metal atoms proximate to and above each of the tetracoordinate O atoms. Since the coordination number of the tetracoordinate O atom is 4, the sum of the number of the metal atoms proximate to and below the O atom and the number of the metal atoms proximate to and above the O atom is 4. Accordingly, when the sum of the number of tetracoordinate O atoms above a metal atom and the number of tetracoordinate O atoms below another metal atom is 4, the two kinds of small groups including the metal atoms can be bonded. For example, in the case where the hexacoordinate metal (In or Sn) atom is bonded through three tetracoordinate O atoms in the lower half, it is bonded to the pentacoordinate metal (Ga or In) atom or the tetracoordinate metal (Zn) atom.

[0098]

A metal atom whose coordination number is 4, 5, or 6 is bonded to another metal atom through a tetracoordinate O atom in the c-axis direction. In addition to the

above, a medium group can be formed in a different manner by combining a plurality of small groups so that the total electric charge of the stacked structure is 0.

[0099]

FIG. 16A illustrates a model of a medium group included in a stacked structure of an In-Sn-Zn-based oxide. FIG. 16B illustrates a large group including three medium groups. Note that FIG. 16C illustrates an atomic arrangement in the case where the stacked structure in FIG. 16B is observed from the c-axis direction.

[0100]

In FIG. 16A, a tricoordinate O atom is omitted for simplicity, and a tetracoordinate O atom is illustrated by a circle; the number in the circle shows the number of tetracoordinate O atoms. For example, three tetracoordinate O atoms existing in each of an upper half and a lower half with respect to a Sn atom are denoted by circled 3. Similarly, in FIG. 16A, one tetracoordinate O atom existing in each of an upper half and a lower half with respect to an In atom is denoted by circled 1. FIG. 16A also illustrates a Zn atom proximate to one tetracoordinate O atom in a lower half and three tetracoordinate O atoms in an upper half, and a Zn atom proximate to one tetracoordinate O atom in an upper half and three tetracoordinate O atoms in a lower half.

[0101]

In the medium group included in the stacked structure of the In-Sn-Zn-based oxide in FIG. 16A, in the order starting from the top, a Sn atom proximate to three tetracoordinate O atoms in each of an upper half and a lower half is bonded to an In atom proximate to one tetracoordinate O atom in each of an upper half and a lower half, the In atom is bonded to a Zn atom proximate to three tetracoordinate O atoms in an upper half, the Zn atom is bonded to an In atom proximate to three tetracoordinate O atoms in each of an upper half and a lower half through one tetracoordinate O atom in a lower half with respect to the Zn atom, the In atom is bonded to a small group that includes two Zn atoms and is proximate to one tetracoordinate O atom in an upper half, and the small group is bonded to a Sn atom proximate to three tetracoordinate O atoms in each of an upper half and a lower half through one tetracoordinate O atom in a lower half with respect to the small group. A plurality of such medium groups is bonded, so that a large group is formed.

[0102]

Here, electric charge for one bond of a tricoordinate O atom and electric charge for one bond of a tetracoordinate O atom can be assumed to be −0.667 and −0.5, respectively. For example, electric charge of a (hexacoordinate or pentacoordinate) In atom, electric charge of a (tetracoordinate) Zn atom, and electric charge of a (pentacoordinate or hexacoordinate) Sn atom are +3, +2, and +4, respectively. Accordingly, electric charge in a small group including a Sn atom is +1. Therefore, electric charge of −1, which cancels +1, is needed to form a stacked structure including a Sn atom. As a structure having electric charge of −1, the small group including two Zn atoms as illustrated in FIG. 15E can be given. For example, with one small group including two Zn atoms, electric charge of one small group including a Sn atom can be cancelled, so that the total electric charge of the stacked structure can be 0.

[0103]

When the large group illustrated in FIG. 16B is repeated, an In-Sn-Zn-based oxide crystal ($In_2SnZn_3O_8$) can be obtained. Note that a stacked structure of the obtained In-Sn-Zn-based oxide can be expressed as a composition formula, $In_2SnZn_2O_7(ZnO)_m$ (m is 0 or a natural number).

[0104]

The above-described rule also applies to the following oxides: a four-component metal oxide such as an In-Sn-Ga-Zn-based oxide; a three-component metal oxide such as an In-Ga-Zn-based oxide (also referred to as IGZO), an In-Al-Zn-based oxide, a Sn-Ga-Zn-based oxide, an Al-Ga-Zn-based oxide, a Sn-Al-Zn-based oxide, an In-Hf-Zn-based oxide, an In-La-Zn-based oxide, an In-Ce-Zn-based oxide, an In-Pr-Zn-based oxide, an In-Nd-Zn-based oxide, an In-Sm-Zn-based oxide, an In-Eu-Zn-based oxide, an In-Gd-Zn-based oxide, an In-Tb-Zn-based oxide, an In-Dy-Zn-based oxide, an In-Ho-Zn-based oxide, an In-Er-Zn-based oxide, an In-Tm-Zn-based oxide, an In-Yb-Zn-based oxide, or an In-Lu-Zn-based oxide; a two-component metal oxide such as an In-Zn-based oxide, a Sn-Zn-based oxide, an Al-Zn-based oxide, a Zn-Mg-based oxide, a Sn-Mg-based oxide, an In-Mg-based oxide, or an In-Ga-based oxide; and the like.

[0105]

As an example, FIG. 17A illustrates a model of a medium group included in a stacked structure of an In-Ga-Zn-based oxide.

[0106]

In the medium group included in the stacked structure of the In-Ga-Zn-based oxide in FIG. 17A, in the order starting from the top, an In atom proximate to three tetracoordinate O atoms in each of an upper half and a lower half is bonded to a Zn atom proximate to one tetracoordinate O atom in an upper half, the Zn atom is bonded to a Ga atom proximate to one tetracoordinate O atom in each of an upper half and a lower half through three tetracoordinate O atoms in a lower half with respect to the Zn atom, and the Ga atom is bonded to an In atom proximate to three tetracoordinate O atoms in each of an upper half and a lower half through one tetracoordinate O atom in a lower half with respect to the Ga atom. A plurality of such medium groups is bonded, so that a large group is formed.

[0107]

FIG. 17B illustrates a large group including three medium groups. Note that FIG. 17C illustrates an atomic arrangement in the case where the stacked structure in FIG. 17B is observed from the c-axis direction.

[0108]

Here, since electric charge of a (hexacoordinate or pentacoordinate) In atom, electric charge of a (tetracoordinate) Zn atom, and electric charge of a (pentacoordinate) Ga atom are +3, +2, and +3, respectively, electric charge of a small group including any of an In atom, a Zn atom, and a Ga atom is 0. As a result, the total electric charge of a medium group having a combination of such small groups is always 0.

[0109]

In order to form the stacked structure of the In-Ga-Zn-based oxide, a large group can be formed using not only the medium group illustrated in FIG. 17A but also a medium group in which the arrangement of the In atom, the Ga atom, and the Zn atom is different from that in FIG. 17A.

[0110]

When the large group illustrated in FIG. 17B is repeated, a crystal of an In-Ga-Zn-based oxide can be obtained. Note that a stacked structure of the obtained In-Ga-Zn-based oxide can be expressed as a composition formula, $InGaO_3(ZnO)_n$ (n is

a natural number).

[0111]

In the case where $n = 1$ ($InGaZnO_4$), a crystal structure illustrated in FIG. 18A can be obtained, for example. Note that in the crystal structure in FIG. 18A, since a Ga atom and an In atom each have five ligands as illustrated in FIG. 15B, a structure in which Ga is replaced with In can be obtained.

[0112]

In the case where $n = 2$ ($InGaZn_2O_5$), a crystal structure illustrated in FIG. 18B can be obtained, for example. Note that in the crystal structure in FIG. 18B, since a Ga atom and an In atom each have five ligands as described in FIG. 15B, a structure in which Ga is replaced with In can be obtained.

[0113]

In this embodiment, first, a first oxide semiconductor having a thickness of greater than or equal to 1 nm and less than or equal to 10 nm is formed over the base layer 102 by a sputtering method. The substrate temperature is set to higher than or equal to 200 °C and lower than or equal to 400 °C in formation of the first oxide semiconductor.

[0114]

A sputtering apparatus used for formation of the oxide semiconductor will be described in detail below.

[0115]

The leakage rate of a deposition chamber used for forming an oxide semiconductor is preferably lower than or equal to 1×10^{-10} $Pa \cdot m^3$/second. Thus, entry of an impurity into a film to be formed by a sputtering method can be decreased.

[0116]

In order to decrease the leakage rate, internal leakage as well as external leakage needs to be reduced. The external leakage refers to inflow of gas from the outside of a vacuum system through a minute hole, a sealing defect, or the like. The internal leakage is due to leakage through a partition, such as a valve, in a vacuum system or due to released gas from an internal member. Measures need to be taken from both aspects of external leakage and internal leakage in order that the leakage rate

be lower than or equal to 1×10^{-10} Pa·m^3/second.

[0117]

In order to decrease external leakage, an open/close portion of the deposition chamber is preferably sealed with a metal gasket. For the metal gasket, a metal material covered with iron fluoride, aluminum oxide, or chromium oxide is preferably used. The metal gasket realizes higher adhesion than an O-ring, and can reduce the external leakage. Further, by use of a metal material covered with iron fluoride, aluminum oxide, chromium oxide, or the like which is in the passive state, released gas containing hydrogen generated from the metal gasket is suppressed, so that the internal leakage can also be reduced.

[0118]

As a member forming an inner wall of the deposition chamber, aluminum, chromium, titanium, zirconium, nickel, or vanadium, from which the amount of a released gas containing hydrogen is smaller, is used. An alloy material containing iron, chromium, nickel, and the like covered with the above-mentioned material may be used. The alloy material containing iron, chromium, nickel, and the like is rigid, resistant to heat, and suitable for processing. Here, when surface unevenness of the member is decreased by polishing or the like to reduce the surface area, the released gas can be reduced. Alternatively, the above-mentioned member of the deposition apparatus may be covered with iron fluoride, aluminum oxide, chromium oxide, or the like which is in the passive state.

[0119]

Furthermore, it is preferable to provide a gas refiner for a sputtering gas just in front of the deposition chamber. At this time, the length of a pipe between the gas refiner and the deposition chamber is less than or equal to 5 m, preferably less than or equal to 1 m. When the length of the pipe is less than or equal to 5 m or less than or equal to 1 m, the effect of the released gas from the pipe can be reduced accordingly.

[0120]

Evacuation of the deposition chamber is preferably performed with a rough vacuum pump, such as a dry pump, and a high vacuum pump, such as a sputter ion pump, a turbo molecular pump, or a cryopump, in appropriate combination. In order

to remove moisture remaining in the deposition chamber, an entrapment vacuum pump such as a cryopump, an ion pump, or a titanium sublimation pump is preferably used. The turbo molecular pump has an outstanding capability in evacuating a large-sized molecule, whereas it has a low capability in evacuating hydrogen or water. Hence, combination of a cryopump having a high capability in evacuating water and a sputter ion pump having a high capability in evacuating hydrogen is effective. The evacuation unit may be a turbo molecular pump provided with a cold trap. In the deposition chamber which is evacuated with an entrapment vacuum pump such as a cryopump, a hydrogen atom, a compound containing a hydrogen atom such as water (H_2O) (more preferably, also a compound containing a carbon atom), and the like are removed, whereby the impurity concentration in the oxide semiconductor layer formed in the deposition chamber can be reduced.

[0121]

An adsorbate present at the inner wall of the deposition chamber does not affect the pressure in the deposition chamber because it is adsorbed on the inner wall, but the adsorbate leads to release of gas at the time of the evacuation of the deposition chamber. Therefore, although the leakage rate and the evacuation rate do not have a correlation, it is important that the adsorbate present in the deposition chamber be desorbed as much as possible and evacuation be performed in advance with the use of a pump having high evacuation capability. Note that the deposition chamber may be subjected to baking for promotion of desorption of the adsorbate. By the baking, the rate of desorption of the adsorbate can be increased about tenfold. The baking should be performed at a temperature greater than or equal to 100 °C and less than or equal to 450 °C. At this time, when the adsorbate is removed while an inert gas is introduced, the rate of desorption of water or the like, which is difficult to desorb only by evacuation, can be further increased.

[0122]

In a sputtering method, an RF power supply device, an AC power supply device, a DC power supply device, or the like can be used as a power supply device for generating plasma as appropriate.

[0123]

As an In-Ga-Zn-based oxide target for forming an In-Ga-Zn-based oxide material as an oxide semiconductor by a sputtering method, for example, a target having a composition ratio of In_2O_3:Ga_2O_3:ZnO = 1:1:1 [molar ratio] can be used. Alternatively, a target having a composition ratio of In_2O_3:Ga_2O_3:ZnO = 1:1:2 [molar ratio], a target having a composition ratio of In_2O_3:Ga_2O_3:ZnO = 1:1:4 [molar ratio], or a target having a composition ratio of In_2O_3:Ga_2O_3:ZnO = 2:1:8 [molar ratio] can be used. Further, an In-Ga-Zn-based oxide target having an atomic ratio of In:Ga:Zn = 1:1:1, 4:2:3, 3:1:2, 1:1:2, 2:1:3, or 3:1:4 can be used. When an oxide semiconductor is formed using an In-Ga-Zn-based oxide target having any of the aforementioned atomic ratios, a polycrystal or CAAC-OS is easily formed.

[0124]

An In-Sn-Zn-based oxide can be referred to as ITZO. In the case of forming an oxide semiconductor using an In-Sn-Zn-based oxide by a sputtering method, it is preferable to use an In-Sn-Zn-based oxide target having an atomic ratio of In:Sn:Zn = 1:1:1, 2:1:3, 1:2:2, or 20:45:35. When an oxide semiconductor is formed using a target of an In-Sn-Zn-based oxide having the above atomic ratio, a polycrystal or a CAAC-OS is likely to be formed.

[0125]

The relative density of the metal oxide target used for forming an oxide semiconductor is higher than or equal to 90 % and lower than or equal to 100 %, preferably higher than or equal to 95 % and lower than or equal to 99.9 %. With the use of a metal oxide target with a high relative density, a dense oxide semiconductor layer can be deposited.

[0126]

As a sputtering gas, a rare gas (typically argon) atmosphere, an oxygen atmosphere, or a mixed gas of a rare gas and oxygen is used as appropriate. It is preferable that a high-purity gas from which impurities such as hydrogen, water, a hydroxyl group, and hydride are removed be used as a sputtering gas. For example, when argon is used as a sputtering gas, it is preferable that the purity be 9N, the dew point be −121 °C, the content of H_2O be 0.1 ppb or lower, and the content of H_2 be 0.5 ppb or lower. When oxygen is used as a sputtering gas, it is preferable that the purity

be 8N, the dew point be −112 °C, the content of H_2O be 1 ppb or lower, and the content of H_2 be 1 ppb or lower.

[0127]

The substrate temperature in deposition is set to higher than or equal to 150 °C and lower than or equal to 450 °C, preferably higher than or equal to 200 °C and lower than or equal to 350 °C. The deposition is performed while the substrate is heated to higher than or equal to 150 °C and lower than or equal to 450 °C, preferably higher than or equal to 200 °C and lower than or equal to 350 °C, whereby moisture (including hydrogen) or the like is prevented from entering a film.

[0128]

By heating the substrate during deposition, the concentration of an impurity such as hydrogen, moisture, hydride, or a hydroxide in the formed oxide semiconductor can be reduced. In addition, damage by sputtering can be reduced. Then, a sputtering gas from which hydrogen and moisture are removed is introduced into the deposition chamber while moisture remaining therein is removed, and the first oxide semiconductor having a thickness of greater than or equal to 1 nm and less than or equal to 10 nm, preferably greater than or equal to 2 nm and less than or equal to 5 nm is formed with the use of the above target.

[0129]

In this embodiment, the first oxide semiconductor is formed to a thickness of 5 nm in an oxygen atmosphere, an argon atmosphere, or an atmosphere including argon and oxygen under conditions where as a target for an oxide semiconductor, a target for an In-Ga-Zn-based oxide semiconductor (In_2O_3:Ga_2O_3:ZnO = 1:1:2 [molar ratio]) is used, the distance between the substrate and the target is 170 mm, the substrate temperature is 250 °C, the pressure is 0.4 Pa, and the direct current (DC) power is 0.5 kW.

[0130]

Next, first heat treatment is performed under a condition where the atmosphere of a chamber in which the substrate is set is an atmosphere of nitrogen or dry air. The temperature of the first heat treatment is higher than or equal to 400 °C and lower than or equal to 750 °C. The first oxide semiconductor is crystallized by the first heat

treatment to be a first crystalline oxide semiconductor.

[0131]

Depending on the temperature of the first heat treatment, the first heat treatment causes crystallization from a film surface and crystal grows from the film surface toward the inside of the film; thus, c-axis aligned crystal is obtained. By the first heat treatment, large amounts of zinc and oxygen gather to the film surface, and one or more layers of graphene-type two-dimensional crystal including zinc and oxygen and having a hexagonal upper plane are formed at the outermost surface; the layer(s) at the outermost surface grow in the thickness direction to form a stack of layers. By increasing the temperature of the heat treatment, crystal growth proceeds from the surface to the inside and further from the inside to the bottom.

[0132]

By the first heat treatment, oxygen in the base layer 102 is diffused to an interface between the base layer and the first crystalline oxide semiconductor layer or the vicinity of the interface (within ±5 nm from the interface), whereby oxygen deficiency in the first crystalline oxide semiconductor is reduced. Therefore, it is preferable that oxygen be included in (in a bulk of) the base layer 102 or at the interface between the first crystalline oxide semiconductor and the base layer 102 at an amount that exceeds at least the stoichiometric proportion.

[0133]

Then, a second oxide semiconductor with a thickness greater than 10 nm is formed over the first crystalline oxide semiconductor. The second oxide semiconductor is formed by a sputtering method, and the substrate temperature in the deposition is set to be higher than or equal to 200 °C and lower than or equal to 400 °C. By setting the substrate temperature in the deposition to be higher than or equal to 200 °C and lower than or equal to 400 °C, precursors can be arranged in the oxide semiconductor formed on and in contact with the surface of the first crystalline oxide semiconductor and so-called orderliness can be obtained.

[0134]

In this embodiment, the second oxide semiconductor is formed to a thickness of 25 nm in an oxygen atmosphere, an argon atmosphere, or an atmosphere including

argon and oxygen under conditions where as a target for an oxide semiconductor, a target for an In-Ga-Zn-based oxide semiconductor (In_2O_3:Ga_2O_3:ZnO = 1:1:2 [molar ratio]) is used, the distance between the substrate and the target is 170 mm, the substrate temperature is 400 °C, the pressure is 0.4 Pa, and the direct current (DC) power is 0.5 kW.

[0135]

Next, second heat treatment is performed under a condition where the atmosphere of a chamber in which the substrate is set is an atmosphere of nitrogen or dry air. The temperature of the second heat treatment is higher than or equal to 400 °C and lower than or equal to 750 °C. A second crystalline oxide semiconductor is formed by the second heat treatment. The second heat treatment is performed in a nitrogen atmosphere, an oxygen atmosphere, or a mixed atmosphere of nitrogen and oxygen, whereby the density of the second crystalline oxide semiconductor is increased and the number of defects therein is reduced. By the second heat treatment, crystal growth proceeds in the thickness direction with the use of the first crystalline oxide semiconductor as a nucleus, that is, crystal growth proceeds from the bottom to the inside; thus, the second crystalline oxide semiconductor is formed. At this time, to compose the first crystalline oxide semiconductor and the second crystalline oxide semiconductor using the same kind of element is referred to as "homo-growth". Alternatively, to compose the first crystalline oxide semiconductor and the second crystalline oxide semiconductor using elements, at least one kind of which differs between the first crystalline oxide semiconductor and the second crystalline oxide semiconductor, is referred to as "hetero-growth".

[0136]

Thus, in a formation step of an oxide semiconductor, entry of an impurity is prevented as much as possible by pressure of a deposition chamber, leakage rate of the deposition chamber, or the like, so that an impurity such as hydrogen or moisture is prevented from entering the oxide semiconductor. Hydrogen contained in the oxide semiconductor is reacted with oxygen bonded to a metal atom to be water, and in addition, a defect is formed in a lattice from which the oxygen is detached (or a portion from which the oxygen is removed).

[0137]

Thus, the impurity is reduced as much as possible in the formation process of the oxide semiconductor, whereby defects in the oxide semiconductor can be reduced. From the above, a transistor in which a channel region is formed in the oxide semiconductor including purified CAAC-OS obtained by removing an impurity as much as possible has a small amount of change in threshold voltage between before and after light irradiation or the BT test against the transistor and thus has stable electric characteristics.

[0138]

Further, after the second heat treatment, it is preferable to perform additional heat treatment in which atmosphere is changed to an oxidizing atmosphere while the temperature is kept. The oxygen defects in the oxide semiconductor can be reduced by the heat treatment in an oxidizing atmosphere.

[0139]

Note that a metal oxide which can be used for the oxide semiconductor has band gap of 2 eV or more, preferably 2.5 eV or more, further preferably 3 eV or more. In this manner, off-state current of a transistor can be reduced by using a metal oxide having a wide band gap.

[0140]

It is preferable to perform the steps from the formation of the base layer 102 to the second heat treatment successively without exposure to the air. The steps from the formation of the base layer 102 to the second heat treatment are preferably performed in an atmosphere which is controlled to include little hydrogen and moisture (such as an inert gas atmosphere, a reduced-pressure atmosphere, or a dry-air atmosphere); in terms of moisture, for example, a dry nitrogen atmosphere with a dew point of −40 °C or lower, preferably a dew point of −50 °C or lower may be employed.

[0141]

Next, the stack of oxide semiconductor layers including the first crystalline oxide semiconductor and the second crystalline oxide semiconductor is processed to form an island-shaped oxide semiconductor layer 103 (see FIG. 3A).

[0142]

The oxide semiconductor can be processed by being etched after a mask having a desired shape is formed over the oxide semiconductor. The mask can be formed by a method such as photolithography. Alternatively, the mask may be formed by a method such as an inkjet method or a printing method.

[0143]

For the etching of the oxide semiconductor, either a dry etching method or a wet etching method may be employed. It is needless to say that both of them may be employed in combination.

[0144]

One of features of the first and second crystalline oxide semiconductors obtained by the above formation method is that they have c-axis alignment. Note that the first crystalline oxide semiconductor and the second crystalline oxide semiconductor have neither a single crystal structure nor an amorphous structure and are crystalline oxide semiconductors having c-axis alignment (CAAC-OS).

[0145]

Without limitation to the two-layer structure in which the second crystalline oxide semiconductor is formed over the first crystalline oxide semiconductor, a stacked structure including three or more layers may be formed by repeatedly performing a process of deposition and heat treatment for forming a third crystalline oxide semiconductor after the second crystalline oxide semiconductor is formed.

[0146]

By forming a transistor with the use of a stack of a first crystalline oxide semiconductor and a second crystalline oxide semiconductor, like the oxide semiconductor layer 103, the transistor can have stable electric characteristics and high reliability.

[0147]

Next, the gate insulating layer 104 is formed over the oxide semiconductor layer 103. The gate insulating layer 104 can be formed using a single layer or a stacked layer using one or more of materials selected from aluminum nitride, aluminum oxide, aluminum nitride oxide, aluminum oxynitride, silicon nitride, silicon oxide, silicon nitride oxide, silicon oxynitride, tantalum oxide, and lanthanum oxide.

[0148]

When a high-k material such as hafnium silicate ($HfSiO_x$ ($x > 0$)), hafnium silicate to which nitrogen is added ($HfSi_xO_yN_z$ ($x > 0$, $y > 0$, $z > 0$)), hafnium aluminate to which nitrogen is added ($HfAl_xO_yN_z$ ($x > 0$, $y > 0$, $z > 0$)), hafnium oxide, or yttrium oxide is used as the gate insulating layer 104, while the substantial (e.g., silicon oxide equivalent) thickness of the gate insulating film is not changed, the physical thickness of the gate insulating film can be increased so that gate leakage current can be reduced. Further, a stacked structure can be used in which a high-k material and one or more of silicon oxide, silicon oxynitride, silicon nitride, silicon nitride oxide, aluminum oxide, aluminum oxynitride, and gallium oxide are stacked. For example, the thickness of the gate insulating layer 104 is preferably greater than or equal to 1 nm and less than or equal to 300 nm, and more preferably greater than or equal to 5 nm and less than or equal to 50 nm. When the thickness of the gate insulating film 104 is greater than or equal to 5 nm, gate leakage current can be reduced.

[0149]

The gate insulating layer 104 is formed by a sputtering method, a CVD method, or the like. Other than a sputtering method and a plasma CVD method, the gate insulating layer 104 can be formed by a deposition method such as a high-density plasma CVD method using microwaves (e.g., a frequency of 2.45 GHz). The gate insulating layer 104 is not limited to a single layer, and a stack of different layers may be used. Note that the gate insulating layer 104 is preferably an insulating layer containing oxygen, more preferably an oxide insulating layer from which oxygen is released by heating, in a portion which is in contact with the oxide semiconductor layer 103. Silicon oxide is used for the gate insulating layer 104, whereby oxygen is diffused to the oxide semiconductor layer 103 and oxygen deficiencies in the oxide semiconductor layer 103 are reduced; thus, favorable transistor characteristics can be obtained.

[0150]

In the structure described in this embodiment, only the oxide semiconductor layer 103 causes a depression and a projection over a substrate and the gate insulating layer 104 hardly has a portion which overlaps with a step; therefore, leakage current due to the gate insulating layer 104 can be reduced and withstand voltage of the gate

insulating layer 104 can be increased. Accordingly, a transistor can be operated even when the gate insulating layer 104 is as thin as approximately 5 nm. Note that a reduction in thickness of the gate insulating layer 104 has effects of reducing a short channel effect and increasing the operation speed of the transistor.

[0151]

Before the gate insulating layer 104 is formed, the surface of the oxide semiconductor layer 103 may be exposed to plasma of an oxidizing gas such as oxygen, ozone, or dinitrogen monoxide so as to be oxidized, thereby reducing the oxygen deficiency. In this embodiment, as the gate insulating layer 104, oxide silicon is formed to a thickness of 100 nm over the oxide semiconductor layer 103.

[0152]

Next, a conductive layer is formed over the gate insulating layer 104 by a sputtering method, a vacuum evaporation method, or a plating method, a mask is formed over the conductive layer, and the conductive layer is selectively etched to form the gate electrode 105. The mask formed over the conductive layer can be formed by a printing method, an inkjet method, or a photolithography method as appropriate. The gate electrode 105 includes the gate electrode 105a which is in contact with the gate insulating layer 104 and the gate electrode 105b which is stacked over the gate electrode 105a.

[0153]

As a material of the gate electrode 105a, indium gallium zinc oxide (In-Ga-Zn-O) containing nitrogen, indium tin oxide (In-Sn-O) containing nitrogen, indium gallium oxide (In-Ga-O) containing nitrogen, indium zinc oxide (In-Zn-O) containing nitrogen, tin oxide (Sn-O) containing nitrogen, indium oxide (In-O) containing nitrogen, or a metal nitride (e.g., InN, ZnN) is preferably used.

[0154]

These material each have a work function of 5 eV or higher, preferably 5.5 eV or higher. The gate electrode 105a is provided between the gate electrode 105b and the gate insulating layer 104 and overlaps with the oxide semiconductor layer 103 with the gate insulating layer 104 interposed therebetween, whereby the threshold voltage of the electric characteristics of the transistor can be positive. Accordingly, a so-called normally-off switching element can be obtained. For example, in the case where

In-Ga-Zn-O containing nitrogen is used for the gate electrode 105a, In-Ga-Zn-O having a nitrogen concentration higher than at least that of the oxide semiconductor layer 103, specifically, In-Ga-Zn-O having a nitrogen concentration of higher than or equal to 7 at.% is used.

[0155]

As a material used for forming the gate electrode 105b, a metal element selected from aluminum (Al), chromium (Cr), copper (Cu), tantalum (Ta), titanium (Ti), molybdenum (Mo), tungsten (W), neodymium (Nd), and scandium (Sc), an alloy containing any of these metal elements as a component, an alloy containing these metal elements in combination, a nitride of any of these metal elements, or the like can be used. Further, one or more metal elements selected from manganese (Mn), magnesium (Mg), zirconium (Zr), and beryllium (Be) may be used.

[0156]

Further, the gate electrode 105b may have a single-layer structure or a stacked structure of two or more layers. For example, a single-layer structure of aluminum containing silicon, a two-layer structure in which titanium is stacked over aluminum, a two-layer structure in which titanium is stacked over titanium nitride, a two-layer structure in which tungsten is stacked over titanium nitride, a two-layer structure in which tungsten is stacked over tantalum nitride, a two-layer structure in which Cu is stacked over a Cu-Mg-Al alloy, a three-layer structure in which titanium, aluminum, and titanium are stacked in this order, and the like can be given.

[0157]

The gate electrode 105b can be formed using a light-transmitting conductive material such as indium tin oxide, indium oxide containing tungsten oxide, indium zinc oxide containing tungsten oxide, indium oxide containing titanium oxide, indium tin oxide containing titanium oxide, indium zinc oxide, or indium tin oxide to which silicon oxide is added. It is also possible to have a stacked structure formed using the above light-transmitting conductive material and the above metal element.

[0158]

In this embodiment, for the gate electrode 105a, indium gallium zinc oxide containing nitrogen is used. Further, for the gate electrode 105b, two-layer structure in which tungsten is stacked over titanium nitrogen is used (see FIG. 3B). Note that end

portions of the formed gate electrode 105 preferably have a tapered shape, so that the coverage with a layer formed later is improved.

[0159]

Next, the source region 103a and the drain region 103b are formed by a self-aligned process. Specifically, with the use of the gate electrode 105 as a mask, a dopant 106 is added to the oxide semiconductor layer 103 by an ion doping method or an ion implantation method. As the dopant 106, one or more elements selected from Group 15 (Group 5B) elements such as nitrogen (N) or phosphorus (P) can be used.

[0160]

Further, the dopant 106 is not added to the region in the oxide semiconductor layer 103, which overlaps with the gate electrode 105, due to the gate electrode 105 serving as a mask. The region serves as the channel formation region 103c.

[0161]

The source region 103a and the drain region 103b to which the dopant 106 is added each become an n-type oxide semiconductor, which have lower resistivity than the channel formation region 103c. Therefore, the resistance of the source region 103a and the drain region 103b is low, whereby the transistor 100 can be operated at high speed. In addition, with the use of a self-aligned process, the source region 103a and the drain region 103b hardly overlap with the gate electrode 105, leading to reduction in parasitic capacitance; thus, the transistor 100 can be operated at higher speed.

[0162]

Alternatively, with the use of the gate electrode 105 as a mask, the source region 103a and the drain region 103b may be formed in such a manner that regions of the gate insulating layer 104 over the regions to be the source region and the drain region in the oxide semiconductor layer 103 are removed so that the regions in the oxide semiconductor layer 103 are exposed, and the dopant 106 is added to the exposed regions of the oxide semiconductor layer 103. The regions of the gate insulating layer 104 over the oxide semiconductor layer 103 are removed under such a condition that the oxide semiconductor layer 103 is hardly etched.

[0163]

The dopant 106 can be added to the exposed regions of the oxide semiconductor layer 103 by a method other than an ion doping method and an ion

implantation method. For example, the dopant 106 can be added to the exposed regions of the oxide semiconductor layer 103 in such a manner that plasma is generated in a gas atmosphere including an element to be added and the exposed regions of the oxide semiconductor layer 103 are subjected to plasma treatment. At this time, a bias is preferably applied to the substrate 101. The greater a bias applied to the substrate is, the deeper the dopant 106 can be added to the oxide semiconductor layer 103. As an apparatus for generating plasma, a dry etching apparatus, a plasma CVD apparatus, a high-density plasma CVD apparatus, or the like can be used.

[0164]

However, an oxide semiconductor layer might be etched to be thinned when a dopant is added by plasma treatment. Therefore, the dopant 106 is preferably added to the exposed regions of oxide semiconductor layer 103 by an ion doping method or an ion implantation method.

[0165]

Further, when the gate insulating layer 104 over regions of the oxide semiconductor layer 103, which are to be the source region 103a and the drain region 103b, is removed, the regions of the oxide semiconductor layer 103, which are to be the source region 103a and the drain region 103b, might be etched at the same time to be thinned. As a result, resistance of the source region 103a and the drain region 103b is increased, and defective units due to over-etching caused by unnecessary thinning of a layer are produced more frequently.

[0166]

This phenomenon becomes serious in the case where dry etching is performed under such conditions that the etching selectivity ratio of the gate insulating layer 104 with respect to the oxide semiconductor layer 103 is not sufficiently high. In general, in the case of forming a transistor with a short channel length, a dry etching method with high processing accuracy needs to be employed; otherwise, a source region and a drain region are easily thinned.

[0167]

The problem would not arise when the oxide semiconductor layer 103 is thick enough. In the case where the channel length is less than or equal to 200 nm, the thickness of the portion of the oxide semiconductor layer to be a channel is required to

be less than or equal to 20 nm, preferably less than or equal to 10 nm, in order to prevent a short channel effect. In the case where such a thin oxide semiconductor layer is used, it is not favorable that the thickness of the oxide semiconductor layer is reduced as described above.

[0168]

Therefore, the dopant 106 is preferably added to the oxide semiconductor layer 103 by an ion doping method or an ion implantation method while the gate insulating layer 104 remains so as not to expose the oxide semiconductor layer 103. The dopant 106 is added to the oxide semiconductor layer 103 through the gate insulating layer 104, so that excessive damage to the oxide semiconductor layer 103 can be reduced. Furthermore, the interface between the oxide semiconductor layer 103 and the gate insulating layer 104 is kept clean, so that characteristics and reliability of the transistor are improved. Moreover, the depth to which the dopant 106 is added (addition region) is easily controlled, so that the dopant 106 can be accurately added to the oxide semiconductor layer 103.

[0169]

In this embodiment, nitrogen (N) is used as the dopant 106, which is added to the oxide semiconductor layer 103 by an ion implantation method through the gate insulating layer 104. The concentrations of nitrogen in the source region 103a and the drain region 103b which are formed by adding nitrogen is controlled to be higher than or equal to 5×10^{19} atoms/cm^3 and lower than or equal to 1×10^{22} atoms/cm^3, preferably higher than or equal to 1×10^{20} atoms/cm^3 and lower than 7 at.% (see FIG. 3C).

[0170]

Next, heat treatment is performed at a temperature higher than or equal to 300 °C and lower than or equal to 600 °C in a reduced pressure atmosphere, an inert gas atmosphere such as a nitrogen atmosphere or a rare gas atmosphere, an oxygen gas atmosphere, or an ultra-dry air (with a moisture content of 20 ppm (equivalent to a dew point of −55 °C) or lower, preferably 1 ppm or lower, further preferably 10 ppb or lower when measured with a dew-point meter using cavity ring-down laser spectroscopy (CRDS)). In this embodiment, heat treatment is performed at 450 °C in a nitrogen

atmosphere for one hour with the use of an electric furnace which is a kind of heat treatment apparatuses.

[0171]

Note that a heat treatment apparatus is not limited to an electrical furnace, and may include a device for heating an object to be processed by heat conduction or heat radiation from a heating element such as a resistance heating element. For example, a rapid thermal annealing (RTA) apparatus such as a gas rapid thermal annealing (GRTA) apparatus or a lamp rapid thermal annealing (LRTA) apparatus can be used. An LRTA apparatus is an apparatus for heating an object to be processed by radiation of light (an electromagnetic wave) emitted from a lamp such as a halogen lamp, a metal halide lamp, a xenon arc lamp, a carbon arc lamp, a high pressure sodium lamp, or a high pressure mercury lamp. A GRTA apparatus is an apparatus for heat treatment using a high-temperature gas. As the high-temperature gas, an inert gas which does not react with an object to be processed by heat treatment, such as nitrogen or a rare gas like argon, is used.

[0172]

For example, as the heat treatment, GRTA may be performed in the following manner: the substrate is moved into an inert gas heated to a high temperature and is heated for several minutes, and then the substrate is moved out of the inert gas.

[0173]

When the heat treatment is performed in an atmosphere of an inert gas such as nitrogen or a rare gas or ultra-dry air, it is preferable that the atmosphere do not contain water, hydrogen, and the like. It is also preferable that nitrogen, oxygen, or a rare gas which is introduced into the heat treatment apparatus have a purity of 6N (99.9999 %) or more, preferably 7N (99.99999 %) or more (that is, the impurity concentration is 1 ppm or less, preferably 0.1 ppm or less).

[0174]

Through the above heat treatment, the source region 103a and the drain region 103b can have a wurtzite crystal structure. Further, the low-concentration region 103d and the low-concentration region 103e can also have a wurtzite crystal structure. The heat treatment can be performed at any time after the dopant 106 is added.

[0175]

The dopant 106 is added by an ion doping method, an ion implantation method, or the like while the substrate is heated, whereby a wurtzite crystal structure can be obtained without additional heat treatment.

[0176]

Next, the insulating layer 107 and the insulating layer 108 are formed by a sputtering method, a CVD method, or the like to cover the oxide semiconductor layer 103 and the gate electrode 105. The insulating layer 107 and the insulating layer 108 can each be formed using a material selected from aluminum nitride, aluminum oxide, aluminum nitride oxide, aluminum oxynitride, silicon nitride, silicon oxide, silicon nitride oxide, and silicon oxynitride.

[0177]

The thickness of each of the insulating layer 107 and the insulating layer 108 is greater than or equal to 50 nm, preferably greater than or equal to 200 nm and less than or equal to 500 nm. In this embodiment, silicon oxide is formed to a thickness of 300 nm as the insulating layer 107 and aluminum oxide is formed to a thickness of 100 nm as the insulating layer 108.

[0178]

The insulating layer 108 is preferably formed using silicon nitride or aluminum oxide in order to prevent entry of an impurity or the like from the outside. In this embodiment, aluminum oxide is formed to a thickness of 100 nm as the insulating layer 108 (see FIG. 3D). One of or both of the insulating layer 107 and the insulating layer 108 may be omitted.

[0179]

After the insulating layer 108 is formed, if necessary, heat treatment may be performed in an atmosphere which hardly contains hydrogen and moisture (a nitrogen atmosphere, an oxygen atmosphere, a dry-air atmosphere (for example, as for moisture, a dew point is lower than or equal to −40 °C, preferably lower than or equal to −60 °C), or the like) at a temperature higher than or equal to 150 °C and lower than or equal to 650 °C, preferably higher than or equal to 200 °C and lower than or equal to 500 °C.

[0180]

Next, a mask is formed over the insulating layer 108, part of the gate insulating layer 104, part of the insulating layer 107, and part of the insulating layer 108 are selectively etched with the use of the mask to expose part of the source region 103a and part of the drain region 103b; thus, the contact holes 109 are formed (see FIG. 4A).

[0181]

Next, a conductive layer is formed over the insulating layer 108, a mask is formed over the conductive layer, and the conductive layer is selectively etched to form the source electrode 110a and the drain electrode 110b (see FIG. 4B). A material similar to that of the gate electrode 105b can be used for the conductive layer used for forming the source electrode 110a and the drain electrode 110b.

[0182]

In this embodiment, as the conductive layer used for forming the source electrode 110a and the drain electrode 110b, a conductive layer in which Cu is stacked over a Cu-Mg-Al alloy is used. The Cu-Mg-Al alloy material is provided in contact with the insulating layer 108, whereby adhesion of the conductive layer can be improved.

[0183]

Note that the channel length of the transistor 100 is equal to the length of the channel formation region 103c which is sandwiched between the source region 103a and the drain region 103b in FIG. 1B. Further, the channel length of the transistor 100 is substantially equal to the width of the gate electrode 105.

[0184]

Through the above steps, even when a transistor is miniaturized and the channel length is reduced, the transistor 100 with high reliability which includes an oxide semiconductor and has favorable electric characteristics can be manufactured.

[0185]

The transistor 140 includes the low-concentration region 103d and the low-concentration region 103e in the oxide semiconductor layer 103. The transistor 140 can be manufactured in such a manner that a step of forming the sidewalls 111 is added to the steps of manufacturing the transistor 100 and the dopant 106 is added to the oxide semiconductor layer 103 in two steps.

[0186]

The low-concentration region 103d and the low-concentration region 103e can be formed by a self-aligned process using the gate electrode 105 as a mask. Specifically, after formation of the gate electrode 105, the dopant 106 is added to the oxide semiconductor layer 103 with the use of the gate electrode 105 as a mask by a method similar to that for forming the transistor 100 (also referred to as a first doping step). As the dopant 106 which is added to the oxide semiconductor layer 103 in the first doping step, an element similar to the dopant 106 which is used for forming the transistor 100 can be used. In the first doping step, the dopant 106 is added so that the concentration of the dopant 106 in the oxide semiconductor layer 103 is higher than or equal to 5×10^{18} atoms/cm^3 and lower than 5×10^{19} atoms/cm^3.

[0187]

Next, the sidewalls 111 are formed on the side surfaces of the gate electrode 105. The sidewalls 111 can be formed by a known method.

[0188]

Next, the dopant 106 is added to the oxide semiconductor layer 103 with the use of the gate electrode 105 and the sidewalls 111 as masks (also referred to as a second doping step). As the dopant 106 which is added to the oxide semiconductor layer 103 in the second doping step, an element similar to the dopant 106 which is used for forming the transistor 100 can be used. In the second doping step, the dopant 106 is added so that the concentration of the dopant 106 in the oxide semiconductor layer 103 is higher than or equal to 5×10^{19} atoms/cm^3 and lower than or equal to 1×10^{22} atoms/cm^3, preferably higher than or equal to 1×10^{20} atoms/cm^3 and lower than 7 at.%.

[0189]

Thus, the source region 103a, the drain region 103b, the low-concentration region 103d, and the low-concentration region 103e can be formed in the transistor 140. The low-concentration region 103d and the low-concentration region 103e have lower dopant concentration and higher resistivity than the source region 103a and the drain region 103b

[0190]

The low-concentration region 103d and the low-concentration region 103e are provided, whereby degradation of the transistor characteristics and negative shift in

threshold voltage due to a short channel effect can be reduced; thus, a highly reliable transistor can be manufactured.

[0191]

Note that the channel length of the transistor 140 is equal to the length of the channel formation region 103c which is sandwiched between the low-concentration region 103d and the low-concentration region 103e in FIG. 2B. Further, the channel length of the transistor 140 is substantially equal to the width of the gate electrode 105.

[0192]

This embodiment can be combined as appropriate with any of the other embodiments.

[0193]

(Embodiment 2)

In this embodiment, examples of transistors having structures different from those of the transistors disclosed in Embodiment 1 will be described.

[0194]

FIG. 5A is a top view illustrating a structure of a transistor 150 and FIG. 5B is a cross-sectional view illustrating a stacked structure of a portion indicated by a chain line C1-C2 in FIG. 5A. In FIG. 5A, a substrate and an insulating layer are omitted.

[0195]

In the transistor 150 illustrated in FIG. 5B, the source electrode 110a and the drain electrode 110b are stacked in a different manner from the stacking manner of the transistor 100 disclosed in Embodiment 1. In the transistor 150, the source electrode 110a and the drain electrode 110b are formed over the base layer 102, and the oxide semiconductor layer 103 is formed over the base layer 102, the source electrode 110a, and the drain electrode 110b.

[0196]

In the transistor 150, the source electrode 110a and the drain electrode 110b are connected to the source region 103a and the drain region 103b in the oxide semiconductor layer 103, respectively, without the contact holes 109; therefore, the contact area can be easily increased, and the contact resistance can be easily reduced.

[0197]

Note that the channel length of the transistor 150 is equal to the length of the

channel formation region 103c which is sandwiched between the source region 103a and the drain region 103b in FIG. 5B. Further, the channel length of the transistor 150 is substantially equal to the width of the gate electrode 105.

[0198]

A transistor 160 illustrated in FIGS. 6A and 6B includes, in addition to the structure of the transistor 150, the sidewalls 111 on the side surfaces of the gate electrode 105 and the low-concentration region 103d and the low-concentration region 103e in regions of the oxide semiconductor layer 103, which overlap with the sidewalls 111. The low-concentration region 103d is formed between the channel formation region 103c and the source region 103a, and the low-concentration region 103e is formed between the channel formation region 103c and the drain region 103b. FIG. 6A is a top view illustrating the structure of the transistor 160 and FIG. 6B is a cross-sectional view illustrating a stacked structure of a portion indicated by a chain line D1-D2 in FIG. 6A.

[0199]

The low-concentration region 103d or the low-concentration region 103e is provided in the oxide semiconductor layer 103, whereby an electric field generated between the channel formation region 103c and the source region 103a or between the channel formation region 103c and the drain region 103b can be relieved; thus, degradation of transistor characteristics can be reduced. In particular, relieving an electric field generated between the channel formation region 103c and the drain region 103b is effective for reduction of degradation of transistor characteristics. Further, the low-concentration region 103d or the low-concentration region 103e is provided, so that a short channel effect due to miniaturization of a transistor can be suppressed.

[0200]

Note that the channel length of the transistor 160 is equal to the length of the channel formation region 103c which is sandwiched between the low-concentration region 103d and the low-concentration region 103e in FIG. 6B. Further, the channel length of the transistor 160 is substantially equal to the width of the gate electrode 105.

[0201]

A transistor 170 illustrated in FIG. 7A is one mode of a bottom-gate transistor.

[0202]

FIG. 7A illustrates a cross-sectional structure of the transistor 170. In the transistor 170, the gate electrode 105 is formed over the substrate 101 and the gate insulating layer 104 is formed over the gate electrode 105. The gate electrode 105 has a structure in which the gate electrode 105a is stacked over the gate electrode 105b. The base layer described in Embodiment 1 may be provided between the substrate 101 and the gate electrode 105.

[0203]

Further, the oxide semiconductor layer 103 is formed over the gate insulating layer 104 and a channel protective layer 112, the source electrode 110a, and the drain electrode 110b are formed over the oxide semiconductor layer 103. The oxide semiconductor layer 103 includes the channel formation region 103c which overlaps with the channel protective layer 112, the source region 103a which is electrically connected to the source electrode 110a, and the drain region 103b which is electrically connected to the drain electrode 110b.

[0204]

The channel protective layer 112 can be formed using a material and a method similar to those of the gate insulating layer 104. The thickness of the channel protective layer 112 is greater than or equal to 10 nm and less than or equal to 500 nm, preferably greater than or equal to 100 nm and less than or equal to 300 nm.

[0205]

The source region 103a and the drain region 103b of the transistor 170 can be formed using the channel protective layer 112 as a mask in a manner similar to the source region 103a and the drain region 103b of the transistor 100.

[0206]

Further, the insulating layer 108 is formed over the channel protective layer 112, the source electrode 110a, and the drain electrode 110b. The insulating layer 108 may be a stack of a plurality of insulating layers.

[0207]

Note that the channel length of the transistor 170 is equal to the length of the channel formation region 103c which is sandwiched between the source region 103a and the drain region 103b in FIG. 7A. Further, the channel length of the transistor 170 is substantially equal to the width of the channel protective layer 112.

[0208]

FIG. 7B illustrates a cross-sectional structure of a transistor 180. The transistor 180 has a structure in which a back gate electrode 115 and an insulating layer 113 are provided to the transistor 100. In the transistor 180, the back gate electrode 115 is formed over the base layer 102 and the insulating layer 113 is formed over the back gate electrode 115. The oxide semiconductor layer 103 of the transistor 180 overlaps with the back gate electrode 115 with the insulating layer 113 interposed therebetween.

[0209]

The back gate electrode 115 is positioned so that the channel formation region 103c of the oxide semiconductor layer 103 is interposed between the gate electrode 105 and the back gate electrode 115. The back gate electrode 115 is formed using a conductive layer and can function in a manner similar to that of the gate electrode 105. By changing a potential of the back gate electrode 115, the threshold voltage of the transistor can be changed.

[0210]

The back gate electrode 115 can be formed using a material and a method similar to those of the gate electrode 105b. A layer similar to the gate electrode 105a may be formed between the back gate electrode 115 and the insulating layer 113.

[0211]

The insulating layer 113 can be formed using a material and a method similar to those of the gate insulating layer 104. Formation of the base layer 102 can be omitted when the insulating layer 113 also serves as the base layer 102.

[0212]

Note that the channel length of the transistor 180 is equal to the length of the channel formation region 103c which is sandwiched between the source region 103a and the drain region 103b in FIG. 7B. Further, the channel length of the transistor 180 is substantially equal to the width of the gate electrode 105.

[0213]

This embodiment can be combined as appropriate with any of the other embodiments.

[0214]

(Embodiment 3)

In this embodiment, a method for forming an oxide semiconductor film including CAAC-OS, which is different from the method disclosed in Embodiment 1, will be described.

[0215]

First, an oxide semiconductor film having a thickness of greater than or equal to 1 nm and less than or equal to 50 nm is formed over the base layer 102.

[0216]

The substrate temperature in deposition is set to higher than or equal to 150 °C and lower than or equal to 450 °C, preferably higher than or equal to 200 °C and lower than or equal to 350 °C. The deposition is performed while the substrate is heated to higher than or equal to 150 °C and lower than or equal to 450 °C, preferably higher than or equal to 200 °C and lower than or equal to 350 °C, whereby moisture (including hydrogen) or the like is prevented from entering a film. Further, CAAC-OS which is an oxide semiconductor layer having crystallinity can be formed.

[0217]

Further, it is preferable that hydrogen be further released from the oxide semiconductor and part of oxygen contained in the base layer 102 be diffused into the oxide semiconductor and the vicinity of the interface of the oxide semiconductor in the base layer 102 by performing heat treatment on the substrate 101 after formation of the oxide semiconductor. An oxide semiconductor including more highly crystalline CAAC-OS can be formed by the heat treatment.

[0218]

The temperature of the heat treatment is preferably a temperature at which hydrogen is released from the oxide semiconductor and part of oxygen contained in the base layer 102 is released and diffused into the oxide semiconductor. The temperature is typically higher than or equal to 200 °C and lower than the strain point of the substrate 101, preferably higher than or equal to 250 °C and lower than or equal to 450 °C. By diffusion of oxygen into the oxide semiconductor, oxygen deficiency in the oxide semiconductor can be reduced.

[0219]

A rapid thermal annealing (RTA) apparatus can be used in the heat treatment. With the use of the RTA apparatus, heat treatment can be performed at a temperature of higher than or equal to the strain point of a substrate if the heating time is short. Therefore, time for forming an oxide semiconductor in which the proportion of a crystalline region is higher than that of an amorphous region can be shortened.

[0220]

The heat treatment can be performed in an inert gas atmosphere; typically the heat treatment is preferably performed in a rare gas (such as helium, neon, argon, xenon, or krypton) atmosphere or a nitrogen atmosphere. Alternatively, the heat treatment may be performed in an oxygen atmosphere or a reduced-pressure atmosphere. The treatment time is from three minutes to 24 hours. The proportion of a crystalline region to an amorphous region in the oxide semiconductor can be increased as the treatment time is increased. However, heat treatment for longer than 24 hours is not preferable because the productivity is reduced.

[0221]

Through the above steps, an oxide semiconductor including CAAC-OS can be formed.

[0222]

This embodiment can be combined as appropriate with any of the other embodiments.

[0223]

(Embodiment 4)

In this embodiment, an influence on the electric characteristics of the transistor including an oxide semiconductor described in Embodiment 1 or 2 will be described with reference to band diagrams.

[0224]

FIG. 8 is a cross-sectional view of a transistor having a stacked structure similar to the transistor 100 illustrated in FIGS. 1A and 1B. FIGS. 9A and 9B are energy band diagrams (schematic diagrams) of cross section X1-X2 in FIG. 8. FIG. 9B shows the case where a voltage of a source and a voltage of a drain are equal to each other (VD = 0V). FIG. 8 illustrates the transistor provided with an oxide semiconductor layer including a first oxide semiconductor region (OS1) and a pair of second oxide

semiconductor regions (OS2) and source and drain electrodes (metal).

[0225]

In FIG. 8, a channel formation region of the transistor is formed using OS1. OS1 is an oxide semiconductor which is made to be intrinsic (i-type) or as close to intrinsic as possible by highly purifying the film through removal or elimination of impurities such as moisture (including hydrogen) as much as possible and further by reducing oxygen deficiency in the film. Thus, the Fermi level (Ef) can be the same as the intrinsic Fermi level (Ei).

[0226]

In addition, in FIG. 8, a source region and a drain region of the transistor are formed using the pair of OS2. OS2 is formed in such a manner that an oxide semiconductor is made to be intrinsic (i-type) or as close to intrinsic as possible as in the case of OS1 by highly purifying the film through removal or elimination of impurities such as moisture (including hydrogen) as much as possible and further by reducing oxygen deficiency in the film, and after that, at least one element selected from Group 15 elements such as nitrogen, phosphorus, or arsenic is added. OS2 has thus higher carrier density than OS1 and the position of its Fermi level is close to the conduction band.

[0227]

FIG. 9A shows a relation of band structures of the vacuum level (Evac), the first oxide semiconductor region (OS1), the second oxide semiconductor region (OS2), and the source and drain electrodes (metal). Here, IP represents the ionization potential; Ea, the electron affinity; Eg, the energy gap; and Wf, the work function. In addition, Ec represents the bottom of the conduction band; Ev, the top of the valence band; and Ef, the Fermi level. As for a sign at the end of each symbol, 1 denotes OS1; 2, OS2; and m, metal. Here, a metal material having Wf_m of 4.1 eV (such as titanium) is assumed as the metal.

[0228]

OS1 is an oxide semiconductor which is made to be i-type or as close to i-type as possible and thus has extremely low carrier density; therefore, Ef_1 is around the middle point between Ec and Ev. OS2 is an n-type oxide semiconductor having high

carrier density, and thus Ec_2 substantially corresponds to Ef_2.

[0229]

It is said that the energy gap (Eg) of the oxide semiconductor denoted by OS1 is 3.15 eV and the electron affinity (Ea) thereof is 4.3 eV. The energy gap (Eg) of the oxide semiconductor denoted by OS2 can be smaller than 3.15 in accordance with the amount of a dopant to be added. In that case, ionization potential hardly changes, so that an electron affinity and work function become large. FIGS. 9A and 9B illustrate the case where Eg of OS2 is smaller than Eg of OS1 (i.e., Eg_1 > Eg_2).

[0230]

As shown in FIG. 9B, in the case where OS1 that is the channel formation region and OS2 that is the source or drain region are in contact with each other, transfer of carriers occurs so that the Fermi levels can be equal to each other; thus, the band edge of OS1 curves. Further, in the case where OS2 is in contact with the metal that is the source or drain electrode, transfer of carriers occurs so that the Fermi levels can be equal to each other; thus, the band edge of OS2 curves.

[0231]

By forming OS2 that is an n-type oxide semiconductor between OS1 that is the channel formation region and the metal that is the source or drain electrode, contact between the oxide semiconductor and the metal can be an ohmic junction, and contact resistance can be reduced. As a result, the on-state current of the transistor can be increased. Moreover, the curve at the band edge of OS1 can be suppressed, and thus a short-channel effect of the transistor can be suppressed.

[0232]

This embodiment can be combined as appropriate with any of the other embodiments.

[0233]

(Embodiment 5)

An example of a circuit diagram of a memory element (hereinafter also referred to as a memory cell) included in a semiconductor device is illustrated in FIG. 10A. The memory cell includes a transistor 1160 in which a channel formation region is formed using a material other than an oxide semiconductor and a transistor 1162 in

which a channel formation region is formed using an oxide semiconductor.

[0234]

The transistor 1162 in which the channel formation region is formed using an oxide semiconductor can be manufactured in accordance with Embodiment 1.

[0235]

As illustrated in FIG. 10A, a gate electrode of the transistor 1160 is electrically connected to one of a source electrode and a drain electrode of the transistor 1162. A first wiring (a 1st line, also referred to as a source line) is electrically connected to a source electrode of the transistor 1160. A second wiring (a 2nd line, also referred to as a bit line) is electrically connected to a drain electrode of the transistor 1160. A third wiring (a 3rd line, also referred to as a first signal line) is electrically connected to the other of the source electrode and the drain electrode of the transistor 1162. A fourth wiring (a 4th line, also referred to as a second signal line) is electrically connected to a gate electrode of the transistor 1162.

[0236]

The transistor 1160 in which the channel formation region is formed using a material other than an oxide semiconductor, e.g., single crystal silicon can operate at sufficiently high speed. Therefore, with the use of the transistor 1160, high-speed reading of stored contents and the like are possible. The transistor 1162 in which the channel formation region is formed using an oxide semiconductor is characterized by its off-state current which is smaller than the off-state current of the transistor 1160. Therefore, when the transistor 1162 is turned off, a potential of the gate electrode of the transistor 1160 can be held for a very long time.

[0237]

By utilizing a characteristic in which the potential of the gate electrode of the transistor 1160 can be held, writing, holding, and reading of data are possible as described below.

[0238]

First, writing and holding of data are described. First, a potential of the fourth wiring is set to a potential at which the transistor 1162 is turned on, so that the transistor 1162 is turned on. Thus, a potential of the third wiring is supplied to the gate electrode of the transistor 1160 (writing). After that, the potential of the fourth wiring is set to a

potential at which the transistor 1162 is turned off, so that the transistor 1162 is turned off, and thus, the potential of the gate electrode of the transistor 1160 is held (holding).

[0239]

Since the off-state current of the transistor 1162 is smaller than the off-state current of the transistor 1160, the potential of the gate electrode of the transistor 1160 is held for a long time. For example, when the potential of the gate electrode of the transistor 1160 is a potential at which the transistor 1160 is in an on state, the on state of the transistor 1160 is held for a long time. In addition, when the potential of the gate electrode of the transistor 1160 is a potential at which the transistor 1160 is an off state, the off state of the transistor 1160 is held for a long time.

[0240]

Then, reading of data is described. When a predetermined potential (a low potential) is supplied to the first wiring in a state where the on state or the off state of the transistor 1160 is held as described above, a potential of the second wiring varies depending on the on state or the off state of the transistor 1160. For example, when the transistor 1160 is in the on state, the potential of the second wiring becomes lower than the potential of the first wiring. On the other hand, when the transistor 1160 is in the off state, the potential of the second wiring does not vary.

[0241]

In such a manner, the potential of the second wiring and a predetermined potential are compared with each other in a state where data is held, whereby the data can be read out.

[0242]

Then, rewriting of data is described. Rewriting of data is performed in a manner similar to that of the writing and holding of data. That is, a potential of the fourth wiring is set to a potential at which the transistor 1162 is turned on, so that the transistor 1162 is turned on. Thus, a potential of the third wiring (a potential for new data) is supplied to the gate electrode of the transistor 1160. After that, the potential of the fourth wiring is set to be a potential at which the transistor 1162 is turned off, so that the transistor 1162 is turned off, and thus, the new data is held.

[0243]

In the memory cell according to the disclosed invention, data can be directly

rewritten by another writing of data as described above. For that reason, erasing operation which is necessary for a flash memory or the like is not needed, so that a reduction in operation speed because of erasing operation can be suppressed. In other words, high-speed operation of the semiconductor device including the memory cell can be realized.

[0244]

FIG. 10B is a circuit diagram illustrating an application example of the memory cell illustrated in FIG. 10A.

[0245]

A memory cell 1100 illustrated in FIG. 10B includes a first wiring SL (a source line), a second wiring BL (a bit line), a third wiring S1 (a first signal line), a fourth wiring S2 (a second signal line), a fifth wiring WL (a word line), a transistor 1164 (a first transistor), a transistor 1161 (a second transistor), and a transistor 1163 (a third transistor). In each of the transistors 1164 and 1163, a channel formation region is formed using a material other than an oxide semiconductor, and in the transistor 1161, a channel formation region is formed using an oxide semiconductor.

[0246]

Here, a gate electrode of the transistor 1164 is electrically connected to one of a source electrode and a drain electrode of the transistor 1161. In addition, the first wiring SL is electrically connected to a source electrode of the transistor 1164, and a drain electrode of the transistor 1164 is electrically connected to a source electrode of the transistor 1163. The second wiring BL is electrically connected to a drain electrode of the transistor 1163, and the third wiring S1 is electrically connected to the other of the source electrode and the drain electrode of the transistor 1161. The fourth wiring S2 is electrically connected to a gate electrode of the transistor 1161, and the fifth wiring WL is electrically connected to a gate electrode of the transistor 1163.

[0247]

Next, operation of the circuit is specifically described.

[0248]

When data is written into the memory cell 1100, the first wiring SL is set to 0 V, the fifth wiring WL is set to 0 V, the second wiring BL is set to 0 V, and the fourth wiring S2 is set to 2 V. The third wiring S1 is set to 2 V in order to write data "1" and

set to 0 V in order to write data "0". At this time, the transistor 1163 is in an off state and the transistor 1161 is in an on state. Note that at the end of the writing, before the potential of the third wiring S1 is changed, the fourth wiring S2 is set to 0 V so that the transistor 1161 is turned off.

[0249]

As a result, a potential of a node (referred to as a node A) connected to the gate electrode of the transistor 1164 is set to approximately 2 V after the writing of the data "1" and set to approximately 0 V after the writing of the data "0". Electric charge corresponding to a potential of the third wiring S1 is accumulated at the node A; since the off-state current of the transistor 1161 is smaller than that of a transistor in which a channel formation region is formed using single crystal silicon, the potential of the gate electrode of the transistor 1164 is held for a long time.

[0250]

When data is read from the memory cell, the first wiring SL is set to 0 V, the fifth wiring WL is set to 2 V, the fourth wiring S2 and the third wiring S1 are set to 0 V, and a reading circuit connected to the second wiring BL is set in an operation state. At this time, the transistor 1163 is in an on state and the transistor 1161 is in an off state.

[0251]

The transistor 1164 is in an off state when the data "0" has been written, that is, the node A is set to approximately 0 V, so that the resistance between the second wiring BL and the first wiring SL is high. On the other hand, the transistor 1164 is in an on state when the data "1" has been written, that is, the node A is set to approximately 2 V, so that the resistance between the second wiring BL and the first wiring SL is low. A reading circuit can read the data "0" or the data "1" in accordance with the difference in resistance state of the memory cell. The second wiring BL at the time of the writing is set to 0 V; however, it may be in a floating state or may be charged to have a potential higher than 0 V. The third wiring S1 at the time of the reading is set to 0 V; however, it may be in a floating state or may be charged to have a potential higher than 0 V.

[0252]

Note that the data "1" and the data "0" are defined for convenience and can be reversed. In addition, the above operation voltages are examples. The operation voltages are set so that the transistor 1164 is turned off in the case of data "0" and turned

on in the case of data "1", the transistor 1161 is turned on at the time of writing and turned off in periods except the time of writing, and the transistor 1163 is turned on at the time of reading. In particular, a power supply potential VDD of a peripheral logic circuit may also be used instead of 2 V.

[0253]

In this embodiment, the memory cell with a minimum storage unit (one bit) is described for easy understanding; however, the structure of the memory cell is not limited thereto. It is also possible to make a more developed semiconductor device with a plurality of memory cells connected to each other as appropriate. For example, it is possible to make a NAND-type or NOR-type semiconductor device by using more than one of the above memory cells. The wiring structure is not limited to that in FIG. 10A or 10B and can be changed as appropriate.

[0254]

FIG. 11 is a block circuit diagram of a semiconductor device according to an embodiment of the present invention. The semiconductor device includes $m \times n$ bits of memory capacitance.

[0255]

The semiconductor device illustrated in FIG. 11 includes m fourth wirings, m fifth wirings, n second wirings, n third wirings, a memory cell array 1110 in which a plurality of memory cells 1100(1,1) to 1100(m,n) is arranged in a matrix of m rows by n columns (m and n are each a natural number), and peripheral circuits such as a wiring driver circuit 1111 for driving the second wirings and the third wirings, a wiring driver circuit 1113 for driving the fourth wirings and the fifth wirings, and a reading circuit 1112. A refresh circuit or the like may be provided as another peripheral circuit.

[0256]

A memory cell 1100(i,j) is considered as a typical example of the memory cell. Here, the memory cell 1100(i,j) (i is an integer of greater than or equal to 1 and less than or equal to m and j is an integer of greater than or equal to 1 and less than or equal to n) is connected to a second wiring BL(j), a third wiring S1(j), a fourth wiring S2(i), a fifth wiring WL(i), and a first wiring. A first wiring potential Vs is supplied to the first wiring. The second wirings BL(1) to BL(n) and the third wirings S1(1) to S1(n) are connected to the wiring driver circuit 1111 for driving the second wirings and the third

wirings and the reading circuit 1112. The fifth wirings WL(1) to WL(*m*) and the fourth wirings S2(1) to S2(*m*) are connected to the wiring driver circuit 1113 for driving the fourth wirings and the fifth wirings.

[0257]

The operation of the semiconductor device illustrated in FIG. 11 is described. In this structure, data is written and read per row.

[0258]

When data is written into memory cells 1100(*i*,1) to 1100(*i*,*n*) of an *i*-th row, the first wiring potential Vs is set to 0 V, a fifth wiring WL(*i*) and the second wirings BL(1) to BL(*n*) are set to 0 V, and a fourth wiring S2(*i*) is set to 2 V. At this time, the transistors 1161 are turned on. Among the third wirings S1(1) to S1(*n*), the third wiring in a column in which data "1" is to be written is set to 2 V and the third wiring in a column in which data "0" is to be written is set to 0 V. Note that, to finish writing, the fourth wiring S2(*i*) is set to 0 V before the potentials of the third wirings S1(1) to S1(*n*) are changed, so that the transistors 1161 are turned off. Moreover, a non-selected fifth wiring WL and a non-selected fourth wiring S2 are set to 0 V.

[0259]

As a result, the potential of the node (referred to as the node A) connected to the gate electrode of the transistor 1164 in the memory cell into which data "1" has been written is set to approximately 2 V, and the potential of the node A in the memory cell into which data "0" has been written is set to approximately 0 V. The potential of the node A of the non-selected memory cell is not changed.

[0260]

When data is read from the memory cells 1100(*i*,1) to 1100(*i*,*n*) of the *i*-th row, the first wiring potential Vs is set to 0 V, the fifth wiring WL(*i*) is set to 2 V, the fourth wiring S2(*i*) and the third wirings S1(1) to S1(*n*) are set to 0 V, and the reading circuit connected to the second wirings BL(1) to BL(*n*) is set in an operation state. The reading circuit can read data "0" or data "1" in accordance with the difference in resistance state of the memory cell, for example. Note that the non-selected fifth wiring WL and the non-selected fourth wiring S2 are set to 0 V. The second wiring BL at the time of the writing is set to 0 V; however, it may be in a floating state or may be charged to have a potential higher than 0 V. The third wiring S1 at the time of the

reading is set to 0 V; however, it may be in a floating state or may be charged to have a potential higher than 0 V.

[0261]

Note that the data "1" and the data "0" are defined for convenience and can be reversed. In addition, the above operation voltages are examples. The operation voltages are set so that the transistor 1164 is turned off in the case of data "0" and turned on in the case of data "1", the transistor 1161 is turned on at the time of writing and turned off in periods except the time of writing, and the transistor 1163 is turned on at the time of reading. A power supply potential VDD of a peripheral logic circuit may also be used instead of 2 V.

[0262]

This embodiment can be combined as appropriate with any of the other embodiments.

[0263]

(Embodiment 6)

In this embodiment, an example of a circuit diagram of a memory cell including a capacitor will be shown. A memory cell 1170 illustrated in FIG. 12A includes a first wiring SL, a second wiring BL, a third wiring S1, a fourth wiring S2, a fifth wiring WL, a transistor 1171 (a first transistor), a transistor 1172 (a second transistor), and a capacitor 1173. In the transistor 1171, a channel formation region is formed using a material other than an oxide semiconductor, and in the transistor 1172, a channel formation region is formed using an oxide semiconductor.

[0264]

Here, a gate electrode of the transistor 1171, one of a source electrode and a drain electrode of the transistor 1172, and one electrode of the capacitor 1173 are electrically connected to each other. In addition, the first wiring SL and a source electrode of the transistor 1171 are electrically connected to each other. The second wiring BL and a drain electrode of the transistor 1171 are electrically connected to each other. The third wiring S1 and the other of the source electrode and the drain electrode of the transistor 1172 are electrically connected to each other. The fourth wiring S2 and a gate electrode of the transistor 1172 are electrically connected to each other. The fifth wiring WL and the other electrode of the capacitor 1173 are electrically connected

to each other.

[0265]

Next, operation of the circuit will be specifically described.

[0266]

When data is written into the memory cell 1170, the first wiring SL is set to 0 V, the fifth wiring WL is set to 0 V, the second wiring BL is set to 0 V, and the fourth wiring S2 is set to 2 V. The third wiring S1 is set to 2 V in order to write data "1" and set to 0 V in order to write data "0". At this time, the transistor 1172 is turned on. Note that, to finish writing, the fourth wiring S2 is supplied with 0 V before the potential of the third wiring S1 is changed, so that the transistor 1172 is turned off.

[0267]

As a result, the potential of a node (referred to as a node A) connected to the gate electrode of the transistor 1171 is set to approximately 2 V after the writing of data "1" and is set to approximately 0 V after the writing of data "0".

[0268]

When data is read from the memory cell 1170, the first wiring SL is set to 0 V, the fifth wiring WL is set to 2 V, the fourth wiring S2 is set to 0 V, the third wiring S1 is set to 0 V, and a reading circuit connected to the second wiring BL is operated. At this time, the transistor 1172 is turned off.

[0269]

The state of the transistor 1171 in the case where the fifth wiring WL is set to 2 V will be described. The potential of the node A which determines the state of the transistor 1171 depends on capacitance C1 between the fifth wiring WL and the node A, and capacitance C2 between the gate electrode of the transistor 1171 and the source and drain electrodes of the transistor 1171.

[0270]

Note that the third wiring S1 at the time of reading is set to 0 V; however, the third wiring S1 may be in a floating state or may be charged to have a potential higher than 0 V. Data "1" and data "0" are defined for convenience and may be reversed.

[0271]

The potential of the third wiring S1 at the time of writing may be selected from the potentials of data "0" and data "1" as long as the transistor 1172 is turned off after

the writing and the transistor 1171 is off in the case where the potential of the fifth wiring WL is set to 0 V. The potential of the fifth wiring WL at the time of reading may be selected so that the transistor 1171 is turned off in the case where data "0" has been written and is turned on in the case where data "1" has been written. Furthermore, the threshold voltage of the transistor 1171 is an example. The transistor 1171 can have any threshold voltage as long as the transistor 1171 operates in the above-described manner.

[0272]

An example of a NOR semiconductor memory device in which a memory cell including a capacitor and a selection transistor having a first gate electrode and a second gate electrode is used will be described with reference to FIG. 12B.

[0273]

A semiconductor device illustrated in FIG. 12B according to an embodiment of the present invention includes a memory cell array including a plurality of memory cells arranged in a matrix of *I* rows (*I* is a natural number of 2 or more) and *J* columns (*J* is a natural number).

[0274]

The memory cell array illustrated in FIG. 12B includes a plurality of memory cells 1180 arranged in a matrix of *i* rows (*i* is a natural number of 3 or more) and *j* columns (*j* is a natural number of 3 or more), *i* word lines WL (word lines WL_1 to WL_*i*), *i* capacitor lines CL (capacitor lines CL_1 to CL_*i*), *i* gate lines BGL (gate lines BGL_1 to BGL_*i*), *j* bit lines BL (bit lines BL_1 to BL_*j*), and a source line SL.

[0275]

Further, each of the plurality of memory cells 1180 (also referred to as a memory cell 1180(*M*,*N*) (note that *M* is a natural number greater than or equal to 1 and less than or equal to *i* and that *N* is a natural number greater than or equal to 1 and less than or equal to *j*)) includes a transistor 1181(*M*,*N*), a capacitor 1183(*M*,*N*), and a transistor 1182(*M*,*N*).

[0276]

Note that in the semiconductor memory device, the capacitor includes a first capacitor electrode, a second capacitor electrode, and a dielectric layer overlapping with

the first capacitor electrode and the second capacitor electrode. Electric charge is accumulated in the capacitor in accordance with voltage applied between the first capacitor electrode and the second capacitor electrode.

[0277]

The transistor 1181(*M*,*N*) is an n-channel transistor which has a source electrode, a drain electrode, a first gate electrode, and a second gate electrode. Note that in the semiconductor memory device in this embodiment, the transistor 1181 does not necessarily need to be an n-channel transistor.

[0278]

One of the source electrode and the drain electrode of the transistor 1181(*M*,*N*) is connected to the bit line BL_*N*. The first gate electrode of the transistor 1181(*M*,*N*) is connected to the word line WL_*M*. The second gate electrode of the transistor 1181(*M*,*N*) is connected to the gate line BGL_*M*. With the structure in which the one of the source electrode and the drain electrode of the transistor 1181(*M*,*N*) is connected to the bit line BL_*N*, data can be selectively read from memory cells.

[0279]

The transistor 1181(*M*,*N*) serves as a selection transistor in the memory cell 1180(*M*,*N*).

[0280]

As the transistor 1181(*M*,*N*), a transistor in which a channel formation region is formed using an oxide semiconductor can be used.

[0281]

The transistor 1182(*M*,*N*) is a p-channel transistor. Note that in the semiconductor memory device in this embodiment, the transistor 1182 does not necessarily need to be a p-channel transistor.

[0282]

One of a source electrode and a drain electrode of the transistor 1182(*M*,*N*) is connected to the source line SL. The other of the source electrode and the drain electrode of the transistor 1182(*M*,*N*) is connected to the bit line BL_*N*. A gate electrode of the transistor 1182(*M*,*N*) is connected to the other of the source electrode and the drain electrode of the transistor 1181(*M*,*N*).

[0283]

The transistor 1182(*M*,*N*) serves as an output transistor in the memory cell 1180(*M*,*N*). As the transistor 1182(*M*,*N*), for example, a transistor in which a channel formation region is formed using single crystal silicon can be used.

[0284]

A first capacitor electrode of the capacitor 1183(*M*,*N*) is connected to the capacitor line CL_*M*. A second capacitor electrode of the capacitor 1183(*M*,*N*) is connected to the other of the source electrode and the drain electrode of the transistor 1181(*M*,*N*). Note that the capacitor 1183(*M*,*N*) serves as a storage capacitor.

[0285]

The voltage of the word lines WL_1 to WL_*i* is controlled by, for example, a driver circuit including a decoder.

[0286]

The voltage of the bit lines BL_1 to BL_*j* is controlled by, for example, a driver circuit including a decoder.

[0287]

The voltage of the capacitor lines CL_1 to CL_*i* is controlled by, for example, a driver circuit including a decoder.

[0288]

The voltage of the gate lines BGL_1 to BGL_*i* is controlled by, for example, a gate line driver circuit.

[0289]

The gate line driver circuit is formed using a circuit which includes a diode and a capacitor whose first capacitor electrode is electrically connected to an anode of the diode and the gate line BGL, for example.

[0290]

By adjustment of the voltage of the second gate electrode of the transistor 1181, the threshold voltage of the transistor 1181 can be adjusted. Accordingly, by adjustment of the threshold voltage of the transistor 1181 functioning as a selection transistor, current flowing between the source electrode and the drain electrode of the transistor 1181 in an off state can be extremely small. Thus, a data retention period in

the memory circuit can be longer. In addition, voltage necessary for writing and reading data can be made lower than that of a conventional semiconductor device; thus, power consumption can be reduced.

[0291]

This embodiment can be combined as appropriate with any of the other embodiments.

[0292]

(Embodiment 7)

In this embodiment, examples of a semiconductor device using the transistor described in any of the above embodiments will be described with reference to FIGS. 13A and 13B.

[0293]

FIG. 13A illustrates an example of a semiconductor device whose structure corresponds to that of a so-called dynamic random access memory (DRAM). A memory cell array 1120 illustrated in FIG. 13A has a structure in which a plurality of memory cells 1130 is arranged in a matrix. Further, the memory cell array 1120 includes *m* first wirings and *n* second wirings. Note that in this embodiment, the first wiring and the second wiring are referred to as a bit line BL and a word line WL, respectively.

[0294]

The memory cell 1130 includes a transistor 1131 and a capacitor 1132. A gate electrode of the transistor 1131 is connected to the first wiring (the word line WL). Further, one of a source electrode and a drain electrode of the transistor 1131 is connected to the second wiring (the bit line BL). The other of the source electrode and the drain electrode of the transistor 1131 is connected to one electrode of the capacitor. The other electrode of the capacitor is connected to a capacitor line CL and is supplied with a predetermined potential. The transistor described in any of the above embodiments is applied to the transistor 1131.

[0295]

The transistor in which a channel formation region is formed using an oxide semiconductor, which is described in any of the above embodiments, is characterized by having smaller off-state current than a transistor in which a channel formation region is

formed using single crystal silicon. Accordingly, when the transistor is applied to the semiconductor device illustrated in FIG. 13A, which is regarded as a so-called DRAM, a substantially nonvolatile memory can be obtained.

[0296]

FIG. 13B illustrates an example of a semiconductor device whose structure corresponds to that of a so-called static random access memory (SRAM). A memory cell array 1140 illustrated in FIG. 13B can have a structure in which a plurality of memory cells 1150 is arranged in a matrix. Further, the memory cell array 1140 includes a plurality of first wirings (word lines WL), a plurality of second wirings (bit lines BL), and a plurality of third wirings (inverted bit lines /BL).

[0297]

The memory cell 1150 includes a first transistor 1151, a second transistor 1152, a third transistor 1153, a fourth transistor 1154, a fifth transistor 1155, and a sixth transistor 1156. The first transistor 1151 and the second transistor 1152 function as selection transistors. One of the third transistor 1153 and the fourth transistor 1154 is an n-channel transistor (here, the fourth transistor 1154 is an n-channel transistor), and the other of the third transistor 1153 and the fourth transistor 1154 is a p-channel transistor (here, the third transistor 1153 is a p-channel transistor). In other words, the third transistor 1153 and the fourth transistor 1154 form a CMOS circuit. Similarly, the fifth transistor 1155 and the sixth transistor 1156 form a CMOS circuit.

[0298]

The first transistor 1151, the second transistor 1152, the fourth transistor 1154, and the sixth transistor 1156 are n-channel transistors and the transistor described in any of the above embodiments can be applied to these transistors. Each of the third transistor 1153 and the fifth transistor 1155 is a p-channel transistor in which a channel formation region is formed using a material (e.g., single crystal silicon) other than an oxide semiconductor.

[0299]

The methods and structures described in this embodiment can be combined as appropriate with any of the structures and the methods described in the other embodiments.

[0300]

This embodiment can be combined as appropriate with any of the other embodiments.

[0301]

(Embodiment 8)

A central processing unit (CPU) can be formed using a transistor including an oxide semiconductor in a channel formation region for at least part of the CPU.

[0302]

FIG. 14A is a block diagram illustrating a specific structure of a CPU. The CPU illustrated in FIG. 14A includes an arithmetic logic unit (ALU) 1191, an ALU controller 1192, an instruction decoder 1193, an interrupt controller 1194, a timing controller 1195, a register 1196, a register controller 1197, a bus interface (Bus I/F) 1198, a rewritable ROM 1199, and an ROM interface (ROM I/F) 1189 over a substrate 1190. A semiconductor substrate, an SOI substrate, a glass substrate, or the like is used as the substrate 1190. The ROM 1199 and the ROM interface 1189 may be provided over a separate chip. Obviously, the CPU illustrated in FIG. 14A is only an example in which the configuration is simplified, and an actual CPU may have various configurations depending on the application.

[0303]

An instruction that is input to the CPU through the bus interface 1198 is input to the instruction decoder 1193 and decoded therein, and then, input to the ALU controller 1192, the interrupt controller 1194, the register controller 1197, and the timing controller 1195.

[0304]

The ALU controller 1192, the interrupt controller 1194, the register controller 1197, and the timing controller 1195 conduct various controls in accordance with the decoded instruction. Specifically, the ALU controller 1192 generates signals for controlling the operation of the ALU 1191. While the CPU is executing a program, the interrupt controller 1194 judges an interrupt request from an external input/output device or a peripheral circuit on the basis of its priority or a mask state, and processes the request. The register controller 1197 generates an address of the register 1196, and reads/writes data from/to the register 1196 in accordance with the state of the CPU.

[0305]

The timing controller 1195 generates signals for controlling operation timings of the ALU 1191, the ALU controller 1192, the instruction decoder 1193, the interrupt controller 1194, and the register controller 1197. For example, the timing controller 1195 includes an internal clock generator for generating an internal clock signal CLK2 based on a reference clock signal CLK1, and supplies the clock signal CLK2 to the above circuits.

[0306]

In the CPU illustrated in FIG. 14A, a memory element is provided in the register 1196. The memory element described in Embodiment 5 can be used as the memory element provided in the register 1196.

[0307]

In the CPU illustrated in FIG. 14A, the register controller 1197 selects an operation of holding data in the register 1196 in accordance with an instruction from the ALU 1191. That is, the register controller 1197 selects whether data is held by a phase-inversion element or a capacitor in the memory element included in the register 1196. When data holding by the phase-inversion element is selected, power supply voltage is supplied to the memory element in the register 1196. When data holding by the capacitor is selected, the data is rewritten in the capacitor, and supply of power supply voltage to the memory element in the register 1196 can be stopped.

[0308]

The power supply can be stopped by providing a switching element between a memory element group and a node to which a power supply potential VDD or a power supply potential VSS is supplied, as illustrated in FIG. 14B or FIG. 14C. Circuits illustrated in FIGS. 14B and 14C are described below.

[0309]

FIGS. 14B and 14C each illustrate an example of a configuration of a memory circuit including a transistor including an oxide semiconductor in a channel formation region as a switching element for controlling supply of a power supply potential to a memory element.

[0310]

The memory device illustrated in FIG. 14B includes a switching element 1141 and a memory element group 1143 including a plurality of memory elements 1142.

Specifically, as each of the memory elements 1142, the memory element described in Embodiment 5 can be used. Each of the memory elements 1142 included in the memory element group 1143 is supplied with the high-level power supply potential VDD via the switching element 1141. Further, each of the memory elements 1142 included in the memory element group 1143 is supplied with a potential of a signal IN and the low-level power supply potential VSS.

[0311]

In FIG. 14B, a transistor including an oxide semiconductor in a channel formation region is used for the switching element 1141, and the switching of the transistor is controlled by a signal Sig A supplied to a gate electrode thereof.

[0312]

Note that FIG. 14B illustrates the configuration in which the switching element 1141 includes only one transistor; however, without limitation thereto, the switching element 1141 may include a plurality of transistors. In the case where the switching element 1141 includes a plurality of transistors which serve as switching elements, the plurality of transistors may be connected to each other in parallel, in series, or in combination of parallel connection and series connection.

[0313]

Although the switching element 1141 controls the supply of the high-level power supply potential VDD to each of the memory elements 1142 included in the memory element group 1143 in FIG. 14B, the switching element 1141 may control the supply of the low-level power supply potential VSS.

[0314]

In FIG. 14C, an example of a memory device in which each of the memory elements 1142 included in the memory element group 1143 is supplied with the low-level power supply potential VSS via the switching element 1141 is illustrated. The supply of the low-level power supply potential VSS to each of the memory elements 1142 included in the memory element group 1143 can be controlled by the switching element 1141.

[0315]

When a switching element is provided between a memory element group and a node to which the power supply potential VDD or the power supply potential VSS is

supplied, data can be held even in the case where an operation of a CPU is temporarily stopped and the supply of the power supply voltage is stopped; accordingly, power consumption can be reduced. Specifically, for example, while a user of a personal computer does not input data to an input device such as a keyboard, the operation of the CPU can be stopped, so that the power consumption can be reduced.

[0316]

Although the CPU is given as an example, the transistor can also be applied to an LSI such as a digital signal processor (DSP), a custom LSI, or a field programmable gate array (FPGA).

[0317]

This embodiment can be combined as appropriate with any of the other embodiments.

EXPLANATION OF REFERENCE

[0318]

100: transistor, 101: substrate, 102: base layer, 103: oxide semiconductor layer, 104: gate insulating layer, 105: gate electrode, 106: dopant, 107: insulating layer, 108: insulating layer, 109: contact hole, 111: sidewall, 112: channel protective layer, 113: insulating layer, 115: back gate electrode, 140: transistor, 150: transistor, 160: transistor, 170: transistor, 180: transistor, 190: transistor, 1100: memory cell, 1110: memory cell array, 1111: wiring driver circuit, 1112: reading circuit, 1113: wiring driver circuit, 1120: memory cell array, 1130: memory cell, 1131: transistor, 1132: capacitor, 1140: memory cell array, 1141: switching element, 1142: memory element, 1143: memory element group, 1150: memory cell, 1151: transistor, 1152: transistor, 1153: transistor, 1154: transistor, 1155: transistor, 1156: transistor, 1160: transistor, 1161: transistor, 1162: transistor, 1163: transistor, 1164: transistor, 1170: memory cell, 1171: transistor, 1172: transistor, 1173: capacitor, 1180: memory cell, 1181: transistor, 1182: transistor, 1183: capacitor, 1189: ROM interface, 1190: substrate, 1191: ALU, 1192: ALU controller, 1193: instruction decoder, 1194: interrupt controller, 1195: timing controller, 1196: register, 1197: register controller, 1198: bus interface, 1199: ROM, 103a: source region, 103b: drain region, 103c: channel formation region, 103d: low-concentration region, 103e: low-concentration region, 105a: gate electrode, 105b: gate electrode,

110a: source electrode, and 110b: drain electrode.

This application is based on Japanese Patent Application serial no. 2010-293047 filed with Japan Patent Office on December 28, 2010, the entire contents of which are hereby incorporated by reference.

CLAIMS

1. A semiconductor device comprising:

an oxide semiconductor layer having crystallinity;

a gate insulating layer; and

a gate electrode,

wherein the oxide semiconductor layer comprises a first oxide semiconductor region and a pair of second oxide semiconductor regions,

wherein the first oxide semiconductor region is sandwiched between the pair of second oxide semiconductor regions,

wherein the first oxide semiconductor region overlaps with the gate electrode with the gate insulating layer interposed therebetween, and

wherein each of the pair of second oxide semiconductor regions comprises at least one element selected from Group 15 elements.

2. The semiconductor device according to claim 1, wherein the oxide semiconductor layer comprises zinc, indium, or gallium

3. The semiconductor device according to claim 1, wherein the oxide semiconductor layer comprises a non-single-crystal semiconductor.

4. The semiconductor device according to claim 1, wherein a concentration of the element included in the pair of second oxide semiconductor regions is higher than or equal to 5×10^{19} atoms/cm^3 and lower than or equal to 1×10^{22} atoms/cm^3.

5. The semiconductor device according to claim 1, wherein the pair of second oxide semiconductor regions has a wurtzite crystal structure.

6. The semiconductor device according to claim 1, wherein the element included in the pair of second oxide semiconductor regions is nitrogen.

7. A semiconductor device comprising:

an oxide semiconductor layer having crystallinity;

a gate insulating layer; and

a gate electrode,

wherein the oxide semiconductor layer comprises a first oxide semiconductor region, a pair of second oxide semiconductor regions, and a pair of third oxide semiconductor regions,

wherein the first oxide semiconductor region is sandwiched between the pair of third oxide semiconductor regions,

wherein the pair of third oxide semiconductor regions is sandwiched between the pair of second oxide semiconductor regions,

wherein the first oxide semiconductor region overlaps with the gate electrode with the gate insulating layer interposed therebetween, and

wherein each of the pair of second oxide semiconductor regions and the pair of third oxide semiconductor regions comprises at least one element selected from Group 15 elements.

8. The semiconductor device according to claim 7, wherein the oxide semiconductor layer comprises zinc, indium, or gallium

9. The semiconductor device according to claim 7, wherein the oxide semiconductor layer comprises a non-single-crystal semiconductor.

10. The semiconductor device according to claim 7, wherein a concentration of the element included in the pair of second oxide semiconductor regions is higher than or equal to 5×10^{19} atoms/cm^3 and lower than or equal to 1×10^{22} atoms/cm^3.

11. The semiconductor device according to claim 7, wherein the pair of second oxide semiconductor regions and the pair of third oxide semiconductor regions have a wurtzite crystal structure.

12. The semiconductor device according to claim 7, wherein the element included in the pair of second oxide semiconductor regions and the pair of third oxide

semiconductor regions is nitrogen.

13. The semiconductor device according to claim 7, wherein a concentration of the element included in the pair of third oxide semiconductor regions is higher than or equal to 5×10^{19} atoms/cm^3 and lower than or equal to 1×10^{22} atoms/cm^3.

14. A method for manufacturing a semiconductor device, comprising the steps of:

forming an oxide semiconductor layer having crystallinity;

forming a gate insulating layer over the oxide semiconductor layer;

forming a gate electrode over the gate insulating layer; and

adding at least one element selected from Group 15 elements to a region of the oxide semiconductor layer with the use of the gate electrode as a mask,

wherein the region of the oxide semiconductor layer to which the element is added obtains a wurtzite crystal structure by heat treatment.

15. The method for manufacturing a semiconductor device, according to claim 14, wherein the oxide semiconductor layer comprises zinc, indium, or gallium.

16. The method for manufacturing a semiconductor device, according to claim 14, wherein the oxide semiconductor layer comprises a non-single-crystal semiconductor.

17. The method for manufacturing a semiconductor device, according to claim 14, wherein the element is nitrogen.

18. A method for manufacturing a semiconductor device, comprising the steps of:

forming an oxide semiconductor layer having crystallinity;

forming a gate insulating layer over the oxide semiconductor layer;

forming a gate electrode over the gate insulating layer;

adding at least one element selected from Group 15 elements to a first region of

the oxide semiconductor layer with the use of the gate electrode as a mask in a first addition step;

forming a sidewall on a side surface of the gate electrode; and

adding at least one element selected from Group 15 elements to a second region of the oxide semiconductor layer with the use of the gate electrode and the sidewall as masks in a second addition step; and

heating the oxide semiconductor layer thereby forming a wurtzite crystal structure in the first region and the second region of the oxide semiconductor layer.

19. The method for manufacturing a semiconductor device, according to claim 18, wherein the oxide semiconductor layer comprises zinc, indium, or gallium.

20. The method for manufacturing a semiconductor device, according to claim 18, wherein the oxide semiconductor layer comprises a non-single-crystal semiconductor.

21. The method for manufacturing a semiconductor device, according to claim 18, wherein the element is nitrogen.

22. The method for manufacturing a semiconductor device, according to claim 18, wherein the element added to the first region of the oxide semiconductor layer in the first addition step is nitrogen.

23. The method for manufacturing a semiconductor device, according to claim 18, wherein the element added to the second region of the oxide semiconductor layer in the second addition step is nitrogen.

24. The method for manufacturing a semiconductor device, according to claim 18, wherein a concentration of the element added to the first region of the oxide semiconductor layer in the first addition step is lower than a concentration of the element added to the second region of the oxide semiconductor layer in the second addition step.

FIG. 1A

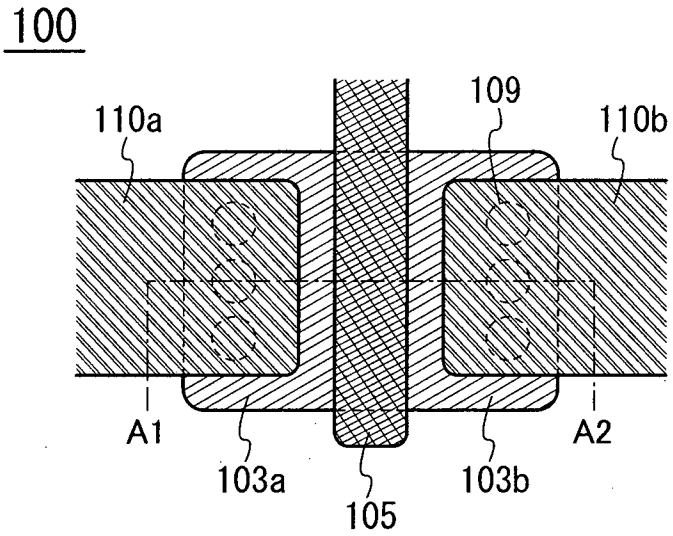

100
110a
109
110b
A1
A2
103a
105
103b

FIG. 1B

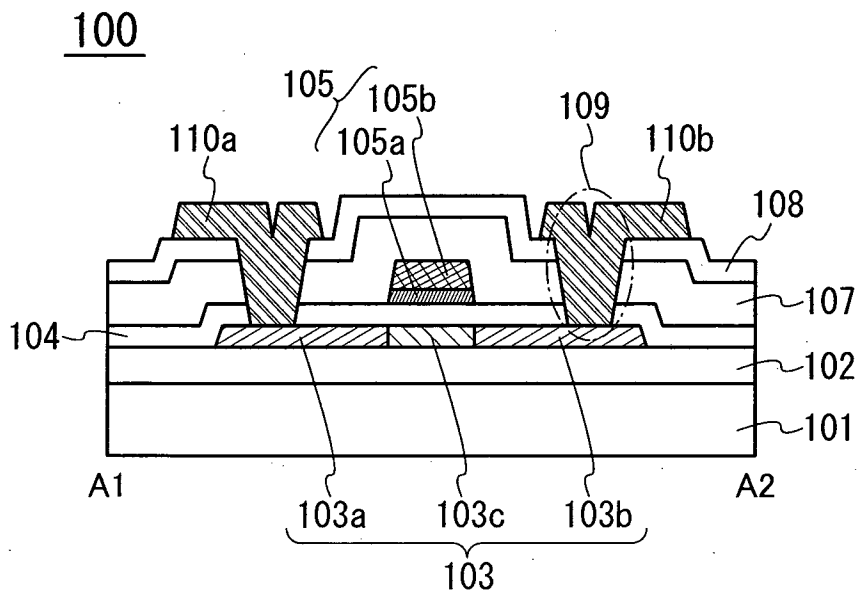

100
105
105b
105a
109
110a
110b
108
107
104
102
101
A1
A2
103a
103c
103b
103

FIG. 2A

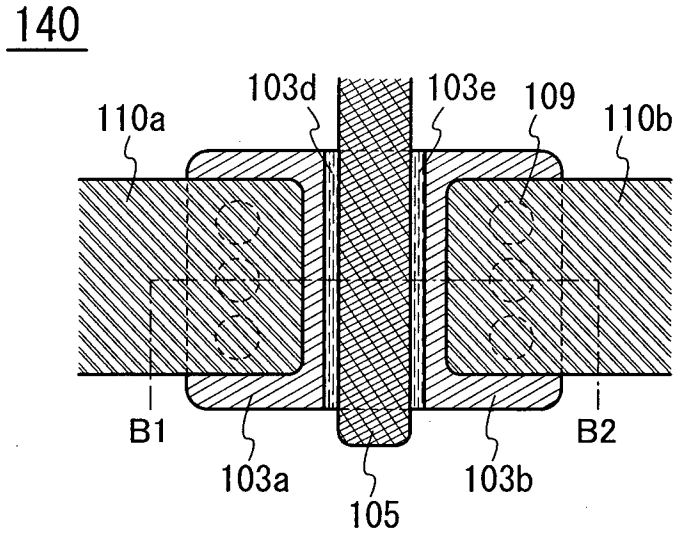

140
103d
103e
109
110a
110b
B1
B2
103a
103b
105

FIG. 2B

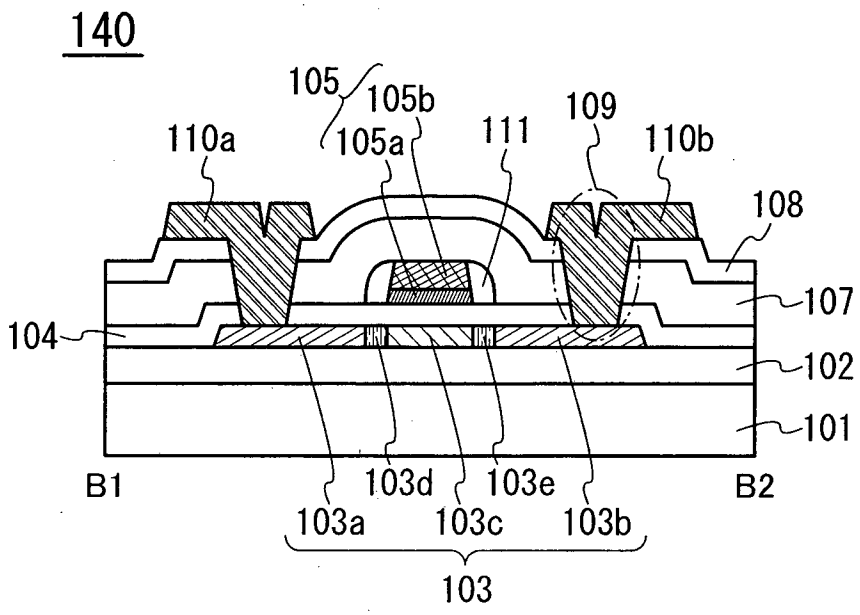

140
105
105b
105a
111
109
110a
110b
108
107
104
102
101
B1
103d
103e
B2
103a
103c
103b
103

FIG. 3A

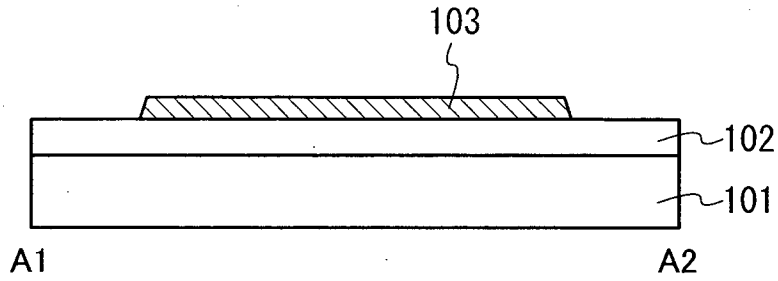


FIG. 3B

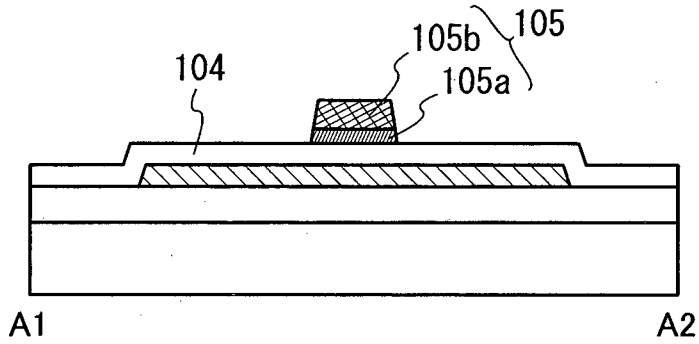


FIG. 3C

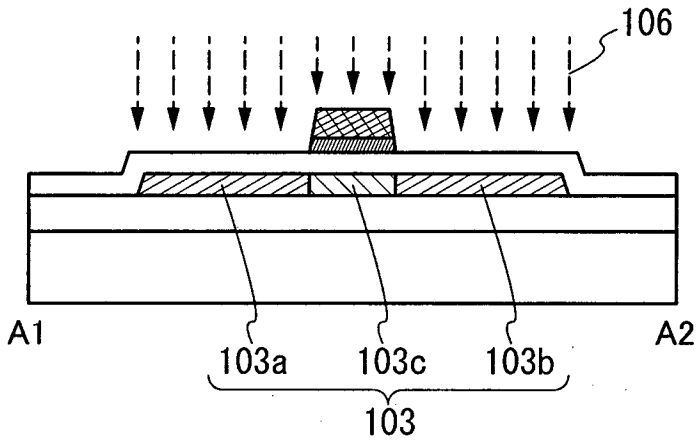


FIG. 3D

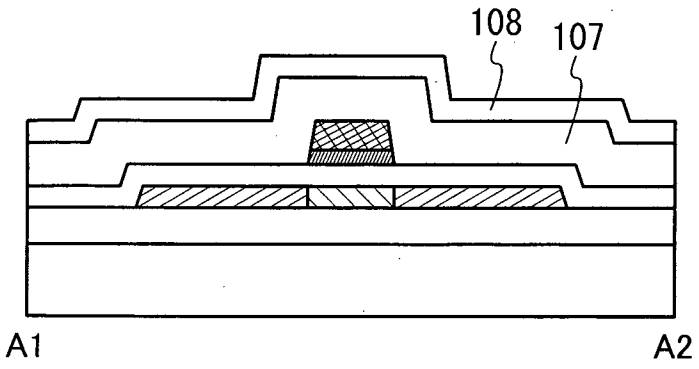


FIG. 4A

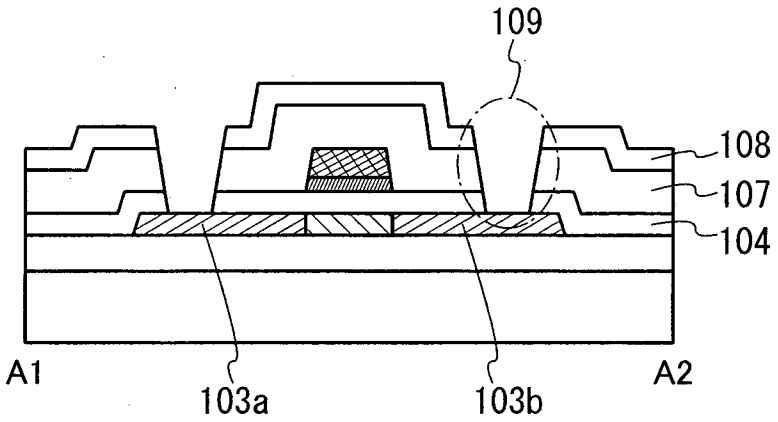


FIG. 4B

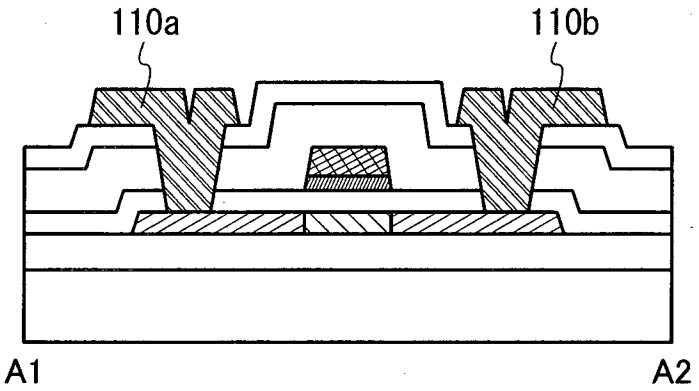


FIG. 5A

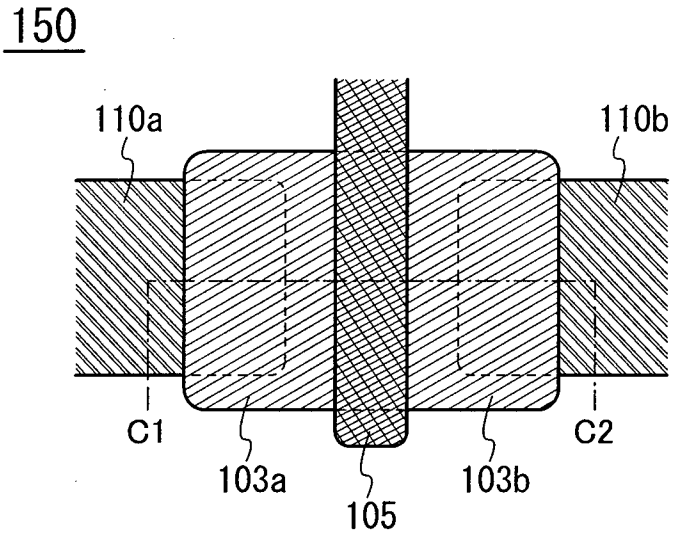

150
110a
110b
C1
C2
103a
105
103b

FIG. 5B

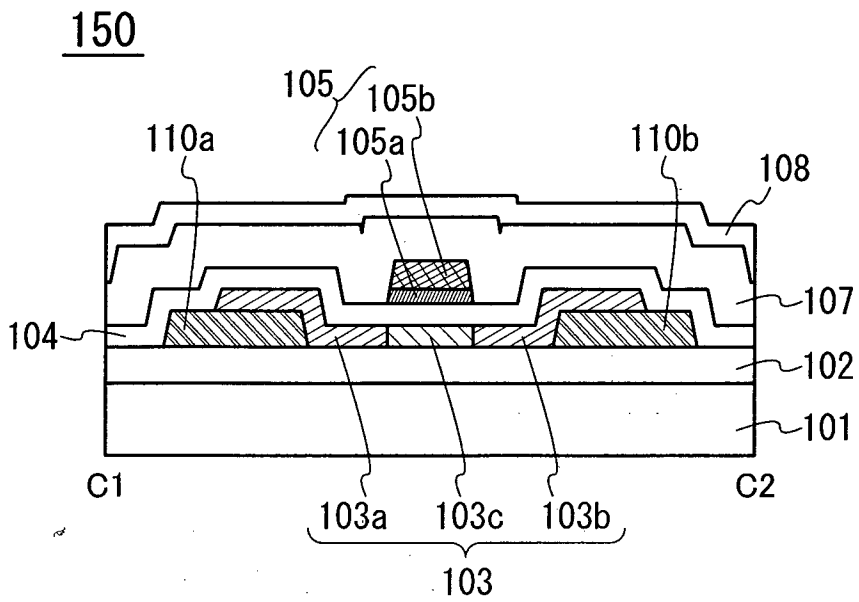

150
105
105b
105a
110a
110b
108
107
104
102
101
C1
C2
103a
103c
103b
103

FIG. 6A

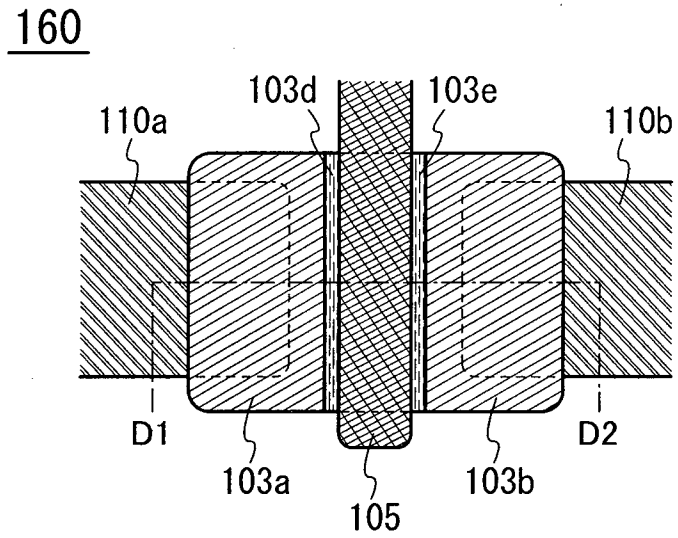

160
103d
103e
110a
110b
D1
D2
103a
105
103b

FIG. 6B

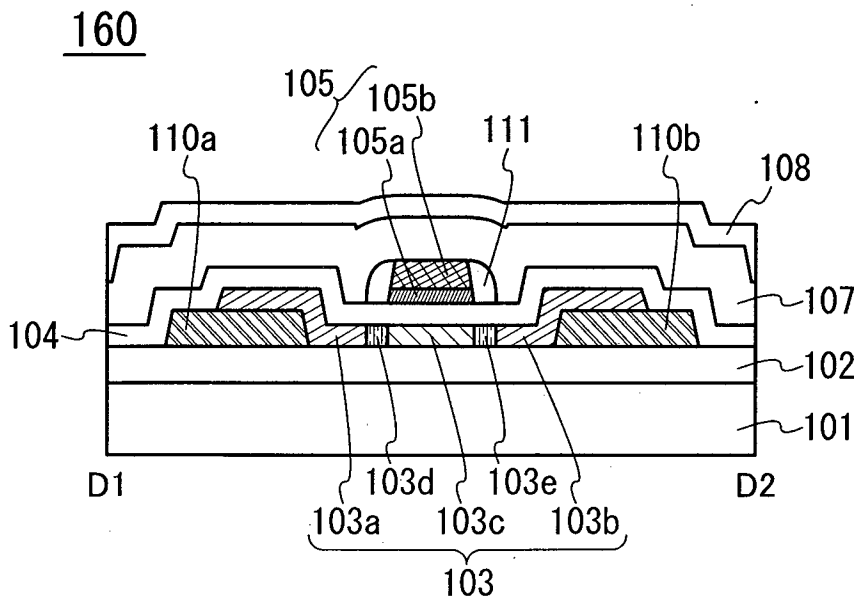

160
105
105b
110a
105a
111
110b
108
107
104
102
101
D1
103d
103e
D2
103a
103c
103b
103

FIG. 7A

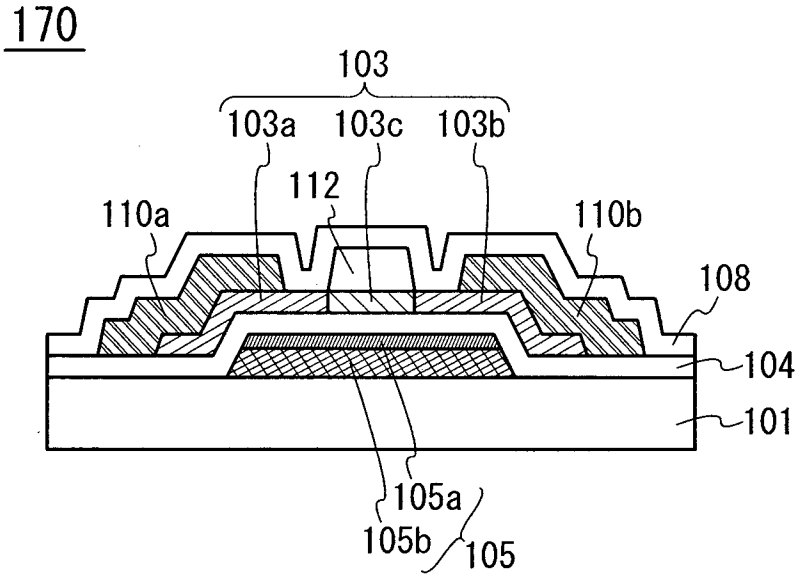

170
103
103a
103c
103b
112
110a
110b
108
104
101
105a
105b
105

FIG. 7B

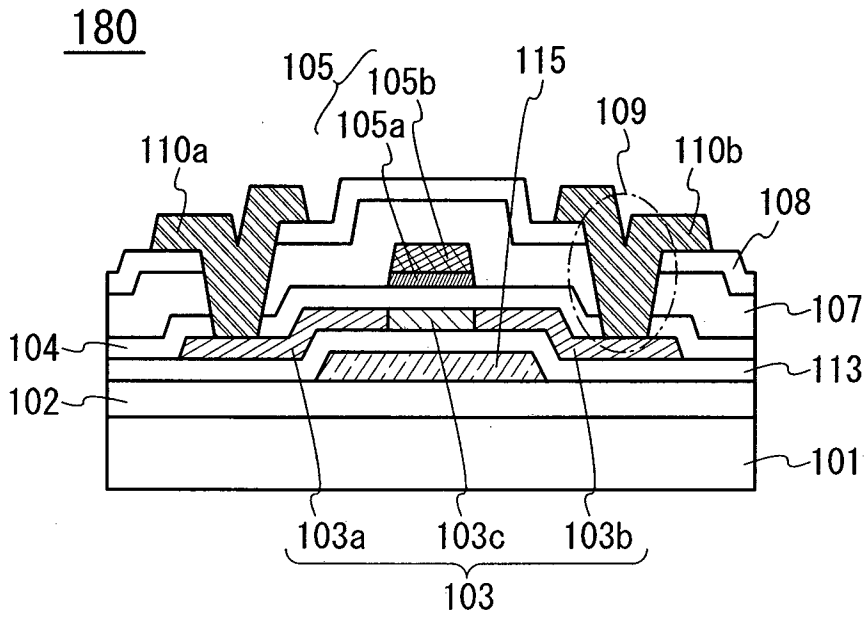

180
105
105b
105a
115
109
110a
110b
108
107
104
113
102
101
103a
103c
103b
103

FIG. 8

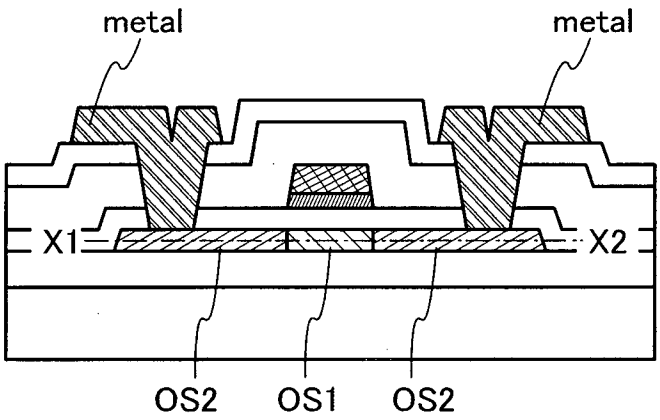


FIG. 9A

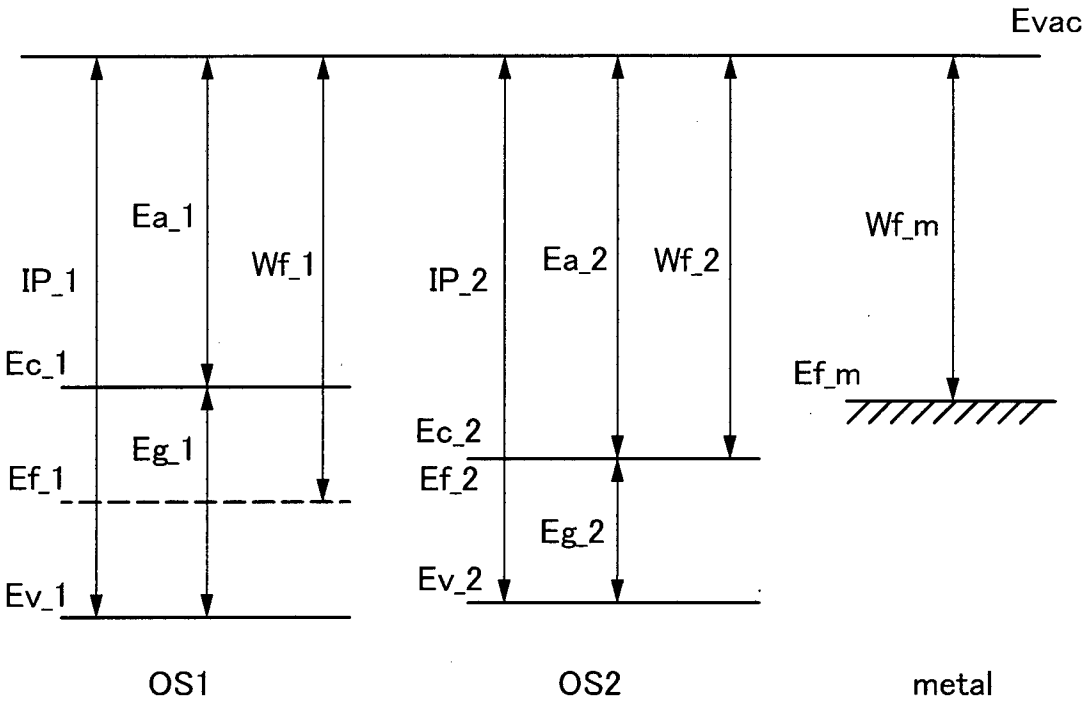

Evac
Ea_1
Wf_1
IP_1
Ec_1
Eg_1
Ef_1
Ev_1
IP_2
Ea_2
Wf_2
Ec_2
Ef_2
Eg_2
Ev_2
Wf_m
Ef_m
OS1
OS2
metal

FIG. 9B

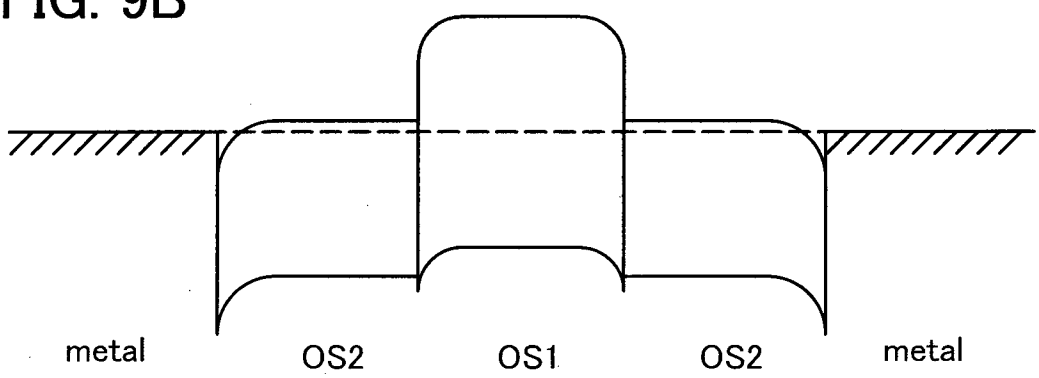

metal
OS2
OS1
OS2
metal

FIG. 10A

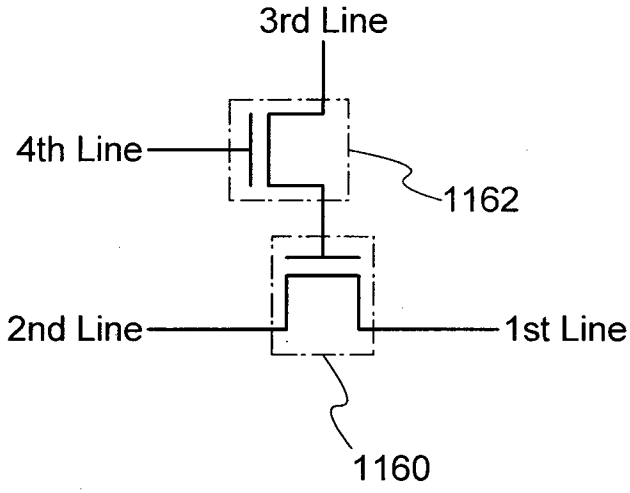


FIG. 10B

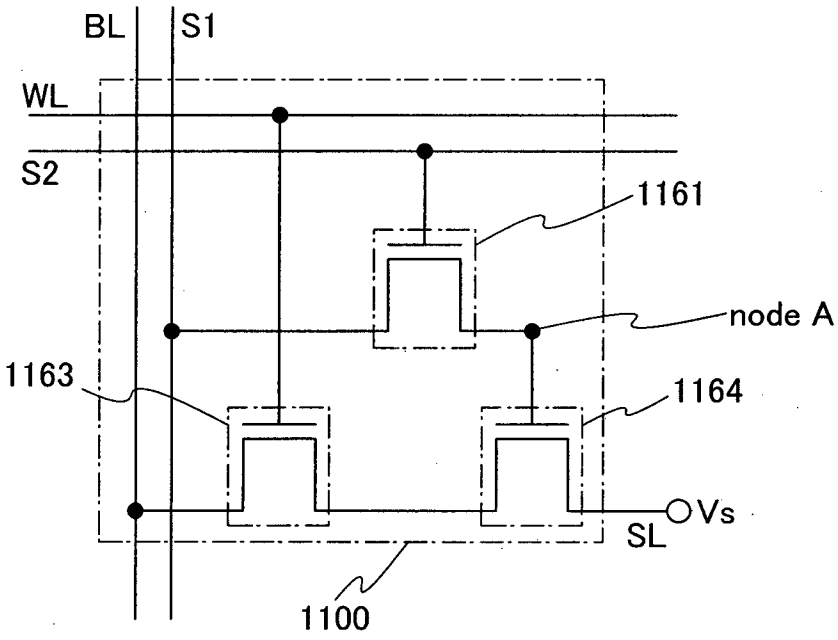


FIG. 11

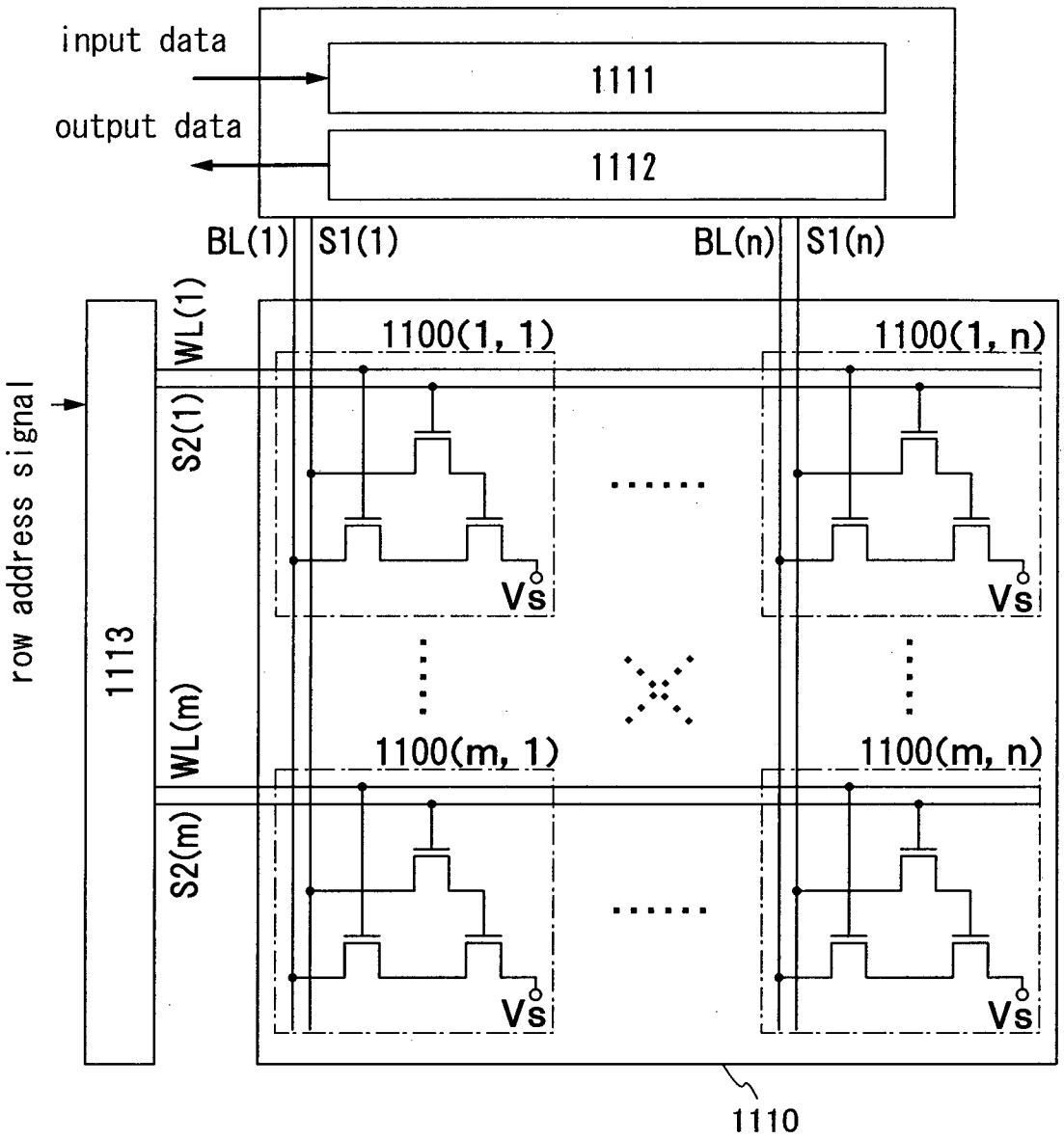

input data
output data
1111
1112
BL(1)
S1(1)
BL(n)
S1(n)
WL(1)
S2(1)
WL(m)
S2(m)
row address signal
1113
1100(1, 1)
1100(1, n)
1100(m, 1)
1100(m, n)
Vs
1110

FIG. 12A

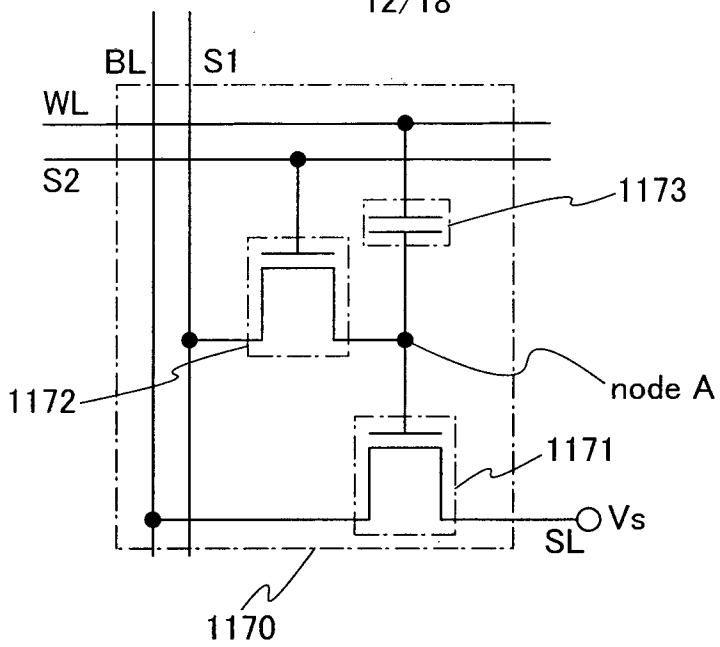


FIG. 12B

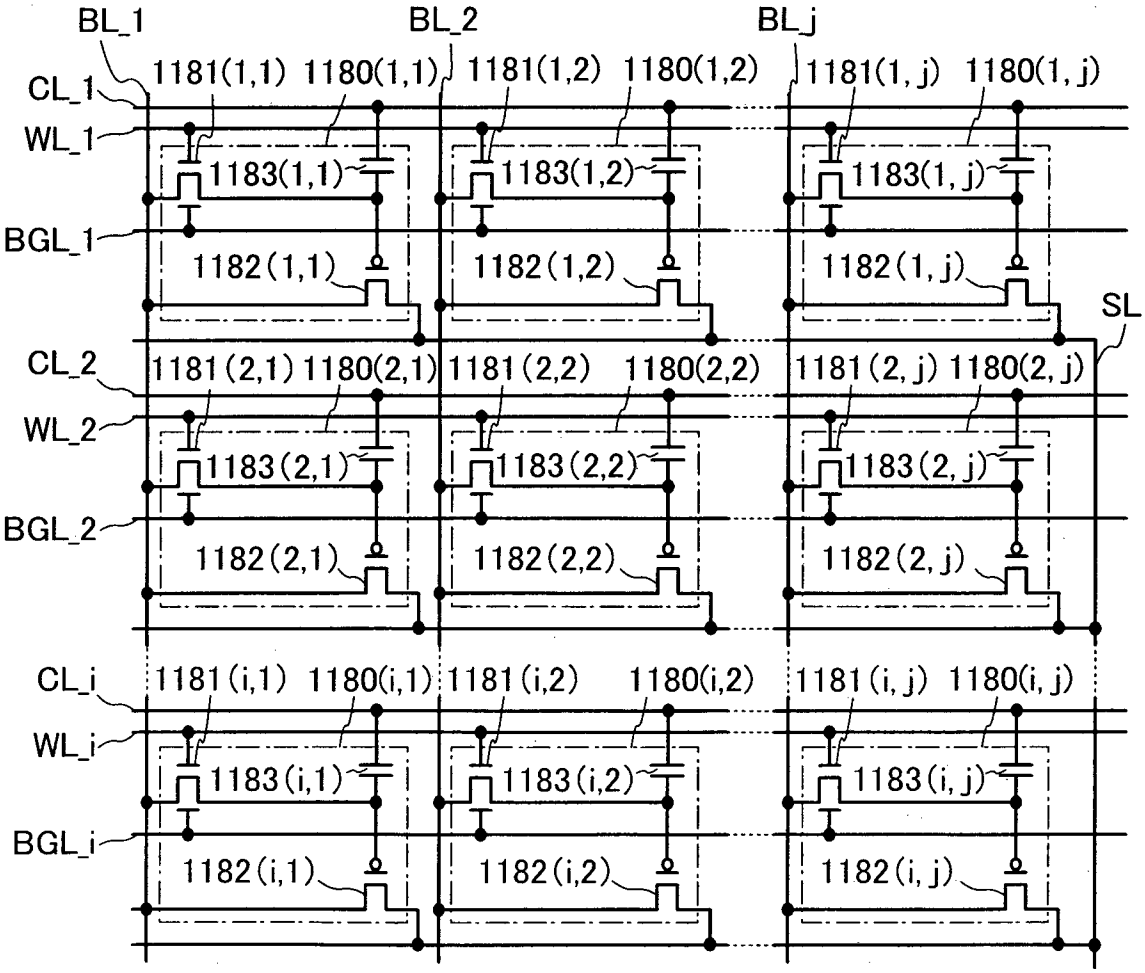


FIG. 13A

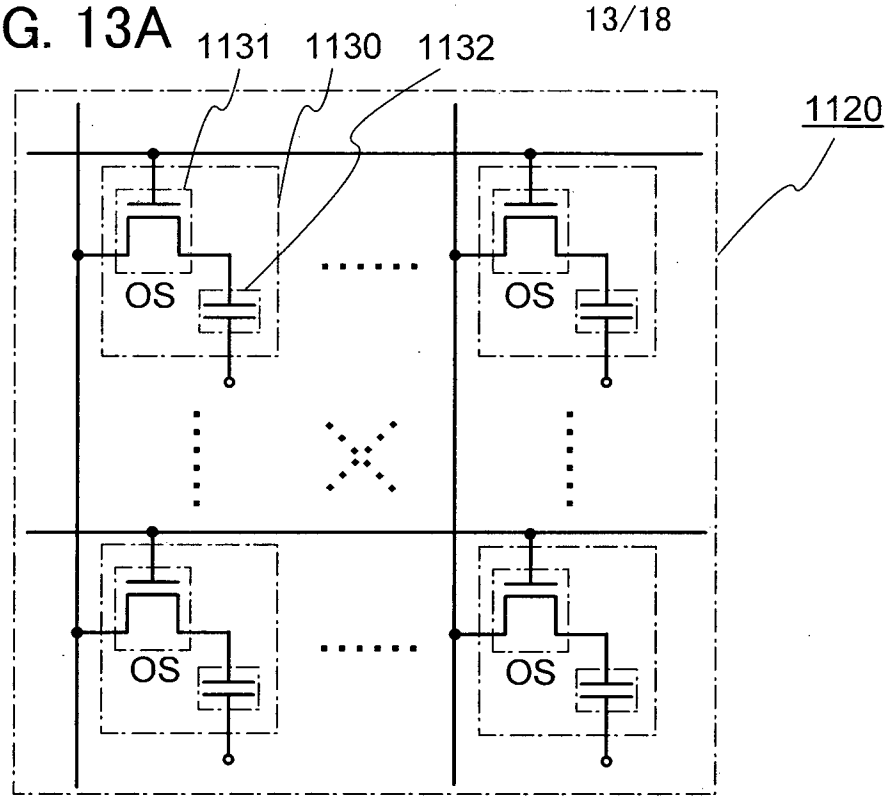


FIG. 13B

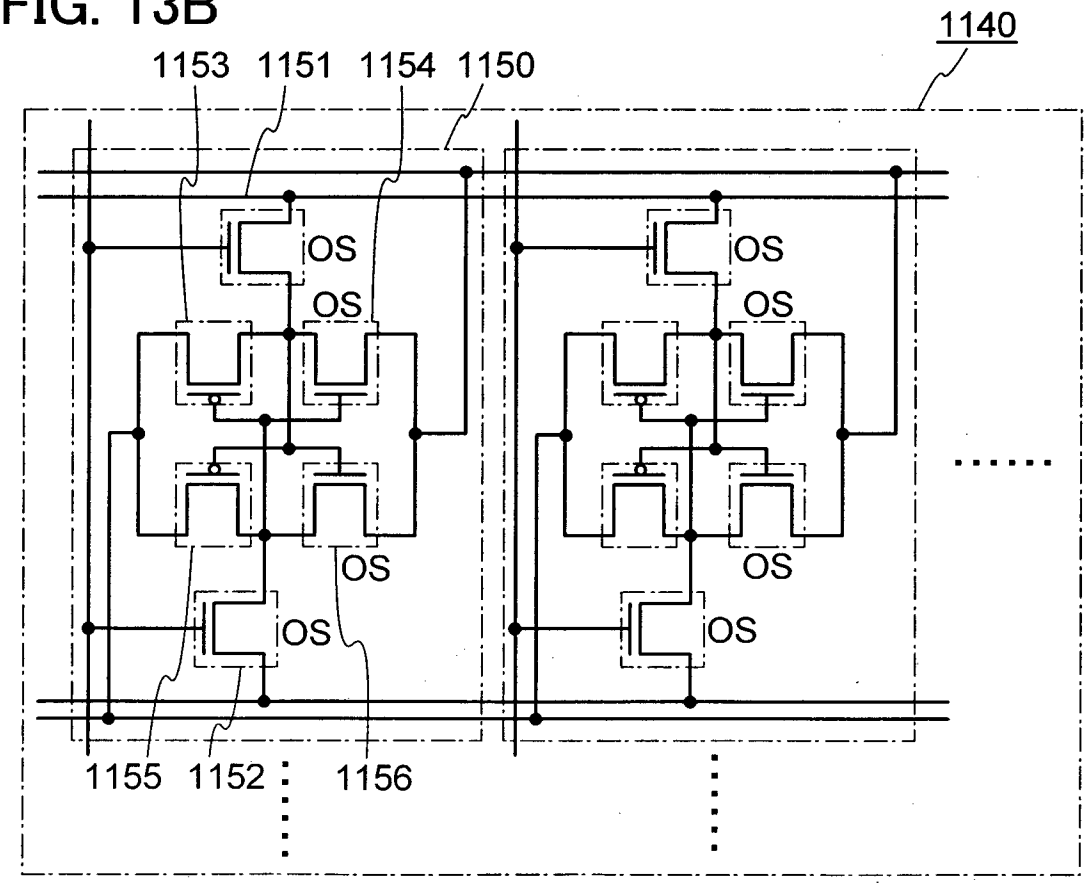


FIG. 14A

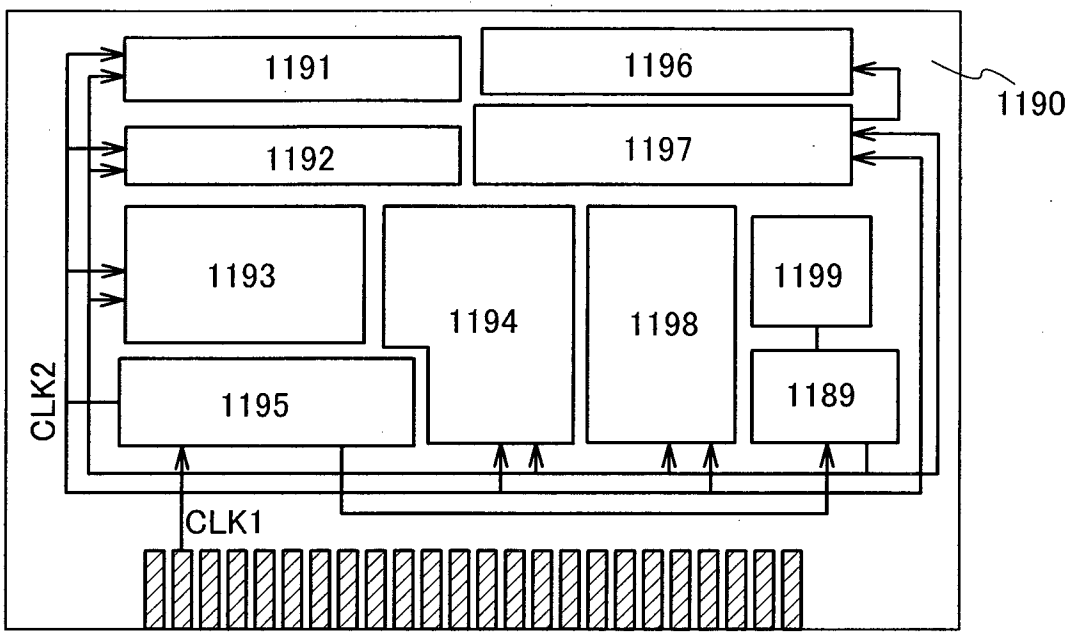


FIG. 14B

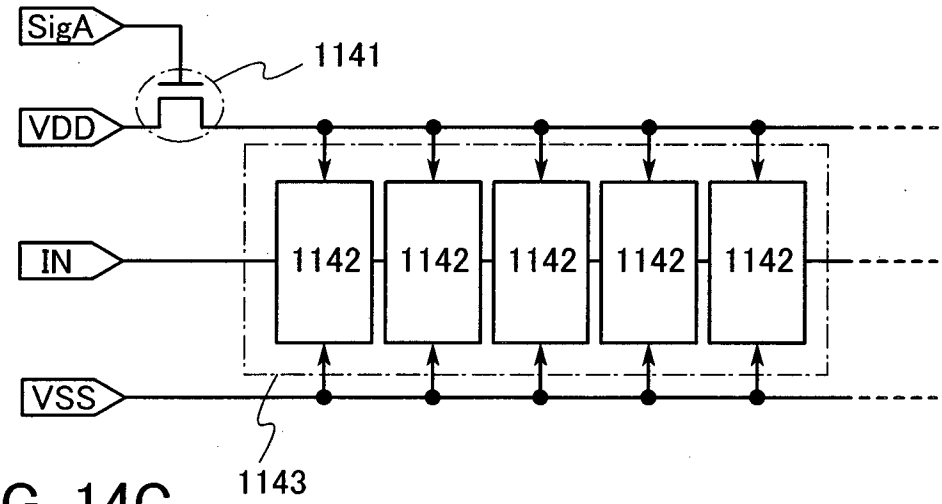


FIG. 14C

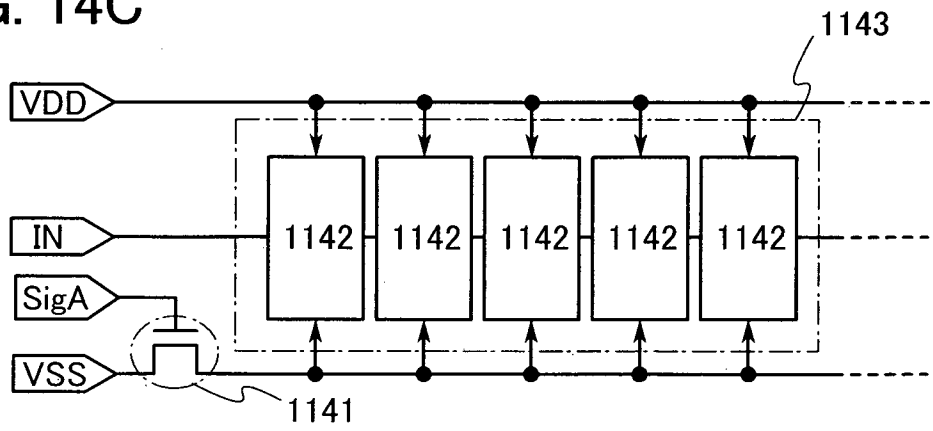


FIG. 15A

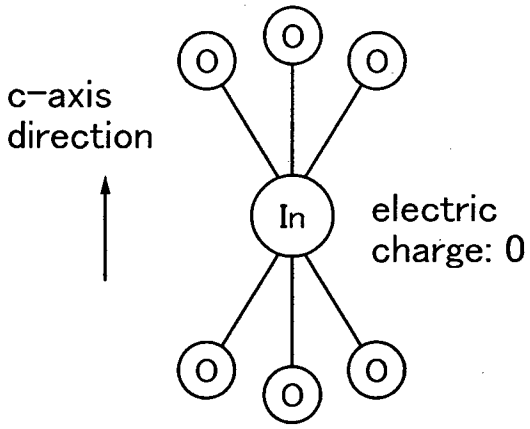


FIG. 15B

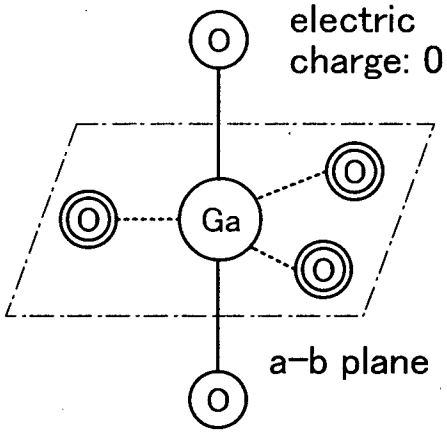


FIG. 15C

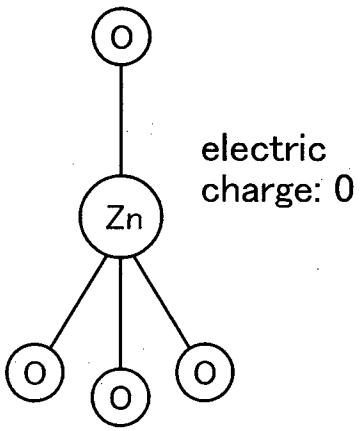


FIG. 15D

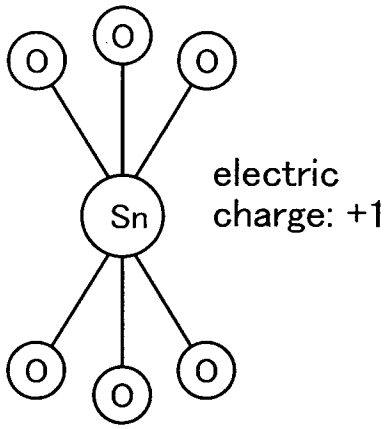


FIG. 15E

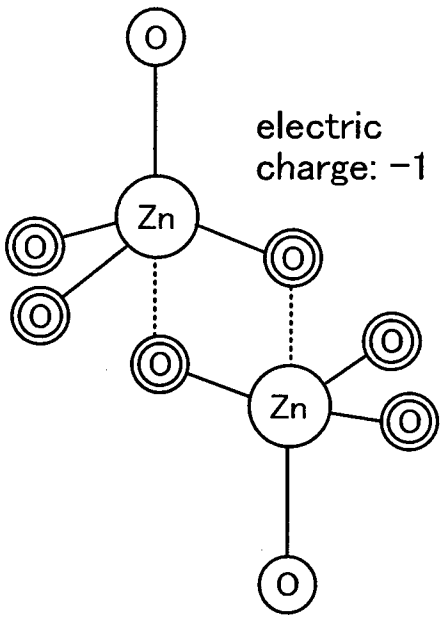


FIG. 16A

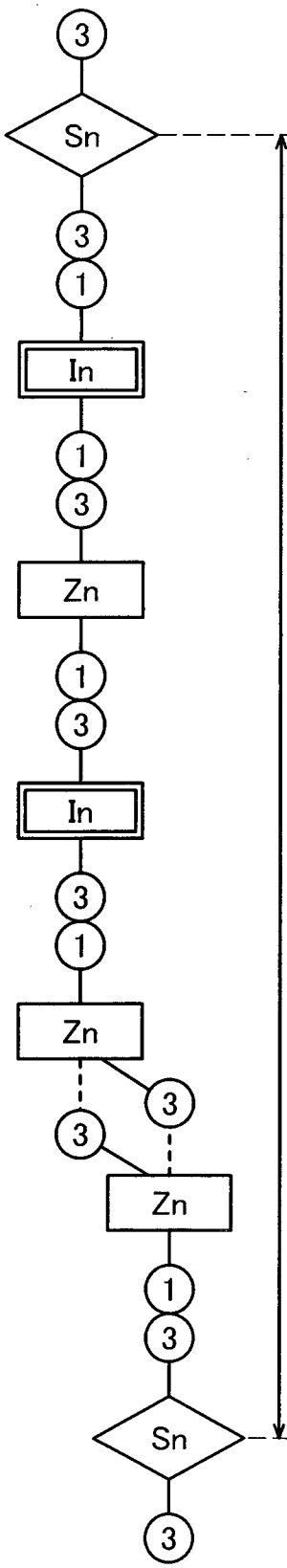


FIG. 16B

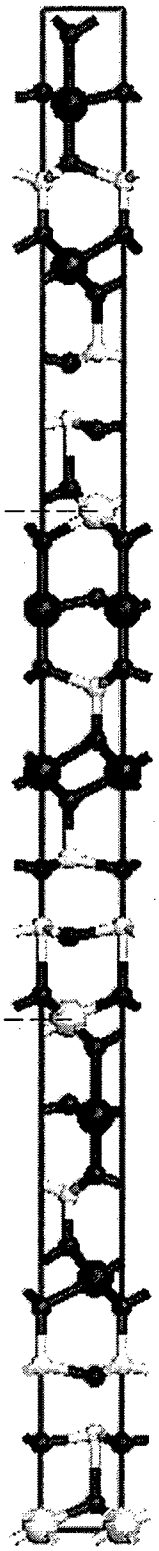
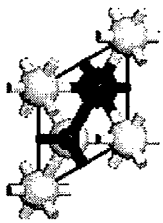


FIG. 16C



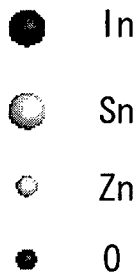


FIG. 17A

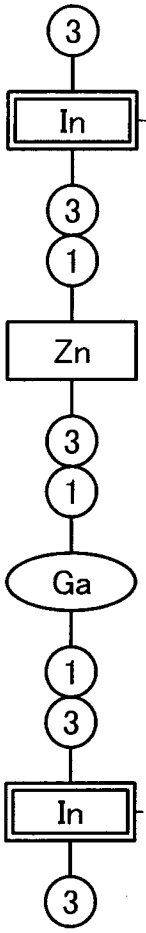


FIG. 17B

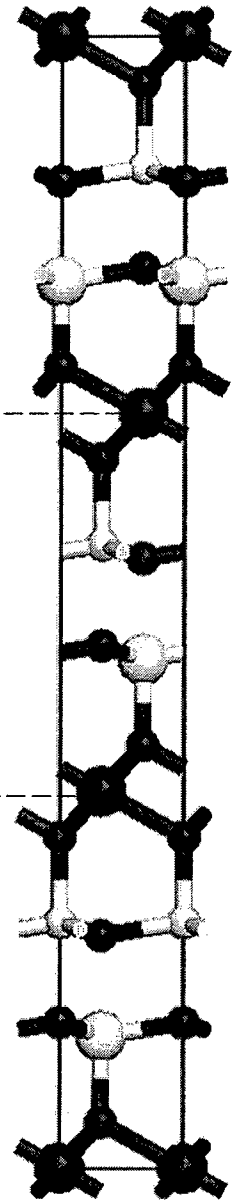


FIG. 17C

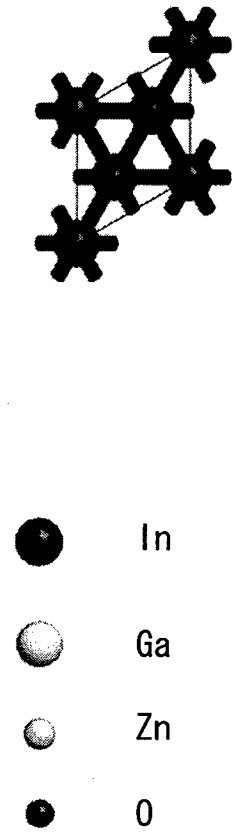


FIG. 18A

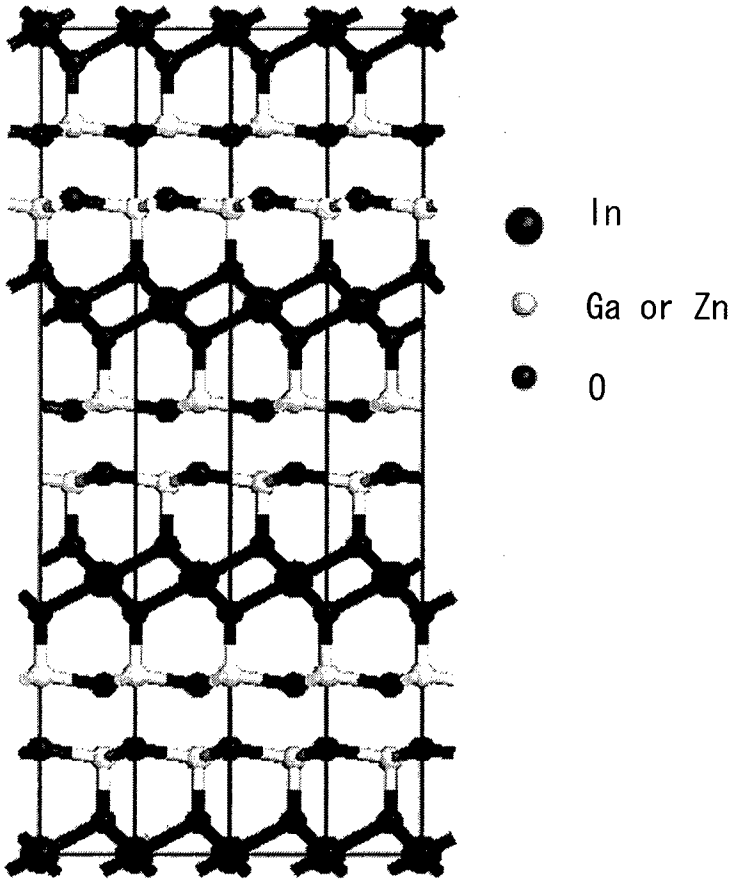
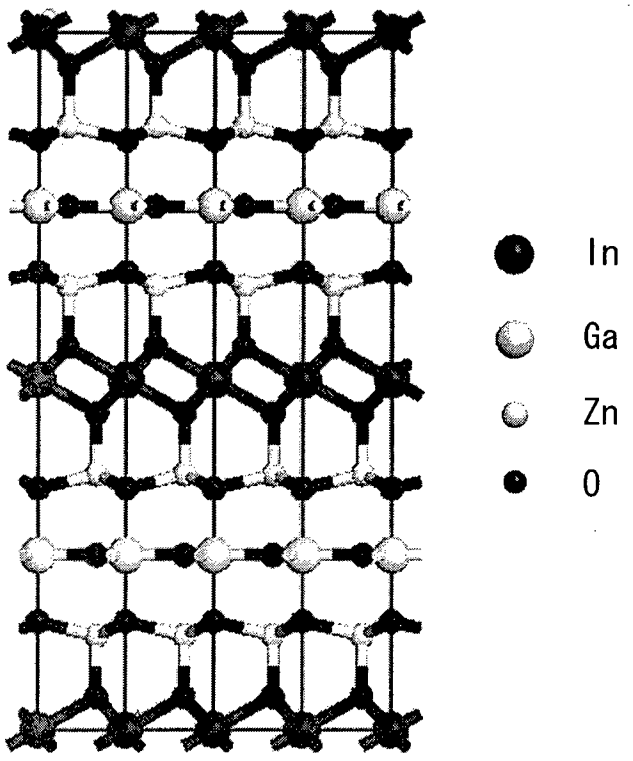


FIG. 18B



INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2011/080141

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl. H01L29/786(2006.01)i, H01L21/336(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl. H01L29/786, H01L21/336

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan 1922-1996
Published unexamined utility model applications of Japan 1971-2012
Registered utility model specifications of Japan 1996-2012
Published registered utility model applications of Japan 1994-2012

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	US 2010/0084655 A1 (CANON KABUSHIKI KAISHA) 2010.04.08, Par. Nos. [0016] to [0100], Figs.1A to 7C & JP 2010-93070 A & EP 2175493 A1 & KR 10-2010-0039806 A & CN 101719514 A	1-6, 14-17 7-13, 18-24
Y	US 2009/0283763 A1 (Samsung Electronics Co., Ltd.) 2009.11.19, Par. Nos. [0066] to [0090], Figs.1 to 3F & JP 2009-278115 A & EP 2120267 A1 & CN 101582453 A & KR 10-2009-0119666 A	7-13, 18-24

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
15.03.2012	27.03.2012

Name and mailing address of the ISA/JP	Authorized officer
Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan	Takashi WATAHIKI 4M 2934 Telephone No. +81-3-3581-1101 Ext. 3462