

ANTI-MUC1 ANTIBODIES AND USES THEREOF

CLAIM OF PRIORITY

This application claims the benefit of PCT Application No. PCT/CN2023/070283, filed on January 4, 2023, and PCT/CN2023/082744, filed on March 21, 2023. The entire contents of the foregoing are incorporated herein by reference.

TECHNICAL FIELD

This disclosure relates to anti-MUC1 (Mucin 1, cell surface associated) antibodies, antigen-binding fragments thereof, antibody-drug conjugate (ADC) derived therefrom, and the uses thereof.

BACKGROUND

Mucin 1, cell surface associated (MUC1), also called polymorphic epithelial mucin (PEM) or epithelial membrane antigen or EMA, is a mucin encoded by the MUC1 gene in humans. The transmembrane glycoprotein Mucin 1 (MUC1) is aberrantly glycosylated and overexpressed in a variety of epithelial cancers, and plays a crucial role in progression of the disease.

Considering the important role of MUC1 in autoimmune diseases, inflammatory responses and tumors, there is a need to develop a therapeutic agent targeting MUC1.

SUMMARY

This disclosure relates to anti-MUC1 antibodies, antigen-binding fragment thereof, and the uses thereof.

In one aspect, the disclosure is related to an antibody or antigen-binding fragment thereof that binds to MUC1 (The transmembrane glycoprotein Mucin 1) comprising:

a heavy chain variable region (VH) comprising complementarity determining regions (CDRs) 1, 2, and 3, wherein the VH CDR1 region comprises an amino acid sequence that is at least 80% identical to a selected VH CDR1 amino acid sequence, the VH CDR2 region comprises an amino acid sequence that is at least 80% identical to a selected VH CDR2 amino acid sequence, and the VH CDR3 region comprises an amino acid sequence that is at least 80% identical to a selected VH CDR3 amino acid sequence; and

a light chain variable region (VL) comprising CDRs 1, 2, and 3, wherein the VL CDR1 region comprises an amino acid sequence that is at least 80% identical to a selected VL CDR1 amino acid sequence, the VL CDR2 region comprises an amino acid sequence that is at least 80% identical to a selected VL CDR2 amino acid sequence, and the VL CDR3 region comprises an amino acid sequence that is at least 80% identical to a selected VL CDR3 amino acid sequence,

wherein the selected VH CDRs 1, 2, and 3 amino acid sequences and the selected VL CDRs, 1, 2, and 3 amino acid sequences are one of the following:

1

(1) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 4-6, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 1-3, respectively;

5 (2) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 7-9, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 1-3, respectively;

(3) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 10-12, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 1-3, respectively;

10 (4) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 13-15, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 1-3, respectively;

(5) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 16-18, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 1-3, respectively;

(6) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 19-21, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 1-3, respectively;

20 (7) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 25-27, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 22-24, respectively;

(8) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 28-30, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 22-24, respectively;

25 (9) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 31-33, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 22-24, respectively;

(10) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 34-36, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 22-24, respectively;

30 (11) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 37-39, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 22-24, respectively;

35 (12) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 40-42, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 22-24, respectively;

(13) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 7, 63, 9, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 1-3, respectively;

40 (14) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 7, 64, 9, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 1-3, respectively;

(15) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 28, 65, 30, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 22-24, respectively;

5 (16) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 28, 66, 30, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 22-24, respectively.

In some embodiments, the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 4-6, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, according to Kabat definition.

10 In some embodiments, the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 7-9, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, according to Kabat definition.

In some embodiments, the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 10-12, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, according to Kabat definition.

15 In some embodiments, the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 13-15, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, according to Kabat definition.

In some embodiments, the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 16-18, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, according to Kabat definition.

20 In some embodiments, the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 19-21, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, according to Kabat definition.

25 In some embodiments, the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 25-27, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, according to Chothia definition.

In some embodiments, the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 28-30, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, according to Chothia definition.

30 In some embodiments, the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 31-33, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, according to Chothia definition.

In some embodiments, the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 34-36, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, according to Chothia definition.

35 In some embodiments, the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 37-39, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, according to Chothia definition.

40 In some embodiments, the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 40-42, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, according to Chothia definition.

In some embodiments, the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 7, 63, 9, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, according to Kabat definition.

5 In some embodiments, the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 7, 64, 9, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, according to Kabat definition.

10 In some embodiments, the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 28, 65, 30, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, according to Chothia definition.

In some embodiments, the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 28, 66, 30, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, according to Chothia definition.

15 In some embodiments, the antibody or antigen-binding fragment thereof specifically binds to human, mouse, monkey, or dog MUC1.

In some embodiments, the antibody or antigen-binding fragment thereof is a human antibody or antigen-binding fragment thereof, a single-chain variable fragment (scFv), a one-armed antibody, and/or a multi-specific antibody (e.g., a bispecific antibody).

20 In some embodiments, the antibody or antigen-binding fragment thereof is a human IgG1 antibody or antigen-binding fragment thereof or a human IgG4 antibody or antigen-binding fragment thereof.

In one aspect, the disclosure is related to a nucleic acid comprising a polynucleotide encoding a polypeptide comprising:

25 an immunoglobulin heavy chain or a fragment thereof comprising a heavy chain variable region (VH) comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 4-6, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;

30 an immunoglobulin light chain or a fragment thereof comprising a VL comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 44 binds to MUC1;

35 an immunoglobulin heavy chain or a fragment thereof comprising a heavy chain variable region (VH) comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 7-9, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;

40 an immunoglobulin light chain or a fragment thereof comprising a VL comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 45 binds to MUC1;

- an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 10-12, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;
- an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 46 binds to MUC1;
- an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 13-15, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;
- an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 47 binds to MUC1;
- an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 16-18, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;
- an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 48 binds to MUC1;
- an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 19-21, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;
- an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 49 binds to MUC1;
- an immunoglobulin heavy chain or a fragment thereof comprising a heavy chain variable region (VH) comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 25-27, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;
- an immunoglobulin light chain or a fragment thereof comprising a VL comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences

set forth in SEQ ID NOs: 22-24, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 44 binds to MUC1;

an immunoglobulin heavy chain or a fragment thereof comprising a heavy chain variable region (VH) comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 28-30, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;

an immunoglobulin light chain or a fragment thereof comprising a VL comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 45 binds to MUC1;

an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 31-33, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;

an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 46 binds to MUC1;

an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 34-36, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;

an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 47 binds to MUC1;

an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 37-39, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;

an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 48 binds to MUC1;

an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 40-42, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;

an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 49 binds to MUC1;

5 an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 7, 63, 9, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;

10 an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 61 binds to MUC1;

15 an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 7, 64, 9, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;

20 an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 62 binds to MUC1;

25 an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 28, 65, 30, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;

30 an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 61 binds to MUC1;

35 an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 28, 66, 30, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1; or

40 an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 62 binds to MUC1.

In some embodiments, the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin light chain or a fragment thereof comprising a VL

comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively.

5 In some embodiments, the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively.

10 In some embodiments, the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 4-6, respectively.

In some embodiments, the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 7-9, respectively.

15 In some embodiments, the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 10-12, respectively.

20 In some embodiments, the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 13-15, respectively.

25 In some embodiments, the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 16-18, respectively.

30 In some embodiments, the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 19-21, respectively.

In some embodiments, the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 25-27, respectively.

35 In some embodiments, the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 28-30, respectively.

40 In some embodiments, the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 31-33, respectively.

In some embodiments, the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 34-36, respectively.

5 In some embodiments, the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 37-39, respectively.

10 In some embodiments, the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 40-42, respectively.

15 In some embodiments, the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 7, 63, 9, respectively.

20 In some embodiments, the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 7, 64, 9, respectively.

In some embodiments, the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 28, 65, 30, respectively.

25 In some embodiments, the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 28, 66, 30, respectively.

30 In some embodiments, the VH when paired with a VL specifically binds to human, mouse, monkey, or dog MUC1, or the VL when paired with a VH specifically binds to human, mouse, monkey, or dog MUC1.

35 In some embodiments, the immunoglobulin heavy chain or the fragment thereof is a human immunoglobulin heavy chain or a fragment thereof (e.g., a human IgG1 heavy chain or a fragment thereof; or a human IgG4 heavy chain or a fragment thereof), and the immunoglobulin light chain or the fragment thereof is a human immunoglobulin light chain or a fragment thereof.

In some embodiments, the nucleic acid encodes a single-chain variable fragment (scFv), a one-armed antibody, a multi-specific antibody (e.g., a bispecific antibody), or a chimeric antigen receptor (CAR).

In some embodiments, the nucleic acid is cDNA.

40 In one aspect, the disclosure is related to a vector comprising one or more of the nucleic acids described herein.

In one aspect, the disclosure is related to a vector comprising two of the nucleic acids described herein, wherein the vector encodes the VL region and the VH region that together bind to MUC1.

5 In one aspect, the disclosure is related to a pair of vectors, wherein each vector comprises one of the nucleic acids described herein, wherein together the pair of vectors encodes the VL region and the VH region that together bind to MUC1.

In one aspect, the disclosure is related to a cell comprising the vector described herein, or the pair of vectors described herein.

In some embodiments, the cell is a CHO cell.

10 In one aspect, the disclosure is related to a cell comprising one or more of the nucleic acids described herein.

In one aspect, the disclosure is related to a cell comprising two of the nucleic acids described herein.

15 In some embodiments, the two nucleic acids together encode the VL region and the VH region that together bind to MUC1.

In one aspect, the disclosure is related to a method of producing an antibody or an antigen-binding fragment thereof, the method comprising culturing the cell described herein under conditions sufficient for the cell to produce the antibody or the antigen-binding fragment; and collecting the antibody or the antigen-binding fragment produced by the cell.

20 In one aspect, the disclosure is related to an antibody or antigen-binding fragment thereof that binds to MUC1 comprising

a heavy chain variable region (VH) comprising an amino acid sequence that is at least 90% identical to a selected VH sequence, and a light chain variable region (VL) comprising an amino acid sequence that is at least 90% identical to a selected VL sequence, wherein the
25 selected VH sequence and the selected VL sequence are one of the following:

the selected VH sequence is SEQ ID NO: 44, and the selected VL sequence is SEQ ID NO: 43;

the selected VH sequence is SEQ ID NO: 45, and the selected VL sequence is SEQ ID NO: 43;

30 the selected VH sequence is SEQ ID NO: 46, and the selected VL sequence is SEQ ID NO: 43;

the selected VH sequence is SEQ ID NO: 47, and the selected VL sequence is SEQ ID NO: 43;

35 the selected VH sequence is SEQ ID NO: 48, and the selected VL sequence is SEQ ID NO: 43;

the selected VH sequence is SEQ ID NO: 49, and the selected VL sequence is SEQ ID NO: 43;

the selected VH sequence is SEQ ID NO: 61, and the selected VL sequence is SEQ ID NO: 43; and

40 the selected VH sequence is SEQ ID NO: 62, and the selected VL sequence is SEQ ID NO: 43.

In some embodiments, the VH comprises the sequence of SEQ ID NO: 44 and the VL comprises the sequence of SEQ ID NO: 43.

In some embodiments, the VH comprises the sequence of SEQ ID NO: 45 and the VL comprises the sequence of SEQ ID NO: 43.

In some embodiments, the VH comprises the sequence of SEQ ID NO: 46 and the VL comprises the sequence of SEQ ID NO: 43.

5 In some embodiments, the VH comprises the sequence of SEQ ID NO: 47 and the VL comprises the sequence of SEQ ID NO: 43.

In some embodiments, the VH comprises the sequence of SEQ ID NO: 48 and the VL comprises the sequence of SEQ ID NO: 43.

10 In some embodiments, the VH comprises the sequence of SEQ ID NO: 49 and the VL comprises the sequence of SEQ ID NO: 43.

In some embodiments, the VH comprises the sequence of SEQ ID NO: 61 and the VL comprises the sequence of SEQ ID NO: 43.

In some embodiments, the VH comprises the sequence of SEQ ID NO: 62 and the VL comprises the sequence of SEQ ID NO: 43.

15 In one aspect, the disclosure is related to an antibody or antigen-binding fragment thereof that binds to MUC1 comprising

a heavy chain variable region (VH) comprising VH CDR1, VH CDR2, and VH CDR3 that are identical to VH CDR1, VH CDR2, and VH CDR3 of a selected VH sequence; and a light chain variable region (VL) comprising VL CDR1, VL CDR2, and VL CDR3 that are identical to
20 VL CDR1, VL CDR2, and VL CDR3 of a selected VL sequence, wherein the selected VH sequence and the selected VL sequence are one of the following:

the selected VH sequence is SEQ ID NO: 44, and the selected VL sequence is SEQ ID NO: 43;

25 the selected VH sequence is SEQ ID NO: 45, and the selected VL sequence is SEQ ID NO: 43;

the selected VH sequence is SEQ ID NO: 46, and the selected VL sequence is SEQ ID NO: 43;

the selected VH sequence is SEQ ID NO: 47, and the selected VL sequence is SEQ ID NO: 43;

30 the selected VH sequence is SEQ ID NO: 48, and the selected VL sequence is SEQ ID NO: 43;

the selected VH sequence is SEQ ID NO: 49, and the selected VL sequence is SEQ ID NO: 43;

35 the selected VH sequence is SEQ ID NO: 61, and the selected VL sequence is SEQ ID NO: 43; and

the selected VH sequence is SEQ ID NO: 62, and the selected VL sequence is SEQ ID NO: 43.

In some embodiments, the antibody or antigen-binding fragment thereof specifically binds to human, mouse, monkey, or dog MUC1.

40 In some embodiments, the antibody or antigen-binding fragment thereof is a human antibody or antigen-binding fragment thereof, a single-chain variable fragment (scFv), a one-armed antibody, and/or a multi-specific antibody (e.g., a bispecific antibody).

In some embodiments, the antibody or antigen-binding fragment is a human IgG1 antibody or antigen-binding fragment thereof or a human IgG4 antibody or antigen-binding fragment thereof.

In one aspect, the disclosure is related to an antibody or antigen-binding fragment thereof that cross-competes with the antibody or antigen-binding fragment thereof described herein.

In some embodiments, the antibody or antigen-binding fragment thereof comprises a fragment crystallizable region (Fc region).

In some embodiments, the Fc region has increased complement-dependent cytotoxicity (CDC) or antibody-dependent cellular cytotoxicity (ADCC).

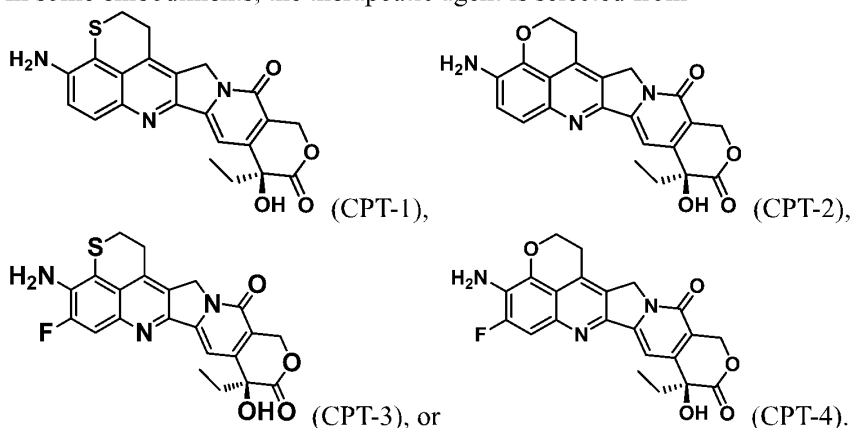
In one aspect, the disclosure is related to a chimeric antigen receptor (CAR) comprising the antibody or antigen-binding fragment thereof described herein.

In one aspect, the disclosure is related to an antibody-drug conjugate comprising the antibody or antigen-binding fragment thereof described herein covalently bound to a therapeutic agent.

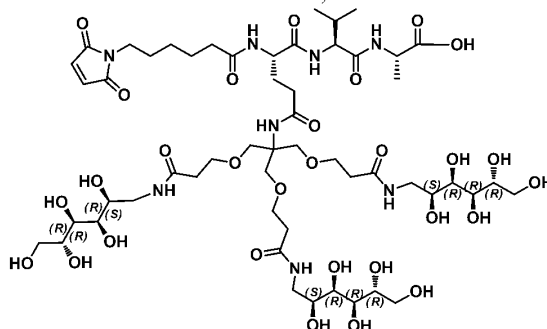
In some embodiments, the therapeutic agent is a cytotoxic or cytostatic agent.

In some embodiments, the therapeutic agent is MMAE or MMAF.

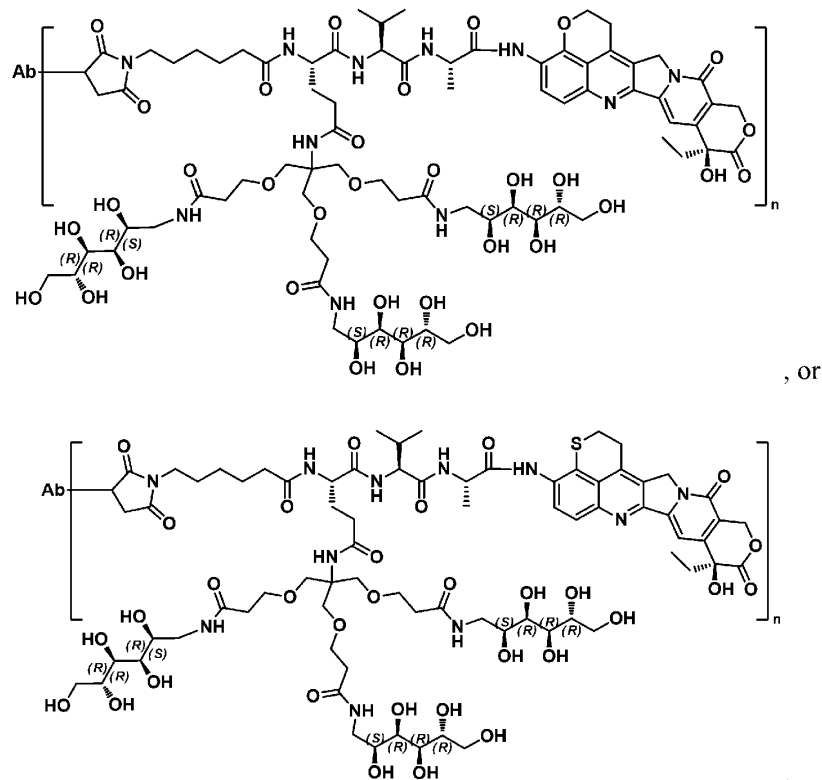
In some embodiments, the therapeutic agent is selected from



In some embodiments, the therapeutic agent is linked to the antibody or antigen-binding fragment thereof via a linker. In some embodiments, the linker has a structure of:



In some embodiments, the antibody-drug conjugate has a structure of:



5 in some embodiments, $n = 1-8$; in some embodiments, "Ab" represents the antibody or antigen-binding fragment thereof.

In some embodiments, the drug-to-antibody ratio (DAR) is about 4 or 8.

10 In one aspect, the disclosure is related to a method of treating a subject having cancer, the method comprising administering a therapeutically effective amount of a composition comprising the antibody or antigen-binding fragment thereof described herein, the CAR described herein, or the antibody-drug conjugate described herein, to the subject.

In some embodiments, the subject has a solid tumor.

15 In some embodiments, the cancer is oesophageal cancer, gastric cancer, breast cancer, ovarian cancer, bladder cancer, pancreatic cancer, lung cancer, colorectal cancer, stomach cancer, or Non-Hodgkin Lymphoma.

In some embodiments, the subject is further treated with an effective amount of an anti-4-1BB antibody, an anti-OX40 antibody, an anti-PD-1 antibody, an anti-CTLA4 antibody, an anti-CD40 antibody, or an anti-PD-L1 antibody.

20 In one aspect, the disclosure is related to a method of decreasing the rate of tumor growth, the method comprising contacting a tumor cell with an effective amount of a composition comprising an antibody or antigen-binding fragment thereof described herein, the CAR described herein, or the antibody-drug conjugate described herein.

In one aspect, the disclosure is related to a method of killing a tumor cell, the method comprising contacting a tumor cell with an effective amount of a composition comprising the antibody or antigen-binding fragment thereof described herein, the CAR described herein, or the antibody-drug conjugate described herein.

5 In one aspect, the disclosure is related to a method of increasing immune response in a subject, the method comprising administering to the subject an effective amount of a composition comprising the antibody or antigen-binding fragment thereof described herein, the CAR described herein, or the antibody-drug conjugate described herein.

10 In one aspect, the disclosure is related to a pharmaceutical composition comprising the antibody or antigen-binding fragment thereof described herein, and a pharmaceutically acceptable carrier.

In one aspect, the disclosure is related to a pharmaceutical composition comprising the antibody drug conjugate described herein, and a pharmaceutically acceptable carrier.

15 As used herein, the term “cancer” refers to cells having the capacity for autonomous growth. Examples of such cells include cells having an abnormal state or condition characterized by rapidly proliferating cell growth. The term is meant to include cancerous growths, e.g., tumors; oncogenic processes, metastatic tissues, and malignantly transformed cells, tissues, or organs, irrespective of histopathologic type or stage of invasiveness. Also included are malignancies of the various organ systems, such as respiratory, cardiovascular, renal, reproductive, hematological, neurological, hepatic, gastrointestinal, and endocrine systems; as
20 well as adenocarcinomas which include malignancies such as most colon cancers, renal-cell carcinoma, prostate cancer and/or testicular tumors, non-small cell carcinoma of the lung, and cancer of the small intestine. Cancer that is “naturally arising” includes any cancer that is not experimentally induced by implantation of cancer cells into a subject, and includes, for example, spontaneously arising cancer, cancer caused by exposure of a patient to a carcinogen(s), cancer resulting from insertion of a transgenic oncogene or knockout of a tumor suppressor gene, and cancer caused by infections, e.g., viral infections. The term “carcinoma” is art recognized and refers to malignancies of epithelial or endocrine tissues. The term also includes carcinosarcomas, which include malignant tumors composed of carcinomatous and sarcomatous tissues. An
30 “adenocarcinoma” refers to a carcinoma derived from glandular tissue or in which the tumor cells form recognizable glandular structures. The term “sarcoma” is art recognized and refers to malignant tumors of mesenchymal derivation. The term “hematopoietic neoplastic disorders” includes diseases involving hyperplastic/neoplastic cells of hematopoietic origin. A hematopoietic neoplastic disorder can arise from myeloid, lymphoid or erythroid lineages, or
35 precursor cells thereof.

As used herein, the term “antibody” refers to any antigen-binding molecule that contains at least one (e.g., one, two, three, four, five, or six) complementary determining region (CDR) (e.g., any of the three CDRs from an immunoglobulin light chain or any of the three CDRs from an immunoglobulin heavy chain) and is capable of specifically binding to an epitope. Non-
40 limiting examples of antibodies include: monoclonal antibodies, polyclonal antibodies, multi-specific antibodies (e.g., bi-specific antibodies), single-chain antibodies, chimeric antibodies, human antibodies, and humanized antibodies. In some embodiments, an antibody can contain an Fc region of a human antibody. The term antibody also includes derivatives, e.g., bi-specific

antibodies, single-chain antibodies, diabodies, linear antibodies, and multi-specific antibodies formed from antibody fragments.

As used herein, the term “antigen-binding fragment” refers to a portion of a full-length antibody, wherein the portion of the antibody is capable of specifically binding to an antigen. In some embodiments, the antigen-binding fragment contains at least one variable domain (e.g., a variable domain of a heavy chain or a variable domain of light chain). Non-limiting examples of antibody fragments include, e.g., Fab, Fab', F(ab')₂, and Fv fragments.

As used herein, the term “human antibody” refers to an antibody that is encoded by an endogenous nucleic acid (e.g., rearranged human immunoglobulin heavy or light chain locus) present in a human. In some embodiments, a human antibody is collected from a human or produced in a human cell culture (e.g., human hybridoma cells). In some embodiments, a human antibody is produced in a non-human cell (e.g., a mouse or hamster cell line). In some embodiments, a human antibody is produced in a bacterial or yeast cell. In some embodiments, a human antibody is produced in a transgenic non-human animal (e.g., a bovine) containing an unrearranged or rearranged human immunoglobulin locus (e.g., heavy or light chain human immunoglobulin locus).

As used herein, the term “chimeric antibody” refers to an antibody that contains a sequence present in at least two different antibodies (e.g., antibodies from two different mammalian species such as a human and a mouse antibody). A non-limiting example of a chimeric antibody is an antibody containing the variable domain sequences (e.g., all or part of a light chain and/or heavy chain variable domain sequence) of a non-human (e.g., mouse) antibody and the constant domains of a human antibody. Additional examples of chimeric antibodies are described herein and are known in the art.

As used herein, the term “humanized antibody” refers to a non-human antibody which contains minimal sequence derived from a non-human (e.g., mouse) immunoglobulin and contains sequences derived from a human immunoglobulin. In non-limiting examples, humanized antibodies are human antibodies (recipient antibody) in which hypervariable (e.g., CDR) region residues of the recipient antibody are replaced by hypervariable (e.g., CDR) region residues from a non-human antibody (e.g., a donor antibody), e.g., a mouse, rat, or rabbit antibody, having the desired specificity, affinity, and capacity. In some embodiments, the Fv framework residues of the human immunoglobulin are replaced by corresponding non-human (e.g., mouse) immunoglobulin residues. In some embodiments, humanized antibodies may contain residues which are not found in the recipient antibody or in the donor antibody. These modifications can be made to further refine antibody performance. In some embodiments, the humanized antibody contains substantially all of at least one, and typically two, variable domains, in which all or substantially all of the hypervariable loops (CDRs) correspond to those of a non-human (e.g., mouse) immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin. The humanized antibody can also contain at least a portion of an immunoglobulin constant region (Fc), typically, that of a human immunoglobulin. Humanized antibodies can be produced using molecular biology methods known in the art. Non-limiting examples of methods for generating humanized antibodies are described herein.

As used herein, the term “single-chain antibody” refers to a single polypeptide that contains at least two immunoglobulin variable domains (e.g., a variable domain of a mammalian

immunoglobulin heavy chain or light chain) that is capable of specifically binding to an antigen. Non-limiting examples of single-chain antibodies are described herein.

As used herein, the term “multimeric antibody” refers to an antibody that contains four or more (e.g., six, eight, or ten) immunoglobulin variable domains. In some embodiments, the multimeric antibody is able to crosslink one target molecule (e.g., MUC1) to at least one second target molecule on the surface of a mammalian cell (e.g., a human T-cell).

As used herein, the terms “subject” and “patient” are used interchangeably throughout the specification and describe an animal, human or non-human, to whom treatment according to the methods of the present invention is provided. Veterinary and non-veterinary applications are contemplated by the present invention. Human patients can be adult humans or juvenile humans (e.g., humans below the age of 18 years old). In addition to humans, patients include but are not limited to mice, rats, hamsters, guinea-pigs, rabbits, ferrets, cats, dogs, and primates. Included are, for example, non-human primates (e.g., monkey, chimpanzee, gorilla, and the like), rodents (e.g., rats, mice, gerbils, hamsters, ferrets, rabbits), lagomorphs, swine (e.g., pig, miniature pig), equine, canine, feline, bovine, and other domestic, farm, and zoo animals.

As used herein, when referring to an antibody, the phrases “specifically binding” and “specifically binds” mean that the antibody interacts with its target molecule (e.g., MUC1) preferably to other molecules, because the interaction is dependent upon the presence of a particular structure (i.e., the antigenic determinant or epitope) on the target molecule; in other words, the reagent is recognizing and binding to molecules that include a specific structure rather than to all molecules in general. An antibody that specifically binds to the target molecule may be referred to as a target-specific antibody. For example, an antibody that specifically binds to a MUC1 molecule may be referred to as a MUC1-specific antibody or an anti-MUC1 antibody.

As used herein, the terms “polypeptide,” “peptide,” and “protein” are used interchangeably to refer to polymers of amino acids of any length of at least two amino acids.

As used herein, the terms “polynucleotide,” “nucleic acid molecule,” and “nucleic acid sequence” are used interchangeably herein to refer to polymers of nucleotides of any length of at least two nucleotides, and include, without limitation, DNA, RNA, DNA/RNA hybrids, and modifications thereof.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Methods and materials are described herein for use in the present invention; other, suitable methods and materials known in the art can also be used. The materials, methods, and examples are illustrative only and not intended to be limiting. All publications, patent applications, patents, sequences, database entries, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control.

Other features and advantages of the invention will be apparent from the following detailed description and figures, and from the claims.

DESCRIPTION OF DRAWINGS

FIG. 1 shows the average tumor volume in different groups of B-NDG mice that were injected with stomach cancer NUGC-4 cells, and were treated with PBS or ADCs.

5 **FIG. 2** shows the average tumor volume in different groups of B-NDG mice that were engrafted in the right flank with patient-derived lung cancer fragments (2 mm×2 mm×2 mm), and were treated with PBS or ADCs.

FIG. 3 lists heavy chain variable region CDR sequences of anti-MUC1 antibodies as defined by Kabat definition.

10 **FIG. 4** lists heavy chain variable region CDR sequences of anti-MUC1 antibodies as defined by Chothia definition.

FIG. 5 lists selected amino acid sequences discussed in the disclosure.

FIG. 6 shows the binding activities of 10B1, 10B1-AG or 10B1-AA to ZR-75-1 cells.

FIG. 7 shows the endocytosis rates of 10B1, 10B1-AG or 10B1-AA in ZR-75-1 cells.

FIG. 8 shows the binding activities of anti-MUC1 antibodies to NUGC-4 cells.

15 **FIG. 9** shows the average body weights in different groups of B-NDG mice that were engrafted with HCC70 cells, and were treated with PBS or ADCs.

FIG. 10 shows the average tumor volumes in different groups of B-NDG mice that were engrafted with HCC70 cells, and were treated with PBS or ADCs.

20 **FIG. 11** shows the average tumor volumes in different groups of B-NDG mice that were engrafted with patient-derived breast tumor fragments (2 mm×2 mm×2 mm), and were treated with PBS or ADCs.

FIG. 12 shows the average tumor volumes in different groups of B-NDG mice that were engrafted with patient-derived pancreatic tumor fragments (2 mm×2 mm×2 mm), and were treated with PBS or ADCs.

25 **FIG. 13** shows the average tumor volumes in different groups of B-NDG mice that were engrafted with patient-derived colorectal tumor fragments (2 mm×2 mm×2 mm), and were treated with PBS or ADCs.

DETAILED DESCRIPTION

30 The present disclosure provides examples of antibodies, antigen-binding fragment thereof, that bind to MUC1.

MUC1

35 Mucin 1 (MUC1; also known as episialin, PEM, H23Ag, EMA, CA15-3, and MCA) is a single pass type I transmembrane protein with a heavily glycosylated extracellular domain that extends up to 200–500 nm from the cell surface. MUC1 is normally expressed in the glandular or luminal epithelial cells of the mammary gland, esophagus, stomach, duodenum, pancreas, uterus, prostate, and lungs, and to a lesser extent, in hematopoietic cells. It is absent in the skin epithelium and in mesenchymal cells. In healthy tissues, MUC1 provides protection to the
40 underlying epithelia. The extended negatively charged sugar branches of MUC1 create a physical barrier and impart an anti-adhesive property to MUC1, limiting accessibility and

preventing pathogenic colonization. The sugar chains oligomerize to form a mucinous gel that lubricates and protects the underlying epithelia from desiccation, changes in pH, pollutants, and microbes.

Aberrantly glycosylated MUC1 is overexpressed in most human epithelial cancers and has gained remarkable attention as an oncogenic molecule.

The MUC1 gene encodes a single polypeptide chain which, due to conformational stress, is autoproteolytically cleaved immediately after translation at the GSVVV motif, located within the Sea urchin sperm protein enterokinase and agrin (SEA) domain, into two peptide fragments: the longer N-terminal subunit (MUC1-N) and the shorter C-terminal subunit (MUC1-C).

Extracellularly, the two subunits remain associated through stable hydrogen bonds.

MUC1-N is composed of the proline, threonine, and serine-rich (PTS) domain and the SEA domain. The PTS domain, also designated as the variable number tandem repeat (VNTR) region, is encoded by a highly polymorphic exon encoding for multiple 20–21 amino acid sequence repeats. In northern Europeans, the VNTR is composed of 20–120 repeats, with 40–80 repeats being the most common. The amino acid sequence of the VNTR region can vary in different cancer cell lines, consistent with the highly polymorphic nature of this motif. The VNTR region is flanked on both ends by a short degenerate sequence which bears subtle sequence similarity to the VNTR region. Unlike human MUC1, mouse Muc1 contains 16 amino acid long tandem repeats and has only 34% sequence similarity with its human counterpart.

MUC1-C is short, comprising a 58 amino acid ECD, a 28 amino acid transmembrane domain (TMD), and a 72 amino acid cytoplasmic tail (CT). Based on the N-glycosylation state of the ECD, the molecular weight of MUC1-C can vary from 17 to 25 kDa. Under normal conditions, MUC1 exists on the plasma membrane as a heterodimeric complex. However, the complex dissociates following stimulation with the proinflammatory cytokines interferon- γ (IFN- γ) and tumor necrosis factor- α (TNF- α), and this is catalyzed by the sheddase activities of the enzymes including TNF- α converting enzyme (TACE, also called disintegrin and metalloprotease do-main containing protein-17; ADAM17) and matrix metalloproteases (MMPs). These enzymes cause release of MUC1-N from MUC1-C, and also catalyze the cleavage of the ECD of MUC1-C, thereby generating smaller peptide fragments MUC1* and MUC1-CTF15. Unlike MUC1-C, the cleavage products of sheddases contain a shorter ECD. MUC1* consists of a 45 amino acid ECD and promotes tumor growth. MUC1* has also been detected in human embryonic stem cells (hESCs), where it functions as a growth factor receptor for a metastasis-associated protein (NM23-H1). MUC1 CTF15 contains a shorter ECD ‘stub’ of 27 amino acids that is recognized by nicartrin, a substrate receptor for γ -secretase. γ -Secretase further cleaves MUC1-C into shorter peptide fragments with a molecular mass of 8–10 kDa, which are rapidly degraded. In most normal and tumor cells, MUC1 CTF15 is undetectable as it is rapidly degraded.

In cancer cells, a significant amount of MUC1-C is detected intracellularly, particularly within the mitochondria and nucleus. MUC1-C monomers dimerize along the CQC motif to generate functional homodimers that translocate to the nucleus via importin- β and nucleoprotein 62 (Nup62). Within the nucleus, MUC1-C is found mostly in the nuclear matrix and nucleoli.

MUC1 is overexpressed in cancer cells and the loss of cell polarity causes TA-MUC1 to be redistributed over the cell surface and within the cytoplasm. Lack of cell polarity also causes

the redistribution of cell surface growth factors that are normally restricted to the basolateral surface of epithelial cells. Growth factors juxtaposed to MUC1 and intracellular kinases such as ZAP-70, PKC-g, GSK-3b, and c-Src phosphorylate serine, tyrosine, and threonine residues on MUC1 CT. It is also thought that hypoglycosylation unmasks the peptide core of TA-MUC1 allowing MUC1-N cleavage and release by extracellular proteases. MUC1-N release induces conformational changes in MUC1-C that alter its ligand status and subsequently activates downstream cell signaling pathways such as the mitogen-activated protein kinase (MAPK), phosphatidylinositol 3-kinase (PI3K/Akt), and wingless type (Wnt) pathways. As a result, MUC1- positive pancreas, breast, lung, and colon cancer cells commonly display hyperactivation of these critical signaling pathways. MUC1-C also associates with various transcription factors (STAT3, NF-kB, p53, and b-catenin) and binds the target gene promoter region to drive their expression. Several studies have indicated that MUC1 plays a critical role in the transcriptional regulation of genes associated with tumor invasion, metastasis, angiogenesis, proliferation, apoptosis, drug resistance, inflammation, and immune regulation.

A detailed review of MUC1 and its functions can be found in Nath, Sritama, and Pinku Mukherjee. "MUC1: a multifaceted oncoprotein with a key role in cancer progression." Trends in molecular medicine 20.6 (2014): 332-342, which is incorporated by reference in its entirety.

The present disclosure provides several anti-MUC1 antibodies, antigen-binding fragments thereof, and methods of using these anti-MUC1 antibodies and antigen-binding fragments to inhibit tumor growth, treat cancers, and to treat autoimmune diseases.

Antibodies and Antigen Binding Fragments

The present disclosure provides anti-MUC1 antibodies and antigen-binding fragments thereof. In general, antibodies (also called immunoglobulins) are made up of two classes of polypeptide chains, light chains and heavy chains. A non-limiting antibody of the present disclosure can be an intact, four immunoglobulin chain antibody comprising two heavy chains and two light chains. The heavy chain of the antibody can be of any isotype including IgM, IgG, IgE, IgA, or IgD or sub-isotype including IgG1, IgG2, IgG2a, IgG2b, IgG3, IgG4, IgE1, IgE2, etc. The light chain can be a kappa light chain or a lambda light chain. An antibody can comprise two identical copies of a light chain and two identical copies of a heavy chain. The heavy chains, which each contain one variable domain (or variable region, V_H) and multiple constant domains (or constant regions), bind to one another via disulfide bonding within their constant domains to form the "stem" of the antibody. The light chains, which each contain one variable domain (or variable region, V_L) and one constant domain (or constant region), each bind to one heavy chain via disulfide binding. The variable region of each light chain is aligned with the variable region of the heavy chain to which it is bound. The variable regions of both the light chains and heavy chains contain three hypervariable regions sandwiched between more conserved framework regions (FR).

These hypervariable regions, known as the complementary determining regions (CDRs), form loops that comprise the principle antigen binding surface of the antibody. The four framework regions largely adopt a beta-sheet conformation and the CDRs form loops connecting, and in some cases forming part of, the beta-sheet structure. The CDRs in each chain are held in

close proximity by the framework regions and, with the CDRs from the other chain, contribute to the formation of the antigen-binding region.

Methods for identifying the CDR regions of an antibody by analyzing the amino acid sequence of the antibody are well known, and a number of definitions of the CDRs are commonly used. The Kabat definition is based on sequence variability, and the Chothia definition is based on the location of the structural loop regions. These methods and definitions are described in, e.g., Martin, "Protein sequence and structure analysis of antibody variable domains," *Antibody engineering*, Springer Berlin Heidelberg, 2001. 422-439; Abhinandan, et al. "Analysis and improvements to Kabat and structurally correct numbering of antibody variable domains," *Molecular immunology* 45.14 (2008): 3832-3839; Wu, T.T. and Kabat, E.A. (1970) *J. Exp. Med.* 132: 211-250; Martin et al., *Methods Enzymol.* 203:121-53 (1991); Morea et al., *Biophys Chem.* 68(1-3):9-16 (Oct. 1997); Morea et al., *J Mol Biol.* 275(2):269-94 (Jan. 1998); Chothia et al., *Nature* 342(6252):877-83 (Dec. 1989); Ponomarenko and Bourne, *BMC Structural Biology* 7:64 (2007); each of which is incorporated herein by reference in its entirety.

The CDRs are important for recognizing an epitope of an antigen. As used herein, an "epitope" is the smallest portion of a target molecule capable of being specifically bound by the antigen binding domain of an antibody. The minimal size of an epitope may be about three, four, five, six, or seven amino acids, but these amino acids need not be in a consecutive linear sequence of the antigen's primary structure, as the epitope may depend on an antigen's three-dimensional configuration based on the antigen's secondary and tertiary structure.

In some embodiments, the antibody is an intact immunoglobulin molecule (e.g., IgG1, IgG2a, IgG2b, IgG3, IgM, IgD, IgE, IgA). The IgG subclasses (IgG1, IgG2, IgG3, and IgG4) are highly conserved, differ in their constant region, particularly in their hinges and upper CH2 domains. The sequences and differences of the IgG subclasses are known in the art, and are described, e.g., in Vidarsson, et al, "IgG subclasses and allotypes: from structure to effector functions." *Frontiers in immunology* 5 (2014); Irani, et al. "Molecular properties of human IgG subclasses and their implications for designing therapeutic monoclonal antibodies against infectious diseases." *Molecular immunology* 67.2 (2015): 171-182; Shakib, Farouk, ed. *The human IgG subclasses: molecular analysis of structure, function and regulation*. Elsevier, 2016; each of which is incorporated herein by reference in its entirety.

The antibody can also be an immunoglobulin molecule that is derived from any species (e.g., human, rodent, mouse, camelid). Antibodies disclosed herein also include, but are not limited to, polyclonal, monoclonal, monospecific, polyspecific antibodies, and chimeric antibodies that include an immunoglobulin binding domain fused to another polypeptide. The term "antigen binding domain" or "antigen binding fragment" is a portion of an antibody that retains specific binding activity of the intact antibody, i.e., any portion of an antibody that is capable of specific binding to an epitope on the intact antibody's target molecule. It includes, e.g., Fab, Fab', F(ab')₂, and variants of these fragments. Thus, in some embodiments, an antibody or an antigen binding fragment thereof can be, e.g., a scFv, a Fv, a Fd, a dAb, a bispecific antibody, a bispecific scFv, a diabody, a linear antibody, a single-chain antibody molecule, a multi-specific antibody formed from antibody fragments, and any polypeptide that includes a binding domain which is, or is homologous to, an antibody binding domain. Non-limiting examples of antigen binding domains include, e.g., the heavy chain and/or light chain CDRs of an intact antibody, the

heavy and/or light chain variable regions of an intact antibody, full length heavy or light chains of an intact antibody, or an individual CDR from either the heavy chain or the light chain of an intact antibody.

In some embodiments, the antigen binding fragment can form a part of a chimeric antigen receptor (CAR). In some embodiments, the chimeric antigen receptor are fusions of single-chain variable fragments (scFv) as described herein, fused to CD3-zeta transmembrane- and endodomain. In some embodiments, the chimeric antigen receptor also comprises intracellular signaling domains from various costimulatory protein receptors (e.g., CD28, 41BB, ICOS). In some embodiments, the chimeric antigen receptor comprises multiple signaling domains, e.g., CD3z-CD28-41BB or CD3z-CD28-OX40, to increase potency. Thus, in one aspect, the disclosure further provides cells (e.g., T cells) that express the chimeric antigen receptors as described herein.

In some embodiments, the scFV has one heavy chain variable domain, and one light chain variable domain. In some embodiments, the scFV has two heavy chain variable domains, and two light chain variable domains.

In some embodiments, sequences (e.g., CDRs or VH/VL sequences) of the antibody or antigen-binding fragment thereof described herein can be used to generate a bispecific antibody targeting MUC1 and an addition antigen (e.g., PD-1).

20 **Anti-MUC1 Antibodies and Antigen-Binding Fragments**

The disclosure provides antibodies and antigen-binding fragments thereof that specifically bind to MUC1 (e.g., human MUC1). The antibodies and antigen-binding fragments described herein are capable of binding to MUC1. These antibodies can be agonists or antagonists. In some embodiments, these antibodies can increase immune response. In some embodiments, these antibodies can block MUC1 activity.

The disclosure provides e.g., anti-MUC1 antibodies 9H4, 10B1, 10G12, 11A6, 11D6, 11E8, the chimeric antibodies thereof, and the human or humanized antibodies thereof.

The CDR sequences for 9H4, and 9H4 derived antibodies (e.g., human or humanized antibodies) include CDRs of the heavy chain variable domain, SEQ ID NOs: 4-6, and CDRs of the light chain variable domain, SEQ ID NOs: 1-3 as defined by Kabat definition. The CDRs can also be defined by Chothia definition. Under the Chothia definition, the CDR sequences of the heavy chain variable domain are set forth in SEQ ID NOs: 25-27 and CDR sequences of the light chain variable domain are set forth in SEQ ID NOs: 22-24.

Similarly, the CDR sequences for 10B1, and 10B1 derived antibodies include CDRs of the heavy chain variable domain, SEQ ID NOs: 7-9, and CDRs of the light chain variable domain, SEQ ID NOs: 1-3, as defined by Kabat definition. Under Chothia definition, the CDR sequences of the heavy chain variable domain are set forth in SEQ ID NOs: 28-30, and CDRs of the light chain variable domain are set forth in SEQ ID NOs: 22-24.

The CDR sequences for 10G12, and 10G12 derived antibodies include CDRs of the heavy chain variable domain, SEQ ID NOs: 10-12, and CDRs of the light chain variable domain, SEQ ID NOs: 1-3, as defined by Kabat definition. Under Chothia definition, the CDR sequences

of the heavy chain variable domain are set forth in SEQ ID NOs: 31-33, and CDRs of the light chain variable domain are set forth in SEQ ID NOs: 22-24.

The CDR sequences for 11A6, and 11A6 derived antibodies include CDRs of the heavy chain variable domain, SEQ ID NOs: 13-15, and CDRs of the light chain variable domain, SEQ ID NOs: 1-3, as defined by Kabat definition. Under Chothia definition, the CDR sequences of the heavy chain variable domain are set forth in SEQ ID NOs: 34-36, and CDRs of the light chain variable domain are set forth in SEQ ID NOs: 22-24.

The CDR sequences for 11D6, and 11D6 derived antibodies include CDRs of the heavy chain variable domain, SEQ ID NOs: 16-18, and CDRs of the light chain variable domain, SEQ ID NOs: 1-3, as defined by Kabat definition. Under Chothia definition, the CDR sequences of the heavy chain variable domain are set forth in SEQ ID NOs: 37-39, and CDRs of the light chain variable domain are set forth in SEQ ID NOs: 22-24.

The CDR sequences for 11E8, and 11E8 derived antibodies include CDRs of the heavy chain variable domain, SEQ ID NOs: 19-21, and CDRs of the light chain variable domain, SEQ ID NOs: 1-3, as defined by Kabat definition. Under Chothia definition, the CDR sequences of the heavy chain variable domain are set forth in SEQ ID NOs: 40-42, and CDRs of the light chain variable domain are set forth in SEQ ID NOs: 22-24.

The amino acid sequence for the heavy chain variable region of 9H4 antibody is set forth in SEQ ID NO: 44. The amino acid sequence for the light chain variable region of 9H4 antibody is set forth in SEQ ID NO: 43.

The amino acid sequence for the heavy chain variable region of 10B1 antibody is set forth in SEQ ID NO: 45. The amino acid sequence for the light chain variable region of 10B1 antibody is set forth in SEQ ID NO: 43.

The amino acid sequence for the heavy chain variable region of 10G12 antibody is set forth in SEQ ID NO: 46. The amino acid sequence for the light chain variable region of 10G12 antibody is set forth in SEQ ID NO: 43.

The amino acid sequence for the heavy chain variable region of 11A6 antibody is set forth in SEQ ID NO: 47. The amino acid sequence for the light chain variable region of 10G12 antibody is set forth in SEQ ID NO: 43.

The amino acid sequence for the heavy chain variable region of 11D6 antibody is set forth in SEQ ID NO: 48. The amino acid sequence for the light chain variable region of 10G12 antibody is set forth in SEQ ID NO: 43.

The amino acid sequence for the heavy chain variable region of 11E8 antibody is set forth in SEQ ID NO: 49. The amino acid sequence for the light chain variable region of 10G12 antibody is set forth in SEQ ID NO: 43.

In some embodiments, the disclosure relates to CDR-modified VH or VL. In some embodiments, the “DG” in VH CDR2 of an antibody described herein is mutated to “AG” or “AA” to create CDR-modified VH. For example, in one embodiment, the “DG” in VH CDR2 of 10B1 is mutated to “AG” and the resulting CDR-modified VH is named 10B1-AG VH. The sequences of some of these CDR-modified VH can be found in **FIG. 5**.

The amino acid sequences for heavy chain variable regions and light variable regions of the modified antibodies are also provided. In some embodiments, the heavy chain variable region is at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to any

one of SEQ ID NOs: 44-49 and SEQ ID NOs: 61-62. In some embodiments, the light chain variable region is at least 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to SEQ ID NO: 43. The heavy chain variable region sequence can be paired with the corresponding light chain variable region sequence, and together they bind to MUC1.

5 Humanization percentage means the percentage identity of the heavy chain or light chain variable region sequence as compared to human antibody sequences in International Immunogenetics Information System (IMGT) database. The top hit means that the heavy chain or light chain variable region sequence is closer to a particular species than to other species. In some embodiments, humanization percentage is greater than 80%, 81%, 82%, 83%, 84%, 85%,
10 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, or 95%. A detailed description regarding how to determine humanization percentage and how to determine top hits is known in the art, and is described, e.g., in Jones, et al. "The INNs and outs of antibody nonproprietary names." MAbs. Vol. 8. No. 1. Taylor & Francis, 2016, which is incorporated herein by reference in its entirety. A high humanization percentage often has various advantages, e.g., more safe and more
15 effective in humans, more likely to be tolerated by a human subject, and/or less likely to have side effects. In some embodiments, the variable regions are fully human, e.g., derived from human heavy chain immunoglobulin locus sequences (e.g., recombination of human IGHV, human IGHD, and human IGHJ genes), and/or human kappa chain immunoglobulin locus sequences (e.g., recombination of human IGKV and human IGKJ genes).

20 Furthermore, in some embodiments, the antibodies or antigen-binding fragments thereof described herein can also contain one, two, or three heavy chain variable region CDRs selected from the group of SEQ ID NOs: 4-6, SEQ ID NOs: 7-9, SEQ ID NOs: 7, 63, 9, SEQ ID NOs: 7, 64, 9, SEQ ID NOs: 10-12, SEQ ID NOs: 13-15, SEQ ID NOs: 16-18, SEQ ID NOs: 19-21, SEQ ID NOs: 25-27, SEQ ID NOs: 28-30, SEQ ID NOs: 28, 65, 30, SEQ ID NOs: 28, 66, 30, SEQ ID
25 NOs: 31-33, SEQ ID NOs: 34-36, SEQ ID NOs: 37-39, and SEQ ID NOs: 40-42, and/or one, two, or three light chain variable region CDRs selected from the group of SEQ ID NOs: 1-3 and SEQ ID NOs: 22-24.

In some embodiments, the antibodies can have a heavy chain variable region (VH) comprising complementarity determining regions (CDRs) 1, 2, 3, wherein the CDR1 region
30 comprises or consists of an amino acid sequence that is at least 80%, 85%, 90%, or 95% identical to a selected VH CDR1 amino acid sequence, the CDR2 region comprises or consists of an amino acid sequence that is at least 80%, 85%, 90%, or 95% identical to a selected VH CDR2 amino acid sequence, and the CDR3 region comprises or consists of an amino acid sequence that is at least 80%, 85%, 90%, or 95% identical to a selected VH CDR3 amino acid sequence. In
35 some embodiments, the antibody can have a light chain variable region (VL) comprising CDRs 1, 2, 3, wherein the CDR1 region comprises or consists of an amino acid sequence that is at least 80%, 85%, 90%, or 95% identical to a selected VL CDR1 amino acid sequence, the CDR2 region comprises or consists of an amino acid sequence that is at least 80%, 85%, 90%, or 95% identical to a selected VL CDR2 amino acid sequence, and the CDR3 region comprises or
40 consists of an amino acid sequence that is at least 80%, 85%, 90%, or 95% identical to a selected VL CDR3 amino acid sequence. The selected VH CDRs 1, 2, 3 amino acid sequences and the selected VL CDRs, 1, 2, 3 amino acid sequences are shown in **FIG. 3** (Kabat CDR) and **FIG. 4** (Chothia CDR).

In some embodiments, the antibody or an antigen-binding fragment described herein can contain a heavy chain variable domain containing one, two, or three of the CDRs of SEQ ID NO: 4 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 5 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 6 with zero, one or two amino acid insertions, deletions, or substitutions.

In some embodiments, the antibody or an antigen-binding fragment described herein can contain a heavy chain variable domain containing one, two, or three of the CDRs of SEQ ID NO: 7 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 8 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 9 with zero, one or two amino acid insertions, deletions, or substitutions.

In some embodiments, the antibody or an antigen-binding fragment described herein can contain a heavy chain variable domain containing one, two, or three of the CDRs of SEQ ID NO: 10 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 11 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 12 with zero, one or two amino acid insertions, deletions, or substitutions.

In some embodiments, the antibody or an antigen-binding fragment described herein can contain a heavy chain variable domain containing one, two, or three of the CDRs of SEQ ID NO: 13 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 14 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 15 with zero, one or two amino acid insertions, deletions, or substitutions.

In some embodiments, the antibody or an antigen-binding fragment described herein can contain a heavy chain variable domain containing one, two, or three of the CDRs of SEQ ID NO: 16 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 17 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 18 with zero, one or two amino acid insertions, deletions, or substitutions.

In some embodiments, the antibody or an antigen-binding fragment described herein can contain a heavy chain variable domain containing one, two, or three of the CDRs of SEQ ID NO: 19 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 20 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 21 with zero, one or two amino acid insertions, deletions, or substitutions.

In some embodiments, the antibody or an antigen-binding fragment described herein can contain a heavy chain variable domain containing one, two, or three of the CDRs of SEQ ID NO: 25 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 26 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 27 with zero, one or two amino acid insertions, deletions, or substitutions.

In some embodiments, the antibody or an antigen-binding fragment described herein can contain a heavy chain variable domain containing one, two, or three of the CDRs of SEQ ID NO: 28 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 29 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 30 with zero, one or two amino acid insertions, deletions, or substitutions.

In some embodiments, the antibody or an antigen-binding fragment described herein can contain a heavy chain variable domain containing one, two, or three of the CDRs of SEQ ID NO: 31 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 32 with

zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 33 with zero, one or two amino acid insertions, deletions, or substitutions.

5 In some embodiments, the antibody or an antigen-binding fragment described herein can contain a heavy chain variable domain containing one, two, or three of the CDRs of SEQ ID NO: 34 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 35 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 36 with zero, one or two amino acid insertions, deletions, or substitutions.

10 In some embodiments, the antibody or an antigen-binding fragment described herein can contain a heavy chain variable domain containing one, two, or three of the CDRs of SEQ ID NO: 37 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 38 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 39 with zero, one or two amino acid insertions, deletions, or substitutions.

15 In some embodiments, the antibody or an antigen-binding fragment described herein can contain a heavy chain variable domain containing one, two, or three of the CDRs of SEQ ID NO: 40 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 41 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 42 with zero, one or two amino acid insertions, deletions, or substitutions.

20 In some embodiments, the antibody or an antigen-binding fragment described herein can contain a heavy chain variable domain containing one, two, or three of the CDRs of SEQ ID NO: 7 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 63 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 9 with zero, one or two amino acid insertions, deletions, or substitutions.

25 In some embodiments, the antibody or an antigen-binding fragment described herein can contain a heavy chain variable domain containing one, two, or three of the CDRs of SEQ ID NO: 7 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 64 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 9 with zero, one or two amino acid insertions, deletions, or substitutions.

30 In some embodiments, the antibody or an antigen-binding fragment described herein can contain a heavy chain variable domain containing one, two, or three of the CDRs of SEQ ID NO: 28 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 65 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 30 with zero, one or two amino acid insertions, deletions, or substitutions.

35 In some embodiments, the antibody or an antigen-binding fragment described herein can contain a heavy chain variable domain containing one, two, or three of the CDRs of SEQ ID NO: 28 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 66 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 30 with zero, one or two amino acid insertions, deletions, or substitutions.

40 In some embodiments, the antibody or an antigen-binding fragment described herein can contain a light chain variable domain containing one, two, or three of the CDRs of SEQ ID NO: 1 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 2 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 3 with zero, one or two amino acid insertions, deletions, or substitutions.

In some embodiments, the antibody or an antigen-binding fragment described herein can contain a light chain variable domain containing one, two, or three of the CDRs of SEQ ID NO: 22 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 23 with zero, one or two amino acid insertions, deletions, or substitutions; SEQ ID NO: 24 with zero, one or two amino acid insertions, deletions, or substitutions.

The insertions, deletions, and substitutions can be within the CDR sequence, or at one or both terminal ends of the CDR sequence. In some embodiments, the CDR is determined based on Kabat definition. In some embodiments, the CDR is determined based on Chothia definition. In some embodiments, the CDR is determined based on a combination of Kabat and Chothia definition.

The disclosure also provides antibodies or antigen-binding fragments thereof that bind to MUC1. The antibodies or antigen-binding fragments thereof contain a heavy chain variable region (VH) comprising or consisting of an amino acid sequence that is at least 80%, 85%, 90%, or 95% identical to a selected VH sequence, and a light chain variable region (VL) comprising or consisting of an amino acid sequence that is at least 80%, 85%, 90%, or 95% identical to a selected VL sequence. In some embodiments, the selected VH sequence is SEQ ID NO: 44, and the selected VL sequence is SEQ ID NO: 43. In some embodiments, the selected VH sequence is SEQ ID NO: 45 and the selected VL sequence is SEQ ID NO: 43. In some embodiments, the selected VH sequence is SEQ ID NO: 46, and the selected VL sequence is SEQ ID NO: 43. In some embodiments, the selected VH sequence is SEQ ID NO: 47, and the selected VL sequence is SEQ ID NO: 43. In some embodiments, the selected VH sequence is SEQ ID NO: 48, and the selected VL sequence is SEQ ID NO: 43. In some embodiments, the selected VH sequence is SEQ ID NO: 49, and the selected VL sequence is SEQ ID NO: 43. In some embodiments, the selected VH sequence is SEQ ID NO: 61, and the selected VL sequence is SEQ ID NO: 43. In some embodiments, the selected VH sequence is SEQ ID NO: 62, and the selected VL sequence is SEQ ID NO: 43.

The disclosure also provides antibodies or antigen-binding fragments thereof that can compete with the antibodies described herein. In some aspects, the antibodies or antigen-binding fragments can bind to the same epitope as the antibodies described herein.

The present disclosure also provides an antibody or antigen-binding fragment thereof that cross-competes with any antibody or antigen-binding fragment as described herein. The cross-competing assay is known in the art, and is described e.g., in Moore et al., "Antibody cross-competition analysis of the human immunodeficiency virus type 1 gp120 exterior envelope glycoprotein." *Journal of virology* 70.3 (1996): 1863-1872, which is incorporated herein reference in its entirety. In one aspect, the present disclosure also provides an antibody or antigen-binding fragment thereof that binds to the same epitope or region as any antibody or antigen-binding fragment as described herein. The epitope binning assay is known in the art, and is described e.g., in Estep et al. "High throughput solution-based measurement of antibody-antigen affinity and epitope binning." *MAbs*. Vol. 5. No. 2. Taylor & Francis, 2013, which is incorporated herein reference in its entirety.

To determine the percent identity of two amino acid sequences, or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal

alignment and non-homologous sequences can be disregarded for comparison purposes). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences. For example, the comparison of sequences and determination of percent identity between two sequences can be accomplished using a Blossum 62 scoring matrix with a gap penalty of 12, a gap extend penalty of 4, and a frameshift gap penalty of 5.

The disclosure also provides nucleic acid comprising a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or an immunoglobulin light chain. The immunoglobulin heavy chain or immunoglobulin light chain comprises CDRs as shown in FIG. 3 or FIG. 4, or have sequences as shown in FIG. 5. When the polypeptides are paired with corresponding polypeptide (e.g., a corresponding heavy chain variable region or a corresponding light chain variable region), the paired polypeptides bind to MUC1 (e.g., human MUC1).

The anti-MUC1 antibodies and antigen-binding fragments can also be antibody variants (including derivatives and conjugates) of antibodies or antibody fragments and multi-specific (e.g., bi-specific) antibodies or antibody fragments. Additional antibodies provided herein are polyclonal, monoclonal, multi-specific (multimeric, e.g., bi-specific), human antibodies, chimeric antibodies (e.g., human-mouse chimera), single-chain antibodies, intracellularly-made antibodies (i.e., intrabodies), and antigen-binding fragments thereof. The antibodies or antigen-binding fragments thereof can be of any type (e.g., IgG, IgE, IgM, IgD, IgA, and IgY), class (e.g., IgG1, IgG2, IgG3, IgG4, IgA1, and IgA2), or subclass. In some embodiments, the antibody or antigen-binding fragment thereof is an IgG antibody or antigen-binding fragment thereof.

Fragments of antibodies are suitable for use in the methods provided so long as they retain the desired affinity and specificity of the full-length antibody. Thus, a fragment of an antibody that binds to MUC1 will retain an ability to bind to MUC1. An Fv fragment is an antibody fragment which contains a complete antigen recognition and binding site. This region consists of a dimer of one heavy and one light chain variable domain in tight association, which can be covalent in nature, for example in scFv. It is in this configuration that the three CDRs of each variable domain interact to define an antigen binding site on the surface of the VH-VL dimer. Collectively, the six CDRs or a subset thereof confer antigen binding specificity to the antibody. However, even a single variable domain (or half of an Fv comprising only three CDRs specific for an antigen) can have the ability to recognize and bind antigen, although usually at a lower affinity than the entire binding site.

Single-chain Fv or (scFv) antibody fragments comprise the VH and VL domains (or regions) of antibody, wherein these domains are present in a single polypeptide chain. Generally, the scFv polypeptide further comprises a polypeptide linker between the VH and VL domains, which enables the scFv to form the desired structure for antigen binding.

The Fab fragment contains a variable and constant domain of the light chain and a variable domain and the first constant domain (CH1) of the heavy chain. F(ab')₂ antibody fragments comprise a pair of Fab fragments which are generally covalently linked near their

carboxy termini by hinge cysteines between them. Other chemical couplings of antibody fragments are also known in the art.

Diabodies are small antibody fragments with two antigen-binding sites, which fragments comprise a VH connected to a VL in the same polypeptide chain (VH and VL). By using a linker that is too short to allow pairing between the two domains on the same chain, the domains are forced to pair with the complementary domains of another chain and create two antigen-binding sites.

Linear antibodies comprise a pair of tandem Fd segments (VH-CH1-VH-CH1) which, together with complementary light chain polypeptides, form a pair of antigen binding regions.

Linear antibodies can be bispecific or monospecific.

One-armed antibodies can have a heavy chain and a light chain, and a heavy chain fragment comprising CH2 and CH3 domains of IgG. In some embodiments, a one-armed antibody is an antibody that only has one of the two antigen binding arms in a typical antibody. In some embodiments, a one-armed antibody comprises an antigen binding arm (e.g., VH+ CH1 and VL+CL), and a Fc.

Antibodies and antibody fragments of the present disclosure can be modified in the Fc region to provide desired effector functions or serum half-life. In some embodiments, the Fc region can be modified to silence or decrease complement-dependent cytotoxicity (CDC) or antibody-dependent cellular cytotoxicity (ADCC). In some embodiments, the Fc region can be modified to increase complement-dependent cytotoxicity (CDC) or antibody-dependent cellular cytotoxicity (ADCC).

Multimerization of antibodies may be accomplished through natural aggregation of antibodies or through chemical or recombinant linking techniques known in the art. For example, some percentage of purified antibody preparations (e.g., purified IgG1 molecules) spontaneously form protein aggregates containing antibody homodimers and other higher-order antibody multimers.

Alternatively, antibody homodimers may be formed through chemical linkage techniques known in the art. For example, heterobifunctional crosslinking agents including, but not limited to SMCC (succinimidyl 4-(maleimidomethyl)cyclohexane-1-carboxylate) and SATA (N-succinimidyl S-acetylthio-acetate) can be used to form antibody multimers. An exemplary protocol for the formation of antibody homodimers is described in Ghetie et al. (*Proc. Natl. Acad. Sci. U.S.A.* 94: 7509-7514, 1997). Antibody homodimers can be converted to Fab₂ homodimers through digestion with pepsin. Another way to form antibody homodimers is through the use of the autophilic T15 peptide described in Zhao et al. (*J. Immunol.* 25:396-404, 2002).

In some embodiments, the multi-specific antibody is a bi-specific antibody. Bi-specific antibodies can be made by engineering the interface between a pair of antibody molecules to maximize the percentage of heterodimers that are recovered from recombinant cell culture. For example, the interface can contain at least a part of the CH3 domain of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g., tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g., alanine or threonine). This provides a mechanism for increasing the yield of the

heterodimer over other unwanted end-products such as homodimers. This method is described, e.g., in WO 96/27011, which is incorporated by reference in its entirety.

Bi-specific antibodies include cross-linked or “heteroconjugate” antibodies. For example, one of the antibodies in the heteroconjugate can be coupled to avidin and the other to biotin.

Heteroconjugate antibodies can also be made using any convenient cross-linking methods. Suitable cross-linking agents and cross-linking techniques are well known in the art and are disclosed in U.S. Patent No. 4,676,980, which is incorporated herein by reference in its entirety.

Methods for generating bi-specific antibodies from antibody fragments are also known in the art. For example, bi-specific antibodies can be prepared using chemical linkage. Brennan et al. (*Science* 229:81, 1985) describes a procedure where intact antibodies are proteolytically cleaved to generate F(ab')₂ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab' TNB derivatives is then reconverted to the Fab' thiol by reduction with mercaptoethylamine, and is mixed with an equimolar amount of another Fab' TNB derivative to form the bi-specific antibody.

Any of the antibodies or antigen-binding fragments described herein may be conjugated to a stabilizing molecule (e.g., a molecule that increases the half-life of the antibody or antigen-binding fragment thereof in a subject or in solution). Non-limiting examples of stabilizing molecules include: a polymer (e.g., a polyethylene glycol) or a protein (e.g., serum albumin, such as human serum albumin). The conjugation of a stabilizing molecule can increase the half-life or extend the biological activity of an antibody or an antigen-binding fragment *in vitro* (e.g., in tissue culture or when stored as a pharmaceutical composition) or *in vivo* (e.g., in a human).

In some embodiments, the antibodies or antigen-binding fragments described herein can be conjugated to a therapeutic agent. The antibody-drug conjugate comprising the antibody or antigen-binding fragment thereof can covalently or non-covalently bind to a therapeutic agent. In some embodiments, the therapeutic agent is a cytotoxic or cytostatic agent (e.g., cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin, maytansinoids such as DM-1 and DM-4, dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, puromycin, epirubicin, and cyclophosphamide and analogs).

In some embodiments, the antibody or antigen-binding fragment thereof described herein recognizes an endogenous MUC1 or a recombinant MUC1. In some embodiments, the antibody or antigen-binding fragment thereof described herein recognizes human MUC1.

In some embodiments, the half-life of the antibody or antigen-binding fragment thereof described herein in wild-type mice (e.g., C57BL/6 mice) is at least 4 days, at least 5 days, at least 6 days, at least 7 days, at least 8 days, at least 9 days, at least 10 days, at least 11 days, at least 12 days, at least 13 days, at least 14 days, at least 15 days, at least 16 days, at least 17 days, or at least 18 days. In some embodiments, the half-life of the antibody or antigen-binding fragment thereof described herein in MUC1 gene humanized mice (e.g., hMUC1 mice) is at least 1 day, at least 2 days, at least 3 days, at least 4 days, or at least 5 days, at least 6 days, at least 7 days, at least 8 days, at least 9 days, at least 10 days, at least 11 days, at least 12 days, at least 13 days, at

least 14 days, at least 15 days, at least 16 days, at least 17 days, or at least 18 days. In some embodiments, the max concentration (C_{max}) of the antibody or antigen-binding fragment thereof described herein in MUC1 gene humanized mice (e.g., hMUC1 mice) is at least 50 µg/ml, at least 75 µg/ml, at least 100 µg/ml, at least 125 µg/ml, at least 150 µg/ml, at least 175 µg/ml, or at least 200 µg/ml. In some embodiments, the serum clearance rate (CL) of the antibody or antigen-binding fragment thereof described herein in MUC1 gene humanized mice (e.g., hMUC1 mice) is at least 7.5 ml/day/kg, at least 10 ml/day/kg, at least 12.5 ml/day/kg, at least 15 ml/day/kg, at least 17.5 ml/day/kg, at least 20 ml/day/kg, at least 25 ml/day/kg, at least 30 ml/day/kg, at least 35 ml/day/kg, or at least 40 ml/day/kg.

In some embodiments, the clearance rate (CL) of the antibody or antigen-binding fragment thereof described herein in wild-type mice (e.g., C57BL/6 mice) is less than 7 ml/day/kg, less than 6 ml/day/kg, less than 5 ml/day/kg, or less than 4 ml/day/kg. In some embodiments, the clearance rate (CL) of the antibody or antigen-binding fragment thereof described herein in MUC1 gene humanized mice (e.g., hMUC1 mice) is less than 15 ml/day/kg, less than 14 ml/day/kg, less than 13 ml/day/kg, or less than 12 ml/day/kg.

In some embodiments, the half-life of the antibody or antigen-binding fragment thereof described herein (e.g., in FcRn gene humanized mice) is at least 6 days, at least 7 days, at least 8 days, at least 9 days, at least 10 days, at least 11 days, at least 12 days, at least 13 days, or at least 14 days. In some embodiments, the clearance rate (CL) of the antibody or antigen-binding fragment thereof described herein (e.g., in FcRn gene humanized mice) is less than 16 ml/day/kg, less than 15 ml/day/kg, less than 14 ml/day/kg, less than 13 ml/day/kg, less than 12 ml/day/kg, less than 11 ml/day/kg, less than 10 ml/day/kg, less than 9 ml/day/kg, less than 8 ml/day/kg, or less than 7 ml/day/kg.

Antibody Drug Conjugates (ADC)

The antibodies or the antigen-binding fragments thereof described herein can be conjugated to a therapeutic agent (a drug). The therapeutic agent can be covalently or non-covalently bind to the antibody or antigen-binding fragment.

In some embodiments, the therapeutic agent is a cytotoxic or cytostatic agent (e.g., monomethyl auristatin E, monomethyl auristatin F, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, teniposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin, maytansinoids such as DM-1 and DM-4, dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, puromycin, epirubicin, and cyclophosphamide and analogs).

Useful classes of cytotoxic, cytostatic, or immunomodulatory agents include, for example, antitubulin agents, DNA minor groove binders, DNA replication inhibitors, and alkylating agents.

In some embodiments, the therapeutic agent can include, but not limited to, cytotoxic reagents, such as chemo-therapeutic agents, immunotherapeutic agents and the like, antiviral agents or antimicrobial agents. In some embodiments, the therapeutic agent to be conjugated can be selected from, but not limited to, MMAE (monomethyl auristatin E), MMAD (monomethyl auristatin D), or MMAF (monomethyl auristatin F).

Definitions of specific functional groups and chemical terms are described in more detail below. For purpose of this invention, the chemical elements are identified in accordance with the Periodic Table of the Elements, CAS version, Handbook of Chemistry and Physics, 75th Edition, inside cover, and specific functional groups are generally defined as described therein.

5 Additionally, general principles of organic chemistry, as well as specific functional moieties and reactivity, are described in Organic Chemistry, Thomas Sorrell, University Science Books, Sausalito, 1999; Smith and March, March's Advanced Organic Chemistry, 5th Edition, John Wiley & Sons, Inc., New York, 2001; Larock, Comprehensive Organic Transformations, VCH Publishers, Inc., New York, 1989; Carruthers, Some Modern Methods of Organic Synthesis, 3rd
10 Edition, Cambridge University Press, Cambridge, 1987.

All ranges cited herein are inclusive, unless expressly stated to the contrary. When a range of values is listed, it is intended to encompass each value and sub-range within the range. For example, "C₁₋₆" is intended to encompass, C₁, C₂, C₃, C₄, C₅, C₆, C₁₋₆, C₁₋₅, C₁₋₄, C₁₋₃, C₁₋₂, C₂₋₆, C₂₋₅, C₂₋₄, C₂₋₃, C₃₋₆, C₃₋₅, C₃₋₄, C₄₋₆, C₄₋₅, and C₅₋₆.

15 The compounds or any formula depicting and describing the compounds of the present disclosure may have one or more chiral (asymmetric) centers. The present invention encompasses all stereoisomeric forms of the compounds or any formula depicting and describing the compounds of the present invention. Centers of asymmetry that are present in the compounds or any formula depicting and describing the compounds of the present invention can all
20 independently of one another have (R) or (S) configuration. When bonds to a chiral carbon are depicted as straight lines in the structural formulas, or when a compound name is recited without an (R) or (S) chiral designation for a chiral carbon, it is understood that both the (R) and (S) configurations of each such chiral carbon, and hence each enantiomer or diastereomer and mixtures thereof, are embraced within the formula or by the name.

25 The disclosure includes all possible enantiomers and diastereomers and mixtures of two or more stereoisomers, for example mixtures of enantiomers and/or diastereomers, in all ratios. Thus, enantiomers are a subject of the disclosure in enantiomerically pure form, both as levorotatory and as dextrorotatory antipodes, in the form of racemates and in the form of mixtures of the two enantiomers in all ratios. In the case of a *cis/trans* isomerism the disclosure
30 includes both the *cis* form and the *trans* form as well as mixtures of these forms in all ratios. The preparation of individual stereoisomers can be carried out, if desired, by separation of a mixture by customary methods, for example by chromatography or crystallization, by the use of stereochemically uniform starting materials for the synthesis or by stereoselective synthesis. Optionally a derivatization can be carried out before a separation of stereoisomers. The
35 separation of a mixture of stereoisomers can be carried out at an intermediate step during the synthesis of a compound or it can be done on a final racemic product. Absolute stereochemistry may be determined by X-ray crystallography of crystalline products or crystalline intermediates which are derivatized, if necessary, with a reagent containing a stereogenic center of known configuration. Alternatively, absolute stereochemistry may be determined by Vibrational
40 Circular Dichroism (VCD) spectroscopy analysis.

Unless otherwise stated, the structures depicted herein are also meant to include the compounds that differ only in the presence of one or more isotopically enriched atoms, in other words, the compounds wherein one or more atoms are replaced by atoms having the same atomic

number, but an atomic mass or mass number different from the atomic mass or mass number which predominates in nature. Such compounds are referred to as a "isotopic variant". The present disclosure is intended to include all pharmaceutically acceptable isotopic variants of the compounds or any formula depicting and describing the compounds of the present invention.

5 Examples of isotopes suitable for inclusion in the compounds of the present invention include, but not limited to, isotopes of hydrogen, such as ^2H (i.e., D) and ^3H ; carbon, such as ^{11}C , ^{13}C , and ^{14}C ; chlorine, such as ^{36}Cl ; fluorine, such as ^{18}F ; iodine, such as ^{123}I and ^{125}I ; nitrogen, such as ^{13}N and ^{15}N ; oxygen, such as ^{15}O , ^{17}O , and ^{18}O ; phosphorus, such as ^{32}P ; and sulfur, such as ^{35}S . Certain isotopic variants of the compounds or any formula depicting and describing the
10 compounds of the present disclosure, for example those incorporating a radioactive isotope, may be useful in drug and/or substrate tissue distribution studies. Particularly, compounds having the depicted structures that differ only in the replacement with heavier isotopes, such as the replacement of hydrogen by deuterium (^2H , or D), can afford certain therapeutic advantages, for example, resulting from greater metabolic stability, increased *in vivo* half-life, or reduced dosage
15 requirements and, hence, may be utilized in some particular circumstances. Isotopic variants of compounds or any formula depicting and describing the compounds of the present disclosure can generally be prepared by techniques known to those skilled in the art or by processes analogous to those described in the accompanying examples and synthesis using an appropriate isotopically-labeled reagent in place of the non-labeled reagent previously employed.

20 The compounds as provided herein are described with reference to both generic formulas and specific compounds. In addition, the compounds of the present disclosure may exist in a number of different forms or derivatives, all within the scope of the disclosure. These include, for example, pharmaceutically acceptable salts, tautomers, stereoisomers, racemic mixtures, regioisomers, prodrugs, solvated forms, different crystal forms or polymorphs, and active
25 metabolites, etc.

As used herein, the term "pharmaceutically acceptable salt", unless otherwise stated, includes salts that retain the biological effectiveness of the free acid/base form of the specified compound and that are not biologically or otherwise undesirable. Pharmaceutically acceptable salts may include salts formed with inorganic bases or acids and organic bases or acids. In cases
30 where the compounds of the present disclosure contain one or more acidic or basic groups, the disclosure also comprises their corresponding pharmaceutically acceptable salts. Thus, the compounds of the present invention which contain acidic groups, such as carboxyl groups, can be present in salt form, and can be used according to the invention, for example, as alkali metal salts, alkaline earth metal salts, aluminum salts or as ammonium salts. More non-limiting
35 examples of such salts include lithium salts, sodium salts, potassium salts, calcium salts, magnesium salts, barium salts, or salts with ammonia or organic amines such as ethylamine, ethanolamine, diethanolamine, triethanolamine, piperidine, N-methylglutamine, or amino acids. These salts are readily available, for instance, by reacting the compound having an acidic group with a suitable base, e.g., lithium hydroxide, sodium hydroxide, sodium propoxide, potassium
40 hydroxide, potassium ethoxide, magnesium hydroxide, calcium hydroxide, or barium hydroxide. Other base salts of compounds of the present disclosure include but are not limited to copper (I), copper (II), iron (II), iron (III), manganese (II), and zinc salts. Compounds of the present disclosure which contain one or more basic groups, e.g., groups which can be protonated, can be

present in salt form, and can be used according to the disclosure in the form of their addition salts with inorganic or organic acids. Examples of suitable acids include hydrogen chloride, hydrogen bromide, hydrogen iodide, phosphoric acid, sulfuric acid, nitric acid, methanesulfonic acid, p-toluenesulfonic acid, naphthalenedisulfonic acid, sulfoacetic acid, trifluoroacetic acid, oxalic acid, acetic acid, tartaric acid, lactic acid, salicylic acid, benzoic acid, carbonic acid, formic acid, propionic acid, pivalic acid, diethylacetic acid, malonic acid, succinic acid, pimelic acid, fumaric acid, malonic acid, maleic acid, malic acid, embonic acid, mandelic acid, sulfaminic acid, phenylpropionic acid, gluconic acid, ascorbic acid, isonicotinic acid, citric acid, adipic acid, taurocholic acid, glutaric acid, stearic acid, glutamic acid, or aspartic acid, and other acids known to those skilled in the art. The salts which are formed are, *inter alia*, hydrochlorides, chlorides, hydrobromides, bromides, iodides, sulfates, phosphates, methanesulfonates (mesylates), tosylates, carbonates, bicarbonates, formates, acetates, sulfoacetates, triflates, oxalates, malonates, maleates, succinates, tartrates, malates, embonates, mandelates, fumarates, lactates, citrates, glutarates, stearates, aspartates, and glutamates. The stoichiometry of the salts formed from the compounds of the disclosure may moreover be an integral or non-integral multiple of one.

Compounds of the present disclosure which contain basic nitrogen-containing groups can be quaternized using agents such as C₁₋₄alkyl halides, for example, methyl, ethyl, isopropyl, and tert-butyl chloride, bromide, and iodide; diC₁₋₄alkyl sulfates, for example, dimethyl, diethyl, and diamyl sulfate; C₁₀₋₁₈alkyl halides, for example, decyl, dodecyl, lauryl, myristyl, and stearyl chloride, bromide, and iodide; and arylC₁₋₄alkyl halides, for example, benzyl chloride and phenethyl bromide.

If the compounds of the present disclosure simultaneously contain acidic and basic groups in the molecule, the disclosure also includes, in addition to the salt forms mentioned, inner salts or betaines (zwitterions). The respective salts can be obtained by customary methods which are known to those skilled in the art, for example by contacting these with an organic or inorganic acid or base in a solvent or dispersant, or by anion exchange or cation exchange with other salts. The present disclosure also includes all salts of the compounds of the present disclosure which, owing to low physiological compatibility, are not directly suitable for use in pharmaceuticals but which can be used, for example, as intermediates for chemical reactions or for the preparation of pharmaceutically acceptable salts. For a review on more suitable salts, see Stahl and Wermuth, *Handbook of Pharmaceutical Salts: Properties, Selection, and Use* (Wiley-VCH, 2002).

The compound or any formula depicting and describing the compounds of the present disclosure and pharmaceutically acceptable salts thereof may exist in unsolvated and solvated forms. As used herein, the term "solvate" refers to a molecular complex comprising the compound of Formula (I), or a pharmaceutically acceptable salt thereof, and one or more pharmaceutically acceptable solvent molecules. For example, the term "hydrate" is employed when the solvent is water.

Pharmaceutically acceptable solvates in accordance with the present disclosure may include those wherein the solvent of crystallization may be isotopically substituted, e.g., D₂O, d₆-acetone, d₆-DMSO.

Linker (linking agent compound)

In some embodiments, the therapeutic agent is conjugated via a linker (or a linking agent compound). As used herein, the term “linker” or “linking agent compound” refers to a compound that can connect a ligand (e.g., the antibodies or the antigen-binding fragments thereof described herein) and a therapeutic agent (e.g., any of the therapeutic agents described herein) together to form a ligand-drug conjugate by reacting with a group of the ligand compound and the therapeutic agent compound respectively by, for example, a coupling reaction.

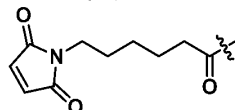
In some embodiments, the linker described herein is a compound having the following formula:



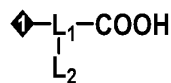
Formula (I),

or a pharmaceutically acceptable salt, solvate, stereoisomer, or isotopic variant thereof, wherein Q denotes to a junction moiety capable of being coupled to a ligand via a bond selected from the group consisting of carbonyl, thioether, amide, disulfide and hydrazone bond; L denotes to a linker moiety capable of connecting Q to a therapeutic agent.

In some embodiments, the junction moiety (Q in Formula (I)) has the following structure:



In some embodiments, the linker moiety (L in Formula (I)) has the following formula:

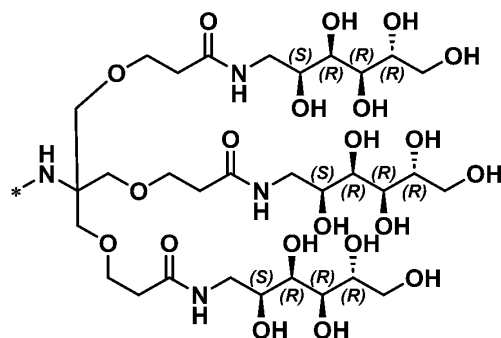


where L_1 is a polypeptide residue consisting of three to eight amino acid residues which comprises at least one amino acid residue with a side chain carboxyl group, for example, glutamic acid residue or aspartic acid residue, where “-COOH” denotes carboxyl group of an amino acid residue at C-terminal of the polypeptide residue;

L_2 is absent or a monodentate, bidentate or tridentate hydrophilic group attached to the side chain carboxyl group on the amino acid residue of the polypeptide residue L_1 , and L_2 has a structure of $-NHC(R^{L2a})(R^{L2b})(R^{L2c})$, where R^{L2a} , R^{L2b} , and R^{L2c} are each independently selected from the group consisting of H, $-(CH_2O)(CH_2CH_2O)_m(CH_2)_pC(O)OH$, and $-(CH_2O)(CH_2CH_2O)_m(CH_2)_pC(O)NHR^{L2d}$, R^{L2d} is H or C_{1-6} alkyl optionally substituted with 1 to 6 hydroxy groups, each m is independently an integer from 0 to 10, preferably 0 to 4, for example 0, 1, 2, 3, or 4, especially preferably m is 0, and each p is independent an integer from 1 to 4, for example, 1, 2, 3, or 4; and

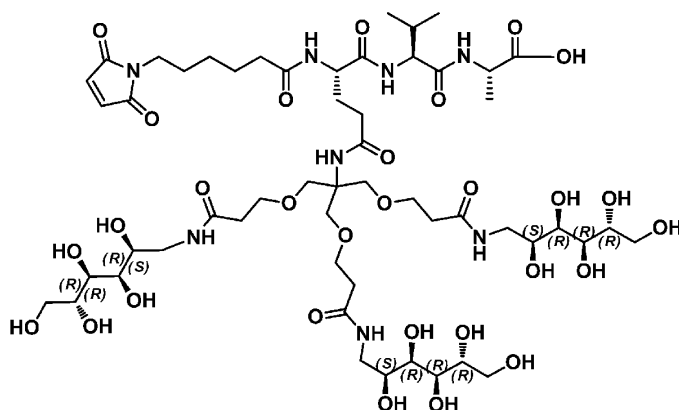
♦ denotes to the N-terminal side of the polypeptide residue covalently attached to the junction moiety Q.

In some embodiments, the polypeptide residue L_1 is $^{NH}-Glu-Val-Ala-COOH$. In some embodiments, the hydrophilic group L_2 has the following structure:



wherein “*” denotes the site covalently attached to polypeptide residue L₁, e.g., side chain of the Glu residue in ^{NH}-Glu-Val-Ala-^{COOH}.

In some embodiments, the linker described herein is a compound having the following structure:



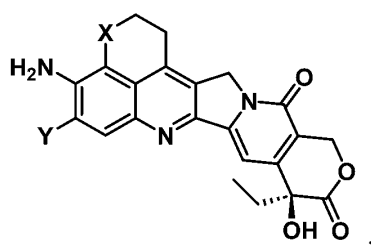
CPT-L.

In some embodiments, the linker is a VC linker. Details of the linkers used for ADCs can be found, e.g., in Su, Z. et al. "Antibody–drug conjugates: Recent advances in linker chemistry." *Acta Pharmaceutica Sinica B* (2021), which is incorporated herein by reference in its entirety.

10 **Therapeutic agent**

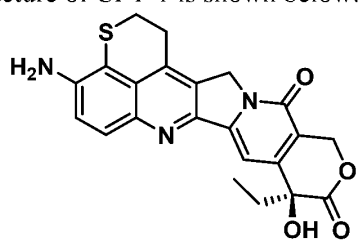
In some embodiments, the therapeutic agent that is conjugated to the antibodies or the antigen-binding fragments thereof described herein is discussed as follows.

In some embodiments, the therapeutic agent described herein is a cytotoxic agent. In some embodiments, the cytotoxic agent is a camptothecin compound, an analogue or a derivative thereof. In some preferred embodiments, the camptothecin compound is a compound having the following structure:



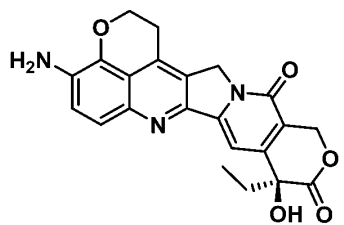
wherein X is selected from the group consisting of -CH₂-, O and S; Y is selected from the group consisting of H, D, and F.

In some embodiments, the therapeutic agent is (S)-4-amino-9-ethyl-9-hydroxy-1,9,12,15-tetrahydro-13H-pyrano[3',4':6,7]indolizino[1,2-b]thiopyrano[4,3,2-de]quinoline-10,13(2H)-dione (CPT-1). The structure of CPT-1 is shown below:



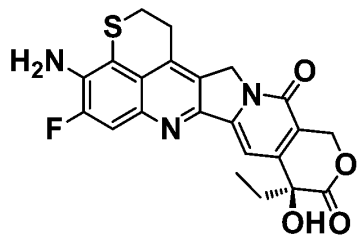
CPT-1

In some embodiments, the therapeutic agent is (S)-4-amino-9-ethyl-9-hydroxy-1,9,12,15-tetrahydro-13H-pyrano[4,3,2-de]pyrano[3',4':6,7]indolizino[1,2-b]quinoline-10,13(2H)-dione (CPT-2). The structure of CPT-2 is shown below:



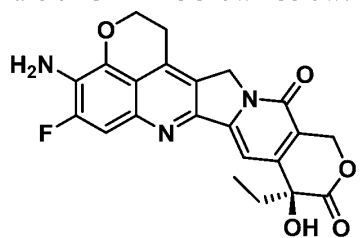
CPT-2

In some embodiments, the therapeutic agent is CPT3. The structure of CPT-3 is shown below:



CPT-3

In some embodiments, the therapeutic agent is (S)-4-amino-9-ethyl-5-fluoro-9-hydroxy-1,9,12,15-tetrahydro-13H-pyrano[4,3,2-de]pyrano[3',4':6,7]indolizino[1,2-b]quinoline-10,13(2H)-dione (CPT-4). The structure of CPT-4 is shown below:



CPT-4

In some embodiments, the therapeutic agent is an auristatin, such as auristatin E (also known in the art as a derivative of dolastatin-10) or a derivative thereof. The auristatin can be, for example, an ester formed between auristatin E and a keto acid. For example, auristatin E can be reacted with paraacetyl benzoic acid or benzoylvaleric acid to produce AEB and AEVB, respectively. Other typical auristatins include AFP, MMAF, and MMAE. The synthesis and structure of exemplary auristatins are described in U.S. Patent Application Publication No. 2003-0083263; International Patent Publication No. WO 04/010957, International Patent Publication No. WO 02/088172, and U.S. Pat. Nos. 7,498,298, 6,884,869, 6,323,315; 6,239,104; 6,034,065; 5,780,588; 5,665,860; 5,663,149; 5,635,483; 5,599,902; 5,554,725; 5,530,097; 5,521,284; 5,504,191; 5,410,024; 5,138,036; 5,076,973; 4,986,988; 4,978,744; 4,879,278; 4,816,444; and 4,486,414, each of which is incorporated by reference herein in its entirety and for all purposes.

Auristatins have been shown to interfere with microtubule dynamics and nuclear and cellular division and have anticancer activity. Auristatins bind tubulin and can exert a cytotoxic or cytostatic effect on cancer cell. There are a number of different assays, known in the art, which can be used for determining whether an auristatin or resultant antibody-drug conjugate exerts a cytostatic or cytotoxic effect on a desired cell.

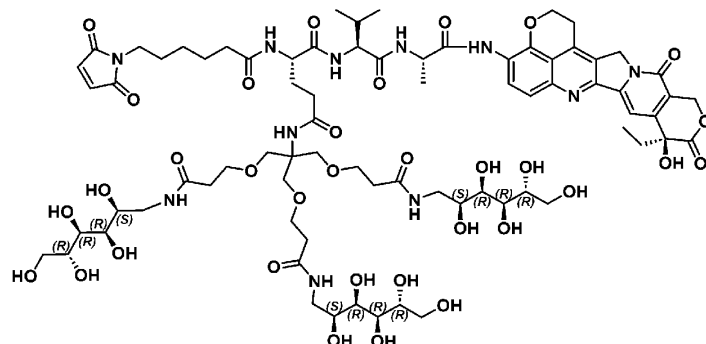
In some embodiments, the therapeutic agent is a chemotherapeutic agent. Examples of chemotherapeutic agents include alkylating agents such as thiotepa and cyclophosphamide (CYTOXANTM); alkyl sulfonates such as busulfan, improsulfan and piposulfan; aziridines such as benzodopa, carboquone, meturedopa, and uredopa; ethylenimines and methylamelamines including altretamine, triethylenemelamine, triethylenephosphoramidate, triethylenethiophosphoramidate and trimethylolomelamine; nitrogen mustards such as chlorambucil, chlornaphazine, cholophosphamide, estramustine, ifosfamide, mechlorethamine, mechlorethamine oxide hydrochloride, melphalan, novembichin, phenesterine, prednimustine, trofosfamide, uracil mustard; nitrosureas such as carmustine, chlorozotocin, fotemustine, lomustine, nimustine, ranimustine; antibiotics such as aclacinomysins, actinomycin, anthramycin, azaserine, bleomycins, cactinomycin, calicheamicin, carabycin, carminomycin, carzinophilin, chromomycins, dactinomycin, daunorubicin, detorubicin, 6-diazo-5-oxo-L-norleucine, doxorubicin, epirubicin, esorubicin, idarubicin, marcellomycin, mitomycins, mycophenolic acid, nogalamycin, olivomycins, peplomycin, potfiromycin, puromycin, quelamycin, rodorubicin,

streptonigrin, streptozocin, tubercidin, ubenimex, zinostatin, zorubicin; anti-metabolites such as methotrexate and 5-fluorouracil (5-FU); folic acid analogues such as denopterin, methotrexate, pteropterin, trimetrexate; purine analogs such as fludarabine, 6-mercaptopurine, thiamiprine, thioguanine; pyrimidine analogs such as ancitabine, azacitidine, 6-azauridine, carmofur, cytarabine, dideoxyuridine, doxifluridine, enocitabine, floxuridine, 5-FU; androgens such as calusterone, dromostanolone propionate, epitio stanol, mepitio stanane, testolactone; anti-adrenals such as aminogluthethimide, mitotane, trilostane; folic acid replenisher such as frolinic acid; aceglatone; aldophosphamide glycoside; aminolevulinic acid; amsacrine; bestrabucil; bisantrene; edatraxate; defofamine; demecolcine; diaziquone; elfornithine; elliptinium acetate; etoglucid; gallium nitrate; hydroxyurea; lentinan; lonidamine; mitoguazone; mitoxantrone; mopidamol; nitracrine; pentostatin; phenamet; pirarubicin; podophyllinic acid; 2-ethylhydrazide; procarbazine; PSK7; razoxane; sizofiran; spirogermanium; tenuazonic acid; triaziquone; 2',2',2'-trichlorotriethylamine; urethan; vindesine; dacarbazine; mannomustine; mitobronitol; mitolactol; pipobroman; gacytosine; arabinoside ("Ara-C"); cyclophosphamide; taxanes, e.g. paclitaxel (TAXOL®, Bristol-Myers Squibb Oncology, Princeton, N.J.) and doxetaxel (TAXOTERE®, Rhone-Poulenc Rorer, Antony, France); chlorambucil; gemcitabine; 6-thioguanine; platinum analogs such as cisplatin and carboplatin; vinblastine; platinum; etoposide (VP-16); ifosfamide; mitomycin C; mitoxantrone; vincristine; vinorelbine; navelbine; novantrone; teniposide; daunomycin; aminopterin; xeloda; ibandronate; CPT-11; topoisomerase inhibitor RFS 2000; difluoromethylornithine (DMFO); retinoic acid; esperamicins; capecitabine; and pharmaceutically acceptable salts, acids or derivatives of any of the above. Also included in this definition are anti-hormonal agents that act to regulate or inhibit hormone action on tumors such as anti-estrogens including for example tamoxifen, raloxifene, aromatase inhibiting 4(5)-imidazoles, 4-hydroxytamoxifen, trioxifene, keoxifene, LY117018, onapristone, and toremifene (Fareston); and anti-androgens such as flutamide, nilutamide, bicalutamide, leuprolide, and goserelin; and pharmaceutically acceptable salts, acids or derivatives of any of the above. A detailed description of the chemotherapeutic agents can be found in, e.g., US20180193477A1, which is incorporated by reference in its entirety.

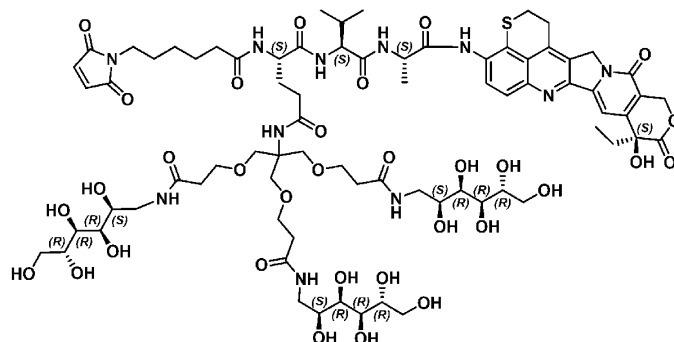
Linker-Therapeutic agent compound

In some embodiments, a linker (e.g., any of the linkers described herein) and a therapeutic agent (e.g., any of the therapeutic agents described herein) can be linked to form a "linker-therapeutic agent" compound.

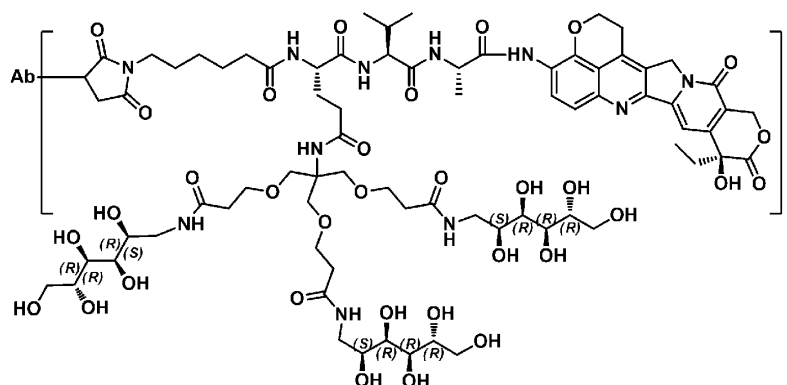
In some embodiments, the linker-therapeutic agent compound has the following structure:



In some embodiments, the linker-therapeutic agent compound has the following structure:



- 5 In some embodiments, an antibody ("Ab"), e.g., any of the antibodies or the antigen-binding fragments thereof described herein, can be linked to a linker-therapeutic agent compound (e.g., any of the linker-therapeutic agent compounds described herein) to generate an antibody-drug conjugate. In some embodiments, the antibody-drug conjugate has the following structure:



10

wherein $n = 1-8$. In some embodiments, $n = 1-8$. In some embodiments, n is about 1, about 2, about 3, about 4, about 5, about 6, about 7, about 8. In some embodiments, n is about 1-8, 1-7, 1-6, 1-5, 1-4, 1-3, 1-2, 2-8, 2-7, 2-6, 2-5, 2-4, 2-3, 3-8, 3-7, 3-6, 3-5, 3-4, 4-8, 4-7, 4-6, 4-5,

5-8, 5-7, 5-6, 6-8, 6-7, or 7-8. In some embodiments, n is an integral or non-integral multiple of one.

In some embodiments, the antibody or the antigen-binding fragments thereof described herein is coupled to the drug via a cleavable linker e.g. a SPBD linker or a maleimidocaproyl-valine-citrulline-p-aminobenzyloxycarbonyl (VC) linker. In some embodiments, the antibody or the antigen-binding fragments thereof is coupled to the drug via a non-cleavable linker e.g. a MCC linker formed using SMCC or sulfo-SMCC. Selection of an appropriate linker for a given ADC can be readily made by the skilled person having knowledge of the art and taking into account relevant factors, such as the site of attachment to the antibody or the antigen-binding fragments thereof, any structural constraints of the drug and the hydrophobicity of the drug (see, for example, review in Nolting, Chapter 5, *Antibody-Drug Conjugates: Methods in Molecular Biology*, 2013, Ducry (Ed.), Springer). A number of specific linker-toxin combinations have been described and may be used with the antibody or the antigen-binding fragments thereof described herein to prepare ADCs in certain embodiments. Examples include, but are not limited to, cleavable peptide-based linkers with auristatins such as MMAE and MMAF, camptothecins such as SN-38, duocarmycins and PBD dimers; non-cleavable MC-based linkers with auristatins MMAF and MMAE; acid-labile hydrazone-based linkers with calicheamicins and doxorubicin; disulfide-based linkers with maytansinoids such as DM1 and DM4, and bis-maleimido-trioxyethylene glycol (BMPEO)-based linkers with maytansinoid DM1. Some these therapeutic agents and linkers are described, e.g., in Peters & Brown, (2015) *Biosci. Rep.* e00225; Dosio et al., (2014) *Recent Patents on Anti-Cancer Drug Discovery* 9:35-65; US Patent Publication No. US 2015/0374847, and US20180193477A1; which are incorporated herein by reference in the entirety.

Depending on the desired drug and selected linker, those skilled in the art can select suitable method for coupling them together. For example, some conventional coupling methods, such as amine coupling methods, can be used to form the desired drug-linker complex which still contains reactive groups for conjugating to the antibodies through covalent linkage. In some embodiments, a drug-maleimide complex (i.e., maleimide linking drug) can be used for the payload bearing reactive group in the present disclosure. Most common reactive group capable of bonding to thiol group in ADC preparation is maleimide. Additionally, organic bromides, iodides also are frequently used.

The ADC can be prepared by one of several routes known in the art, employing organic chemistry reactions, conditions, and reagents known to those skilled in the art (see, for example, *Bioconjugate Techniques* (G. T. Hermanson, 2013, Academic Press). For example, conjugation can be achieved by (1) reaction of a nucleophilic group or an electrophilic group of an antibody with a bivalent linker reagent, to form antibody-linker intermediate Ab-L, via a covalent bond, followed by reaction with an activated drug moiety D; or (2) reaction of a nucleophilic group or an electrophilic group of a drug moiety with a linker reagent, to form drug-linker intermediate D-L, via a covalent bond, followed by reaction with the nucleophilic group or an electrophilic group of an antibody. Conjugation methods (1) and (2) can be employed with a variety of antibodies, drug moieties, and linkers to prepare the ADCs described here. Various prepared linkers, linker components and toxins are commercially available or may be prepared using standard synthetic organic chemistry techniques. These methods are described e.g., in March's

Advanced Organic Chemistry (Smith & March, 2006, Sixth Ed., Wiley); Toki et al., (2002) J. Org. Chem. 67:1866-1872; Frisch et al., (1997) Bioconj. Chem. 7:180-186; Bioconjugate Techniques (G. T. Hermanson, 2013, Academic Press); US20210379193A1, and US20180193477A1, which are incorporated herein by reference in the entirety. In addition, a number of pre-formed drug-linkers suitable for reaction with a selected antibody or the antigen-binding fragments thereof are also available commercially, for example, linker-toxins comprising DM1, DM4, MMAE, MMAF or Duocarmycin SA are available from Creative BioLabs (Shirley, N.Y.).

Several specific examples of methods of preparing ADCs are known in the art and are described in U.S. Pat. No. 8,624,003 (pot method), U.S. Pat. No. 8,163,888 (one-step), and U.S. Pat. No. 5,208,020 (two-step method), and US20180193477A1, which are incorporated herein by reference in the entirety. Other methods are known in the art and include those described in Antibody-Drug Conjugates: Methods in Molecular Biology, 2013, Ducry (Ed.), Springer.

Drug loading is represented by the number of drug moieties per antibody in a molecule of ADC. For some antibody-drug conjugates, the drug loading may be limited by the number of attachment sites on the antibody. For example, where the attachment is a cysteine thiol, as in certain exemplary embodiments described herein, the drug loading may range from 0 to 8 drug moieties per antibody. In certain embodiments, higher drug loading, e.g. $p \geq 5$, may cause aggregation, insolubility, toxicity, or loss of cellular permeability of certain antibody-drug conjugates. In certain embodiments, the average drug loading for an antibody-drug conjugate ranges from 1 to about 8; from about 2 to about 6; or from about 3 to about 5. Indeed, it has been shown that for certain antibody-drug conjugates, the optimal ratio of drug moieties per antibody can be around 4 or 8. In some embodiments, the DAR is about or at least 1, 2, 3, 4, 5, 6, 7, or 8. In some embodiments, the average DAR in the composition is about 1~ about 2, about 2~ about 3, about 3~ about 4, about 3~ about 5, about 4~ about 5, about 5~ about 6, about 6~ about 7, or about 7~ about 8.

Antibody and ADC Characteristics

The antibodies or antigen-binding fragments thereof described herein or ADC derived therefrom can block the binding between MUC1 and MUC1 ligands (e.g., selectins and/or intercellular adhesion molecule-1).

The antibodies or antigen-binding fragments thereof as described herein or ADC derived therefrom can be MUC1 agonist or antagonist. In some embodiments, by binding to MUC1, the antibody can inhibit MUC1 activity. In some embodiments, the antibody can upregulate immune response or downregulate immune response.

In some implementations, the antibody (or antigen-binding fragments thereof) or ADC derived therefrom specifically binds to MUC1 (e.g., human MUC1, monkey MUC1 (e.g., rhesus macaques, *Macaca fascicularis*), dog MUC1, mouse MUC1, human MUC1b, human MUC18, and/or human MUC16) with a dissociation rate (k_{off}) of less than 0.1 s^{-1} , less than 0.01 s^{-1} , less than 0.001 s^{-1} , less than 0.0001 s^{-1} , less than 0.00001 s^{-1} , less than 0.000001 s^{-1} or less than 0.0000001 s^{-1} . In some embodiments, the dissociation rate (k_{off}) is greater than 0.01 s^{-1} , greater

than 0.001 s^{-1} , greater than 0.0001 s^{-1} , greater than 0.00001 s^{-1} , greater than 0.000001 s^{-1} , greater than 0.0000001 s^{-1} or greater than $0.00000001 \text{ s}^{-1}$.

In some embodiments, kinetic association rates (k_{on}) is greater than $1 \times 10^2/\text{Ms}$, greater than $1 \times 10^3/\text{Ms}$, greater than $1 \times 10^4/\text{Ms}$, greater than $1 \times 10^5/\text{Ms}$, or greater than $1 \times 10^6/\text{Ms}$. In some embodiments, kinetic association rates (k_{on}) is less than $1 \times 10^5/\text{Ms}$, less than $1 \times 10^6/\text{Ms}$, or less than $1 \times 10^7/\text{Ms}$.

Affinities can be deduced from the quotient of the kinetic rate constants ($KD=k_{off}/k_{on}$). In some embodiments, KD is less than $1 \times 10^{-6} \text{ M}$, less than $1 \times 10^{-7} \text{ M}$, less than $1 \times 10^{-8} \text{ M}$, less than $1 \times 10^{-9} \text{ M}$, less than $1 \times 10^{-10} \text{ M}$, less than $1 \times 10^{-11} \text{ M}$, less than $1 \times 10^{-12} \text{ M}$, less than $1 \times 10^{-13} \text{ M}$ or less than $1 \times 10^{-14} \text{ M}$. In some embodiments, the KD is less than 50 nM , 30 nM , 20 nM , 15 nM , 10 nM , 9 nM , 8 nM , 7 nM , 6 nM , 5 nM , 4 nM , 3 nM , 2 nM , or 1 nM . In some embodiments, KD is greater than $1 \times 10^{-7} \text{ M}$, greater than $1 \times 10^{-8} \text{ M}$, greater than $1 \times 10^{-9} \text{ M}$, greater than $1 \times 10^{-10} \text{ M}$, greater than $1 \times 10^{-11} \text{ M}$, greater than $1 \times 10^{-12} \text{ M}$, greater than $1 \times 10^{-13} \text{ M}$, greater than $1 \times 10^{-14} \text{ M}$.

General techniques for measuring the affinity of an antibody for an antigen include, e.g., ELISA, RIA, and surface plasmon resonance (SPR). In some embodiments, the antibody binds to human MUC1 (SEQ ID NO: 50), monkey MUC1 (SEQ ID NO: 51), dog MUC1 (SEQ ID NO: 52), mouse MUC1 (SEQ ID NO: 53), human MUC1b (hMUC1b) (SEQ ID NO: 54), human MUC18 (hMUC18) (SEQ ID NO: 55) and/or human MUC16 (hMUC16) (SEQ ID NO: 56). In some embodiments, the antibody does not bind to human MUC1, monkey MUC1, dog MUC1, mouse MUC1, human MUC1b (hMUC1b), human MUC18 (hMUC18) and/or human MUC16 (hMUC16).

In some embodiments, the antibody or antigen-binding fragment thereof as described herein or ADC derived therefrom is added to ZR-75-1 cells or CHO-hMUC1(961-1152) cells (obtained by transfecting CHO-S cells with hMUC1(961-1152)) to test the endocytosis rate. In some embodiments, the antibody or antigen-binding fragment thereof as described herein or ADC derived therefrom has an endocytosis rate of above 60%, above 65%, above 70%, above 75%, above 80%, above 85%, above 90%, above 91%, above 92%, above 93%, above 94%, above 95%, above 96%, above 97%, or above 98%.

In some embodiments, thermal stabilities are determined. The antibodies or antigen binding fragments as described herein or ADC derived therefrom can have a T_m greater than 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, or 95°C . In some embodiments, T_m is less than 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, or 95°C .

In some embodiments, the antibodies or antigen-binding fragments thereof as described herein or ADC derived therefrom can bind to the same epitope of MUC1. In some embodiments, the antibodies or antigen-binding fragments thereof as described herein or ADC derived therefrom can bind to different epitopes of MUC1.

In some embodiments, the antibodies or antigen-binding fragments thereof as described herein have a purity of above 90%, above 91%, above 92%, above 93%, above 94%, above 95%, above 96%, above 97%, or above 98%, as determined by size exclusion chromatography (SEC). In some embodiments, the antibodies or antigen-binding fragments thereof as described herein

have a hydrophobic interaction chromatography (HIC) retention time that is higher than 2 minutes, higher than 3 minutes, higher than 4 minutes, higher than 5 minutes, higher than 6 minutes, higher than 7 minutes, higher than 8 minutes, higher than 9 minutes, higher than 10 minutes, higher than 11 minutes, higher than 12 minutes, higher than 13 minutes, higher than 14 minutes, higher than 15 minutes, higher than 16 minutes, higher than 17 minutes, higher than 18 minutes, higher than 19 minutes, higher than 20 minutes, higher than 21 minutes or higher than 22 minutes.

In some embodiments, the antibodies or antigen-binding fragments thereof as described herein have a purity of above 90%, above 91%, above 92%, above 93%, above 94%, above 95%, above 96%, above 97%, above 98%, or above 99%, as determined by capillary electrophoresis-sodium dodecyl sulphate (CE-SDS).

In some embodiments, the antibodies or antigen-binding fragments thereof as described herein have a main peak that constitutes more than 40%, more than 50%, more than 55%, more than 60%, more than 65%, more than 70%, more than 75%, more than 80%, more than 85%, or more than 90%, as determined by capillary isoelectric focusing (cIEF).

In some embodiments, the ADC described herein has an average drug-to-antibody ratio (DAR) of higher than 3, higher than 3.2, higher than 3.4, higher than 3.6, higher than 3.8, higher than 4, higher than 4.2, higher than 4.4, or higher than 4.6, as determined by HPLC. In some embodiments, the ADC described herein has an average DAR of lower than 3, lower than 3.2, lower than 3.4, lower than 3.6, lower than 3.8, lower than 4, lower than 4.2, lower than 4.4 or lower than 4.6, as determined by HPLC.

In some embodiments, the antibody or antigen-binding fragment thereof as described herein or ADC derived therefrom has a tumor growth inhibition percentage (TGI%) that is greater than 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, or 200%. In some embodiments, the antibody or antigen-binding fragment thereof as described herein or ADC derived therefrom has a tumor growth inhibition percentage that is less than 60%, 70%, 80%, 90%, 100%, 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, or 200%. The TGI% can be determined, e.g., at 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 days after the treatment starts, or 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 months after the treatment starts. As used herein, the tumor growth inhibition percentage (TGI%) is calculated using the following formula:

$$\text{TGI (\%)} = [1 - (\text{Ti} - \text{T0}) / (\text{Vi} - \text{V0})] \times 100\%$$

Ti is the average tumor volume in the treatment group on day i. T0 is the average tumor volume in the treatment group on day zero. Vi is the average tumor volume in the control group on day i. V0 is the average tumor volume in the control group on day zero.

In some embodiments, the antibodies or antigen-binding fragments thereof as described herein or ADC derived therefrom are MUC1 antagonist. In some embodiments, the antibodies or antigen binding fragments as described herein or ADC derived therefrom decrease MUC1 signal transduction in a target cell that expresses MUC1.

In some embodiments, the antibodies or antigen binding fragments as described herein or ADC derived therefrom can enhance APC (e.g., DC cell) function, for example, inducing surface

expression of costimulatory and MHC molecules, inducing production of proinflammatory cytokines, and/or enhancing T cell triggering function.

In some embodiments, the antibodies or antigen binding fragments as described herein or ADC derived therefrom can bind to tumor cells that express MUC1. In some embodiments, the antibodies or antigen binding fragments as described herein or ADC derived therefrom can induce complement-dependent cytotoxicity (CDC) and/or antibody dependent cellular cytotoxicity (ADCC), and kill the tumor cell.

In some embodiments, the antibodies or antigen binding fragments as described herein or ADC derived therefrom have a functional Fc region. In some embodiments, effector function of a functional Fc region is antibody-dependent cell-mediated cytotoxicity (ADCC). In some embodiments, effector function of a functional Fc region is phagocytosis. In some embodiments, effector function of a functional Fc region is ADCC and phagocytosis.

In some embodiments, the antibodies or antigen binding fragments as described herein or ADC derived therefrom can induce complement complement-dependent cytotoxicity (CDC).

In some embodiments, the Fc region is human IgG1, human IgG2, human IgG3, or human IgG4. In some embodiments, the antibody is a human IgG1 antibody, optionally with SI mutations, LALA mutations, N297A mutation, YTE mutations, and/or FLAA mutations. In some embodiments, the antibody is a human IgG4 antibody, optionally with SI mutations, LALA mutations, N297A mutation, YTE mutations, and/or FLAA mutations.

In some embodiments, the antibodies or antigen binding fragments as described herein or ADC derived therefrom do not have a functional Fc region. For example, the antibodies or antigen binding fragments are Fab, Fab', F(ab')₂, and Fv fragments. In some embodiments, the Fc region has LALA mutations (L234A and L235A mutations according to EU numbering), or LALA-PG mutations (L234A, L235A, P329G mutations according to EU numbering). In some embodiments, the Fc region has FLAA mutations (F234A and L235A according to EU numbering). In some embodiments, the Fc has SI mutations (S239D and I332E mutations according to EU numbering). In some embodiments, the Fc has N297A mutation according to EU numbering. In some embodiments, the Fc has YTE mutations (M252Y, S254T and T256E according to EU numbering).

Methods of Making Anti-MUC1 Antibodies

An isolated fragment of human MUC1 can be used as an immunogen to generate antibodies using standard techniques for polyclonal and monoclonal antibody preparation. Polyclonal antibodies can be raised in animals by multiple injections (e.g., subcutaneous or intraperitoneal injections) of an antigenic peptide or protein. In some embodiments, the antigenic peptide or protein is injected with at least one adjuvant. In some embodiments, the antigenic peptide or protein can be conjugated to an agent that is immunogenic in the species to be immunized. Animals can be injected with the antigenic peptide or protein more than one time (e.g., twice, three times, or four times).

The full-length polypeptide or protein can be used or, alternatively, antigenic peptide fragments thereof can be used as immunogens. The antigenic peptide of a protein comprises at least 8 (e.g., at least 10, 15, 20, or 30) amino acid residues of the amino acid sequence of MUC1

and encompasses an epitope of the protein such that an antibody raised against the peptide forms a specific immune complex with the protein. As described above, the full length sequence of human MUC1 (SEQ ID NO: 50) is known in the art. In some embodiments, an Fc-tagged or His-tagged human MUC1 protein is used as the immunogen.

5 An immunogen typically is used to prepare antibodies by immunizing a suitable subject (e.g., human or transgenic animal expressing at least one human immunoglobulin locus). An appropriate immunogenic preparation can contain, for example, a recombinantly-expressed or a chemically-synthesized polypeptide (e.g., a fragment of human MUC1). The preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or a similar

10 immunostimulatory agent.

Polyclonal antibodies can be prepared as described above by immunizing a suitable subject with a MUC1 polypeptide, or an antigenic peptide thereof (e.g., part of MUC1) as an immunogen. The antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme-linked immunosorbent assay (ELISA) using the immobilized

15 MUC1 polypeptide or peptide. If desired, the antibody molecules can be isolated from the mammal (e.g., from the blood) and further purified by well-known techniques, such as protein A or protein G chromatography to obtain the IgG fraction. At an appropriate time after immunization, e.g., when the specific antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques,

20 such as the hybridoma technique originally described by Kohler et al. (*Nature* 256:495-497, 1975), the human B cell hybridoma technique (Kozbor et al., *Immunol. Today* 4:72, 1983), the EBV-hybridoma technique (Cole et al., *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96, 1985), or trioma techniques. The technology for producing hybridomas is well known (see, generally, *Current Protocols in Immunology*, 1994, Coligan et al. (Eds.), John

25 Wiley & Sons, Inc., New York, NY). Hybridoma cells producing a monoclonal antibody are detected by screening the hybridoma culture supernatants for antibodies that bind the polypeptide or epitope of interest, e.g., using a standard ELISA assay.

Variants of the antibodies or antigen-binding fragments described herein can be prepared by introducing appropriate nucleotide changes into the DNA encoding a human, humanized, or

30 chimeric antibody, or antigen-binding fragment thereof described herein, or by peptide synthesis. Such variants include, for example, deletions, insertions, or substitutions of residues within the amino acids sequences that make-up the antigen-binding site of the antibody or an antigen-binding domain. In a population of such variants, some antibodies or antigen-binding fragments will have increased affinity for the target protein, e.g., MUC1. Any combination of deletions,

35 insertions, and/or combinations can be made to arrive at an antibody or antigen-binding fragment thereof that has increased binding affinity for the target. The amino acid changes introduced into the antibody or antigen-binding fragment can also alter or introduce new post-translational modifications into the antibody or antigen-binding fragment, such as changing (e.g., increasing or decreasing) the number of glycosylation sites, changing the type of glycosylation site (e.g.,

40 changing the amino acid sequence such that a different sugar is attached by enzymes present in a cell), or introducing new glycosylation sites.

Antibodies disclosed herein can be derived from any species of animal, including mammals. Non-limiting examples of native antibodies include antibodies derived from humans,

primates, e.g., monkeys and apes, cows, pigs, horses, sheep, camelids (e.g., camels and llamas), chicken, goats, and rodents (e.g., rats, mice, hamsters and rabbits), including transgenic rodents genetically engineered to produce human antibodies.

Human and humanized antibodies include antibodies having variable and constant regions derived from (or having the same amino acid sequence as those derived from) human germline immunoglobulin sequences. Human antibodies may include amino acid residues not encoded by human germline immunoglobulin sequences (e.g., mutations introduced by random or site-specific mutagenesis in vitro or by somatic mutation in vivo), for example in the CDRs.

A humanized antibody, typically has a human framework (FR) grafted with non-human CDRs. Thus, a humanized antibody has one or more amino acid sequence introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as “import” residues, which are typically taken from an “import” variable domain. Humanization can be essentially performed by e.g., substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. These methods are described in e.g., Jones et al., *Nature*, 321:522-525 (1986); Riechmann et al., *Nature*, 332:323-327 (1988); Verhoeyen et al., *Science*, 239:1534-1536 (1988); each of which is incorporated by reference herein in its entirety. Accordingly, “humanized” antibodies are chimeric antibodies wherein substantially less than an intact human V domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically mouse antibodies in which some CDR residues and some FR residues are substituted by residues from analogous sites in human antibodies.

The choice of human VH and VL domains to be used in making the humanized antibodies is very important for reducing immunogenicity. According to the so-called “best-fit” method, the sequence of the V domain of a mouse antibody is screened against the entire library of known human-domain sequences. The human sequence which is closest to that of the mouse is then accepted as the human FR for the humanized antibody (Sims et al., *J. Immunol.*, 151:2296 (1993); Chothia et al., *J. Mol. Biol.*, 196:901 (1987)).

It is further important that antibodies be humanized with retention of high specificity and affinity for the antigen and other favorable biological properties. To achieve this goal, humanized antibodies can be prepared by a process of analysis of the parental sequences and various conceptual humanized products using three-dimensional models of the parental and humanized sequences. Three-dimensional immunoglobulin models are commonly available and are familiar to those skilled in the art. Computer programs are available which illustrate and display probable three-dimensional conformational structures of selected candidate immunoglobulin sequences. Inspection of these displays permits analysis of the likely role of the residues in the functioning of the candidate immunoglobulin sequence, i.e., the analysis of residues that influence the ability of the candidate immunoglobulin to bind its antigen. In this way, FR residues can be selected and combined from the recipient and import sequences so that the desired antibody characteristic, such as increased affinity for the target antigen(s), is achieved.

Ordinarily, amino acid sequence variants of the human, humanized, or chimeric anti-MUC1 antibody will contain an amino acid sequence having at least 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% percent identity with a sequence present in the light or heavy chain of the original antibody.

In some embodiments, a mouse (e.g., RenMab™ mouse) with a humanized heavy chain immunoglobulin locus and a humanized kappa chain immunoglobulin locus is used to generate antibodies. The heavy chain immunoglobulin locus is a region on the chromosome that contains genes for the heavy chains of antibodies. The locus can include e.g., human IGHV (variable) genes, human IGHD (diversity) genes, human IGHJ (joining) genes, and mouse heavy chain constant domain genes. The kappa chain immunoglobulin locus is a region on the chromosome that contains genes that encode the light chains of antibodies (kappa chain). The kappa chain immunoglobulin locus can include e.g., human IGKV (variable) genes, human IGKJ (joining) genes, and mouse light chain constant domain genes. A detailed description regarding RenMab™ mice can be found in PCT/CN2020/075698 or US20200390073A1, which is incorporated herein by reference in its entirety.

In some embodiments, a mouse (e.g., RenLite™ mouse) with a humanized heavy chain immunoglobulin locus and a humanized kappa chain immunoglobulin locus is used to generate antibodies. The heavy chain immunoglobulin locus is a region on the chromosome that contains genes for the heavy chains of antibodies. The locus can include e.g., human IGHV (variable) genes, human IGHD (diversity) genes, human IGHJ (joining) genes, and mouse heavy chain constant domain genes. The kappa chain immunoglobulin locus is a region on the chromosome that contains genes that encode a common light chains. The kappa chain immunoglobulin locus can include e.g., a human IGKV (variable) gene, a human IGKJ (joining) gene, and mouse light chain constant domain genes. A detailed description regarding RenLite™ mice can be found in PCT/CN2021/097652, which is incorporated herein by reference in its entirety.

The antibodies generated by the mice have a full human VH, a full human VL, and mouse constant regions. In some embodiments, the human VH and human VL is linked to a human IgG constant regions (e.g., IgG1, IgG2, IgG3, and IgG4).

Identity or homology with respect to an original sequence is usually the percentage of amino acid residues present within the candidate sequence that are identical with a sequence present within the human, humanized, or chimeric anti-MUC1 antibody or fragment, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity.

Additional modifications to the anti-MUC1 antibodies or antigen-binding fragments can be made. For example, a cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated may have any increased half-life *in vitro* and/or *in vivo*. Homodimeric antibodies with increased half-life *in vitro* and/or *in vivo* can also be prepared using heterobifunctional cross-linkers as described, for example, in Wolff et al. (*Cancer Res.* 53:2560-2565, 1993). Alternatively, an antibody can be engineered which has dual Fc regions (see, for example, Stevenson et al., *Anti-Cancer Drug Design* 3:219-230, 1989).

In some embodiments, a covalent modification can be made to the anti-MUC1 antibody or antigen-binding fragment thereof. These covalent modifications can be made by chemical or enzymatic synthesis, or by enzymatic or chemical cleavage. Other types of covalent modifications of the antibody or antibody fragment are introduced into the molecule by reacting

targeted amino acid residues of the antibody or fragment with an organic derivatization agent that is capable of reacting with selected side chains or the N- or C-terminal residues.

In some embodiments, antibody variants are provided having a carbohydrate structure that lacks fucose attached (directly or indirectly) to an Fc region. For example, the amount of fucose in such antibody may be from 1% to 80%, from 1% to 65%, from 5% to 65% or from 20% to 40%. The amount of fucose is determined by calculating the average amount of fucose within the sugar chain at Asn297, relative to the sum of all glycostructures attached to Asn 297 (e.g. complex, hybrid and high mannose structures) as measured by MALDI-TOF mass spectrometry, as described in WO 2008/077546, for example. Asn297 refers to the asparagine residue located at about position 297 in the Fc region (Eu numbering of Fc region residues; or position 314 in Kabat numbering); however, Asn297 may also be located about ± 3 amino acids upstream or downstream of position 297, i.e., between positions 294 and 300, due to minor sequence variations in antibodies. Such fucosylation variants may have improved ADCC function. In some embodiments, to reduce glycan heterogeneity, the Fc region of the antibody can be further engineered to replace the Asparagine at position 297 with Alanine (N297A).

In some embodiments, to facilitate production efficiency by avoiding Fab-arm exchange, the Fc region of the antibodies was further engineered to replace the serine at position 228 (EU numbering) of IgG4 with proline (S228P). A detailed description regarding S228 mutation is described, e.g., in Silva et al. "The S228P mutation prevents in vivo and in vitro IgG4 Fab-arm exchange as demonstrated using a combination of novel quantitative immunoassays and physiological matrix preparation." *Journal of Biological Chemistry* 290.9 (2015): 5462-5469, which is incorporated by reference in its entirety.

Recombinant Vectors

The present disclosure also provides recombinant vectors (e.g., an expression vectors) that include an isolated polynucleotide disclosed herein (e.g., a polynucleotide that encodes a polypeptide disclosed herein), host cells into which are introduced the recombinant vectors (i.e., such that the host cells contain the polynucleotide and/or a vector comprising the polynucleotide), and the production of recombinant antibody polypeptides or fragments thereof by recombinant techniques.

As used herein, a "vector" is any construct capable of delivering one or more polynucleotide(s) of interest to a host cell when the vector is introduced to the host cell. An "expression vector" is capable of delivering and expressing the one or more polynucleotide(s) of interest as an encoded polypeptide in a host cell into which the expression vector has been introduced. Thus, in an expression vector, the polynucleotide of interest is positioned for expression in the vector by being operably linked with regulatory elements such as a promoter, enhancer, and/or a poly-A tail, either within the vector or in the genome of the host cell at or near or flanking the integration site of the polynucleotide of interest such that the polynucleotide of interest will be translated in the host cell introduced with the expression vector.

A vector can be introduced into the host cell by methods known in the art, e.g., electroporation, chemical transfection (e.g., DEAE-dextran), transformation, transfection, and infection and/or transduction (e.g., with recombinant virus). Thus, non-limiting examples of

vectors include viral vectors (which can be used to generate recombinant virus), naked DNA or RNA, plasmids, cosmids, phage vectors, and DNA or RNA expression vectors associated with cationic condensing agents.

In some implementations, a polynucleotide disclosed herein (e.g., a polynucleotide that encodes a polypeptide disclosed herein) is introduced using a viral expression system (e.g., vaccinia or other pox virus, retrovirus, or adenovirus), which may involve the use of a non-pathogenic (defective), replication competent virus, or may use a replication defective virus. In the latter case, viral propagation generally will occur only in complementing virus packaging cells. Suitable systems are disclosed, for example, in Fisher-Hoch et al., 1989, Proc. Natl. Acad. Sci. USA 86:317-321; Flexner et al., 1989, Ann. N.Y. Acad. Sci. 569:86-103; Flexner et al., 1990, Vaccine, 8:17-21; U.S. Pat. Nos. 4,603,112, 4,769,330, and 5,017,487; WO 89/01973; U.S. Pat. No. 4,777,127; GB 2,200,651; EP 0,345,242; WO 91/02805; Berkner-Biotechniques, 6:616-627, 1988; Rosenfeld et al., 1991, Science, 252:431-434; Kolls et al., 1994, Proc. Natl. Acad. Sci. USA, 91:215-219; Kass-Eisler et al., 1993, Proc. Natl. Acad. Sci. USA, 90:11498-11502; Guzman et al., 1993, Circulation, 88:2838-2848; and Guzman et al., 1993, Cir. Res., 73:1202-1207. Techniques for incorporating DNA into such expression systems are well known to those of ordinary skill in the art. The DNA may also be "naked," as described, for example, in Ulmer et al., 1993, Science, 259:1745-1749, and Cohen, 1993, Science, 259:1691-1692. The uptake of naked DNA may be increased by coating the DNA onto biodegradable beads that are efficiently transported into the cells.

For expression, the DNA insert comprising an antibody-encoding or polypeptide-encoding polynucleotide disclosed herein can be operatively linked to an appropriate promoter (e.g., a heterologous promoter), such as the phage lambda PL promoter, the *E. coli* lac, trp and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters are known to the skilled artisan. The expression constructs can further contain sites for transcription initiation, termination and, in the transcribed region, a ribosome binding site for translation. The coding portion of the mature transcripts expressed by the constructs may include a translation initiating at the beginning and a termination codon (UAA, UGA, or UAG) appropriately positioned at the end of the polypeptide to be translated.

As indicated, the expression vectors can include at least one selectable marker. Such markers include dihydrofolate reductase or neomycin resistance for eukaryotic cell culture and tetracycline or ampicillin resistance genes for culturing in *E. coli* and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as *E. coli*, *Streptomyces*, and *Salmonella typhimurium* cells; fungal cells, such as yeast cells; insect cells such as *Drosophila* S2 and *Spodoptera Sf9* cells; animal cells such as CHO, COS, Bowes melanoma, and HK 293 cells; and plant cells. Appropriate culture mediums and conditions for the host cells described herein are known in the art.

Non-limiting vectors for use in bacteria include pQE70, pQE60 and pQE-9, available from Qiagen; pBS vectors, Phagescript vectors, Bluescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene; and ptc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia. Non-limiting eukaryotic vectors include pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

Non-limiting bacterial promoters suitable for use include the *E. coli* lacI and lacZ promoters, the T3 and T7 promoters, the gpt promoter, the lambda PR and PL promoters and the trp promoter. Suitable eukaryotic promoters include the CMV immediate early promoter, the HSV thymidine kinase promoter, the early and late SV40 promoters, the promoters of retroviral LTRs, such as those of the Rous sarcoma virus (RSV), and metallothionein promoters, such as the mouse metallothionein-I promoter.

In the yeast *Saccharomyces cerevisiae*, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH may be used. For reviews, see Ausubel *et al.* (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York, N.Y., and Grant *et al.*, *Methods Enzymol.*, 153: 516-544 (1997).

Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection or other methods. Such methods are described in many standard laboratory manuals, such as Davis *et al.*, Basic Methods In Molecular Biology (1986), which is incorporated herein by reference in its entirety.

Transcription of DNA encoding an antibody of the present disclosure by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp that act to increase transcriptional activity of a promoter in a given host cell-type. Examples of enhancers include the SV40 enhancer, which is located on the late side of the replication origin at base pairs 100 to 270, the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

For secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the expressed polypeptide. The signals may be endogenous to the polypeptide or they may be heterologous signals.

The polypeptide (e.g., antibody) can be expressed in a modified form, such as a fusion protein (e.g., a GST-fusion) or with a histidine-tag, and may include not only secretion signals, but also additional heterologous functional regions. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence in the host cell, during purification, or during subsequent handling and storage. Also, peptide moieties can be added to the polypeptide to facilitate purification. Such regions can be removed prior to final preparation of the polypeptide. The addition of peptide moieties to polypeptides to engender secretion or excretion, to improve stability and to facilitate purification, among others, are familiar and routine techniques in the art.

Methods of Treatment

The antibodies or antigen-binding fragments thereof of the present disclosure can be used for various therapeutic purposes.

In one aspect, the disclosure provides methods for treating a cancer in a subject, methods of reducing the rate of the increase of volume of a tumor in a subject over time, methods of reducing the risk of developing a metastasis, or methods of reducing the risk of developing an

additional metastasis in a subject. In some embodiments, the treatment can halt, slow, retard, or inhibit progression of a cancer. In some embodiments, the treatment can result in the reduction of in the number, severity, and/or duration of one or more symptoms of the cancer in a subject.

In one aspect, the disclosure features methods that include administering a therapeutically effective amount of an antibody or antigen-binding fragment thereof disclosed herein to a subject in need thereof (e.g., a subject having, or identified or diagnosed as having, a cancer), e.g., breast cancer (e.g., triple-negative breast cancer), carcinoid cancer, cervical cancer, endometrial cancer, glioma, head and neck cancer, liver cancer, lung cancer, small cell lung cancer, lymphoma, melanoma, ovarian cancer, pancreatic cancer, prostate cancer, renal cancer, colorectal cancer, gastric cancer, testicular cancer, thyroid cancer, bladder cancer, urethral cancer, or hematologic malignancy. In some embodiments, the cancer is unresectable melanoma or metastatic melanoma, non-small cell lung carcinoma (NSCLC), small cell lung cancer (SCLC), bladder cancer, or metastatic hormone-refractory prostate cancer. In some embodiments, the cancer is NSCLC, ovarian cancer, melanoma, colorectal cancer, breast cancer, a hematological malignancy, head and neck cancer, gastrointestinal cancer, bladder cancer, or bone cancer. In some embodiments, the subject has a solid tumor. In some embodiments, the cancer is squamous cell carcinoma of the head and neck (SCCHN), renal cell carcinoma (RCC), triple-negative breast cancer (TNBC), or colorectal carcinoma. In some embodiments, the subject has Hodgkin's lymphoma. In some embodiments, the subject has triple-negative breast cancer (TNBC), gastric cancer, urothelial cancer, Merkel-cell carcinoma, or head and neck cancer. In some embodiments, the cancer is melanoma, pancreatic carcinoma, mesothelioma, hematological malignancies, especially Non-Hodgkin's lymphoma, lymphoma, chronic lymphocytic leukemia, or advanced solid tumors. In some embodiments, the cancer is oesophageal cancer, gastric cancer, breast cancer, ovarian cancer, bladder cancer, pancreatic cancer, lung cancer, colorectal cancer, stomach cancer, prostate cancer, kidney cancer, multiple myeloma, or Non-Hodgkin Lymphoma.

In some embodiments, the compositions and methods disclosed herein can be used for treatment of patients at risk for a cancer. Patients with cancer can be identified with various methods known in the art.

In one aspect, the disclosure provides methods for treating, preventing, or reducing the risk of developing disorders associated with an abnormal or unwanted immune response, e.g., an autoimmune disorder, e.g., by affecting the functional properties of the APC cells (e.g., by blocking the interaction between MUC1 and MUC1 ligands). These autoimmune disorders include, but are not limited to, Alopecia areata, lupus, ankylosing spondylitis, Meniere's disease, antiphospholipid syndrome, mixed connective tissue disease, autoimmune Addison's disease, multiple sclerosis, autoimmune hemolytic anemia, myasthenia gravis, autoimmune hepatitis, pemphigus vulgaris, Behcet's disease, pernicious anemia, bullous pemphigoid, polyarthritis nodosa, cardiomyopathy, polychondritis, celiac sprue-dermatitis, polyglandular syndromes, chronic fatigue syndrome (CFIDS), polymyalgia rheumatica, chronic inflammatory demyelinating, polymyositis and dermatomyositis, chronic inflammatory polyneuropathy, primary agammaglobulinemia, Churg-Strauss syndrome, primary biliary cirrhosis, cicatricial pemphigoid, psoriasis, CREST syndrome, Raynaud's phenomenon, cold agglutinin disease, Reiter's syndrome, Crohn's disease, Rheumatic fever, discoid lupus, rheumatoid arthritis, Cryoglobulinemia sarcoidosis, fibromyalgia, scleroderma, Grave's disease, Sjögren's syndrome,

Guillain-Barre, stiff-man syndrome, Hashimoto's thyroiditis, Takayasu arteritis, idiopathic pulmonary fibrosis, temporal arteritis/giant cell arteritis, idiopathic thrombocytopenia purpura (ITP), ulcerative colitis, IgA nephropathy, uveitis, diabetes (e.g., Type I), vasculitis, lichen planus, and vitiligo. The anti-MUC1 antibodies or antigen-binding fragments thereof can also be administered to a subject to treat, prevent, or reduce the risk of developing disorders associated with an abnormal or unwanted immune response associated with cell, tissue or organ transplantation, e.g., renal, hepatic, and cardiac transplantation, e.g., graft versus host disease (GVHD), or to prevent allograft rejection. In some embodiments, the subject has dermatological disorders, liver disease (e.g., cirrhosis), Hidradenitis, experimental autoimmune encephalomyelitis. In some embodiments, the subject has renal disease, lupus, Sjogren's syndrome, ulcerative colitis, psoriasis, Hidradenitis suppurativa, Immune Thrombocytopenia (ITP), or other inflammatory arthritis. In some embodiments, the subject has multiple sclerosis or myasthenia gravis. In some embodiments, the subject has Crohn's disease, ulcerative colitis or type 1 diabetes. In some embodiments, the subject has autoimmune thyroid disease, Grave's disease, multiple sclerosis, psoriasis, inflammatory bowel disease (e.g., Crohn's Disease (CD) and ulcerative colitis), rheumatoid arthritis, Sjögren's syndrome, autoimmune nephritis, or systemic lupus erythematosus. In some embodiments, the methods involve administering to the subject an effective amount of a composition comprising the antibody or antigen-binding fragment thereof as described herein.

As used herein, by an "effective amount" is meant an amount or dosage sufficient to effect beneficial or desired results including halting, slowing, retarding, or inhibiting progression of a disease, e.g., an autoimmune disease or a cancer. An effective amount will vary depending upon, e.g., an age and a body weight of a subject to which the antibody, antigen binding fragment, antibody-encoding polynucleotide, vector comprising the polynucleotide, and/or compositions thereof is to be administered, a severity of symptoms and a route of administration, and thus administration can be determined on an individual basis.

An effective amount can be administered in one or more administrations. By way of example, an effective amount of an antibody or an antigen binding fragment is an amount sufficient to ameliorate, stop, stabilize, reverse, inhibit, slow and/or delay progression of an autoimmune disease or a cancer in a patient or is an amount sufficient to ameliorate, stop, stabilize, reverse, slow and/or delay proliferation of a cell (e.g., a biopsied cell, any of the cancer cells described herein, or cell line (e.g., a cancer cell line)) *in vitro*. As is understood in the art, an effective amount of an antibody or antigen binding fragment may vary, depending on, *inter alia*, patient history as well as other factors such as the type (and/or dosage) of antibody used.

Effective amounts and schedules for administering the antibodies, antibody-encoding polynucleotides, and/or compositions disclosed herein may be determined empirically, and making such determinations is within the skill in the art. Those skilled in the art will understand that the dosage that must be administered will vary depending on, for example, the mammal that will receive the antibodies, antibody-encoding polynucleotides, and/or compositions disclosed herein, the route of administration, the particular type of antibodies, antibody-encoding polynucleotides, antigen binding fragments, and/or compositions disclosed herein used and other drugs being administered to the mammal. Guidance in selecting appropriate doses for antibody or antigen binding fragment can be found in the literature on therapeutic uses of antibodies and

antigen binding fragments, e.g., Handbook of Monoclonal Antibodies, Ferrone et al., eds., Nokes Publications, Park Ridge, N.J., 1985, ch. 22 and pp. 303-357; Smith et al., Antibodies in Human Diagnosis and Therapy, Haber et al., eds., Raven Press, New York, 1977, pp. 365-389.

A typical daily dosage of an effective amount of an antibody is 0.01 mg/kg to 100 mg/kg.

5 In some embodiments, the dosage can be less than 100 mg/kg, 10 mg/kg, 9 mg/kg, 8 mg/kg, 7 mg/kg, 6 mg/kg, 5 mg/kg, 4 mg/kg, 3 mg/kg, 2 mg/kg, 1 mg/kg, 0.5 mg/kg, or 0.1 mg/kg. In some embodiments, the dosage can be greater than 10 mg/kg, 9 mg/kg, 8 mg/kg, 7 mg/kg, 6 mg/kg, 5 mg/kg, 4 mg/kg, 3 mg/kg, 2 mg/kg, 1 mg/kg, 0.5 mg/kg, 0.1 mg/kg, 0.05 mg/kg, or 0.01 mg/kg. In some embodiments, the dosage is about 10 mg/kg, 9 mg/kg, 8 mg/kg, 7 mg/kg, 6
10 mg/kg, 5 mg/kg, 4 mg/kg, 3 mg/kg, 2 mg/kg, 1 mg/kg, 0.9 mg/kg, 0.8 mg/kg, 0.7 mg/kg, 0.6 mg/kg, 0.5 mg/kg, 0.4 mg/kg, 0.3 mg/kg, 0.2 mg/kg, or 0.1 mg/kg.

In any of the methods described herein, the at least one antibody, antigen-binding fragment thereof, or pharmaceutical composition (e.g., any of the antibodies, antigen-binding fragments, or pharmaceutical compositions described herein) and, optionally, at least one
15 additional therapeutic agent can be administered to the subject at least once a week (e.g., once a week, twice a week, three times a week, four times a week, once a day, twice a day, or three times a day). In some embodiments, at least two different antibodies and/or antigen-binding fragments are administered in the same composition (e.g., a liquid composition). In some embodiments, at least one antibody or antigen-binding fragment and at least one additional
20 therapeutic agent are administered in the same composition (e.g., a liquid composition). In some embodiments, the at least one antibody or antigen-binding fragment and the at least one additional therapeutic agent are administered in two different compositions (e.g., a liquid composition containing at least one antibody or antigen-binding fragment and a solid oral composition containing at least one additional therapeutic agent). In some embodiments, the at
25 least one additional therapeutic agent is administered as a pill, tablet, or capsule. In some embodiments, the at least one additional therapeutic agent is administered in a sustained-release oral formulation.

In some embodiments, the one or more additional therapeutic agents can be administered to the subject prior to, or after administering the at least one antibody, antigen-binding antibody
30 fragment, or pharmaceutical composition (e.g., any of the antibodies, antigen-binding antibody fragments, or pharmaceutical compositions described herein). In some embodiments, the one or more additional therapeutic agents and the at least one antibody, antigen-binding antibody fragment, or pharmaceutical composition (e.g., any of the antibodies, antigen-binding antibody fragments, or pharmaceutical compositions described herein) are administered to the subject such
35 that there is an overlap in the bioactive period of the one or more additional therapeutic agents and the at least one antibody or antigen-binding fragment (e.g., any of the antibodies or antigen-binding fragments described herein) in the subject.

In some embodiments, the subject can be administered the at least one antibody, antigen-binding antibody fragment, or pharmaceutical composition (e.g., any of the antibodies, antigen-binding antibody fragments, or pharmaceutical compositions described herein) over an extended
40 period of time (e.g., over a period of at least 1 week, 2 weeks, 3 weeks, 1 month, 2 months, 3 months, 4 months, 5 months, 6 months, 7 months, 8 months, 9 months, 10 months, 11 months, 12 months, 1 year, 2 years, 3 years, 4 years, or 5 years). A skilled medical professional may

determine the length of the treatment period using any of the methods described herein for diagnosing or following the effectiveness of treatment (e.g., the observation of at least one symptom of cancer). As described herein, a skilled medical professional can also change the identity and number (e.g., increase or decrease) of antibodies or antigen-binding antibody fragments (and/or one or more additional therapeutic agents) administered to the subject and can also adjust (e.g., increase or decrease) the dosage or frequency of administration of at least one antibody or antigen-binding antibody fragment (and/or one or more additional therapeutic agents) to the subject based on an assessment of the effectiveness of the treatment (e.g., using any of the methods described herein and known in the art).

In some embodiments, one or more additional therapeutic agents can be administered to the subject. The additional therapeutic agent can comprise one or more inhibitors selected from the group consisting of an inhibitor of B-Raf, an EGFR inhibitor, an inhibitor of a MEK, an inhibitor of ERK, an inhibitor of K-Ras, an inhibitor of c-Met, an inhibitor of anaplastic lymphoma kinase (ALK), an inhibitor of a phosphatidylinositol 3-kinase (PI3K), an inhibitor of an Akt, an inhibitor of mTOR, a dual PI3K/mTOR inhibitor, an inhibitor of Bruton's tyrosine kinase (BTK), and an inhibitor of Isocitrate dehydrogenase 1 (IDH1) and/or Isocitrate dehydrogenase 2 (IDH2). In some embodiments, the additional therapeutic agent is an inhibitor of indoleamine 2,3-dioxygenase-1 (IDO1) (e.g., epacadostat).

In some embodiments, the additional therapeutic agent can comprise one or more inhibitors selected from the group consisting of an inhibitor of HER3, an inhibitor of LSD1, an inhibitor of MDM2, an inhibitor of BCL2, an inhibitor of CHK1, an inhibitor of activated hedgehog signaling pathway, and an agent that selectively degrades the estrogen receptor.

In some embodiments, the additional therapeutic agent can comprise one or more therapeutic agents selected from the group consisting of Trabectedin, nab-paclitaxel, Trabectedin, Pazopanib, Cediranib, Palbociclib, everolimus, fluoropyrimidine, IFL, regorafenib, Reolysin, Alimta, Zykadia, Sutent, temsirolimus, axitinib, everolimus, sorafenib, Votrient, Pazopanib, IMA-901, AGS-003, cabozantinib, Vinflunine, an Hsp90 inhibitor, Ad-GM-CSF, Temazolomide, IL-2, IFNa, vinblastine, Thalomid, dacarbazine, cyclophosphamide, lenalidomide, azacytidine, lenalidomide, bortezomid, amrubicine, carfilzomib, pralatrexate, and enzastaurin.

In some embodiments, the additional therapeutic agent can comprise one or more therapeutic agents selected from the group consisting of an adjuvant, a TLR agonist, tumor necrosis factor (TNF) alpha, IL-1, HMGB1, an IL-10 antagonist, an IL-4 antagonist, an IL-13 antagonist, an IL-17 antagonist, an HVEM antagonist, an ICOS agonist, a treatment targeting CX3CL1, a treatment targeting CXCL9, a treatment targeting CXCL10, a treatment targeting CCL5, an LFA-1 agonist, an ICAM1 agonist, and a Selectin agonist.

In some embodiments, carboplatin, nab-paclitaxel, paclitaxel, cisplatin, pemetrexed, gemcitabine, FOLFOX, or FOLFIRI are administered to the subject.

In some embodiments, the additional therapeutic agent is an anti-OX40 antibody, an anti-PD-1 antibody, an anti-PD-L1 antibody, an anti-PD-L2 antibody, an anti-LAG-3 antibody, an anti-TIGIT antibody, an anti-BTLA antibody, an anti-CTLA-4 antibody, anti-ICOS antibody, anti-CD27 antibody, anti-OX40 antibody, anti-4-1BB antibody, anti-CD40 antibody, and/or an anti-GITR antibody.

In one aspect, the disclosure provides a combination therapy. In some embodiments, the anti-MUC1 antibody or antigen-binding fragment thereof (e.g., any antibody described herein) can be administered together with an anti-PD-L1 antibody. In some embodiments, the anti-MUC1 antibody or antigen-binding fragment thereof (e.g., any antibody described herein) can be administered together with an anti-OX40 antibody. In some embodiments, the anti-MUC1 antibody or antigen-binding fragment thereof (e.g., any antibody described herein) can be administered together with an anti-4-1-BB antibody.

Pharmaceutical Compositions and Routes of Administration

Also provided herein are pharmaceutical compositions that contain at least one (e.g., one, two, three, or four) of the antibodies or antigen-binding fragments described herein. Two or more (e.g., two, three, or four) of any of the antibodies or antigen-binding fragments described herein can be present in a pharmaceutical composition in any combination. The pharmaceutical compositions may be formulated in any manner known in the art.

Pharmaceutical compositions are formulated to be compatible with their intended route of administration (e.g., intravenous, intraarterial, intramuscular, intradermal, subcutaneous, or intraperitoneal). The compositions can include a sterile diluent (e.g., sterile water or saline), a fixed oil, polyethylene glycol, glycerine, propylene glycol or other synthetic solvents, antibacterial or antifungal agents, such as benzyl alcohol or methyl parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like, antioxidants, such as ascorbic acid or sodium bisulfite, chelating agents, such as ethylenediaminetetraacetic acid, buffers, such as acetates, citrates, or phosphates, and isotonic agents, such as sugars (e.g., dextrose), polyalcohols (e.g., mannitol or sorbitol), or salts (e.g., sodium chloride), or any combination thereof. Liposomal suspensions can also be used as pharmaceutically acceptable carriers (see, e.g., U.S. Patent No. 4,522,811). Preparations of the compositions can be formulated and enclosed in ampules, disposable syringes, or multiple dose vials. Where required (as in, for example, injectable formulations), proper fluidity can be maintained by, for example, the use of a coating, such as lecithin, or a surfactant. Absorption of the antibody or antigen-binding fragment thereof can be prolonged by including an agent that delays absorption (e.g., aluminum monostearate and gelatin). Alternatively, controlled release can be achieved by implants and microencapsulated delivery systems, which can include biodegradable, biocompatible polymers (e.g., ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid; Alza Corporation and Nova Pharmaceutical, Inc.).

Compositions containing one or more of any of the antibodies or antigen-binding fragments described herein can be formulated for parenteral (e.g., intravenous, intraarterial, intramuscular, intradermal, subcutaneous, or intraperitoneal) administration in dosage unit form (i.e., physically discrete units containing a predetermined quantity of active compound for ease of administration and uniformity of dosage).

Pharmaceutical compositions for parenteral administration are preferably sterile and substantially isotonic and manufactured under Good Manufacturing Practice (GMP) conditions. Pharmaceutical compositions can be provided in unit dosage form (i.e., the dosage for a single administration). Pharmaceutical compositions can be formulated using one or more

physiologically acceptable carriers, diluents, excipients or auxiliaries. The formulation depends on the route of administration chosen. For injection, antibodies can be formulated in aqueous solutions, preferably in physiologically-compatible buffers to reduce discomfort at the site of injection. The solution can contain formulatory agents such as suspending, stabilizing and/or dispersing agents. Alternatively antibodies can be in lyophilized form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

Toxicity and therapeutic efficacy of compositions can be determined by standard pharmaceutical procedures in cell cultures or experimental animals (e.g., monkeys). One can, for example, determine the LD₅₀ (the dose lethal to 50% of the population) and the ED₅₀ (the dose therapeutically effective in 50% of the population): the therapeutic index being the ratio of LD₅₀:ED₅₀. Agents that exhibit high therapeutic indices are preferred. Where an agent exhibits an undesirable side effect, care should be taken to minimize potential damage (i.e., reduce unwanted side effects). Toxicity and therapeutic efficacy can be determined by other standard pharmaceutical procedures.

Data obtained from cell culture assays and animal studies can be used in formulating an appropriate dosage of any given agent for use in a subject (e.g., a human). A therapeutically effective amount of the one or more (e.g., one, two, three, or four) antibodies or antigen-binding fragments thereof (e.g., any of the antibodies or antibody fragments described herein) will be an amount that treats the disease in a subject (e.g., kills cancer cells) in a subject (e.g., a human subject identified as having cancer), or a subject identified as being at risk of developing the disease (e.g., a subject who has previously developed cancer but now has been cured), decreases the severity, frequency, and/or duration of one or more symptoms of a disease in a subject (e.g., a human). The effectiveness and dosing of any of the antibodies or antigen-binding fragments described herein can be determined by a health care professional or veterinary professional using methods known in the art, as well as by the observation of one or more symptoms of disease in a subject (e.g., a human). Certain factors may influence the dosage and timing required to effectively treat a subject (e.g., the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and the presence of other diseases).

Exemplary doses include milligram or microgram amounts of any of the antibodies or antigen-binding fragments described herein per kilogram of the subject's weight (e.g., about 1 µg/kg to about 500 mg/kg; about 100 µg/kg to about 500 mg/kg; about 100 µg/kg to about 50 mg/kg; about 10 µg/kg to about 5 mg/kg; about 10 µg/kg to about 0.5 mg/kg; or about 1 µg/kg to about 50 µg/kg). While these doses cover a broad range, one of ordinary skill in the art will understand that therapeutic agents, including antibodies and antigen-binding fragments thereof, vary in their potency, and effective amounts can be determined by methods known in the art. Typically, relatively low doses are administered at first, and the attending health care professional or veterinary professional (in the case of therapeutic application) or a researcher (when still working at the development stage) can subsequently and gradually increase the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular subject will depend upon a variety of factors including the activity of the specific compound employed, the age, body weight, general health, gender, and diet of the subject, the time of administration, the route of administration, the rate of excretion, and the half-life of the antibody or antibody fragment *in vivo*.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration. The disclosure also provides methods of manufacturing the antibodies or antigen binding fragments thereof for various uses as described herein.

5

EXAMPLES

The invention is further described in the following examples, which do not limit the scope of the invention described in the claims.

10 Example 1. Generating human anti-MUC1 antibodies

Human MUC1 protein (hMUC1-his, position 24-1158 of SEQ NO ID: 50) or DNA encoding this protein was emulsified with adjuvants, and was used to immunize RenLite™ mice (Biocytogen, complete human heavy chain variable domain combined with a common light chain substitution *in situ*). The RenLite™ mice are described, e.g., in PCT/CN2021/097652, which is incorporated herein by reference in its entirety. Before immunization, retro-orbital blood was collected as a negative control.

Freund's complete adjuvant (CFA) was used for the first immunization and Freund's incomplete adjuvant (IFA) was used for the second and third immunizations. A total of three immunizations were performed. The first and second immunizations were separated by two weeks, and the second and third immunizations were also separated by two weeks. One week after the third immunization, retro-orbital blood was collected, and the antibody titer of serum was detected by fluorescence activated cell sorting (FACS). One week later, mice with high titer were further injected with hMUC1-his protein and DNA through intraperitoneal injection and intramuscular injection respectively for impulse immunization.

Antigen-specific immune cells were isolated from the immunized mice to further obtain anti-MUC1 antibodies or to obtain the light chain and heavy chain variable region sequences of the anti-MUC1 antibodies. For example, single cell technology (e.g., Beacon® Optofluidic System, Berkeley Lights Inc.) was used to screen and find plasma cells secreting antigen-specific monoclonal antibodies, and then reverse transcription and PCR sequencing were used to obtain antibody variable region sequences. The obtained variable region sequences were cloned into a vector containing a sequence encoding the human IgG1 constant region for antibody expression. Binding of the expressed antibodies to MUC1 was verified by FACS. Exemplary antibodies obtained by this method included: 9H4, 10B1, 10G12, 11A6, 11D6 and 11E8.

These antibodies contained substantially the same light chain, and their VH and VL CDR 1-3 sequences are shown in **FIG. 3 or FIG. 4**. The VH and VL regions of these antibodies are shown in **FIG. 5**.

Example 2. Cross-species binding of anti-MUC1 antibodies

H293F-hMUC1(24-1158) cells, H293F-hMUC1(961-1152) cells, CHO-fasMUC1 cells, CHO-mMUC1 cells, CHO-dMUC1 cells or ZR-75-1 cells were transferred to a 96-well plate at a density of 1×10^5 cells/well respectively. The anti-MUC1 antibodies were added to the 96-well

plate, and incubated at 4°C for 30 min. Then, the cells were incubated with the secondary antibody anti-hIgG-Fc-Alex Flour 647 (RL1-H) (Jackson ImmunoResearch Laboratories, Inc., Cat#: 109-606-170) at 4°C in the dark for 15 minutes before flow cytometry analysis. For the negative control (NC), only cells were added and no antibodies were added.

5 H293F-hMUC1(24-1158) cells and H293F-hMUC1(961-1152) cells were obtained by transfecting H293F cells with vectors expressing position 24-1158 (hMUC1(24-1158)) or 961-1152 (hMUC1(961-1152)) of human MUC1 (SEQ ID NO: 50), respectively. CHO-fasMUC1 cells, CHO-dMUC1 cells and CHO-mMUC1 cells were obtained by transfecting CHO-S cells with vectors expressing monkey (*Macaca fascicularis*) MUC1 (fasMUC1, SEQ ID NO: 51), dog
10 MUC1 (dMUC1, SEQ ID NO: 52) or mouse MUC1 (mMUC1, SEQ ID NO: 53) respectively. ZR-75-1 is a human breast cancer cell line with MUC1 high-expression.

The test results are shown in the table below. All anti-MUC1 antibodies 9H4, 10B1, 10G12, 11A6, 11D6 and 11E8 can bind to human MUC1 and monkey MUC1. Among these antibodies, 11E8 can also bind to mouse MUC1 and dog MUC1.

15

Table 1

Antibody	MUC1	Percentage of Positive Cells	Evaluation
NC	hMUC1(24-1158)	0.1%	No Binding
	hMUC1(961-1152)	0.2%	No Binding
	fasMUC1	0.1	No Binding
	mMUC1	0	No Binding
	dMUC1	0	No Binding
	ZR-75-1	0.7%	No Binding
NC+ secondary antibody	hMUC1(24-1158)	0.1%	No Binding
	hMUC1(961-1152)	0.2%	No Binding
	fasMUC1	0.3%	No Binding
	mMUC1	0.1%	No Binding
	dMUC1	0	No Binding
	ZR-75-1	1.0%	No Binding
9H4	hMUC1(24-1158)	37.3%	Binding
	hMUC1(961-1152)	64.3%	Binding
	fasMUC1	43.8%	Binding
	mMUC1	1.0%	No Binding
	dMUC1	2.1%	No Binding
	ZR-75-1	100%	Binding
10B1	hMUC1(24-1158)	36.0%	Binding
	hMUC1(961-1152)	64.3%	Binding
	fasMUC1	45.8%	Binding
	mMUC1	0.2%	No Binding
	dMUC1	0.2%	No Binding
	ZR-75-1	99.9%	Binding
10G12	hMUC1(24-1158)	36.6%	Binding
	hMUC1(961-1152)	63.3%	Binding
	fasMUC1	39.9%	Binding
	mMUC1	0.4%	No Binding

	dMUC1	0.2%	No Binding
	ZR-75-1	99.7%	Binding
11A6	hMUC1(24-1158)	36.8%	Binding
	hMUC1(961-1152)	63.3%	Binding
	fasMUC1	40.1%	Binding
	mMUC1	0.1%	No Binding
	dMUC1	0.1%	No Binding
	ZR-75-1	99.9%	Binding
11D6	hMUC1(24-1158)	36.4%	Binding
	hMUC1(961-1152)	61.2%	Binding
	fasMUC1	42.6%	Binding
	mMUC1	3.6%	No Binding
	dMUC1	0.1%	No Binding
	ZR-75-1	99.9%	Binding
11E8	hMUC1(24-1158)	36.8%	Binding
	hMUC1(961-1152)	67.2%	Binding
	fasMUC1	49.0%	Binding
	mMUC1	30.7%	Binding
	dMUC1	33.5%	Binding
	ZR-75-1	99.8%	Binding

Example 3. Cross binding activity of anti-MUC1 antibody to MUC family proteins

Similar to the cross-species binding experiment, the cross binding activities of anti-MUC1 antibodies to MUC family proteins were also tested by flow cytometry analysis. The CHO-S cells were transfected with vectors expressing human MUC1b (hMUC1b, SEQ ID NO: 54), human MUC18 (hMUC18, SEQ ID NO: 55), position 12067-13286 of human MUC16 (hMUC16(12067-13286), SEQ ID NO: 56) or position 13287-14451 of human MUC16 (hMUC16(13287-14451), SEQ ID NO: 67), respectively.

The above cells and H293F-hMUC1 (961-1152) cells were plated in a 96-well plate at a density of 1×10^5 cells/well. Anti-MUC1 antibodies (10 ug/mL) were added to the 96-well plate, and incubated at 4°C for 30 min. Then, the cells were incubated with the secondary antibody Anti-hIgG-Fc-Alexa Fluor 647 (RL1-H) at 4°C in the dark for 15 minutes before flow cytometry analysis.

The test results are shown in the table below. None of the anti-MUC1 antibodies can bind human MUC1b, human MUC16 or human MUC18.

Table 2

Antibody	MUC Protein	Percentage of Positive Cells	Evaluation
NC	hMUC1(961-1152)	0.4%	No binding
	hMUC1b	0.1%	No binding
	hMUC16(12067-13286)	0.1%	No binding
	hMUC16(13287-14451)	0.1%	No binding
	hMUC18	0.1%	No binding
9H4	hMUC1(961-1152)	70.2%	Binding

	hMUC1b	1.8%	No binding
	hMUC16(12067-13286)	0.8%	No binding
	hMUC16(13287-14451)	1.1%	No binding
	hMUC18	1.3%	No binding
10B1	hMUC1(961-1152)	66.8%	Binding
	hMUC1b	1.1%	No binding
	hMUC16(12067-13286)	0.2%	No binding
	hMUC16(13287-14451)	0.2%	No binding
	hMUC18	0.5%	No binding
10G12	hMUC1(961-1152)	67.0%	Binding
	hMUC1b	0.4%	No binding
	hMUC16(12067-13286)	0.3%	No binding
	hMUC16(13287-14451)	0.4%	No binding
	hMUC18	0.5%	No binding
11A6	hMUC1(961-1152)	72.0%	Binding
	hMUC1b	0.5%	No binding
	hMUC16(12067-13286)	0.1%	No binding
	hMUC16(13287-14451)	0.4%	No binding
	hMUC18	0.3%	No binding
11D6	hMUC1(961-1152)	71.3%	Binding
	hMUC1b	0.6%	No binding
	hMUC16(12067-13286)	0.5%	No binding
	hMUC16(13287-14451)	1.3%	No binding
	hMUC18	0.7%	No binding
11E8	hMUC1(961-1152)	73.7%	Binding
	hMUC1b	1.0%	No binding
	hMUC16(12067-13286)	0.3%	No binding
	hMUC16(13287-14451)	0.7%	No binding
	hMUC18	0.5%	No binding

Furthermore, Western Blot was used to detect the binding epitope of the antibodies. Briefly, sample proteins was loaded on SDS-PAGE gels ((Dakewe (Shenzhen) Medical Equipment Co., Ltd., Cat#: 8012031)) and ran at 80 V for 28 min or 100 V for 60 min respectively, followed by transfer to PVDF membranes according to the manufacturer's recommendations. The SDS-PAGE gels make it possible to analyze protein in the sizes ranging from 10 to 250 kDa. For the analysis of the binding epitope, 5 μ g human MUC1(1036-1155) protein (ACRO Biosystems Inc., Cat#: MU1-H52H5) and human MUC1(1098-1158) protein (ACRO Biosystems Inc., Cat#: MU1-H5259) were applied on the SDS-PAGE gradient gel. 1H7 analog was used as a positive control. PVDF membrane was activated in methanol before blocking (5% dry milk in TBST buffer) for 1 hour at room temperature during constant shaking. The membrane was then incubated with antibody 10B1 or 1H7 analog overnight at 4°C. Next day, the membrane was washed by TBST buffer and incubated with the secondary antibody horseradish peroxidase labeled goat anti rabbit IgG (H+L) (Beyotime Biotechnology, Cat#:

A0208). A CCD-camera (Clinx Science Instruments, ChemiScope6100) was used for detection of signal from the substrate.

1H7 is a murine monoclonal IgG antibody targeting the extracellular region of human MUC1 C-terminal subunit, developed by Chungbuk National University and Peptron Co. Ltd. The heavy chain variable region and light chain variable region of 1H7 are shown in SEQ ID NO: 57 and SEQ ID NO: 58 respectively. The heavy chain variable region and light chain variable region was linked to human IgG1 constant region to form the positive control 1H7 analog.

Western Blot results indicated that 10B1 targets post-cleavage membrane bounded MUC1. The previous flow cytometry results showed that 10B1 didn't bind to hMUC1b, potentially due to incorrect expression of hMUC1b on CHO-S cells.

Example 4. Internalization of anti-MUC1 antibodies

Anti-MUC1 antibodies (2.5µg/mL) and pHAb-AffiniPure Fab Goat Anti-Human IgG Secondary Antibody were added to the ZR-75-1 cells or CHO-hMUC1(961-1152) cells (obtained by transfecting CHO-S cells with hMUC1(961-1152)). After incubating for 6 hours, the cells were centrifuged and washed in FACS buffer. MFI was detected on a flow cytometer, and the endocytosis rates of anti-MUC1 antibodies were calculated. Human IgG1 was used as ISO control. The results are shown in the table below, which showed that all anti-MUC1 antibodies 9H4, 10B1, 10G12, 11A6, 11D6 and 11E8 had better endocytosis rates in ZR-75-1 cells and CHO-hMUC1(961-1152) cells than positive control 1H7 analog. For the negative control (NC), only cells were added and no antibodies were added.

Table 3

Antibody	ZR-75-1		CHO-hMUC1(961-1152)	
	MFI	Percentage of Positive	MFI	Percentage of Positive
NC	5924	0.9%	6923	0.4%
NC+ secondary antibody	6850	0.8%	8740	0.9%
ISO	7013	1.0%	8488	0.7%
1H7 analog	45123	94.9%	165025	64.4%
9H4	69518	98.0%	568000	75.9%
10B1	76272	99.2%	402000	74.9%
10G12	66777	97.6%	470000	72.5%
11A6	79078	98.2%	410000	70.9%
11D6	75184	98.0%	456000	70.2%
11E8	72136	99.0%	303857	68.8%

Example 5. Binding affinity of anti-MUC1 antibodies

The binding affinity of the anti-MUC1 antibodies to his-tagged human MUC1 proteins (including hMUC1(24-1158) and hMUC1(961-1152)) and his-tagged monkey MUC1 protein (fasMUC1) were verified using Biacore™ (Biacore, Inc., Piscataway N.J.) 8K biosensor equipped with pre-immobilized Protein A sensor chips.

Purified anti-MUC1 antibodies was captured on the Protein A chip for the detection. 1 µg/mL purified antibodies was loaded at 10 µL/min to bind to a gradient concentration of

recombinant his-tagged human MUC1 proteins (200 nM, 100 nM, 50 nM, 25 nM, 12.5 nM, 6.25 nM, 3.125 nM, 1.5625 nM or 0 nM). For ISO control, an antibody targeting an irrelevant target protein was used. The flow rate was 30 μ L/min, the binding and dissociation time were set to 180 s and 400 s, respectively. The chip was regenerated after the last injection of each titration with a glycine solution (pH 2.0) at 30 μ L/min for 30 seconds.

Kinetic association rates (k_{on}) and dissociation rates (k_{off}) were obtained simultaneously by fitting the data globally to a 1:1 Langmuir binding model (Karlsson, R. Roos, H. Fagerstam, L. Petersson, B., 1994. *Methods Enzymology* 6. 99-110) using Biacore™ 8K Evaluation Software 3.0. Affinities were deduced from the quotient of the kinetic rate constants ($K_D = k_{off}/k_{on}$).

As a person of ordinary skill in the art would understand, the same method with appropriate adjustments for parameters (e.g., antibody concentration) was performed for each tested antibody. The results for the tested antibodies are summarized in the table below.

Table 4

Antibody	Analyte 1 Solution	k_{on} (1/Ms)	k_{off} (1/s)	K_D (M)
ISO	hMUC1(24-1158)	—	—	No binding
	hMUC1(961-1152)	—	—	No binding
	fasMUC1	—	—	No binding
9H4	hMUC1(24-1158)	1.12E+05	1.07E-02	9.54E-08
	hMUC1(961-1152)	9.60E+05	1.70E-02	1.77E-08
	fasMUC1	1.03E+05	3.98E-03	3.88E-08
10B1	hMUC1(24-1158)	4.12E+04	5.97E-04	1.45E-08
	hMUC1(961-1152)	1.39E+05	8.26E-04	5.94E-09
	fasMUC1	2.14E+04	4.56E-04	2.13E-08
10G12	hMUC1(24-1158)	1.95E+05	7.24E-03	3.71E-08
	hMUC1(961-1152)	1.18E+06	8.45E-03	7.15E-09
	fasMUC1	1.14E+05	8.81E-04	7.70E-09
11A6	hMUC1(24-1158)	1.84E+05	1.12E-03	6.08E-09
	hMUC1(961-1152)	7.46E+05	1.23E-03	1.64E-09
	fasMUC1	—	—	—
11D6	hMUC1(24-1158)	7.51E+04	5.14E-03	6.85E-08
	hMUC1(961-1152)	3.36E+05	6.79E-03	2.02E-08
	fasMUC1	1.58E+04	2.69E-03	1.70E-07
11E8	hMUC1(24-1158)	7.35E+04	4.91E-04	6.68E-09
	hMUC1(961-1152)	2.43E+05	8.15E-04	3.35E-09
	fasMUC1	4.78E+04	6.19E-03	1.30E-07

The results show that all anti-MUC1 antibodies 9H4, 10B1, 10G12, 11A6, 11D6 and 11E8 can bind to human MUC1 with high affinity. In addition, 9H4, 10B1, 10G12, 11D6 and 11E8 can also bind to monkey MUC1.

In another experiment, the binding affinities of the anti-MUC1 antibodies to his-tagged human MUC1 proteins (including hMUC1(890-1158) and hMUC1(1036-1155)) and his-tagged monkey MUC1 protein (cMUC1) were verified using Biacore™ 8K biosensor equipped with pre-immobilized Protein A sensor chips.

Specifically, hMUC1(890-1158) (ACRO Biosystems Inc., Cat#: MU1-H52H9, containing position 890-1158 of human MUC1 (SEQ ID NO: 50)), hMUC1(1036-1155) (ACRO Biosystems Inc., Cat#: MU1-H52H5, containing position 1036-1155 of human MUC1 (SEQ ID NO: 50)) and cMUC1 (ACRO Biosystems Inc., Cat#: MU1-C52H5, containing position 254-373 of monkey MUC1 XP_005541632.1) were diluted to 200 nM with 1× HBS-EP+ buffer (pH 7.4). For ISO control, an antibody targeting an irrelevant target protein was used. The protein was injected into the Biacore™ 8K biosensor at 10 µL/min for about 50 seconds to achieve a desired protein density (e.g., about 200 response units (RU)). Purified antibodies at a concentration of 2 µg/mL were then injected at 10 µL/min for 50 seconds. the binding and dissociation time were set to 180 s and 600 s, respectively. The chip was regenerated after the last injection of each titration with a glycine solution (pH 2.0) at 30 µL/min for 30 seconds.

Kon and koff were obtained simultaneously by fitting the data globally to a 1:1 Langmuir binding model (Karlsson, R. Roos, H. Fagerstam, L. Petersson, B., 1994. *Methods Enzymology* 6. 99-110) using Biacore™ 8K Evaluation Software 3.0. Affinities were deduced from the quotient of the kinetic rate constants ($KD=koff/kon$).

The results are shown in the table below, which indicated that both 10B1 and 11E8 can bind to hMUC1(890-1158), hMUC1(1036-1155) and cMUC1 with high affinity, while the positive control Gatipotuzumab analog only binds to hMUC1(890-1158), not binds to hMUC1(1036-1155) or cMUC1; and the positive control 1H7 analog showed lower binding affinity to monkey MUC1, as compared with 10B1 and 11E8.

Table 5

Antibody	Analyte 1 Solution	kon (1/Ms)	koff (1/s)	KD (M)
ISO	hMUC1(890-1158)	—	—	No binding
	hMUC1(1036-1155)	—	—	No binding
	cMUC1	—	—	No binding
Gatipotuzumab analog	hMUC1(890-1158)	6.38E+04	1.97E-02	3.08E-07
	hMUC1(1036-1155)	—	—	No binding
	cMUC1	—	—	No binding
1H7 analog	hMUC1(890-1158)	4.70E+04	2.69E-04	5.73E-09
	hMUC1(1036-1155)	7.01E+04	1.65E-04	2.36E-09
	cMUC1	3.72E+03	1.58E-02	4.25E-06
10B1	hMUC1(890-1158)	1.36E+05	8.12E-04	5.97E-09
	hMUC1(1036-1155)	3.04E+05	1.00E-03	3.29E-09
	cMUC1	4.72E+05	6.16E-04	1.31E-09
11E8	hMUC1(890-1158)	2.60E+05	7.55E-04	2.90E-09
	hMUC1(1036-1155)	7.17E+05	1.14E-03	1.59E-09
	cMUC1	8.22E+05	9.06E-03	1.10E-08

Gatipotuzumab, a glyco-engineered humanized monoclonal antibody that recognizes the tumor-associated epitope of MUC1, is in phase II clinical development at Glycotope for the treatment of recurrent epithelial ovarian, fallopian tube or primary peritoneal cancers. The heavy

chain variable region and light chain variable region of Gatipotuzumab are shown in SEQ ID NO: 59 and SEQ ID NO: 60 respectively.

In another experiment, the binding activity of anti-MUC1 antibody to tumor cells NUGC-4 was measured by flow cytometry. The secondary antibody used in the experiment was: Alexa Fluor® 647 anti-human IgG Fcγ (Jackson ImmunoResearch Laboratories, Inc., Cat #:109-606-170). Human IgG1 was used as a isotype control (ISO). MFI was determined using serially diluted sample antibodies (maximum concentration: 10 µg/mL, 3-fold dilutions). A fitting curve was obtained using concentration (µg/mL) as the X-axis and MFI as the Y-axis. Results are shown in **FIG. 8**, which indicated that 10B1 exhibited better binding activity than that of the positive control 1H7 analog.

Example 6. Stability analysis of anti-MUC1 antibodies

The biophysical properties and stability of the anti-MUC1 antibodies were evaluated.

Anti-MUC1 antibodies were prepared in a buffer solution containing 3 mg/ml histidine, 80 mg/ml sucrose, and 0.2 mg/ml Tween 80 at pH 6.0, and kept in sealed Eppendorf tubes at $40 \pm 2^\circ\text{C}$, $60\% \pm 5\%$ relative humidity (RH) (hereinafter referred to as 40°C) for 7 days, and their thermal stability were evaluated.

The antibodies were loaded into protein A columns and eluted with a buffer (0.1mol/L HAC) at pH 3.5. Half of the antibodies were added into a 2M Tris buffer so that the pH is adjusted to 7.5. The remaining half were kept at pH 3.5 for 6 hours and before adjusting the pH to 7.5. As such, half of the diluted antibodies were kept in sealed Eppendorf tubes at $\text{pH } 3.5 \pm 0.1$, $25 \pm 2^\circ\text{C}$ (hereinafter referred to as pH 3.5) for 6h to test stability at low pH.

The antibodies were placed in a buffer solution (3 mg/ml histidine, 80 mg/ml sucrose, and 0.2 mg/ml Tween 80), and a H_2O_2 stock solution was added to reach a final H_2O_2 concentration of 0.05% and 0.5% (hereinafter referred to as H_2O_2 0.05% and H_2O_2 0.5%). Samples were kept in sealed Eppendorf tubes at $40 \pm 3^\circ\text{C}$, $60\% \pm 5\%$ RH for 30 min to test the stability of the antibodies under strong oxidizing conditions.

The antibodies were prepared in a buffer solution (3 mg/ml histidine, 80 mg/ml sucrose, and 0.2 mg/ml Tween 80) and a NH_4HCO_3 stock solution was added to reach a final NH_4HCO_3 concentration of 0.94%. Samples were kept in sealed Eppendorf tubes at $40 \pm 2^\circ\text{C}$, $60\% \pm 5\%$ RH for 6h (hereinafter referred to as NH_4HCO_3 6h), 24h (hereinafter referred to as NH_4HCO_3 24h) to test the stability of the antibodies under strong deamidation conditions.

Specifically, the following tests were performed: (1) detecting the purity of the antibodies by Size-Exclusion High Performance Liquid Chromatography (SEC-HPLC) (indicated as the percentage of the main peak area to the sum of all peak areas (Purity, %)); (2) detecting the apparent hydrophobicity of the antibodies using the Hydrophobic Interaction Chromatography-High Performance Liquid Chromatography (HIC-HPLC) method (indicated as the retention time of the main peak (HIC, min)); (3) detecting the purity changes of the antibodies by capillary electrophoresis-sodium dodecyl sulphate (CE-SDS) under non-reducing (CE-SDS(NR)) conditions (indicated as the percentage of the main peak area to the sum of all peak areas (Purity, %)); (4) detecting the pI (isoelectric point) and charge variants of the antibodies by the

Capillary Isoelectric Focusing (cIEF) method (indicated as the percentages of the main component, acidic component, and alkaline component).

In the SEC-HPLC experiments, the antibody samples were diluted to 1 mg/mL with purified water and an Agilent 1290 chromatograph system (connected with XBridge Protein BEH SEC column (200 Å, Waters Corporation)) was used. The following parameters were used: mobile phase: 0.1M phosphate buffer (PB) +10% ACN, pH7.4; flow rate: 1.8 ml/min; column temperature: 25 °C; detection wavelength: 280 nm, 220nm; injection volume: 10 µL; sample tray temperature: about 8°C; and running time: 7 minutes.

In the HIC-HPLC experiments, an Agilent 1260 chromatograph system (connected with ProPac HIC-10 column (4.6 x 250 mm, Thermo Scientific)) was used, and samples were 10 times diluted using mobile phase A. The following parameters were used: mobile phase A: 0.9 M ammonium sulfate, 0.1M phosphate buffer (PB), 10% acetonitrile pH 6.5; mobile phase B: 0.1M phosphate buffer (PB), 10% acetonitrile pH 6.5; flow rate: 0.8 ml/min; gradient: 0 min 100% A, 2 min 100% A, 32 min 100% B, 34 min 100% B, 35 min 100% A, and 45 min 100% A; column temperature: 30 °C; detection wavelength: 280 nm, 220nm; injection volume: 10 µg; sample tray temperature: about 10 °C; and running time: 50 minutes.

In the CE-SDS(NR) experiments, Maurice CE-SDS Size Application Kit (Protein simple, Cat#: PS-MAK02-S) was used. 60 µL antibody sample (0.5mg/mL), 2.4 µL 25× internal standard, 3 µL 250 mM Iodoacetamide (SIGMA, Cat#: 16125) were added to a microcentrifuge tube, followed by centrifugation at 3000 rpm for 1 minute and heating in a 70 °C water bath for 10 min. The samples were then cooled to room temperature followed by centrifugation at 10000 rpm for 3 minutes. The supernatants were then transferred to a 96-well plate and tested in Maurice analyzer. The following parameters were used: injection voltage: 4.6 kV; injection time: 20 seconds; separation voltage 5.75 kV, and separation time: 45 minutes.

In the cIEF experiments, a Maurice cIEF Method Development Kit (Protein Simple, Cat #: PS-MDK01-C) was used for sample preparation. Specifically, 8 µL, 30ug protein sample was mixed with the following reagents in the kit: 1 µL Maurice cIEF pI Marker-7.05, 1 µL Maurice cIEF pI Marker-10.10, 35 µL 1% Methyl Cellulose Solution, 2 µL Maurice cIEF 500 mM Arginine, 1.33µL Ampholytes (Pharmalyte pH ranges 3-10), 6.66 µL Ampholytes (Pharmalyte pH ranges 8-10.5), and water (added to make a final volume of 100 µL). On the Maurice analyzer (Protein Simple, Santa Clara, CA), Maurice cIEF Cartridges (PS-MC02-C) were used to generate imaging capillary isoelectric focusing spectra. The sample was focused for a total of 10 minutes.

Detailed results are shown in the table below. The results showed that anti-MUC1 antibodies 9H4, 10B1, 10G12, 11A6, 11D6 and 11E8 have good stability and physical and chemical properties.

Table 6

Antibody	Treatment	SEC Purity (%)	HIC (min)	CE-SDS (NR)			cIEF		
				Purity (%)	HMW (%)	LMW (%)	Acidic peak (%)	Main peak (%)	Alkali peak (%)
9H4	0d	97.20	7.62	—	—	—	7.2	57.4	35.3
	40°C 7d	98.73	7.60	—	—	—	13.2	54.2	32.7

	pH3.5 0h	98.69	7.53	—	—	—	7.4	57.6	34.9
	pH3.5 6h	98.85	7.52	—	—	—	8.3	56.1	35.6
	H ₂ O ₂ 0.05%	97.92	3.66	—	—	—	8.6	58.0	33.4
	H ₂ O ₂ 0.5%	98.06	3.47	—	—	—	9.4	57.9	32.6
	NH ₄ HCO ₃ 6h	98.05	7.56	—	—	—	11.4	55.2	33.4
	NH ₄ HCO ₃ 24h	98.69	7.52	—	—	—	19.2	50.4	30.4
10B1	0d	91.97	7.84	97.63	0	2.37	12.5	48.1	39.4
	40°C 7d	93.40	7.82	97.42	0	2.58	16.2	52.1	31.7
	pH3.5 0h	93.21	7.84	95.88	0	4.12	10.9	52.3	36.8
	pH3.5 6h	93.30	7.87	95.46	0	4.54	10.8	52.8	36.5
	H ₂ O ₂ 0.05%	91.91	7.13	96.56	0	3.44	12.1	50.0	37.9
	H ₂ O ₂ 0.5%	91.85	6.72	93.84	0	6.16	11.5	47.2	41.4
	NH ₄ HCO ₃ 6h	92.61	7.85	94.79	0	5.21	12.1	56.4	31.5
	NH ₄ HCO ₃ 24h	92.79	7.84	92.89	0	7.11	26.2	43.8	29.9
10G12	0d	98.90	20.25	100	0	0	7.2	58.8	34.0
	40°C 7d	98.88	20.20	99.78	0	0.22	11.5	56.0	32.5
	pH3.5 0h	98.91	20.49	100	0	0	7.4	58.9	33.7
	pH3.5 6h	98.89	20.50	100	0	0	7.2	58.4	34.4
	H ₂ O ₂ 0.05%	98.83	19.83	97.94	0	2.06	6.0	60.8	33.2
	H ₂ O ₂ 0.5%	97.92	19.64	96.71	0	3.29	0	65.8	34.2
	NH ₄ HCO ₃ 6h	98.88	20.47	100	0	0	10.6	56.8	32.6
	NH ₄ HCO ₃ 24h	98.72	20.46	100	0	0	19.0	51.3	29.7
11A6	0d	98.30	9.75	98.99	0	1.01	4.9	65.3	29.8
	40°C 7d	98.31	9.74	99.21	0	0.79	9.8	62.0	28.2
	pH3.5 0h	98.39	9.70	98.55	0	1.46	5.6	64.0	30.3
	pH3.5 6h	98.54	9.69	98.49	0	1.51	4.1	66.6	29.4
	H ₂ O ₂ 0.05%	98.37	9.29	98.79	0	1.21	4.1	69.7	26.1
	H ₂ O ₂ 0.5%	97.57	9.83	96.06	0	3.94	7.7	63.8	28.6
	NH ₄ HCO ₃ 6h	98.68	9.70	99.09	0	0.91	7.5	63.6	28.9
	NH ₄ HCO ₃ 24h	98.63	9.67	98.82	0	1.18	17.7	56.0	26.4
11D6	0d	98.61	14.64	—	—	—	7.4	58.5	34.1
	40°C 7d	98.72	14.67	—	—	—	12.5	57.7	29.9
	pH3.5 0h	98.34	14.64	—	—	—	7.5	58.3	34.2
	pH3.5 6h	98.70	14.70	—	—	—	6.5	59.9	33.7
	H ₂ O ₂ 0.05%	98.16	14.34	—	—	—	6.7	60.0	33.3
	H ₂ O ₂ 0.5%	98.06	14.12	—	—	—	6.8	52.4	30.9
	NH ₄ HCO ₃ 6h	98.56	14.62	—	—	—	7.8	62.9	29.3
	NH ₄ HCO ₃ 24h	98.76	14.61	—	—	—	19.5	54.2	26.3
11E8	0d	98.23	22.50	99.23	0	0.77	10.3	56.3	33.5
	40°C 7d	98.32	22.50	99.06	0	0.94	15.2	53.7	31.2
	pH3.5 0h	95.30	22.49	99.02	0	0.98	10.8	56.1	33.1
	pH3.5 6h	95.40	22.49	99.13	0	0.87	9.6	57.3	32.7
	H ₂ O ₂ 0.05%	97.06	22.35	99.65	0	0.45	11.2	56.3	32.5
	H ₂ O ₂ 0.5%	96.20	20.95	98.35	0	1.65	9.3	59.0	31.7
	NH ₄ HCO ₃ 6h	96.48	22.49	99.43	0	0.57	13.9	54.3	31.9

	NH ₄ HCO ₃ 24h	96.31	22.49	98.44	0	1.56	18.2	52.5	29.4
--	--------------------------------------	-------	-------	-------	---	------	------	------	------

— No detected.

Example 7. Generating anti-MUC1 Antibody Drug Conjugates (ADC)

Each purified antibody (10B1, 10G12, 11A6 or 11E8) was coupled with MMAE (monomethyl auristatin E) or MMAF (monomethyl auristatin F) through a maleimidocaproyl-valine-citrulline-p-aminobenzyloxycarbonyl (VC) linker. For the names of antibody drug conjugates, “ADC” is added directly after the antibody name. For example, if 10B1 is coupled to MMAE, it is named as 10B1-ADC.

HIC-HPLC were used to detect the coupling of antibodies with drug molecules, and the results are shown in **Table 7**. The average drug-to-antibody ratio (DAR) is determined by multiplying PA% (PA% is the peak area percentage as measured by the area under the 280 nm peak) multiplied by the corresponding drug load of 0, 1, 2, 3, 4, 5, 6, 7 or 8 and divided by 100. For example, the average DAR of 10B1-ADC = [(9.603×0)+(38.758×2)+(40.342×4)+(11.296×6)]/100=3.07.

Table 7

ADC	PA% at each DAR Peak									Average DAR
	0	1	2	3	4	5	6	7	8	
10B1-ADC	9.603	—	38.758	—	40.342	—	11.296	—	—	3.07
10G12-ADC	6.069	—	36.656	—	41.589	—	15.686	—	—	3.34
11A6-ADC	4.209	—	26.631	—	41.034	—	19.664	—	8.461	4.03
11E8-ADC	4.836	—	40.347	—	40.257	—	14.56	—	—	3.29

In another experiment, the purified antibodies (10B1, 10G12, 11A6, 11E8, 1H7 analog and Gatipotuzumab analog) were also coupled with CPT-1, CPT-2, CPT-3, or CPT-4 through a CPT-L linker. For the names of antibody-drug conjugates, CPTx (x = 1, 2, 3, or 4) is added directly after the antibody name. For example, when 10B1 is coupled to CPT-1, it is named as 10B1-CPT1. As another example, when 1H7 analog is coupled to CPT-2, it is named as 1H7-CPT2. Exemplary ADCs obtained by this method included: 10B1-CPT2, 10G12-CPT2, 11A6-CPT2, 11E8-CPT2, 1H7-CPT2 and Gatipotuzumab-CPT2.

For comparative purposes, Gatipotuzumab analog were also coupled with Dxd through GGFG linker, and the resulting ADC was named Gatipotuzumab-Dxd.

An IgG1 monoclonal antibody targeting irrelevant target was coupled to CPT-2 through CPT-L linker. This formed isotype-CPT2 (ISO-CPT2), which were used as isotype controls.

MS (Mass Spectrometry) was used to detect the coupling of antibodies with drug molecules. The MS detection results showed that the DAR of the ADCs was about 8.

Example 8. Anti-tumor activity of anti-MUC1 ADC in stomach cancer model

B-NDG mice (Biocytogen Pharmaceuticals (Beijing) Co., Ltd., Cat#: B-CM-002) were used to determine the anti-tumor activity of the anti-MUC1 ADCs. About 5×10⁶ stomach cancer

NUGC-4 cells were injected subcutaneously in B-NDG mice, and when the tumor volume grew to about 200 mm³, the mice were divided to a control group and different treatment groups based on tumor volume (5 mice per group). The treatment groups were randomly selected for 1H7-ADC (1H7 analog coupled with MMAE) treatment or 10B1-ADC treatment. The control group mice were injected with phosphate buffer saline (PBS). The frequency of administration was once a week (2 administrations in total).

The tumor volumes were measured twice a week and the body weights of the mice were determined as well. Euthanasia was performed when the tumor volume of a mouse reached 3000 mm³. Details of the administration scheme are shown in the table below.

Table 8

Group	No. of mice	Antibodies	Dosage (mg/kg)	Route	Frequency	Total No. of administration
G1	5	PBS	-	i.v.	QW	2
G2	5	1H7-ADC	3 mg/kg	i.v.	QW	2
G3	5	10B1-ADC	3 mg/kg	i.v.	QW	2

The lengths of the long axis and the short axis of the tumor were measured and the volume of a tumor was calculated as $0.5 \times (\text{long axis}) \times (\text{short axis})^2$. The tumor growth inhibition percentage (TGI%) was calculated using the following formula: $\text{TGI} (\%) = [1 - (\text{Ti} - \text{T0}) / (\text{Vi} - \text{V0})] \times 100\%$. Ti is the average tumor volume in the treatment group on day i. T0 is the average tumor volume in the treatment group on day zero. Vi is the average tumor volume in the control group on day i. V0 is the average tumor volume in the control group on day zero.

T-test was performed for statistical analysis. $P < 0.05$ is a threshold to indicate a significant difference.

During the experimental period, no significant difference was observed in the body weight of mice among different groups, and all mice survived at the end of the experiment. This indicated that the tested ADCs were well tolerated and were not obviously toxic to the mice.

The tumor volumes in the treatment groups are shown in **FIG. 1**, which showed that 10B1-ADC exhibited significant tumor inhibiting efficiency with a higher TGI% than that of the control group.

Example 9. Anti-tumor activity of anti-MUC1 ADC in lung cancer model

The anti-MUC1 ADCs were tested for their antitumor effects in a human lung cancer PDX model. B-NDG mice were engrafted in the right flank with patient-derived lung cancer fragments (2 mm × 2 mm × 2 mm). When the tumors in the mice reached a volume of about 250-300 mm³, the mice were randomly placed into different groups based on tumor volume. The mice were then injected with PBS, Gatipotuzumab-ADC (Gatipotuzumab analog coupled with MMAE) or 10B1-ADC by i.v. administration. Details of the administration scheme are shown in the table below.

Table 9

Group	No. of mice	Antibodies	Dosage (mg/kg)	Route	Frequency	Total No. of administration
G1	5	PBS	-	i.v.	QW	2
G2	5	Gatipotuzumab-ADC	3 mg/kg	i.v.	QW	2
G3	5	10B1-ADC	3 mg/kg	i.v.	QW	2

The tumor volumes in the treatment groups are shown in **FIG. 2**. **Table 10** below summarizes the results for this experiment, including the tumor volumes on the day of grouping (day 0), 14 days after grouping (day 14) and at the end of the experiment (day 24); the survival rate of the mice; TGI%; and the statistical differences (P value) of tumor volumes between the treatment and control groups.

Table 10

Group	Tumor volume (mm ³)			Survival	TGI%	P value
	Day 0	Day 14	Day 24			
G1	256 ± 21	1701 ± 212	2300 ± 267	5/5	NA	NA
G2	257 ± 28	906 ± 231	2319 ± 518	5/5	-0.9%	0.979
G3	256 ± 31	278 ± 34	873 ± 151	5/5	69.8%	0.002

As shown in **FIG. 2** and **Table 10**, 10B1-ADC inhibited tumor growth with a higher TGI% than that of positive control Gatipotuzumab-ADC.

Example 10. Anti-tumor activity of anti-MUC1 ADC in breast cancer model

Similar to the previous experiments, about 1×10^7 breast cancer HCC70 cells were injected subcutaneously in B-NDG mice, and when the tumor volume grew to about 200 mm³, the mice were divided to a control group and different treatment groups based on tumor volume (5 mice per group). The treatment groups were randomly selected for 3 mg/kg 1H7-ADC treatment (G3), 3 mg/kg Gatipotuzumab-ADC treatment (G4) or 3 mg/kg 10B1-ADC treatment (G5). The control group mice were injected with PBS (G1). For ISO control (G2), human IgG1 coupling with MMAE was used. The frequency of administration was once a week (2 administrations in total).

During the experimental period, no significant difference was observed in the body weight of mice among different groups. The tumor volumes in the treatment groups are shown in the table below, which showed that 10B1-ADC (G5) exhibited significant tumor inhibiting efficiency with a higher TGI% than that of the positive control groups (G3 and G4).

Table 11

Group	Tumor volume (mm ³)				TGI% (Day 35)	P value (tumor)
	Day 0	Day 14	Day 25	Day 35		
G1	204 ± 11	562 ± 55	856 ± 77	1412 ± 144	NA	NA
G2	204 ± 13	378 ± 30	607 ± 54	1033 ± 90	31.4	0.056
G3	204 ± 13	148 ± 31	175 ± 84	336 ± 150	89.0	0.001
G4	204 ± 14	255 ± 50	353 ± 94	591 ± 105	68.0	0.002

G5	204±15	130±26	91±16	175±39	102.4	3.37E-05
----	--------	--------	-------	--------	-------	----------

In a similar experiment, about 1×10^7 breast cancer HCC70 cells were injected subcutaneously in B-NDG mice, and when the tumor volume grew to about 200 mm^3 , the mice were divided into different groups, and treated with 3 mg/kg 10B1-CPT2, 1H7-CPT2, or Gatipotuzumab-Dxd by i.v. administration. The tumor volumes were measured twice a week and body weights of the mice were recorded as well.

As shown in **FIG. 9**, the body weight of mice in different groups all increased, indicating that all the anti-MUC1 ADCs were well tolerated and not toxic to the mice. With regard to the tumor volume, 10B1-CPT2 inhibited tumor growth with a higher TGI% (e.g., on Day 38) than that of the positive control groups (**FIG. 10**), demonstrating that 10B1-CPT2 has good tumor inhibitory effect in breast cancer model.

In another experiment, B-NDG mice were engrafted in the right flank with patient-derived breast cancer fragments ($2\text{ mm}\times2\text{ mm}\times2\text{ mm}$). When the tumors in the mice reached a volume of about $250\text{-}300\text{ mm}^3$, the mice were randomly placed into different groups based on tumor volume, and were injected with PBS (G1), ISO-CPT2 (G2), 10B1-CPT2 (G3) or Gatipotuzumab-CPT2 (G4) by i.v. administration. The frequency of administration was once a week. The tumor volumes were measured twice a week, and the results are shown in **FIG. 11**, in which 10B1-CPT2 inhibited tumor growth with a higher TGI% (e.g., on Day 16) than that of the positive control Gatipotuzumab-CPT2.

These datas indicate that 10B1-CPT2 has potential in the treatment of breast cancer.

Example 11. Anti-tumor activity of anti-MUC1 ADC in pancreatic cancer model

The anti-MUC1 ADCs were tested for their antitumor effects in a human pancreatic cancer PDX model. B-NDG mice were engrafted in the right flank with patient-derived pancreatic cancer fragments ($2\text{ mm}\times2\text{ mm}\times2\text{ mm}$). When the tumors in the mice reached a volume of about $250\text{-}300\text{ mm}^3$, the mice were randomly placed into different groups based on tumor volume. The mice were then injected with PBS (G1), ISO (G2), 3 mg/kg 1H7-ADC (G3), 3 mg/kg Gatipotuzumab-ADC (G4) or 3 mg/kg 10B1-ADC (G5) by i.v. administration. For ISO control (G2), human IgG1 coupling with MMAE was used. The frequency of administration was once a week (2 administrations in total).

Similar results (as shown in **Table 12**) were obtained as compared to the breast cancer model, in which 10B1-ADC exhibited better anti-tumor effects than that of the positive control 1H7-ADC and Gatipotuzumab-ADC.

Table 12

Group	Tumor volume (mm ³)					TGI% (Day 52)	P value
	Day 0	Day 14	Day 28	Day 38	Day 52		
G1	255±17	716±70	1335±105	1863±95	2634±120	NA	NA
G2	255±17	655±81	1052±194	1487±191	2086±246	23.0	0.069
G3	255±31	395±41	399±99	717±153	1557±258	45.3	0.005
G4	255±18	474±75	636±131	742±157	1188±243	60.8	0.001

G5	255±21	354±37	244±34	403±70	891±157	73.3	2.16E-05
----	--------	--------	--------	--------	---------	------	----------

In a similar experiment, patient-derived pancreatic cancer model mice were injected with PBS (G1), ISO-CPT2 (G2), 10B1-CPT2 (G3), 1H7-CPT2 (G4) or Gatipotuzumab-CPT2 (G5) by i.v. administration. The frequency of administration was once a week (2 administrations in total).

5 As shown in **FIG. 12**, 10B1-CPT2 exhibited the best tumor inhibition effect with a TGI% of 75.0% among the three anti-MUC1 ADCs in pancreatic cancer model.

Example 12. Anti-Tumor Activity in colorectal cancer PDX model

10 The anti-tumor activity of 10B1-CPT2 was tested *in vivo* in colorectal cancer PDX models with HER3-high-expression/MUC1-low-expression. Specifically, patient-derived colorectal tumor fragments (2 mm×2 mm×2 mm) were engrafted in the right flank of B-NDG mice. When the tumors in the mice reached a volume of about 250-300 mm³, the mice were randomly placed into different groups based on tumor volumes (5 mice per group), and then injected with PBS, 10B1-CPT2, 1H7-CPT2 or Gatipotuzumab-CPT2 with a concentration of 3

15 mg/kg by i.v. administration. The tumor volumes were measured twice a week.

As shown in **FIG. 13**, 10B1-CPT2 exhibited better anti-tumor activity than that of the positive controls 1H7-CPT2 and Gatipotuzumab-CPT2.

In another similar experiment, colorectal cancer PDX model mice were injected with PBS, 10B1-CPT2 or 1H7-CPT2 with a concentration of 6 mg/kg by i.v. administration. The tumor volumes were measured twice a week. The results showed that 10B1-CPT2 inhibited tumor growth with a higher TGI% (e.g., on Day 25) than that of the positive controls 1H7-CPT2. The TGI% were 63.0% and 39.2% respectively.

Example 13. Anti-MUC1 antibodies with different mutations

25 In another experiment, the “DG” in VH CDR2 of the antibodies described herein was mutated to “AG” or “AA” to create CDR-modified VHs. For example, in one embodiment, the “DG” in VH CDR2 of 10B1 was mutated to “AG,” and the resulting CDR-modified VH is named 10B1-AG VH. The sequences of some of these CDR-modified VHs can be found in **FIG. 5**.

30 The CDR-modified VHs and VL of 10B1 are connected to IgG1 constant regions, and the binding activities of the mutated antibodies to ZR-75-1 cells were verified by flow cytometry. Specifically, ZR-75-1 cells were transferred to a 96-well plate at a density of 1×10⁵ cells/well respectively. Diluted sample antibody 10B1, 10B1-AG or 10B1-AA was added to the 96-well plate, and incubated at 4°C for 30 min. Then, the cells were incubated with the secondary antibody anti-hIgG-Fc-Alexa Fluor 647 (RL1-H) at 4°C in the dark for 15 minutes before flow cytometry analysis. Human IgG1 was used for the isotope control (ISO). The results are shown in **FIG. 6**, which showed that 10B1-AG and 10B1-AA exhibited binding activities that were comparable to that of 10B1.

40 The endocytosis rates of the mutated antibodies in ZR-75-1 cells were also tested. Diluted sample antibodies 10B1, 10B1-AG or 10B1-AA and pHAb-AffiniPure Fab Goat Anti-Human

IgG Secondary Antibody were added to the ZR-75-1 cells. After incubating for 6 hours, the cells were centrifuged and washed in FACS buffer. MFI was detected on a flow cytometer, and the endocytosis rates of the antibodies were calculated. Human IgG1 was used as ISO control. The results are shown in **FIG. 7**, which showed that 10B1-AG and 10B1-AA exhibited a endocytosis rates that were comparable to that of 10B1.

OTHER EMBODIMENTS

It is to be understood that while the invention has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the invention, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

WHAT IS CLAIMED IS:

1. An antibody or antigen-binding fragment thereof that binds to MUC1 (The transmembrane glycoprotein Mucin 1) comprising:
 - a heavy chain variable region (VH) comprising complementarity determining regions (CDRs) 1, 2, and 3, wherein the VH CDR1 region comprises an amino acid sequence that is at least 80% identical to a selected VH CDR1 amino acid sequence, the VH CDR2 region comprises an amino acid sequence that is at least 80% identical to a selected VH CDR2 amino acid sequence, and the VH CDR3 region comprises an amino acid sequence that is at least 80% identical to a selected VH CDR3 amino acid sequence; and
 - a light chain variable region (VL) comprising CDRs 1, 2, and 3, wherein the VL CDR1 region comprises an amino acid sequence that is at least 80% identical to a selected VL CDR1 amino acid sequence, the VL CDR2 region comprises an amino acid sequence that is at least 80% identical to a selected VL CDR2 amino acid sequence, and the VL CDR3 region comprises an amino acid sequence that is at least 80% identical to a selected VL CDR3 amino acid sequence,wherein the selected VH CDRs 1, 2, and 3 amino acid sequences and the selected VL CDRs, 1, 2, and 3 amino acid sequences are one of the following:
 - (1) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 4-6, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 1-3, respectively;
 - (2) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 7-9, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 1-3, respectively;
 - (3) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 10-12, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 1-3, respectively;
 - (4) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 13-15, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 1-3, respectively;
 - (5) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 16-18, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 1-3, respectively;
 - (6) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 19-21, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 1-3, respectively;
 - (7) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 25-27, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 22-24, respectively;
 - (8) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 28-30, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 22-24, respectively;

(9) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 31-33, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 22-24, respectively;

(10) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 34-36, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 22-24, respectively;

(11) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 37-39, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 22-24, respectively;

(12) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 40-42, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 22-24, respectively;

(13) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 7, 63, 9, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 1-3, respectively;

(14) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 7, 64, 9, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 1-3, respectively;

(15) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 28, 65, 30, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 22-24, respectively; and

(16) the selected VH CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 28, 66, 30, respectively, and the selected VL CDRs 1, 2, 3 amino acid sequences are set forth in SEQ ID NOs: 22-24, respectively.

2. The antibody or antigen-binding fragment thereof of claim 1, wherein the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 4-6, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, according to Kabat definition.

3. The antibody or antigen-binding fragment thereof of claim 1, wherein the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 7-9, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, according to Kabat definition.

4. The antibody or antigen-binding fragment thereof of claim 1, wherein the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 10-12, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, according to Kabat definition.

5. The antibody or antigen-binding fragment thereof of claim 1, wherein the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 13-15, respectively, and

the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, according to Kabat definition.

6. The antibody or antigen-binding fragment thereof of claim 1, wherein the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 16-18, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, according to Kabat definition.

7. The antibody or antigen-binding fragment thereof of claim 1, wherein the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 19-21, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, according to Kabat definition.

8. The antibody or antigen-binding fragment thereof of claim 1, wherein the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 25-27, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, according to Chothia definition.

9. The antibody or antigen-binding fragment thereof of claim 1, wherein the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 28-30, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, according to Chothia definition.

10. The antibody or antigen-binding fragment thereof of claim 1, wherein the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 31-33, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, according to Chothia definition.

11. The antibody or antigen-binding fragment thereof of claim 1, wherein the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 34-36, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, according to Chothia definition.

12. The antibody or antigen-binding fragment thereof of claim 1, wherein the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 37-39, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, according to Chothia definition.

13. The antibody or antigen-binding fragment thereof of claim 1, wherein the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 40-42, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, according to Chothia definition.

14. The antibody or antigen-binding fragment thereof of claim 1, wherein the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 7, 63, 9, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, according to Kabat definition.
15. The antibody or antigen-binding fragment thereof of claim 1, wherein the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 7, 64, 9, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, according to Kabat definition.
16. The antibody or antigen-binding fragment thereof of claim 1, wherein the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 28, 65, 30, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, according to Chothia definition.
17. The antibody or antigen-binding fragment thereof of claim 1, wherein the VH comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 28, 66, 30, respectively, and the VL comprises CDRs 1, 2, 3 with the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, according to Chothia definition.
18. The antibody or antigen-binding fragment thereof of any one of claims 1-17, wherein the antibody or antigen-binding fragment thereof specifically binds to human, mouse, monkey, or dog MUC1.
19. The antibody or antigen-binding fragment thereof of any one of claims 1-18, wherein the antibody or antigen-binding fragment thereof is a human antibody or antigen-binding fragment thereof, a single-chain variable fragment (scFv), a one-armed antibody, and/or a multi-specific antibody (e.g., a bispecific antibody).
20. The antibody or antigen-binding fragment thereof of any one of claims 1-19, wherein the antibody or antigen-binding fragment thereof is a human IgG1 antibody or antigen-binding fragment thereof or a human IgG4 antibody or antigen-binding fragment thereof.
21. A nucleic acid comprising a polynucleotide encoding a polypeptide comprising:
- (1) an immunoglobulin heavy chain or a fragment thereof comprising a heavy chain variable region (VH) comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 4-6, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;
 - (2) an immunoglobulin light chain or a fragment thereof comprising a VL comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, and wherein the VL, when paired

- with a VH comprising the amino acid sequence set forth in SEQ ID NO: 44 binds to MUC1;
- (3) an immunoglobulin heavy chain or a fragment thereof comprising a heavy chain variable region (VH) comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 7-9, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;
 - (4) an immunoglobulin light chain or a fragment thereof comprising a VL comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 45 binds to MUC1;
 - (5) an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 10-12, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;
 - (6) an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 46 binds to MUC1;
 - (7) an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 13-15, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;
 - (8) an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 47 binds to MUC1;
 - (9) an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 16-18, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;
 - (10) an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 48 binds to MUC1;
 - (11) an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 19-21, respectively, and wherein the VH,

- when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;
- (12) an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 49 binds to MUC1;
 - (13) an immunoglobulin heavy chain or a fragment thereof comprising a heavy chain variable region (VH) comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 25-27, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;
 - (14) an immunoglobulin light chain or a fragment thereof comprising a VL comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 44 binds to MUC1;
 - (15) an immunoglobulin heavy chain or a fragment thereof comprising a heavy chain variable region (VH) comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 28-30, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;
 - (16) an immunoglobulin light chain or a fragment thereof comprising a VL comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 45 binds to MUC1;
 - (17) an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 31-33, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;
 - (18) an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 46 binds to MUC1;
 - (19) an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 34-36, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;
 - (20) an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 22-24,

- respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 47 binds to MUC1;
- (21) an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 37-39, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;
- (22) an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 48 binds to MUC1;
- (23) an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 40-42, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;
- (24) an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 49 binds to MUC1;
- (25) an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 7, 63, 9, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;
- (26) an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 61 binds to MUC1;
- (27) an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 7, 64, 9, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;
- (28) an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 62 binds to MUC1;
- (29) an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 28, 65, 30, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1;

- (30) an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 61 binds to MUC1;
- (31) an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising complementarity determining regions (CDRs) 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 28, 66, 30, respectively, and wherein the VH, when paired with a light chain variable region (VL) comprising the amino acid sequence set forth in SEQ ID NO: 43 binds to MUC1; or
- (32) an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively, and wherein the VL, when paired with a VH comprising the amino acid sequence set forth in SEQ ID NO: 62 binds to MUC1.

22. The nucleic acid of claim 21, wherein the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 1-3, respectively.

23. The nucleic acid of claim 21, wherein the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin light chain or a fragment thereof comprising a VL comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 22-24, respectively.

24. The nucleic acid of claim 21, wherein the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 4-6, respectively.

25. The nucleic acid of claim 21, wherein the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 7-9, respectively.

26. The nucleic acid of claim 21, wherein the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 10-12, respectively.

27. The nucleic acid of claim 21, wherein the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 13-15, respectively.

28. The nucleic acid of claim 21, wherein the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 16-18, respectively.
29. The nucleic acid of claim 21, wherein the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 19-21, respectively.
30. The nucleic acid of claim 21, wherein the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 25-27, respectively.
31. The nucleic acid of claim 21, wherein the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 28-30, respectively.
32. The nucleic acid of claim 21, wherein the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 31-33, respectively.
33. The nucleic acid of claim 21, wherein the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 34-36, respectively.
34. The nucleic acid of claim 21, wherein the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 37-39, respectively.
35. The nucleic acid of claim 21, wherein the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 40-42, respectively.
36. The nucleic acid of claim 21, wherein the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH

comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 7, 63, 9, respectively.

37. The nucleic acid of claim 21, wherein the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 7, 64, 9, respectively.

38. The nucleic acid of claim 21, wherein the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 28, 65, 30, respectively.

39. The nucleic acid of claim 21, wherein the nucleic acid comprises a polynucleotide encoding a polypeptide comprising an immunoglobulin heavy chain or a fragment thereof comprising a VH comprising CDRs 1, 2, and 3 comprising the amino acid sequences set forth in SEQ ID NOs: 28, 66, 30, respectively.

40. The nucleic acid of any one of claims 21-39, wherein the VH when paired with a VL specifically binds to human, mouse, monkey, or dog MUC1, or the VL when paired with a VH specifically binds to human, mouse, monkey, or dog MUC1.

41. The nucleic acid of any one of claims 21-40, wherein the immunoglobulin heavy chain or the fragment thereof is a human immunoglobulin heavy chain or a fragment thereof (e.g., a human IgG1 heavy chain or a fragment thereof; or a human IgG4 heavy chain or a fragment thereof), and the immunoglobulin light chain or the fragment thereof is a human immunoglobulin light chain or a fragment thereof.

42. The nucleic acid of any one of claims 21-41, wherein the nucleic acid encodes a single-chain variable fragment (scFv), a one-armed antibody, a multi-specific antibody (e.g., a bispecific antibody), or a chimeric antigen receptor (CAR).

43. The nucleic acid of any one of claims 21-42, wherein the nucleic acid is cDNA.

44. A vector comprising one or more of the nucleic acids of any one of claims 21-43.

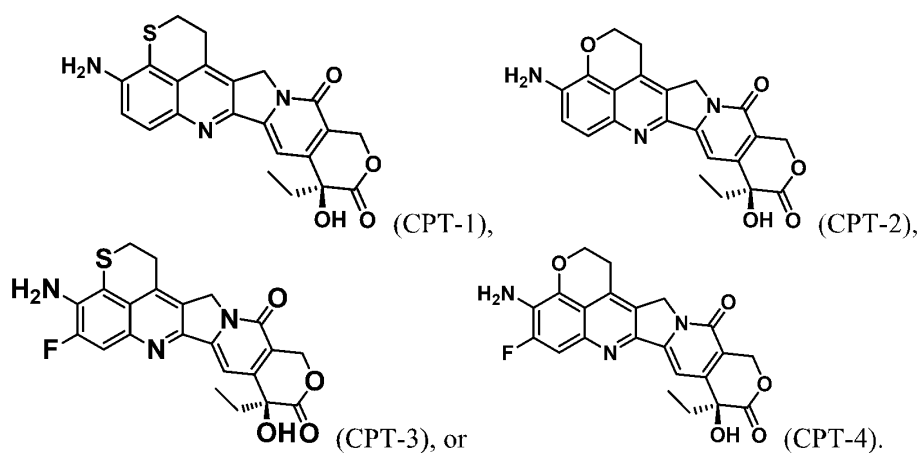
45. A vector comprising two of the nucleic acids of any one of claims 21-43, wherein the vector encodes the VL region and the VH region that together bind to MUC1.

46. A pair of vectors, wherein each vector comprises one of the nucleic acids of any one of claims 21-43, wherein together the pair of vectors encodes the VL region and the VH region that together bind to MUC1.

47. A cell comprising the vector of claim 44 or 45, or the pair of vectors of claim 46.
48. The cell of claim 47, wherein the cell is a CHO cell.
49. A cell comprising one or more of the nucleic acids of any one of claims 21-43.
50. A cell comprising two of the nucleic acids of any one of claims 21-43.
51. The cell of claim 50, wherein the two nucleic acids together encode the VL region and the VH region that together bind to MUC1.
52. A method of producing an antibody or an antigen-binding fragment thereof, the method comprising
- (a) culturing the cell of any one of claims 47-51 under conditions sufficient for the cell to produce the antibody or the antigen-binding fragment; and
 - (b) collecting the antibody or the antigen-binding fragment produced by the cell.
53. An antibody or antigen-binding fragment thereof that binds to MUC1 comprising a heavy chain variable region (VH) comprising an amino acid sequence that is at least 90% identical to a selected VH sequence, and a light chain variable region (VL) comprising an amino acid sequence that is at least 90% identical to a selected VL sequence, wherein the selected VH sequence and the selected VL sequence are one of the following:
- (1) the selected VH sequence is SEQ ID NO: 44, and the selected VL sequence is SEQ ID NO: 43;
 - (2) the selected VH sequence is SEQ ID NO: 45, and the selected VL sequence is SEQ ID NO: 43;
 - (3) the selected VH sequence is SEQ ID NO: 46, and the selected VL sequence is SEQ ID NO: 43;
 - (4) the selected VH sequence is SEQ ID NO: 47, and the selected VL sequence is SEQ ID NO: 43;
 - (5) the selected VH sequence is SEQ ID NO: 48, and the selected VL sequence is SEQ ID NO: 43;
 - (6) the selected VH sequence is SEQ ID NO: 49, and the selected VL sequence is SEQ ID NO: 43;
 - (7) the selected VH sequence is SEQ ID NO: 61, and the selected VL sequence is SEQ ID NO: 43; and
 - (8) the selected VH sequence is SEQ ID NO: 62, and the selected VL sequence is SEQ ID NO: 43.
54. The antibody or antigen-binding fragment thereof of claim 53, wherein the VH comprises the sequence of SEQ ID NO: 44 and the VL comprises the sequence of SEQ ID NO: 43.

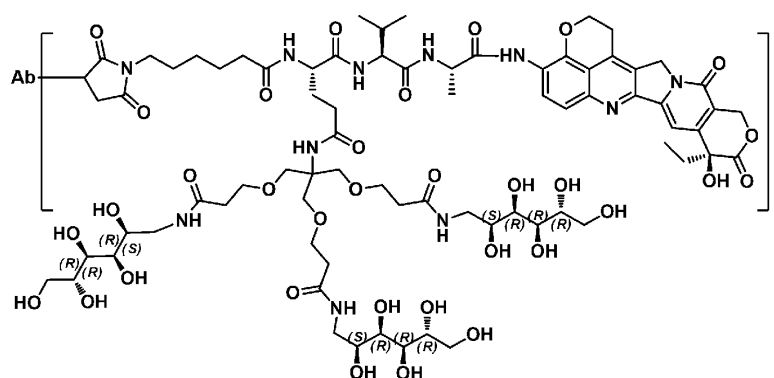
55. The antibody or antigen-binding fragment thereof of claim 53, wherein the VH comprises the sequence of SEQ ID NO: 45 and the VL comprises the sequence of SEQ ID NO: 43.
56. The antibody or antigen-binding fragment thereof of claim 53, wherein the VH comprises the sequence of SEQ ID NO: 46 and the VL comprises the sequence of SEQ ID NO: 43.
57. The antibody or antigen-binding fragment thereof of claim 53, wherein the VH comprises the sequence of SEQ ID NO: 47 and the VL comprises the sequence of SEQ ID NO: 43.
58. The antibody or antigen-binding fragment thereof of claim 53, wherein the VH comprises the sequence of SEQ ID NO: 48 and the VL comprises the sequence of SEQ ID NO: 43.
59. The antibody or antigen-binding fragment thereof of claim 53, wherein the VH comprises the sequence of SEQ ID NO: 49 and the VL comprises the sequence of SEQ ID NO: 43.
60. The antibody or antigen-binding fragment thereof of claim 53, wherein the VH comprises the sequence of SEQ ID NO: 61 and the VL comprises the sequence of SEQ ID NO: 43.
61. The antibody or antigen-binding fragment thereof of claim 53, wherein the VH comprises the sequence of SEQ ID NO: 62 and the VL comprises the sequence of SEQ ID NO: 43.
62. An antibody or antigen-binding fragment thereof that binds to MUC1 comprising
a heavy chain variable region (VH) comprising VH CDR1, VH CDR2, and VH CDR3 that
are identical to VH CDR1, VH CDR2, and VH CDR3 of a selected VH sequence; and a light
chain variable region (VL) comprising VL CDR1, VL CDR2, and VL CDR3 that are identical to
VL CDR1, VL CDR2, and VL CDR3 of a selected VL sequence, wherein the selected VH
sequence and the selected VL sequence are one of the following:
- (1) the selected VH sequence is SEQ ID NO: 44, and the selected VL sequence is SEQ ID NO: 43;
 - (2) the selected VH sequence is SEQ ID NO: 45, and the selected VL sequence is SEQ ID NO: 43;
 - (3) the selected VH sequence is SEQ ID NO: 46, and the selected VL sequence is SEQ ID NO: 43;
 - (4) the selected VH sequence is SEQ ID NO: 47, and the selected VL sequence is SEQ ID NO: 43;
 - (5) the selected VH sequence is SEQ ID NO: 48, and the selected VL sequence is SEQ ID NO: 43;
 - (6) the selected VH sequence is SEQ ID NO: 49, and the selected VL sequence is SEQ ID NO: 43;
 - (7) the selected VH sequence is SEQ ID NO: 61, and the selected VL sequence is SEQ ID NO: 43; and
 - (8) the selected VH sequence is SEQ ID NO: 62, and the selected VL sequence is SEQ ID NO: 43.

63. The antibody or antigen-binding fragment thereof of any one of claims 53-62, wherein the antibody or antigen-binding fragment thereof specifically binds to human, mouse, monkey, or dog MUC1.
64. The antibody or antigen-binding fragment thereof of any one of claims 53-63, wherein the antibody or antigen-binding fragment thereof is a human antibody or antigen-binding fragment thereof, a single-chain variable fragment (scFv), a one-armed antibody, and/or a multi-specific antibody (e.g., a bispecific antibody).
65. The antibody or antigen-binding fragment thereof of any one of claims 53-64, wherein the antibody or antigen-binding fragment is a human IgG1 antibody or antigen-binding fragment thereof or a human IgG4 antibody or antigen-binding fragment thereof.
66. An antibody or antigen-binding fragment thereof that cross-competes with the antibody or antigen-binding fragment thereof of any one of claims 1-20 and 53-65.
67. The antibody or antigen-binding fragment thereof of any one of claims 1-20 and 53-66, wherein the antibody or antigen-binding fragment thereof comprises a fragment crystallizable region (Fc region).
68. The antibody or antigen-binding fragment thereof of claim 67, wherein the Fc region has increased complement-dependent cytotoxicity (CDC) or antibody-dependent cellular cytotoxicity (ADCC).
69. A chimeric antigen receptor (CAR) comprising the antibody or antigen-binding fragment thereof of any one of claims 1-20 and 53-66.
70. An antibody-drug conjugate comprising the antibody or antigen-binding fragment thereof of any one of claims 1-20 and 53-68 covalently bound to a therapeutic agent.
71. The antibody-drug conjugate of claim 70, wherein the therapeutic agent is a cytotoxic or cytostatic agent.
72. The antibody-drug conjugate of claim 70 or 71, wherein the therapeutic agent is MMAE or MMAF.
73. The antibody-drug conjugate of claim 70 or 71, wherein the therapeutic agent is selected from

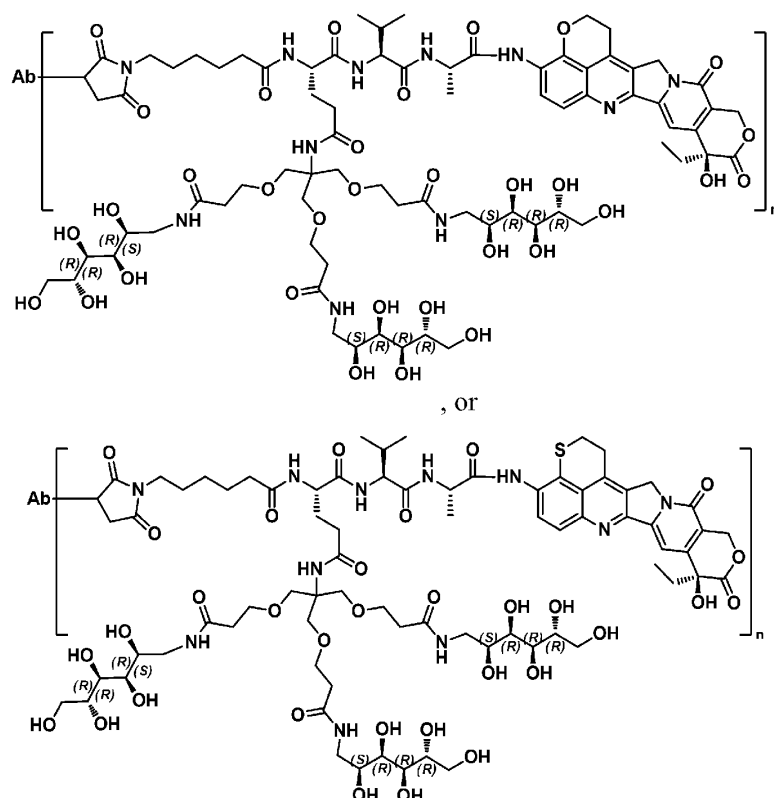


74. The antibody-drug conjugate of any one of claims 70-71 or 73, wherein the therapeutic agent is linked to the antibody or antigen-binding fragment thereof via a linker.

75. The antibody-drug conjugate of claim 74, wherein the linker has a structure of:



76. The antibody-drug conjugate of any one of claims 70-71 and 73-75, wherein the antibody-drug conjugate has a structure of:



wherein $n = 1-8$; wherein “Ab” represents the antibody or antigen-binding fragment thereof.

77. A method of treating a subject having cancer, the method comprising administering a therapeutically effective amount of a composition comprising the antibody or antigen-binding fragment thereof of any one of claims 1-20 and 53-68, the CAR of claim 69, or the antibody-drug conjugate of any one of claims 70-76, to the subject.

78. The method of claim 77, wherein the subject has a solid tumor.

79. The method of claim 77, wherein the cancer is oesophageal cancer, gastric cancer, breast cancer, ovarian cancer, bladder cancer, pancreatic cancer, lung cancer, colorectal cancer, stomach cancer, or Non-Hodgkin Lymphoma.

80. The method of claim 77, wherein the subject is further treated with an effective amount of an anti-4-1BB antibody, an anti-OX40 antibody, an anti-PD-1 antibody, an anti-CTLA4 antibody, an anti-CD40 antibody, or an anti-PD-L1 antibody.

81. A method of decreasing the rate of tumor growth, the method comprising

contacting a tumor cell with an effective amount of a composition comprising an antibody or antigen-binding fragment thereof of any one of claims 1-20 and 53-68, the CAR of claim 69, or the antibody-drug conjugate of any one of claims 70-76.

82. A method of killing a tumor cell, the method comprising contacting a tumor cell with an effective amount of a composition comprising the antibody or antigen-binding fragment thereof of any one of claims 1-20 and 53-68, the CAR of claim 69, or the antibody-drug conjugate of any one of claims 70-76.
83. A method of increasing immune response in a subject, the method comprising administering to the subject an effective amount of a composition comprising the antibody or antigen-binding fragment thereof of any one of claims 1-20 and 53-68, the CAR of claim 69, or the antibody-drug conjugate of any one of claims 70-76.
84. A pharmaceutical composition comprising the antibody or antigen-binding fragment thereof of any one of claims 1-20 and 53-68, and a pharmaceutically acceptable carrier.
85. A pharmaceutical composition comprising the antibody drug conjugate of any one of claims 70-76, and a pharmaceutically acceptable carrier.