

MULTISPECIFIC ANTIGEN BINDING PROTEIN AGAINST EpCAM**TECHNICAL FIELD**

5 The present disclosure relates broadly to antigen binding proteins, variants and fragments thereof specific to Epithelial Cellular Adhesion Molecule (EpCAM). In particular, the present disclosure encompasses the nucleotide and amino-acid sequences of 6 human antibodies, 5 llama antibodies and 1 humanized llama antibody, or the antigen-binding portions thereof, that specifically target human EpCAM, both in
10 solution and on the surface of cells. Also provided are multispecific antigen binding proteins, such as bispecific T cell engagers (BiTEs), which bind to EpCAM and also other targets.

BACKGROUND

15 In 2020, there was an estimated 19.3 million new cancers reported worldwide, and the cancer incidence is expected to rise to 28.4 million cases in 2040. Among all adult human cancer cases, approximately 90% are solid tumors. Solid tumors such as lung cancer, liver cancer, stomach cancer and female breast cancer are among the leading causes of cancer-related death. The global high cancer incidence burden,
20 coupled with poor treatment outcomes and survival rates with standard treatment modalities of surgery, chemotherapy and radiotherapy presents a strong need for the development of more effective treatment options for cancer.

 Advanced therapeutics such as antibody-based or cell-based immunotherapies, offers a new paradigm in oncology, and hold immense potential for more effective
25 treatments for cancers compared to conventional treatment strategies. As a result, antibody and cell-based therapies are increasingly gaining immense attention and investments for its development as a potential more promising standard of care for the treatment of cancer. However, even equipped with powerful cytotoxic killing mechanisms such as bi-specific T-cell engager (BiTE) and antibody-drug conjugates (ADC), antibody-
30 based targeted therapies rarely cure patients with solid tumors. Cell-based therapy such as chimeric antigen receptor (CAR) T achieves great success in haematological malignancy but shows limited efficacy in treating solid tumors. Among the complex difficulties each approach has encountered, one intrinsic problem is with regards to the

heterogeneous distribution of tumor markers in the cell populations of solid tumors and antigen escape upon treatment. To overcome these challenges, there is urgent medical need to explore next-generation immunotherapy to prevent cancer relapse.

Overexpression of cancer antigens such as GPC3, Her2, and Claudin18.2 has been found in various cancer types and known as promising targets for antibody or CAR-T cell-based immunotherapies. Nevertheless, single targeting to a known tumor antigen has been clinically proven to have limited efficacy. Epithelial cell adhesion molecule (EpCAM) is a transmembrane glycoprotein that is widely expressed in almost all epithelia and epithelial-derived neoplasms, and represents a different category of biomarkers. Additional targeting to EpCAM would be an attractive treatment regime to overcome the challenge of tumor heterogeneity.

However, as EpCAM is also widely expressed at low levels on normal epithelia, anti-EpCAM T-cell engagers have been demonstrated to be highly toxic to normal tissues. As a result, high affinity anti-EpCAM BiTE antibodies such as Solitomab and Catumaxomab showed dose-limiting toxicities and were unable to achieve FDA approval.

Hence, developing low-affinity anti-EpCAM clones used for EpCAM-targeting T cell engagers, and which may spare normal epithelial cells with low EpCAM levels, is in urgent need.

20

SUMMARY

Traditional methods in antibody discovery often start with multiple immunization steps in animals, such as mouse and rabbit, with antigen proteins. This normally results in identification of high-affinity binders of non-human origin. After extensive characterization of their properties, an extra step of humanization is usually required to avoid immunogenicity for further development. However, humanization process always involves complicated antibody engineering and the empirically derived humanized clones may greatly lose the binding affinity or the stability of the original clones. Using the present inventors' in-house naïve human Fab phage display library, they were able to identify low-affinity EpCAM binders in a simple and easy way.

30

The presently disclosed fully human anti-EpCAM antibodies (1B5, 1C1, 1C11, 1D4, 1E4 and 1H6), despite their relatively low binding affinity, exhibited potent cytotoxic activity when used in a BiTE format to mediate target specific cell killing. Applications

using the BiTE format can therefore overcome issues of immunogenicity and high toxicity/low tolerability brought by high-affinity non-human anti-EpCAM antibodies as shown by most prior arts.

On the other hand, compared to the conventional scFv formats of BiTEs, which often suffer from poor stability and low expression yield due to difficulty in proper folding, the nanobody format of BiTE have superior advantages:

- i) high stability;
- ii) large production yield; and
- iii) small gene size to be packaged in the lentivirus system to achieve high T cell transduction efficiency

Hence, the present inventors have also generated anti-EpCAM nanobodies of llama-derived single-domain antibody fragments (VHH) against EpCAM, which can be fused to many different anti-CD3 T cell agonist antibodies via a flexible linker. These BiTE antibodies can be used for targeting EpCAM locally. As an example, by using CAR-T cells secreting anti-EpCAM BiTE, the present inventors could achieve not only the specific tumor antigen directed T cell expansion but also tumor restricted expression of BiTE targeting a more widely expressed tumor antigen EpCAM to avoid toxicity brought by the traditional systemic delivery.

In one aspect, there is provided an antigen binding protein, an antigen binding variant, or an antigen binding fragment thereof that binds specifically to Epithelial Cellular Adhesion Molecule (EpCAM), wherein the antigen binding protein, variant or fragment thereof comprises a heavy chain variable region and/or a light chain variable region selected from the group consisting of:

- (i) a heavy chain variable region comprising:
 - a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
 - a CDR-H2 comprising IPIFGTA (**SEQ ID NO: 2**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

- a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto
- (ii) a heavy chain variable region comprising:
- 5 a CDR-H1 comprising GDSISSNSVA (**SEQ ID NO: 5**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- a CDR-H2 comprising TYYRSKWYS (**SEQ ID NO: 6**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
- 10 a CDR-H3 comprising AREVEGSSYDAFDI (**SEQ ID NO: 7**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto
- (iii) a light chain variable region comprising:
- 15 a CDR-L1 comprising:
- QSLHNSNGYNY (**SEQ ID NO: 9**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
 - QSLHNSNRYNY (**SEQ ID NO: 17**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
 - QSISDF (**SEQ ID NO: 19**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto
- a CDR-L2 comprising:
- 25 • LGS (**SEQ ID NO: 10**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
- AAS (**SEQ ID NO: 20**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
- a CDR-L3 comprising:
- 30 • MQALQTPYT (**SEQ ID NO: 11**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

- 5
- MQGLQSPWT (**SEQ ID NO: 15**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
 - QQSYIMPDT (**SEQ ID NO: 21**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
 - MQGLQTPYT (**SEQ ID NO: 23**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and
- 10
- (iv) a heavy chain variable region comprising:
- a CDR-H1 comprising:
- 15
- GSIFSGND (**SEQ ID NO: 25**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
 - GSSERFTS (**SEQ ID NO: 29**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto
- a CDR-H2 comprising:
- 20
- ITSGGST (**SEQ ID NO: 26**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
 - ITNGGST (**SEQ ID NO: 30**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto and
- a CDR-H3 comprising:
- 25
- TNGRWSGDTYYAHH (**SEQ ID NO: 27**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
 - MAGTS (**SEQ ID NO: 31**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- 30
- TNGRWSGDTYYAHL (**SEQ ID NO: 33**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In one embodiment, the antigen binding protein, variant or fragment thereof of comprises a heavy chain variable region selected from the group consisting of:

(i) a heavy chain variable region comprising:

5 a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-H2 comprising IIPFGTA (**SEQ ID NO: 2**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

10 a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto

(ii) a heavy chain variable region comprising:

15 a CDR-H1 comprising GDSISSNSVA (**SEQ ID NO: 5**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-H2 comprising TYRSKWYS (**SEQ ID NO: 6**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

20 a CDR-H3 comprising AREVEGSSYDAFDI (**SEQ ID NO: 7**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and

comprises a light chain variable region comprising:

a CDR-L1 comprising:

25 • QSLHNSNGYNY (**SEQ ID NO: 9**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

• QSLHNSRNYNY (**SEQ ID NO: 17**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or

30 • QSISDF (**SEQ ID NO: 19**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto

a CDR-L2 comprising:

- LGS (**SEQ ID NO: 10**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
 - AAS (**SEQ ID NO: 20**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
- 5 a CDR-L3 comprising:
- MQALQTPYT (**SEQ ID NO: 11**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
 - MQGLQSPWT (**SEQ ID NO: 15**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
 - QQSYIMPDT (**SEQ ID NO: 21**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
 - MQGLQTPYT (**SEQ ID NO: 23**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In one embodiment, the antigen binding protein, variant or binding fragment thereof comprises a heavy chain variable region comprising:

- 20 a CDR-H1 comprising:
- GSIFSGND (**SEQ ID NO: 25**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
 - GSSERFTS (**SEQ ID NO: 29**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto
- 25 a CDR-H2 comprising:
- ITSGGST (**SEQ ID NO: 26**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
 - ITNGGST (**SEQ ID NO: 30**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto;
- 30 and
- a CDR-H3 comprising:

- TNGRWSGDTYYAHH (**SEQ ID NO: 27**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- 5 • MAGTS (**SEQ ID NO: 31**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
- TNGRWSGDTYYAHL (**SEQ ID NO: 33**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

10 In one embodiment, the antigen binding protein, variant or fragment thereof comprises a heavy chain variable region and/or a light chain variable region selected from the group consisting of:

(i) a heavy chain variable region comprising:

15 a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-H2 comprising IPIFGTA (**SEQ ID NO: 2**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

20 a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto

(ii) a heavy chain variable region comprising:

25 a CDR-H1 comprising GDSISSNSVA (**SEQ ID NO: 5**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-H2 comprising TYYSKWYS (**SEQ ID NO: 6**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

30 a CDR-H3 comprising AREVEGSSYDAFDI (**SEQ ID NO: 7**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto

(iii) a light chain variable region comprising:

- a CDR-L1 comprising QSLHNSNGYNY (**SEQ ID NO: 9**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- 5 a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
- a CDR-L3 comprising MQALQTPYT (**SEQ ID NO: 11**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto
- 10 (iv) a light chain variable region comprising:
- a CDR-L1 comprising QSLHNSNGYNY (**SEQ ID NO: 9**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- 15 a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
- a CDR-L3 comprising MQGLQSPWT (**SEQ ID NO: 15**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto
- 20 (v) a light chain variable region comprising:
- a CDR-L1 comprising QSLHNSNRYNY (**SEQ ID NO: 17**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- 25 a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
- a CDR-L3 comprising MQALQTPYT (**SEQ ID NO: 11**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto
- 30 (vi) a light chain variable region comprising:
- a CDR-L1 comprising QSIDF (**SEQ ID NO: 19**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

- a CDR-L2 comprising AAS (**SEQ ID NO: 20**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
- 5 a CDR-L3 comprising QQSYIMPDT (**SEQ ID NO: 21**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto
- (vii) a light chain variable region comprising:
- a CDR-L1 comprising QSLHNSNGYNY (**SEQ ID NO: 9**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- 10 a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
- a CDR-L3 comprising MQGLQTPYT (**SEQ ID NO: 23**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto
- 15 (viii) a heavy chain variable region comprising:
- a CDR-H1 comprising GSIFSGND (**SEQ ID NO: 25**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- 20 a CDR-H2 comprising ITSGGST (**SEQ ID NO: 26**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
- a CDR-H3 comprising TNGRWSGDTYYAHH (**SEQ ID NO: 27**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto
- 25 (ix) a heavy chain variable region comprising:
- a CDR-H1 comprising GSSEFRTS (**SEQ ID NO: 29**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- 30 a CDR-H2 comprising ITNGGST (**SEQ ID NO: 30**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-H3 comprising MAGTS (**SEQ ID NO: 31**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and

(x) a heavy chain variable region comprising:

5 a CDR-H1 comprising GSIFSGND (**SEQ ID NO: 25**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-H2 comprising ITSGGST (**SEQ ID NO: 26**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

10 a CDR-H3 comprising TNGRWSGDTYYAHL (**SEQ ID NO: 33**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In one embodiment, the antigen binding protein, variant or fragment thereof
15 comprises a heavy chain variable region and a light chain variable region selected from the group consisting of:

(i) a heavy chain variable region comprising:

a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

20 a CDR-H2 comprising IIPFGTA (**SEQ ID NO: 2**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**) or a sequence
25 having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and

a light chain variable region comprising:

a CDR-L1 comprising QSLHSNGYNY (**SEQ ID NO: 9**) or a sequence
30 having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-L3 comprising MQALQTPYT (**SEQ ID NO: 11**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto

(ii) a heavy chain variable region comprising:

5 a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-H2 comprising IIPFGTA (**SEQ ID NO: 2**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

10 a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and

a light chain variable region comprising:

15 a CDR-L1 comprising QSLHNSNGYNY (**SEQ ID NO: 9**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

20 a CDR-L3 comprising MQGLQSPWT (**SEQ ID NO: 15**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto

(iii) a heavy chain variable region comprising:

25 a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-H2 comprising IIPFGTA (**SEQ ID NO: 2**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

30 a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and

a light chain variable region comprising:

a CDR-L1 comprising QSLHNSRNYNY (**SEQ ID NO: 17**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

5 a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

10 a CDR-L3 comprising MQALQTPYT (**SEQ ID NO: 11**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto

(iv) a heavy chain variable region comprising:

a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

15 a CDR-H2 comprising IIPFGTA (**SEQ ID NO: 2**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

20 a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and

a light chain variable region comprising:

a CDR-L1 comprising QSLHNSGYNY (**SEQ ID NO: 9**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

25 a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

30 a CDR-L3 comprising MQGLQTPYT (**SEQ ID NO: 23**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and

(v) a heavy chain variable region comprising:

- a CDR-H1 comprising GDSISSNSVA (**SEQ ID NO: 5**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- 5 a CDR-H2 comprising TYYSRWYS (**SEQ ID NO: 6**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
- a CDR-H3 comprising AREVEGSSYDAFDI (**SEQ ID NO: 7**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and
- 10 a light chain variable region comprising:
- a CDR-L1 comprising QSSSDF (**SEQ ID NO: 19**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- 15 a CDR-L2 comprising AAS (**SEQ ID NO: 20**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
- a CDR-L3 comprising QSSYMPDT (**SEQ ID NO: 21**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.
- 20 In one embodiment, the antigen binding protein, variant or fragment thereof comprises a heavy chain variable region selected from the group consisting of:
- (i) a heavy chain variable region comprising:
- a CDR-H1 comprising GSIFSGND (**SEQ ID NO: 25**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- 25 a CDR-H2 comprising ITSGGST (**SEQ ID NO: 26**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
- a CDR-H3 comprising TNGRWSGDTYYAHH (**SEQ ID NO: 27**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto
- 30 (ii) a heavy chain variable region comprising:

- a CDR-H1 comprising GSSERFTS (**SEQ ID NO: 29**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- a CDR-H2 comprising ITNGGST (**SEQ ID NO: 30**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
- a CDR-H3 comprising MAGTS (**SEQ ID NO: 31**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and
- (iii) a heavy chain variable region comprising:
- a CDR-H1 comprising GSIFSGND (**SEQ ID NO: 25**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- a CDR-H2 comprising ITSGGST (**SEQ ID NO: 26**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
- a CDR-H3 comprising TNGRWSGDTYYAHL (**SEQ ID NO: 33**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.
- In one embodiment, the antigen binding protein, variant or fragment thereof comprises a heavy chain variable domain and/or a light chain variable domain selected from the group consisting of:
- (i) a heavy chain variable domain comprising
 EVQLVQSGAEVKKPGSSVKVSCKASGGTFSSYAISWVRQAPGQGLE
 WMGGIPIFGTANYAQNFQGRVTMTADTSISTAYMELSSLRSEDTAVY
 YCARSLGGRFRYWGGTGL (**SEQ ID NO: 4** - 1B6, 1C1, 1C11, 1D4 and 1H6) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions
- (ii) a heavy chain variable domain comprising
 QVQLQQSGPGLVKPSQTLSTCAISGDSISNSVAWNWIRQSPSRGL
 EWLGRITYYRSKWYSDYAIKGRDLINPDTSKNQFSLQLNSVTPEDT
 AVYYCAREVEGSSYDAFDIWGGTGM (**SEQ ID NO: 8**) or a fragment,

variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions,

- (iii) a light chain variable domain comprising
- 5 DVVMTQSPLSLPVTGPGEASISCRSS**QSLHNSG**NYLDWYLQKPGQ
SPQLLIY**LGSN**RASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYY**CM**
QALQTPYTFGQGTK (**SEQ ID NO: 12**) or a fragment, variant or
sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%,
or 100% identical thereto and/or having two or three amino acids
10 substitutions
- (iv) a light chain variable domain comprising
- EIVLTQSPLSLPVTGPGEASISCRSS**QSLHNSG**NYLDWYLQKPGQS
PQLLIY**LGSN**RASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYY**CMQ**
GLQSPWTFGQGTK (**SEQ ID NO: 16**) or a fragment, variant or
15 sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%,
or 100% identical thereto and/or having two or three amino acids
substitutions
- (v) a light chain variable domain comprising
- DVVMTQSPLSLPVTGPGEASISCRSS**QSLHNSR**NYLDWYLQKPGQ
20 SPQLLIY**LGSN**RASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYY**CM**
QALQTPYTFGQGTK (**SEQ ID NO: 18**) or a fragment, variant or
sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%,
or 100% identical thereto and/or having two or three amino acids
substitutions
- (vi) a light chain variable domain comprising
- DIQLTQSPSSLSASVGDRVTITCRAS**QSISD**FLNWWYQQKPGKAPKLLIY
25 **AASSLQ**TGVPSRFGGSGSGTEFTLTISLQPEDLGTY**YCQQS**YIMPDT
FGQGTK (**SEQ ID NO: 22**) or a fragment, variant or sequence thereof
at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical
thereto and/or having two or three amino acids substitutions
- (vii) a light chain variable domain comprising
- DVVMTQSPLSLPVTGPGEASISCRSS**QSLHNSG**NYLDWYLQKPGQ
30 SPQLLIY**LGSN**RASGVPDRFSGSGSGTDFTLQISRVEAEDAGVYY**CM**

- QGLQTPYTFGQGTK (SEQ ID NO: 24)** or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions
- 5 (viii) a heavy chain variable domain comprising
QVQLQESGGGLVQPGGSLRLSCADSGSIFSGNDMAWYRRAPGVERE
LVAVITSGGSTHYADSVKGRFTISRDNAAKTVYLQTNDLPEDTAVYY
CTNGRWSGDTYYAHHWGQGTQ (**SEQ ID NO: 28**) or a fragment,
variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%,
10 98%, 99%, or 100% identical thereto and/or having two or three amino
acids substitutions
- (ix) a heavy chain variable domain comprising
QVQLQESGGGLVQPGGSLRLSCAASGSSEFTSVAWYRQAPGKERE
LVAFITNGGSTRYTDVPVKGRFTISRDNAAKNTVYLMNSLKAEDTAVYY
15 CMAGTSWGQGTQ (**SEQ ID NO: 32**) or a fragment, variant or
sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%,
or 100% identical thereto and/or having two or three amino acids
substitutions
- (x) a heavy chain variable domain comprising
20 QVQLQESGGGLVQPGGSLRLSCADSGSIFSGNDMAWYRRAPGVERE
LVAVITSGGSTHYADSVKGRFTISRDNAAKTVYLQTNDLPEDTAVYY
CTNGRWSGDTYYAHLWGQGTQ (**SEQ ID NO: 34**) or a fragment,
variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%,
25 98%, 99%, or 100% identical thereto and/or having two or three amino
acids substitutions
- (xi) a heavy chain variable domain comprising
QVQLQESGGGLVQAGGSLRLSCADSGSIFSGNDMAWYRRAPGVERE
LVAVITSGGSTHYADSVKGRFTISRDNAAKTVYLQTNDLPEDTAVYY
CTNGRWSGDTYYAHHWGQGTQ (**SEQ ID NO: 35**) or a fragment,
30 variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%,
98%, 99%, or 100% identical thereto and/or having two or three amino
acids substitutions

- (xii) a heavy chain variable domain comprising
 QVQLQESGGGLVQAGDSLRLSCADSGSIFSGNDMAWYRRAPGVERE
 LVAVITSGGSTHYADSVKGRFTISRDNQKTVYLQTNLKPEDTAVYY
 CTNGRWSGDTYYAHHWGQGTQ (**SEQ ID NO: 36**) or a fragment,
 5 variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%,
 98%, 99%, or 100% identical thereto and/or having two or three amino
 acids substitutions; and
 a heavy chain variable domain comprising
 QVQLVESGGGLVQAGGSLRLSCAASGSIFSGNDMSWYRQAPGKGLE
 10 LVAVITSGGSTYYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYY
 CTNGRWSGDTYYAHHWGQGTQ (**SEQ ID NO: 37**) or a fragment,
 variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%,
 98%, 99%, or 100% identical thereto and/or having two or three amino
 acids substitutions.
- 15 In one embodiment, the antigen binding protein, variant or fragment thereof
 comprises a heavy chain variable domain and a light chain variable domain selected
 from the group consisting of:
- (i) a heavy chain variable domain comprising:
 EVQLVQSGAEVKKPGSSVKVCKASGGTFSSYAISWVRQAPGQGLEWM
 20 GGIIPIFGTANYAQNFQGRVTMTADTSISTAYMELSSLRSEDVAVYYCARS
 LGGRFRYWGGQTL (**SEQ ID NO: 4**) or a fragment, variant or sequence
 thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%
 identical thereto and/or having two or three amino acids substitutions, and
 a light chain variable domain comprising
 25 DVVMTQSPLSLPVTPGEPASISCRSSQSLHNSGYNLYLDWYLQKPGQSP
 QLLIYLGSNRASGVPRFSGSGSGTDFTLKISRVEAEDVGVYYCMQALQT
 PYTFGQGTK (**SEQ ID NO: 12**) or a fragment, variant or sequence thereof
 at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical
 thereto and/or having two or three amino acids substitutions
- 30 (ii) a heavy chain variable domain comprising:
 EVQLVQSGAEVKKPGSSVKVCKASGGTFSSYAISWVRQAPGQGLEWM
 GGIIPIFGTANYAQNFQGRVTMTADTSISTAYMELSSLRSEDVAVYYCARS
 LGGRFRYWGGQTL (**SEQ ID NO: 4**) or a fragment, variant or sequence

thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions, and a light chain variable domain comprising
 EIVLTQSPLSLPVTGPGEASISCRSS**QSLH**SNGYNYLDWYLQKPGQSPQ
 5 LLIY**LGSN**RASGVDPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**MQGLQSP**
WTFGQGTK (SEQ ID NO: 16) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions

(iii) a heavy chain variable domain comprising:

10 EVQLVQSGAEVKKPGSSVKV**SCKASGGTFSSYA**ISWVRQAPGQGLEWM
 GGI**PIFGT**ANYAQN**FQGRVTMTADTSISTAYMELSSLR**SEDTAVYYC**CARS**
LGGRFRYWGQGTL (SEQ ID NO: 4) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions, and

15 a light chain variable domain comprising
 DVVMTQSPLSLPVTGPGEASISCRSS**QSLH**SNRYNYLDWYLQKPGQSP
 QLLIY**LGSN**RASGVDPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**MQALQT**
PYTFGQGTK (SEQ ID NO: 18) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical
 20 thereto and/or having two or three amino acids substitutions

(iv) a heavy chain variable domain comprising:

QVQLQQSGPGLVKPSQTL**SLTCAISGDSISSNSVA**WNWIRQSPSRGL
 EWLGR**TYR**SKWYSDY**AI**SVKGRLDINPDTSKNQFSLQLNSVTPEDT
 AVYYC**CAREVEGSSYDAFDI**WGQGT**M (SEQ ID NO: 8)** or a fragment,
 25 variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions, and

a light chain variable domain comprising

DIQLTQSPSSLSASVGDRVTITCRAS**QSISD**FLN**WY**QKPGKAPKLLI**YA**
 30 **ASSLQ**TGVPSR**FGGSGSGTEFTLTI**SSLQPEDLGTY**YCQQSYIMPDTFG**
QGTK (SEQ ID NO: 22) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions, or

(v) a heavy chain variable domain comprising:

EVQLVQSGAEVKKPGSSVKVSCKAS**GGTFSSYA**ISWVRQAPGQGLEWM
GGI**IPIFGT**ANYAQNFQGRVTMTADTSISTAYMELSSLRSEDTAVYYC**CARS**

LGGRFRYWQGQTL (SEQ ID NO: 4) or a fragment, variant or sequence

thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions, and a light chain variable domain

DVVMTQSPLSLPVTGPGEPAISCRSS**QSLLHSNGYNY**LDWYLQKPGQSP
QLLIY**LGSN**RASGVPDRFSGSGSGTDFTLQISRVEAEDAGVYYC**MQGLQ**

PYTFGQGK (SEQ ID NO: 24) or a fragment, variant or sequence thereof

at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions.

In one embodiment, the antigen binding protein, variant or fragment thereof comprises a single domain heavy chain variable domain having a sequence:

(i) QVQLQESGGGLVQPGGSLRLSCADS**GSIFSGND**MAWYRRAPGVERE
LVAVIT**SGGST**HYADSVKGRFTISRDNAAQKTYYLQTNLKPEDTAVYY

CTNGRWSGDTYYAHHWGQGTQ (SEQ ID NO: 28) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions, or

QVQLQESGGGLVQPGGSLRLSCAAS**GSSERFTS**VAWYRQAPGKERE
LVAFIT**NGGST**RYTDPVKGRFTISRDNAKNTVYQLQMNSLKAEDTAVYY

CMAGTSWGQGTQ (SEQ ID NO: 32) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%

identical thereto and/or having two or three amino acids substitutions, or
) QVQLQESGGGLVQPGGSLRLSCADSGSIFSGNDMAWYRRAPGVERE

LVAVIT**SGGST**HYADSVKGRFTISRDNAAQKTYYLQTNDLKPEDTAVYY
CTNGRWSGD**TTYAHL**WGQGTQ (SEQ ID NO: 34) or a fragment

variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino

acids substitutions, or

LVAITSGGSTHYADSVKGRFTISRDNAAQKTYYLQTNLDKPEDTAVYY

CTNGRWSGDTYYAHHWGQGTQ (**SEQ ID NO: 35**) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions, or

5 (v) QVQLQESGGGLVQAGDSLRLSCADSGSIFSGNDMAWYRRAPGVERE
LVAVITSGGSTHYADSVKGRFTISRDNQKTVYLQTNLKPEDTAVYY
CTNGRWSGDTYYAHHWGQGTQ (**SEQ ID NO: 36**) or a fragment,
variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%,
98%, 99%, or 100% identical thereto and/or having two or three amino
10 acids substitutions, or

(vi) QVQLVESGGGLVQAGGSLRLSCAASGSIFSGNDMSWYRQAPGKGLE
LVAVITSGGSTYYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYY
CTNGRWSGDTYYAHHWGQGTQ (**SEQ ID NO: 37**) or a fragment,
variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%,
15 98%, 99%, or 100% identical thereto and/or having two or three amino
acids substitutions.

In one embodiment, the antigen binding protein, variant or fragment thereof comprises a light chain constant domain having a sequence:

(i) RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSG
20 NSQESVAEQDSKDSTYLSSTLTLSKADYEKHKLYACEVTHQGLSSPVTK
SFNRGEC (**SEQ ID NO: 13**) or a fragment, variant or sequence thereof at
least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto
and/or having two or three amino acids substitutions, or

(ii) RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSG
25 NSQESVTEQDSKDSTYLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTK
SFSRGEC (**SEQ ID NO: 14**) or a fragment, variant or sequence thereof at
least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto
and/or having two or three amino acids substitutions.

In one embodiment, the antigen binding protein, variant or fragment thereof is an
30 IgG antibody, in particular an IgG1 antibody.

In one embodiment, the antigen binding protein, variant or fragment thereof is a multi-specific antigen binding protein, variant or fragment thereof, such as a bispecific antibody.

In one embodiment, the multi-specific antigen binding protein, variant or fragment thereof binds to an immune marker selected from the group consisting of CD3, NKG2D, CD4, CD8, CD16 and CD64.

In one embodiment, the antigen binding protein, variant or fragment thereof is a bispecific T cell engager (BiTE).

In one embodiment, the bispecific T cell engager (BiTE) comprises an anti-EpCAM Heavy chain antibody variable region (i.e. VHH) or a single chain variable fragment (scFv).

In one embodiment, the antigen binding protein, variant or fragment thereof comprises a Fc region.

In one aspect, there is provided a polynucleotide encoding the antigen binding protein, variant or fragment thereof of any one of the preceding claims.

In one embodiment, the heavy chain variable domain is encoded by a nucleotide sequence comprising:

- (i) GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
 GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCT
 TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCAA**ACTA
 CGCACAGAACTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
 CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
 CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
CTACTGGGGCCAGGGAACCCCTG (**SEQ ID NO: 41**) or a sequence at
 least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical
 thereto and/or having 10-20 nucleic acid substitutions, or
- (ii) CAGGTACAGCTGCAGCAGTCAGGTCCAGGGCTGGTGAAGCCCTC
 GCAGACCCTCTCACTCACCTGTGCCATCTCC**GGGGACAGTATCTC**
TAGTAACAGTGTGCTTGGAAGTGGATCAGGCAGTCCCCATCGAG
 AGGCCTTGAGTGGCTGGGAAGG**ACATACTACAGGTCCAAGTGGT**
ACAGTGATTATGCAATATCTGTGAAAGTGGATTAGACATCAACCC
 AGACACATCCAAGAACCAGTTCTCCCTGCAGCTGAACTCTGTGACT
 CCCGAGGACACGGCTGTGTATTATTGT**GCAAGAGAAGTTGAGGGC**
AGCAGCTATGATGCTTTTGATATCTGGGGCCAAGGGACAATG
 (**SEQ ID NO: 45**) or a sequence at least 80%, 85%, 90%, 95%, 96%,

97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions;

(iii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
AGGGTCTCTGAGACTCTCCTGTGCAGACTCT**TGGAAGCATCTTCAG**
5 **TGGCAATGAC**ATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTAC**ACACTATG
CAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATGCCC
AGAAGACCGTATATCTGCAAACGAACGACTTGAAACCTGAGGACAC
GGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTA**
10 **CTATGCCCCATCACT**GGGGCCAGGGGACCCAG (**SEQ ID NO: 65**) or a
sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%
identical thereto and/or having 10-20 nucleic acid substitutions,

(iv) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
GGGGTCTCTGAGACTCTCCTGTGCAGCCTCT**TGGAAGCTCCGAAAG**
15 **ATTCACATCA**GTGGCCTGGTACCGCCAGGCTCCAGGAAAGGAGC
GCGAGTTGGTCGCATTT**ATTACTAATGGTGGTAGCACA**AGATATAC
AGACCCCGTGAAGGGCCGATTACCATCTCCAGAGACAACGCCAA
GAACACGGTGTATCTGCAAATGAACAGCCTGAAAGCTGAGGACAC
GGCCGTCTATTATTGT**TATGGCGGGTACGTCCT**GGGGCCAGGGGAC
20 CCAG (**SEQ ID NO: 69**) or a sequence at least 80%, 85%, 90%, 95%,
96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20
nucleic acid substitutions,

(v) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
AGGGTCTCTGAGACTCTCCTGTGCAGACTCT**TGGAAGCATCTTCAG**
25 **TGGCAATGAC**ATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTAC**ACACTATG
CAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATGCCC
AGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACA
CGGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTT**
30 **ACTATGCCCCATCTCT**GGGGCCAGGGGACCCAG (**SEQ ID NO: 71**) or
a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or
100% identical thereto and/or having 10-20 nucleic acid substitutions,

- (vi) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGG
GGGCTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG**
TGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACACACTATG**
5 CAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATGCCC
AGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACA
CGGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTT**
ACTATGCCCATCACTGGGGGCCAGGGGACCCAG (**SEQ ID NO: 72**) or
a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or
10 100% identical thereto and/or having 10-20 nucleic acid substitutions,
- (vii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGG
GGACTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGT**
GGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCG
CGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACACACTATGC**
15 AGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATGCCCA
GAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACAC
GGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTA**
CTATGCCCATCACTGGGGGCCAGGGGACCCAG (**SEQ ID NO: 73**) or a
sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%
20 identical thereto and/or having 10-20 nucleic acid substitutions, or
- (viii) CAGGTGCAGCTGGTGGAGTCTGGGGGAGGATTGGTGCAGG
CTGGGGGCTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCATCT**
TCAGTGGCAATGACATGTCCTGGTACCGCCAGGCTCCAGGGAAG
GGACTCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACATAC**
25 TATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATT
CCaAGAAcACCcTATATCTGCAAATGAACAGCCTGAGAGCTGAGGA
CACGGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATA**
CTTACTATGCCCATCACTGGGGGCCAGGGGACCCTG (**SEQ ID NO:**
30 **74**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%,
or 100% identical thereto and/or having 10-20 nucleic acid substitutions;
and/or
- the light chain variable domain is encoded by a nucleotide sequence comprising:

- (i) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCCGTACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
5 CGGGGTCCCTGACAGGTTCA GTGGCAGTGGATCAGGCACAGATTT
TACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGCAT**GCAAGCTCTACAACTCCGTACACT**TTTGGCCAGGGG
ACCAAG (**SEQ ID NO: 49**) or a sequence at least 80%, 85%, 90%, 95%,
96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20
10 nucleic acid substitutions, or
- (ii) GAAATTGTGCTGACTCAGTCTCCACTCTCCCTGCCCCGTACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
15 CGGGGTCCCTGACAGGTTCA GTGGCAGTGGATCAGGTACAGATTT
TACACTGAAAATAAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGCAT**GCAAGGTCTACAAAGTCCCTGGACG**TTCCGCCAAGGG
ACCAAG (**SEQ ID NO: 53**) or a sequence at least 80%, 85%, 90%, 95%,
96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20
20 nucleic acid substitutions, or
- (iii) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCCGTACCCCTG
GAGAGTCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATAGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
25 CGGGGTCCCTGACAGGTTCA GTGGCAGTGGATCAGGCACAGATTT
TACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGCAT**GCAAGCTCTACAACTCCGTACACT**TTTGGCCAGGGG
ACCAAG (**SEQ ID NO: 55**) or a sequence at least 80%, 85%, 90%, 95%,
96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20
30 nucleic acid substitutions, or
- (iv) GACATCCAGTTGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAG
GAGACAGAGTCACCATCACTTGCCGGGCAAGT**CAGAGTATTAGCG**
ACTTTTTAAATTGGTACCAGCAGAAACCAGGTAAAGCCCCGAAGCT

CCTGATCTAT**GCTGCATCG**AGTTTACAACTGGGGTCCCCCTCAAGA
 TTCGGTGGCAGTGGATCTGGGACAGAATTCACCTCTCACCATAAGCA
 GTCTACAACCTGAAGATTTGGGAAGTTATTACTGT**CAACAGAGTTA**
CATTATGCCCCGACACTTTTTGGCCAGGGGACGAAA (**SEQ ID NO: 59**)
 5 or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or
 100% identical thereto and/or having 10-20 nucleic acid substitutions, or
 GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
 GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
 10 GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
 CGGGGTCCCTGACAGGTTCA GTGGCAGTGGATCAGGCACAGATTT
 TAACTGCAAATCAGCAGAGTGGAGGCTGAGGATGCTGGGGTTTA
 TTA CTGCAT**GCAAGGTCTACAGACTCCGTACACT**TTTGGCCAGGG
 15 GACCAAG (**SEQ ID NO: 61**) or a sequence at least 80%, 85%, 90%,
 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-
 20 nucleic acid substitutions.

In one embodiment, the heavy chain and light chain variable domains are encoded by nucleotide sequences selected from the group consisting of:

(i) a heavy chain variable domain encoded by the nucleotide sequence
 20 comprising:
 GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
 GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCT
 TGAGTGGATGGGAGGGAT**CATCCCTATCTTTGGTACAGCA**AACTA
 25 CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
 CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
 CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
CTACTGGGGCCAGGGAACCCTG (**SEQ ID NO: 41**) or a sequence at
 30 least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical
 thereto and/or having 10-20 nucleic acid substitutions, and
 a light chain variable domain encoded by the nucleotide sequence
 comprising:

5 GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCCGTACCCCTG
 GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
 GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
 CGGGGTCCCTGACAGGTTCA GTGGCAGTGGATCAGGCACAGATTT
 TACTGCAAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
 TACTGCAT**GCAAGCTCTACAACTCCGTACACT**TTTGGCCAGGGG
 ACCAAG (**SEQ ID NO: 49**) or a sequence at least 80%, 85%, 90%, 95%,
 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20
 10 nucleic acid substitutions

(ii) a heavy chain variable domain encoded by the nucleotide sequence comprising:

15 GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
 GTCCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCT
 TGAGTGGATGGGAGGGAT**CATCCCTATCTTTGGTACAGCA**AACTA
 CGCACAGAACTTCCAGGGCAGAGTACCATGACCGCAGACACCTC
 CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
 CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
 20 **CTACTGGGGCCAGGGA**ACCCTG (**SEQ ID NO: 41**) or a sequence at
 least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical
 thereto and/or having 10-20 nucleic acid substitutions, and
 a light chain variable domain encoded by the nucleotide sequence
 comprising:

25 GAAATTGTGCTGACTCAGTCTCCACTCTCCCTGCCCCGTACCCCTG
 GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
 GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
 CGGGGTCCCTGACAGGTTCA GTGGCAGTGGATCAGGTACAGATTT
 30 TACTGCAAAAATAAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
 TACTGCAT**GCAAGGTCTACAAAGTCCCTGGACG**TTTCGGCCAAGGG
 ACCAAG (**SEQ ID NO: 53**) or a sequence at least 80%, 85%, 90%, 95%,

96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions

(iii) a heavy chain variable domain encoded by the nucleotide sequence comprising:

5 GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
 GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCT
 TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCAA**ACTA
 CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
 10 CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
 CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
CTACTGGGGCCAGGGAACCCTG (SEQ ID NO: 41) or a sequence at
 least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical
 thereto and/or having 10-20 nucleic acid substitutions, and

15 a light chain variable domain encoded by the nucleotide sequence comprising:

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
 GAGAGTCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATAGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
 20 GGCAGTCTCCACAGCTCCTGATCTATTT**GGTTCTAATCGGGCCTC**
 CGGGGTCCCTGACAGGTTCA GTGGCAGTGGATCAGGCACAGATTT
 TACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
 TACTGC**ATGCAAGCTCTACAACTCCGTACACTTTTGGCCAGGGG**
ACCAAG (SEQ ID NO: 55) or a sequence at least 80%, 85%, 90%, 95%,
 25 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20
 nucleic acid substitutions

(iv) a heavy chain variable domain encoded by the nucleotide sequence comprising:

30 GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
 GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCT
 TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCAA**ACTA
 CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC

- 5 CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
CTACTGGGGCCAGGGAACCCTG (SEQ ID NO: 41) or a sequence at
least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical
thereto and/or having 10-20 nucleic acid substitutions, and
a light chain variable domain encoded by the nucleotide sequence
comprising:
GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
10 **ATAGTAATGGATACAAC**TATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTATTT**GGGTTCT**AATCGGGCCTC
CGGGGTCCCTGACAGGTTCA GTGGCAGTGGATCAGGCACAGATT
TACACTGCAAATCAGCAGAGTGGAGGCTGAGGATGCTGGGGTTTA
TTACTGCAT**GCAAGGTCTACAGACTCCGTACACT**TTTGGCCAGGG
15 GACCAAG (**SEQ ID NO: 61**) or a sequence at least 80%, 85%, 90%,
95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-
20 nucleic acid substitutions; and
(v) a heavy chain variable domain encoded by the nucleotide sequence
comprising:
20 CAGGTACAGCTGCAGCAGTCAGGTCCAGGGCTGGTGAAGCCCTC
GCAGACCCTCTCACTCACCTGTGCCATCTCC**GGGGACAGTATCTC**
TAGTAACAGTGTTGCTTGGAAGTGGATCAGGCAGTCCCCATCGAG
AGGCCTTGAGTGGCTGGGAAGG**ACATACTACAGGTCCAAGTGGT**
ACAGTGATTATGCAATATCTGTGAAAGGTCGATTAGACATCAACCC
25 AGACACATCCAAGAACCAGTTCTCCCTGCAGCTGAACTCTGTGACT
CCCAGGACACGGCTGTGTATTATTGT**GCAAGAGAAGTTGAGGGC**
AGCAGCTATGATGCTTTTGATATCTGGGGCCAAGGGACAATG
(**SEQ ID NO: 45**) or a sequence at least 80%, 85%, 90%, 95%, 96%,
97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic
30 acid substitutions, and
a light chain variable domain encoded by the nucleotide sequence
comprising:

- GACATCCAGTTGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAG
 GAGACAGAGTCACCATCACTTGCCGGGCAAGT**CAGAGTATTAGCG**
ACTTTTTAAATTGGTACCAGCAGAAACCAGGTAAAGCCCCGAAGCT
 CCTGATCTAT**GCTGCATCG**AGTTTACAACTGGGGTCCCCCTCAAGA
 5 TTCGGTGGCAGTGGATCTGGGACAGAATCACTCTCACCATAAGCA
 GTCTACAACCTGAAGATTTGGGAACCTTACTGT**CAACAGAGTTA**
CATTATGCCCCGACACTTTTGGCCAGGGGACGAAA (SEQ ID NO: 59)
 or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or
 100% identical thereto and/or having 10-20 nucleic acid substitutions.
- 10 In one embodiment, the single domain heavy chain variable domain is encoded
 by a nucleotide sequence comprising
- (i) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
 AGGGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG**
TGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
 15 GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACACACTATG**
 CAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATGCCC
 AGAAGACCGTATATCTGCAAACGAACGACTTGAAACCTGAGGACAC
 GGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTA**
CTATGCCCCATCACTGGGGCCAGGGGACCCAG (SEQ ID NO: 65) or a
 20 sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%
 identical thereto and/or having 10-20 nucleic acid substitutions,
- (ii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
 GGGGTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCTCCGAAAG**
ATTCACATCAGTGGCCTGGTACCGCCAGGCTCCAGGAAAGGAGC
 25 GCGAGTTGGTCGCATTT**ATTACTAATGGTGGTAGCACAA**AGATATAC
 AGACCCCGTGAAGGGCCGATTACCATCTCCAGAGACAACGCCAA
 GAACACGGTGTATCTGCAAATGAACAGCCTGAAAGCTGAGGACAC
 GGCCGTCTATTATTGT**TATGGCGGGTACGTCCTGGGGCCAGGGGAC**
 CCAG (**SEQ ID NO: 69**) or a sequence at least 80%, 85%, 90%, 95%,
 30 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20
 nucleic acid substitutions,
- (iii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
 AGGGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG**

TGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
GCGAGTTGGTCGCGGTTATTACTAGCGGTGGTAGTACACACTATG
CAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATGCCC
AGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACA
 5 **CGGCCGTGTATTACTGCACAAACGGAAGATGGTCAGGCGATACTT**
ACTATGCCCATCTCTGGGGCCAGGGGACCCAG (SEQ ID NO: 71) or
 a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or
 100% identical thereto and/or having 10-20 nucleic acid substitutions,
 (iv) **CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGG**
 10 **GGGCTCTCTGAGACTCTCCTGTGCAGACTCTGGAAGCATCTTCAG**
TGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
GCGAGTTGGTCGCGGTTATTACTAGCGGTGGTAGTACACACTATG
CAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATGCCC
AGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACA
 15 **CGGCCGTGTATTACTGCACAAACGGAAGATGGTCAGGCGATACTT**
ACTATGCCCATCACTGGGGCCAGGGGACCCAG (SEQ ID NO: 72) or
 a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or
 100% identical thereto and/or having 10-20 nucleic acid substitutions,
 (v) **CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGG**
 20 **GGACTCTCTGAGACTCTCCTGTGCAGACTCTGGAAGCATCTTCAGT**
GGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCG
CGAGTTGGTCGCGGTTATTACTAGCGGTGGTAGTACACACTATGC
AGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATGCCCCA
GAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACAC
 25 **GGCCGTGTATTACTGCACAAACGGAAGATGGTCAGGCGATACTTA**
CTATGCCCATCACTGGGGCCAGGGGACCCAG (SEQ ID NO: 73) or a
 sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%
 identical thereto and/or having 10-20 nucleic acid substitutions, or
 (vi) **CAGGTGCAGCTGGTGGAGTCTGGGGGAGGATTGGTGCAGGCTGG**
 30 **GGGCTCTCTGAGACTCTCCTGTGCAGCCTCTGGAAGCATCTTCAG**
TGGCAATGACATGTCCTGGTACCGCCAGGCTCCAGGGAAGGGACT
CGAGTTGGTCGCGGTTATTACTAGCGGTGGTAGTACATACTATGC
AGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCaAG

AAcACcTATATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGG
 CCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTACT**
ATGCCCATCACTGGGGCCAGGGGACCCTG (**SEQ ID NO: 74**) or a
 sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%
 identical thereto and/or having 10-20 nucleic acid substitutions.

5

In one embodiment, the light chain constant domain is encoded by a nucleotide
 sequence comprising

(i) CGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATG
 AGCAGTTGAAATCTGGAAGTGCCTCTGTTGTGTGCCTGCTGAATAA
 CTTCTATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGC
 CCTCCAATCGGGTAACTCCCAGGAGAGTGTGCGAGAGCAGGACAG
 CAAGGACAGCACCTACAGCCTCAGCAGCACCCCTGACGCTGAGCAA
 AGCAGACTACGAGAAACACAACTCTACGCCTGCGAAGTCACCCAT
 CAGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCAACAGGGGAGA
 GTGT (**SEQ ID NO: 50**) or a sequence at least 80%, 85%, 90%, 95%,
 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20
 nucleic acid substitutions, or

10

15

20

25

(ii) CGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATG
 AGCAGTTGAAATCTGGAAGTGCCTCTGTTGTGTGCCTGCTGAATAA
 CTTCTATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGC
 CCTCCAATCGGGTAACTCCCAGGAGAGTGTGACAGAGCAGGACAG
 CAAGGACAGCACCTACAGCCTCAGCAGCACCCCTGACGCTGAGCAA
 AGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCCA
 TCAGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCAACAGGGGAGA
 GTGT (**SEQ ID NO: 51**) or a sequence at least 80%, 85%, 90%, 95%,
 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20
 nucleic acid substitutions.

In one aspect, there is provided a vector expressing the polynucleotide as
 described above.

30

In one aspect, there is provided a host cell comprising the vector as described
 above.

In one aspect, there is provided a cell expressing/secreting the antigen binding
 protein, variant of fragment thereof as described above.

In one aspect, there is provided a cell expressing/secretory an immune cell engager which is specific to EpCAM.

In one embodiment, the immune cell engager is selected from the group comprising a T cell engager, an NK cell engager, a monocyte engager and a macrophage engager.

In one embodiment, the immune cell engager is a bispecific T cell engager (BiTE), such as an inducible BiTE, non-inducible BiTE or a constitutive expression BiTE comprising the antigen binding protein, variant or fragment thereof as described above.

In one embodiment, the cell is an immune cell, for example selected from the group comprising a T cell, a macrophage, a monocyte and an NK cell.

In one embodiment, the immune cell is a T-cell, in particular a CAR T-cell.

In one embodiment, the cell is a stem cell, for example selected from the group comprising a mesenchymal stem cell, a neural stem cell and a pluripotent stem cell, such as an induced pluripotent stem cell (iPSC).

In one aspect, there is provided a composition comprising the antigen binding protein, variant or fragment thereof and/or cell as described above.

In one aspect, there is provided an antigen binding protein, variant or fragment thereof, cell or a composition as described above for use in treatment of a disease.

In one aspect, there is provided a method of treating a disease in a subject in need thereof, comprising administering to the subject an antigen binding protein, variant or fragment thereof, cell or a composition as described above.

In one aspect, there is provided a use of the antigen binding protein, variant or fragment thereof, cell or the composition as described above in the manufacture of a medicament for preventing and/or treating a disease.

In one embodiment, the disease is a proliferative disease such as a tumor or cancer.

In one aspect, there is provided a method of diagnosis / determining the prognosis or presence of a solid tumor originating from epithelium comprising detecting the expression of high EpCAM on the solid tumor using the antigen binding protein, variant or fragment thereof or a composition as described above.

DEFINITIONS

The term "antigen binding protein" herein is used in the broadest sense and encompasses various antibody structures, including but not limited to monoclonal antibodies, polyclonal antibodies, multi-specific antibodies as long as they exhibit the desired antigen-binding activity.

The term "antibody" as used herein relates to whole (i.e., full length) antibodies (i.e., comprising the elements of two heavy chains and two light chains) and functionally active fragments thereof (i.e., molecules that contain an antigen binding domain that specifically binds an antigen, also termed antibody fragments or antigen-binding fragments). Features described herein with respect to antibodies also apply to antibody fragments unless context dictates otherwise. The term "antibody" encompasses monovalent, i.e., antibodies comprising only one antigen binding domain (e.g., one-armed antibodies comprising a full-length heavy chain and a full-length light chain interconnected, also termed "half-antibody"), and multivalent antibodies, i.e., antibodies comprising more than one antigen binding domain, e.g., bivalent.

The term "antigen binding fragment" as employed herein refers to functionally active antibody binding fragments including but not limited to Fab, modified Fab, Fab', modified Fab', F(ab')₂, Fv, single domain antibodies, scFv, Fv, bi, tri or tetra-valent antibodies, Bis-scFv, diabodies, triabodies, tetrabodies and epitope-binding fragments of any of the above.

A "binding fragment" as employed herein refers to a fragment capable of binding a target peptide or antigen with sufficient affinity to characterize the fragment as specific for the peptide or antigen.

The term "monoclonal antibody" (or "mAb") refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e. each individual of a monoclonal antibody preparation are identical except for possible mutations (e.g., naturally occurring mutations), that may be present in minor amounts. Certain differences in the protein sequences linked to post-translational modifications (for example, cleavage of the heavy chain C-terminal lysine, deamidation of asparagine residues and/or isomerization of aspartate residues) may nevertheless exist between the various different antibody molecules present in the composition. Contrary to polyclonal antibody preparations, each monoclonal antibody of a monoclonal antibody preparation is directed against a single determinant on an antigen.

The term "diabody" as employed herein refers to two Fv pairs, a first VH/VL pair and a further VH/VL pair which have two inter-Fv linkers, such that the VH of a first Fv is linked to the VL of the second Fv and the VL of the first Fv is linked to the VH of the second Fv.

5 The term "tribody" (also referred to a Fab(scFv)₂) as employed herein refers to a Fab fragment with a first scFv appended to the C-terminal of the light chain and a second scFv appended to the C-terminal of the heavy chain. The term "tetrabody" as employed herein refers to a format similar to the diabody comprising four Fvs and four inter-Fv linkers.

10 The term "multivalent antibody" refers to an antibody comprising more than one antigen binding domain e.g., bivalent.

 The term "Fv" refers to two variable domains of full-length antibodies, for example co-operative variable domains, such as a cognate pair or affinity matured variable domains, i.e., a VH and VL pair. The term "scFv" refers to single chain variable fragment
15 which is a fusion protein of the variable regions of the heavy and light chains of the immunoglobulins, connected with a short linker peptide of ten to about 25 amino acids. The term "bis-scFv" as described herein refers to a bispecific scFv.

 The term "dsscFv" or "disulphide-stabilised single chain variable fragment" as employed herein refers to a single chain variable fragment which is stabilised by a
20 peptide linker between the VH and VL variable domain and also includes an inter-domain disulphide bond between VH and VL.

 The term "DVD-Ig" (also known as dual V domain IgG) refers to a full-length antibody with 4 additional variable domains, one on the N-terminus of each heavy and each light chain.

25 The term "Fab" refers to as used herein refers to an antibody fragment comprising a light chain fragment comprising a VL (variable light) domain and a constant domain of a light chain (CL), and a VH (variable heavy) domain and a first constant domain (CH1) of a heavy chain. Dimers of a Fab' according to the present disclosure create a F(ab')₂ where, for example, dimerization may be through the hinge. The term "F(ab)" refers to
30 a monovalent fragment of a single light chain homodimer, which is obtained by pepsin digestion of IgG, followed by reduction of the light chain disulfide bond. The term "F(ab')₂" as described herein refers to a fragment of IgG that is prepared by pepsin digestion of IgG. The F(ab')₂ fragment is a disulfide-linked homodimer of the two light

chain dimers, so it retains bivalent epitope binding like whole IgG, but as it lacks the heavy chains, it is smaller in size compared to a whole IgG. F(ab')₂ and F(ab') fragments do not bind to immunoglobulin receptors on cells, which can be useful for achieving specific staining of the primary antibody target.

5 The term "DiFab" as employed herein refers to two Fab molecules linked via their C-terminus of the heavy chains or two Fab' molecules linked via one or more disulfide bonds in the hinge region thereof.

As used herein, the term "nanobody" refers to a single domain antibody (sdAb), with an antibody fragment consisting of a single monomeric variable antibody domain.

10 The term "antigen binding variant" refers to a polypeptide, for example, an antibody possessing the desired characteristics described herein and comprising a VH and/or a VL that has at least about 80% amino acid sequence identity with a VH and/or a VL of the reference antibody. Such antibody variants include, for instance, antibodies wherein one or more amino acid residues are added to or deleted from the VH and/or a
15 VL domain. Ordinarily, an antibody variant will have at least about 80% amino acid sequence identity, alternatively at least about 85%, 90%, 95%, 96%, 97%, 98%, or 99% amino acid sequence identity, to an antibody described herein. Optionally, variant antibodies will have no more than one conservative amino acid substitution as compared to an antibody sequence provided herein, alternatively no more than about any of 2, 3,
20 4, 5, 6, 7, 8, 9, or 10 conservative amino acid substitutions as compared to an antibody sequence provided herein.

The term "specifically" as employed herein in the context of antibodies is intended to refer to an antibody that only recognizes the antigen to which it is specific or an antibody that has significantly higher binding affinity to the antigen to which it is specific
25 compared to binding to antigens to which it is non-specific, for example at least 5, 6, 7, 8, 9, 10 times higher binding affinity.

The term "epitope" or "binding site" in the context of antibodies refers to a site (or a part) on an antigen to which the paratope of an antibody binds or recognizes. Epitopes can be formed both from contiguous amino acids (also often called "linear epitopes") or
30 noncontiguous amino acids formed by tertiary folding of a protein (often called "conformational epitopes"). Epitopes formed from contiguous amino acids are typically retained on exposure to denaturing solvents whereas epitopes formed by folding are typically lost on treatment with denaturing solvents. An epitope typically includes at least

3, and more usually, at least 5-10 amino acids in a unique spatial conformation. Epitopes usually consist of chemically active surface groups of molecules such as amino acids, sugar side chains and usually have specific 3D structural and charge characteristics.

The "class" of an antibody refers to the type of constant domain or constant region possessed by its heavy chain. There are five major classes of antibodies: IgA, IgD, IgE, IgG, and IgM, and several of these may be further divided into subclasses (isotypes), e.g., IgG1, IgG2, IgG3, IgG4, IgA1, and IgA2. The heavy chain constant domains that correspond to the different classes of immunoglobulins are called a, d, e, g, and m, respectively.

10 The term "chimeric antibody" (or antigen-binding fragment thereof) is an antibody molecule (or antigen-binding fragment thereof) in which (a) the constant region, or a portion thereof, is altered, replaced or exchanged so that the antigen binding site (variable region) is linked to a constant region of a different or altered class, effector function and/or species, or an entirely different molecule which confers new properties
15 to the chimeric antibody, e. g., an enzyme, toxin, hormone, growth factor, drug, etc.; or (b) the variable region, or a portion thereof, is altered, replaced or exchanged with a variable region having a different or altered antigen specificity. For example, a mouse antibody can be modified by replacing its constant region with the constant region from a human immunoglobulin. Due to the replacement with a human constant region, the
20 chimeric antibody can retain its specificity in recognizing the antigen while having reduced antigenicity in human as compared to the original mouse antibody.

The term "chimeric antigen receptors" refers to receptor protein that has been engineered to give T cells the new ability to target a specific antigen. The receptors are chimeric in that they combine both antigen-binding and T cell activating functions into a
25 single receptor. CAR T cell therapy uses T cells engineered with CARs to treat cancer. T cells in CAR T immunotherapy are modified to recognize cancer cells in order to more effectively target and destroy them. CAR T cells can be derived either from T cells in a patient's own blood (autologously) or from the T cells of another, healthy, donor (allogeneically). Once isolated from a person, these T cells are genetically engineered
30 to express a specific CAR, which programs them to target an antigen that is present on the surface of tumors. For safety, CAR T cells are engineered to be specific to an antigen that is expressed on a tumor but is not expressed on healthy cells. CAR T cells destroy cells through extensive stimulated cell proliferation, increasing the degree to which they

are toxic to other living cells (cytotoxicity) and by causing the increased secretion of factors that can affect other cells such as cytokines, interleukins and growth factors. The surface of CAR T cells can bear either two types of co-receptors, CD4 and CD8, each with different and interacting cytotoxic effects.

5 The term "human antibody" or "humanized antibody" (or antigen-binding fragment thereof), as used herein, is intended to include antibodies (and antigen-binding fragments thereof) having variable regions in which both the framework and CDR regions are derived from sequences of human origin. Depending on the amino acid sequence of the constant region of their heavy chains, antibodies or immunoglobulins are divided into
10 the classes: IgA, IgD, IgE, IgG and IgM, and several of these may be further divided into subclasses (subtypes), e.g. IgG1, IgG2, IgG3, and IgG4, IgA1, and IgA2. Therefore, human IgG constant region domains may be used, especially of the IgG1 and IgG3 isotypes when the antibody molecule is intended for therapeutic uses and antibody effector functions are required. Alternatively, IgG2 and IgG4 isotypes may be used when
15 the antibody molecule is intended for therapeutic purposes and antibody effector functions are not required. Furthermore, if the antibody contains a constant region, the constant region also is derived from such human sequences. A humanized antibody (or antigen-binding fragment thereof) retains the reactivity of a non-human antibody while being less immunogenic in humans. This can be achieved, for instance, by retaining the
20 non-human CDR regions and replacing the remaining parts of the antibody with their human counterparts (i. e., the constant region as well as the framework portions of the variable region). Additional framework region modifications may be made within the human framework sequences as well as within the CDR sequences derived from the germline of another mammalian species. The humanized antibodies of the present
25 disclosure may include amino acid residues not encoded by human sequences (e. g., mutations introduced by random or site-specific mutagenesis in vitro or by somatic mutation in vivo, or a conservative substitution to promote stability or manufacturing). This definition of a humanized antibody specifically excludes a humanized antibody comprising non-human antigen-binding residues. Human antibodies can be produced
30 using various techniques known in the art, including phage-display libraries, administering the antigen to a transgenic animal that has been modified to produce such antibodies in response to antigenic challenge, but whose endogenous loci have been disabled, e. g., immunized xenomice via a human B-cell hybridoma technology.

The term "recombinant humanized antibody" as used herein, includes all human antibodies that are prepared, expressed, created or isolated by recombinant means, such as antibodies isolated from a host cell transformed to express the humanized antibody, e. g., from a transfectoma, and antibodies prepared, expressed, created or isolated by any other means that involve splicing of all or a portion of a human immunoglobulin gene, sequences to other DNA sequences.

The term "isolated" means, throughout this specification, that the antibody, or polynucleotide, as the case may be, exists in a physical milieu distinct from that in which it may occur in nature. The term "isolated" nucleic acid refers to a nucleic acid molecule that has been isolated from its natural environment or that has been synthetically created. An isolated nucleic acid may comprise synthetic DNA, for instance produced by chemical processing, cDNA, genomic DNA or any combination thereof. An isolated antibody refers to an antibody that is substantially free of other cellular material and/or chemicals.

The term "Complementarity Determining Regions" ("CDRs") refers to amino acid sequences with boundaries determined using any of a number of well-known schemes, including those described by Kabat (i.e., "Kabat" numbering scheme); Al-Lazikani ("Chothia" numbering scheme); ImMunoGenTics (IMGT) numbering ("IMGT" numbering scheme); and the like. The term "Complementarity Determining Regions" ("CDRs") refers to regions of hypervariability that contain the binding domain that interacts with an antigen. Antibodies typically comprise six CDRs: three in the VH (H1, H2, H3), and three in the VL (L1, L2, L3).

As used herein, the term "sequence identity" refers to the percentage sequence identities that are determined with antibody sequences maximally aligned by the Kabat numbering convention. After alignment, if a subject antibody region (e.g., the entire mature variable region of a heavy or light chain) is being compared with the same region of a reference antibody, the percentage sequence identity between the subject and reference antibody regions is the number of positions occupied by the same amino acid in both the subject and reference antibody region divided by the total number of aligned positions of the two regions, with gaps not counted, multiplied by 100 to convert to percentage.

"Conservative substitutions" may be made, for instance, on the basis of similarity in polarity, charge, size, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the amino acid residues involved. The 20 naturally occurring amino acids can

be grouped into the following six standard amino acid groups: (1) hydrophobic: Met, Ala, Val, Leu, Ile; (2) neutral hydrophilic: Cys, Ser, Thr; Asn, Gln; (3) acidic: Asp, Glu; (4) basic: His, Lys, Arg; (5) residues that influence chain orientation: Gly, Pro; and (6) aromatic: Trp, Tyr, Phe.

5 As used herein, "conservative substitutions" are defined as exchanges of an amino acid by another amino acid listed within the same group of the six standard amino acid groups shown above. For example, the exchange of Asp by Glu retains one negative charge in the so modified polypeptide. In addition, glycine and proline may be substituted for one another based on their ability to disrupt α -helices.

10 As used herein, "non-conservative substitutions" are defined as exchanges of an amino acid by another amino acid listed in a different group of the six standard amino acid groups (1) to (6) shown above.

 The term "affinity" refers to the strength of all noncovalent interactions between an antibody thereof and the target protein. Unless indicated otherwise, as used herein, 15 the term "binding affinity" refers to intrinsic binding affinity which reflects a 1:1 interaction between members of a binding pair (e.g., antibody and antigen). The affinity of a molecule for its binding partner can be generally represented by the dissociation constant (K_D). Affinity can be measured by common methods known in the art, including those described herein.

20 The term " K_D " as used herein refers to the constant of dissociation which is obtained from the ratio of K_d to K_a (i.e. K_d / K_a) and is expressed as a molar concentration (M). K_d and K_a refer to the dissociation rate and association rate, respectively, of a particular antigen-antibody interaction. K_D values for antibodies can be determined using methods well established in the art. As used herein, the term "low affinity" refers to K_D 25 of 100 nM or more.

 As used herein, the term "moderate affinity" refers to K_D ranging from 10nM to 100nM.

 As used herein, the term "high affinity" refers to K_D of 1 to 10nM.

 As used herein, the term "very high affinity" refers to K_D of 1 nM or less.

30 The term "EC50," as used herein, refers to the concentration of an antibody or an antigen-binding protein/portion thereof, which induces a response, either in an *in vivo* or an *in vitro* assay, which is 50% of the maximal response (i.e., halfway between the maximal response and the baseline).

The term "multispecific" or "multi-specific antibody" as employed herein refers to an antibody as described herein which has at least two binding domains, i.e. two or more binding domains, for example two or three binding domains, wherein the at least two binding domains independently bind two different antigens or two different epitopes on the same antigen. Multi-specific antibodies are generally monovalent for each specificity (antigen). Multi-specific antibodies described herein encompass monovalent and multivalent, e.g. bivalent, trivalent, tetravalent multi-specific antibodies.

The term "bispecific" or "bispecific antibody" as employed herein refers to an antibody with two antigen specificities or an antibody that has the ability to simultaneously bind to two target antigens/sites.

As used herein, a "bispecific T cell engager" (BiTE) refers to a class of artificial bispecific monoclonal antibodies that direct a host's immune system, such as the T cells' cytotoxic activity against target cells (such as cancer cells). BiTEs are fusion proteins consisting of two single-chain variable fragments (scFvs) of different antibodies, or amino acid sequences from four different genes, on a single peptide chain of about 55 kDa. One of the scFvs binds to an immune cell (such as a T cell via the CD3 receptor), and the other to target of interest (e.g., a tumor cell via a tumor specific molecule). Like other bispecific antibodies, BiTEs form a link between an immune cell (e.g., a T cell) and a target cell (such as a tumor cell). This causes the immune cell (e.g., T cell) to exert cytotoxic activity on tumor cells. For example, if the immune cell is a T cell, the T cell would exert cytotoxic activity by producing proteins like perforin and granzymes that enter tumor cells and initiate the cell's apoptosis.

The term "immune cell" refers to a type of specialized cell that plays a crucial role in the body's defense against infections and foreign substances. They are a part of the immune system, which is responsible for identifying and eliminating harmful pathogens, such as bacteria, viruses, and parasites, as well as abnormal or cancerous cells. As used herein, an 'immune cell' refers to any cell of the immune system, including but not limited to T-cells, helper T-cells, B-cells, natural killer (NK) cells, dendritic cells (DC), granulocytes (such as basophils, eosinophils, neutrophils), mast cells, monocytes, and macrophages.

As used herein, a "nanobody with a heavy chain only" refers to nanobody-based heavy chain antibody. A heavy-chain antibody is an antibody which consists of two heavy chains and lacks the two light chains usually found in antibodies.

As described herein, a "vector" is any molecule or composition that has the ability to carry a nucleic acid sequence into a suitable host cell where e.g., synthesis of the encoded polypeptide can take place. Typically, and preferably, a vector is a nucleic acid that has been engineered, using recombinant DNA techniques that are known in the art, to incorporate a desired nucleic acid sequence (e.g., a nucleic acid of the present disclosure). Expression vectors typically contain one or more of the following components (if they are not already provided by the nucleic acid molecules): a promoter, one or more enhancer sequences, an origin of replication, a transcriptional termination sequence, a complete intron sequence containing a donor and acceptor splice site, a leader sequence for secretion, a ribosome binding site, a polyadenylation sequence, a polylinker region for inserting the nucleic acid encoding the polypeptide to be expressed, and a selectable marker element.

Vectors are typically selected to be functional in the host cell in which the vector will be used (the vector is compatible with the host cell machinery such that amplification of the gene and/or expression of the gene can occur. The vector as described herein may be an expression vector and/or a cloning vector.

The term "host cell," as used herein, is intended to refer to a cell into which an expression vector has been introduced. It should be understood that such terms are intended to refer not only to the particular subject cell but to the progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term "host cell" as used herein.

The terms "treating", "treat" and "therapy", and synonyms thereof refer to both therapeutic treatment and prophylactic or preventative measures, wherein the object is to prevent or slow down (lessen) a medical condition, which includes but is not limited to diseases, symptoms and disorders. A medical condition also includes a body's response to a disease or disorder, e.g. inflammation. Those in need of such treatment include those already with a medical condition as well as those prone to getting the medical condition or those in whom a medical condition is to be prevented.

The term "subject" as used herein includes patients and non-patients. The term "patient" refers to individuals suffering or are likely to suffer from a medical condition, while "non-patients" refer to individuals not suffering and are likely to not suffer from the medical condition. "Non-patients" include healthy individuals, non-diseased individuals

and/or an individual free from the medical condition. The term "subject" includes humans and animals. Animals may include, but is not limited to, mammals (for example non-human primates, canine, murine and the like), and the like. "Murine" refers to any mammal from the family Muridae and / or Leporidae, such as mouse, rat, rabbit, and the like.

The term "preventing" and/or "reducing the severity of symptoms" as used herein refers to process of delaying the onset, reducing the severity of symptoms, reducing and/or preventing weight loss, preventing death, inhibiting deterioration, inhibiting further deterioration, and/or ameliorating at least one sign or symptom of a disease.

The term "and/or", e.g., "X and/or Y" is understood to mean either "X and Y" or "X or Y" and should be taken to provide explicit support for both meanings or for either meaning.

Further, in the description herein, the word "substantially" whenever used is understood to include, but not restricted to, "entirely" or "completely" and the like. In addition, terms such as "comprising", "comprise", and the like whenever used, are intended to be non-restricting descriptive language in that they broadly include elements/components recited after such terms, in addition to other components not explicitly recited. For example, when "comprising" is used, reference to a "one" feature is also intended to be a reference to "at least one" of that feature. Terms such as "consisting", "consist", and the like, may in the appropriate context, be considered as a subset of terms such as "comprising", "comprise", and the like. Therefore, in embodiments disclosed herein using the terms such as "comprising", "comprise", and the like, it will be appreciated that these embodiments provide teaching for corresponding embodiments using terms such as "consisting", "consist", and the like. Further, terms such as "about", "approximately" and the like whenever used, typically means a reasonable variation, for example a variation of $\pm 5\%$ of the disclosed value, or a variance of 4% of the disclosed value, or a variance of 3% of the disclosed value, a variance of 2% of the disclosed value or a variance of 1% of the disclosed value.

Furthermore, in the description herein, certain values may be disclosed in a range. The values showing the end points of a range are intended to illustrate a preferred range. Whenever a range has been described, it is intended that the range covers and teaches all possible sub-ranges as well as individual numerical values within that range. That is, the end points of a range should not be interpreted as inflexible limitations. For

example, a description of a range of 1% to 5% is intended to have specifically disclosed sub-ranges 1% to 2%, 1% to 3%, 1% to 4%, 2% to 3% etc., as well as individually, values within that range such as 1%, 2%, 3%, 4% and 5%. It is to be appreciated that the individual numerical values within the range also include integers, fractions and decimals. Furthermore, whenever a range has been described, it is also intended that the range covers and teaches values of up to 2 additional decimal places or significant figures (where appropriate) from the shown numerical end points. For example, a description of a range of 1% to 5% is intended to have specifically disclosed the ranges 1.00% to 5.00% and also 1.0% to 5.0% and all their intermediate values (such as 1.01%, 1.02% ... 4.98%, 4.99%, 5.00% and 1.1%, 1.2% ... 4.8%, 4.9%, 5.0% etc.) spanning the ranges. The intention of the above specific disclosure is applicable to any depth/breadth of a range.

"At least 95% identical" as employed herein is intended to refer to an amino acid sequence which over its full length is 95% identical or more to a reference sequence, such as 96, 97, 98 or 99% identical. Software programmes can be employed to calculate percentage identity.

Additionally, when describing some embodiments, the disclosure may have disclosed a method and/or process as a particular sequence of steps. However, unless otherwise required, it will be appreciated that the method or process should not be limited to the particular sequence of steps disclosed. Other sequences of steps may be possible. The particular order of the steps disclosed herein should not be construed as undue limitations. Unless otherwise required, a method and/or process disclosed herein should not be limited to the steps being carried out in the order written. The sequence of steps may be varied and still remain within the scope of the disclosure.

Furthermore, it will be appreciated that while the present disclosure provides embodiments having one or more of the features/characteristics discussed herein, one or more of these features/characteristics may also be disclaimed in other alternative embodiments and the present disclosure provides support for such disclaimers and these associated alternative embodiments.

DESCRIPTION OF EMBODIMENTS

It will be appreciated by a person skilled in the art that other variations and/or modifications may be made to the embodiments disclosed herein without departing

from the spirit or scope of the disclosure as broadly described. For example, in the description herein, features of different exemplary embodiments may be mixed, combined, interchanged, incorporated, adopted, modified, included etc. or the like across different exemplary embodiments. The present embodiments are, therefore, to be considered in all respects to be illustrative and not restrictive.

In one aspect, there is provided an antigen binding protein, an antigen binding variant, and/or an antigen binding fragment thereof that binds specifically to Epithelial Cellular Adhesion Molecule (EpCAM).

In one embodiment, the antigen binding protein, variant or fragment thereof is a monoclonal antibody.

In one embodiment, the antigen binding protein or variant thereof is a full length antibody.

In one embodiment, the antigen binding fragment thereof is selected from the group comprising: an Fab, a modified Fab, an Fab', a modified Fab', an F(ab')₂, an Fv, a single domain antibody, a VHH, an scFv, an Fv, an bivalent antibody, a trivalent antibody, a tetra-valent antibody, a Bis-scFv, a diabody, a triabody, a tetrabody, an epitope-binding fragment, and the like.

In one embodiment, the antigen binding protein, variant or fragment thereof is an IgG, IgA, IgD, IgE, or IgM antibody.

In one embodiment, the antigen binding protein, variant or fragment thereof is a chimeric antibody, a human antibody / humanized antibody, recombinant humanized antibody, an animal-derived antibody (such as llama antibody) or the like.

In one embodiment, the antigen binding protein, variant or fragment thereof is a human antibody. In one embodiment, the antigen binding protein, variant or fragment thereof is a llama antibody. In one embodiment, the antigen binding protein, variant or fragment thereof is a humanized llama antibody.

In one embodiment, the antigen binding protein, variant or fragment thereof comprises one or more amino acid sequences selected from the group consisting of SEQ ID NOs: 1 to 37.

In one embodiment, the antigen binding protein, variant or fragment thereof is encoded by one or more nucleic acid sequences selected from the group consisting of SEQ ID NOs: 38 to 74.

In one embodiment, the antigen binding protein, variant or fragment thereof comprises a heavy chain variable region and a light chain variable region encoded by nucleic acid sequences selected from the group consisting of:

- (i) a heavy chain variable region comprising: (1B6 and 1C1)
- 5 a CDR-H1 comprising GGAGGCACCTTCAGCAGCTATGCT (**SEQ ID NO: 38**),
 a CDR-H2 comprising ATCATCCCTATCTTTGGTACAGCA (**SEQ ID NO: 39**), and
 a CDR-H3 comprising GCGAGATCGTTGGGTGGGAGATTTCGCTAC
10 (**SEQ ID NO: 40**); and
 a light chain variable region comprising:
 a CDR-L1 comprising
 CAGAGCCTCCTGCATAGTAATGGATACAACTAT (**SEQ ID NO: 46**),
 a CDR-L2 comprising TTGGGTCT (**SEQ ID NO: 47**), and
15 a CDR-L3 comprising ATGCAAGCTCTACAACTCCGTACACT (**SEQ ID NO: 48**)
- (ii) a heavy chain variable region comprising: (1C11)
- a CDR-H1 comprising GGAGGCACCTTCAGCAGCTATGCT (**SEQ ID NO: 38**),
20 a CDR-H2 comprising ATCATCCCTATCTTTGGTACAGCA (**SEQ ID NO: 39**), and
 a CDR-H3 comprising GCGAGATCGTTGGGTGGGAGATTTCGCTAC
 (**SEQ ID NO: 40**); and
 a light chain variable region comprising:
25 a CDR-L1 comprising
 CAGAGCCTCCTGCATAGTAATGGATACAACTAT (**SEQ ID NO: 46**),
 a CDR-L2 comprising TTGGGTCT (**SEQ ID NO: 47**), and
 a CDR-L3 comprising ATGCAAGGTCTACAAAGTCCCTGGACG (**SEQ ID NO: 52**)
- 30 (iii) a heavy chain variable region comprising: (1D4)
 a CDR-H1 comprising GGAGGCACCTTCAGCAGCTATGCT (**SEQ ID NO: 38**),

- a CDR-H2 comprising ATCATCCCTATCTTTGGTACAGCA (**SEQ ID NO: 39**), and
a CDR-H3 comprising GCGAGATCGTTGGGTGGGAGATTTCGCTAC (**SEQ ID NO: 40**); and
- 5 a light chain variable region comprising:
a CDR-L1 comprising
CAGAGCCTCCTGCATAGTAATAGATACTAT (**SEQ ID NO: 54**),
a CDR-L2 comprising TTGGGTCT (**SEQ ID NO: 47**), and
a CDR-L3 comprising ATGCAAGCTCTACAACTCCGTACACT (**SEQ**
10 **ID NO: 48**)
- (iv) a heavy chain variable region comprising: (1H6)
a CDR-H1 comprising GGAGGCACCTTCAGCAGCTATGCT (**SEQ ID NO: 38**),
a CDR-H2 comprising ATCATCCCTATCTTTGGTACAGCA (**SEQ ID**
15 **NO: 39**), and
a CDR-H3 comprising GCGAGATCGTTGGGTGGGAGATTTCGCTAC (**SEQ ID NO: 40**); and
a light chain variable region comprising:
a CDR-L1 comprising
20 CAGAGCCTCCTGCATAGTAATGGATACTAT (**SEQ ID NO: 46**),
a CDR-L2 comprising TTGGGTCT (**SEQ ID NO: 47**), and
a CDR-L3 comprising ATGCAAGGTCTACAGACTCCGTACACT (**SEQ ID NO: 60**); and
- (v) a heavy chain variable region comprising: (1E4)
25 a CDR-H1 comprising GGGGACAGTATCTCTAGTAACAGTGTGCT (**SEQ ID NO: 42**),
a CDR-H2 comprising ACATACTACAGGTCCAAGTGGTACAGT (**SEQ ID NO: 43**), and
a CDR-H3 comprising
30 GCAAGAGAAGTTGAGGGCAGCAGCTATGATGCTTTTGATATC (**SEQ ID NO: 44**); and
a light chain variable region comprising: (1E4)
a CDR-L1 comprising CAGAGTATTAGCGACTTT (**SEQ ID NO: 56**),

a CDR-L2 comprising GCTGCATCG (**SEQ ID NO: 57**), and
a CDR-L3 comprising TTACATTATGCCCCGACACT (**SEQ ID NO: 58**)
or fragment or variation or sequence at least 80%, 85%, 90%, 95%, 96%, 97%,
98%, 99%, or 100% identity thereto.

5 In one embodiment, the antigen binding protein, variant or fragment thereof
comprises a heavy chain variable region and a light chain variable region encoded by
nucleic acid sequences selected from the group consisting of:

(i) a heavy chain variable region comprising: (1A5-VHH, 2C4-VHH, 2D10-VHH
and hu2C4-VHH)

10 a CDR-H1 comprising GGAAGCATCTTCAGTGGCAATGAC (**SEQ ID**
NO: 62),

a CDR-H2 comprising ATTACTAGCGGTGGTAGTACA (**SEQ ID NO:**
63), and

15 a CDR-H3 comprising
ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCAC (**SEQ**
ID NO: 64)

(ii) a heavy chain variable region comprising: (1B8-VHH)

a CDR-H1 comprising GGAAGCTCCGAAAGATTCACATCA (**SEQ ID**
NO: 66),

20 a CDR-H2 comprising ATTACTAATGGTGGTAGCACA (**SEQ ID NO:**
67), and

a CDR-H3 comprising ATGGCGGGTACGTCC (**SEQ ID NO: 68**); and

(iii) a heavy chain variable region comprising: (2B7-VHH)

25 a CDR-H1 comprising GGAAGCATCTTCAGTGGCAATGAC (**SEQ ID**
NO: 62),

a CDR-H2 comprising ATTACTAGCGGTGGTAGTACA (**SEQ ID NO:**
63), and

30 a CDR-H3 comprising
ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCTC (**SEQ**
ID NO: 70)

or fragment or variation or sequence at least 80%, 85%, 90%, 95%, 96%, 97%,
98%, 99%, or 100% identity thereto.

In one embodiment, the antigen binding protein, variant or fragment thereof comprises a heavy chain variable domain encoded by the nucleotide sequence comprising

- (i) GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGGGTCCT
 5 CGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAGCAGCTATGC**
TATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCTTGAGTGGATGGGA
GGGATCATCCCTATCTTTGGTACAGCAACTACGCACAGAACTCCAGGG
 CAGAGTCACCATGACCGCAGACACCTCCATAAGCACAGCCTACATGGAGC
 TGAGCAGCCTGAGATCTGAGGACACGGCCGTGTATTACTGT**GCGAGATCG**
 10 **TTGGGTGGGAGATTTGCTACT**GGGGCCAGGGAACCCTG (**SEQ ID NO: 41**
 - for clones 1B6, 1C1, 1C11, 1D4, 1H6), or
- (ii) CAGGTACAGCTGCAGCAGTCAGGTCCAGGGCTGGTGAAGCCCTCGCAGA
 CCCTCTCACTCACCTGTGCCATCTCC**GGGGACAGTATCTCTAGTAACAGT**
GTTGCTTGGAAGTGGATCAGGCAGTCCCCATCGAGAGGCCTTGAGTGGCT
 15 GGGAAAG**ACATACTACAGGTCCAAGTGGTACAGT**GATTATGCAATATCTG
 TGAAAGGTCGATTAGACATCAACCCAGACACATCCAAGAACCAGTTCTCCC
 TGCAGCTGAAGTCTGTGACTCCCGAGGACACGGCTGTGTATTATTGT**GCA**
AGAGAAGTTGAGGGCAGCAGCTATGATGCTTTTGATATCTGGGGCCAAG
 GGACAATG (**SEQ ID NO: 45** - clone 1E4),
 20 or a sequence at least 60% identical thereto and/or having 10-20 nucleic acid
 substitutions.

In one embodiment, the antigen binding protein, variant or fragment thereof comprises a light chain variable domain encoded by the nucleotide sequence comprising

- (i) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTGGAGA
 25 GCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGCATAGTAATG**
GATACAACTATTTGGATTGGTACCTGCAGAAGCCAGGGCAGTCTCCACAG
 CTCCTGATCTATTT**GGGTTCT**AATCGGGCCTCCGGGGTCCCTGACAGGTT
 CAGTGGCAGTGGATCAGGCACAGATTTTACACTGAAAATCAGCAGAGTGG
 AGGCTGAGGATGTTGGGGTTTATTACTGC**ATGCAAGCTCTACAACTCCG**
 30 **TACACTTTT**GGCCAGGGGACCAAG (**SEQ ID NO: 49** - clone 1B6 and 1C1), or
- (ii) GAAATTGTGCTGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTGGAGA
 GCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGCATAGTAATG**
GATACAACTATTTGGATTGGTACCTGCAGAAGCCAGGGCAGTCTCCACAG

- CTCCTGATCTATTTGGGTTCTAATCGGGCCTCCGGGGTCCCTGACAGGTT
 CAGTGGCAGTGGATCAGGTACAGATTTTACACTGAAAATAAGCAGAGTGG
 AGGCTGAGGATGTTGGGGTTTATTACTGCATGCAAGGTCTACAAAGTCCC
TGGACGTTCGGCCAAGGGACCAAG (**SEQ ID NO: 53** - clone 1C11), or
- 5 (iii) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTGGAGA
 GTCGGCCTCCATCTCCTGCAGGTCTAGTCAGAGCCTCCTGCATAGTAATA
GATACAACTATTTGGATTGGTACCTGCAGAAGCCAGGGCAGTCTCCACAG
 CTCCTGATCTATTTGGGTTCTAATCGGGCCTCCGGGGTCCCTGACAGGTT
 CAGTGGCAGTGGATCAGGCACAGATTTTACACTGAAAATCAGCAGAGTGG
 10 AGGCTGAGGATGTTGGGGTTTATTACTGCATGCAAGCTCTACAAACTCCG
TACACTTTTGGCCAGGGGACCAAG (**SEQ ID NO: 55** - clone 1D4), or
- (iv) GACATCCAGTTGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAGGAGA
 CAGAGTCACCATCACTTGCCGGGCAAGTCAGAGTATTAGCGACTTTTAA
 ATGGTACCAGCAGAAACCAGGTAAAGCCCCGAAGCTCCTGATCTATGCT
 15 **GCATCG**AGTTTACAACTGGGGTCCCCTCAAGATTCGGTGGCAGTGGATC
 TGGGACAGAATTCACCTCTCACCATAAGCAGTCTACAACTGAAGATTTGGG
 AACTTATTACTGTCAACAGAGTTACATTATGCCCGACACTTTTGGCCAGG
 GGACGAAA (**SEQ ID NO: 59** - clone 1E4), or
- (v) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTGGAGA
 20 GCCGGCCTCCATCTCCTGCAGGTCTAGTCAGAGCCTCCTGCATAGTAATG
GATACAACTATTTGGATTGGTACCTGCAGAAGCCAGGGCAGTCTCCACAG
 CTCCTGATCTATTTGGGTTCTAATCGGGCCTCCGGGGTCCCTGACAGGTT
 CAGTGGCAGTGGATCAGGCACAGATTTTACACTGCAAATCAGCAGAGTGG
 AGGCTGAGGATGCTGGGGTTTATTACTGCATGCAAGGTCTACAGACTCCG
 25 **TACACT**TTTGGCCAGGGGACCAAG (**SEQ ID NO: 61** - clone 1H6),
 or a sequence at least 60% identical thereto and/or having 10-20 nucleic acid
 substitutions.
- In one embodiment, the antigen binding protein, variant or fragment thereof
 comprises a single domain heavy chain variable domain encoded by the nucleotide
 30 sequence comprising:
- (i) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGAGGGT
 CTCTGAGACTCTCCTGTGCAGACTCTGGAAGCATCTTCAGTGGCAATGAC
 ATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCAGTTGGTCGCG

- GTTATTACTAGCGGTGGTAGTACACACTATGCAGACTCCGTGAAGGGCCG
ATTCACCATCTCCAGAGACAATGCCAGAAGACCGTATATCTGCAAACGAA
CGACTTGAAACCTGAGGACACGGCCGTGTATTACTGC**ACAAACGGAAGAT**
GGTCAGGCGATACTTACTATGCCCATCACTGGGGCCAGGGGACCCAG
5 (SEQ ID NO: 65 - clone 1A5),
- (ii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGGGGG
TCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCTCCGAAAGATTACATC**
AGTGGCCTGGTACCGCCAGGCTCCAGGAAAGGAGCGCGAGTTGGTCGCA
TTTATTACTAATGGTGGTAGCACAAGATATACAGACCCCGTGAAGGGCCG
10 ATTCACCATCTCCAGAGACAACGCCAAGAACACGGTGTATCTGCAAATGAA
CAGCCTGAAAGCTGAGGACACGGCCGTCTATTATTGT**ATGGCGGGTACGT**
CCTGGGGCCAGGGGACCCAG (SEQ ID NO: 69 - clone 1B8),
- (iii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGAGGGT
CTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCAATGAC**
15 ATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTTGGTCGCG
GTTATTACTAGCGGTGGTAGTACACACTATGCAGACTCCGTGAAGGGCCG
ATTCACCATCTCCAGAGACAATGCCAGAAGACCGTATATCTGCAAACGAA
CGACCTGAAACCTGAGGACACGGCCGTGTATTACTGC**ACAAACGGAAGA**
TGGTCAGGCGATACTTACTATGCCCATCTCTGGGGCCAGGGGACCCAG
20 (SEQ ID NO: 71 - clone 2B7),
- (iv) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGGGGGC
TCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCAATGA**
CATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTTGGTCGCG
GGTTATTACTAGCGGTGGTAGTACACACTATGCAGACTCCGTGAAGGGCCG
25 GATTACCATCTCCAGAGACAATGCCAGAAGACCGTATATCTGCAAACGA
ACGACCTGAAACCTGAGGACACGGCCGTGTATTACTGC**ACAAACGGAAG**
ATGGTCAGGCGATACTTACTATGCCCATCACTGGGGCCAGGGGACCCAG
(SEQ ID NO: 72 - clone 2C4),
- (v) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGGGGACT
30 CTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCAATGAC**
ATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTTGGTCGCG
GTTATTACTAGCGGTGGTAGTACACACTATGCAGACTCCGTGAAGGGCCG
ATTCACCATCTCCAGAGACAATGCCAGAAGACCGTATATCTGCAAACGAA

CGACCTGAAACCTGAGGACACGGCCGTGTATTACTGC**ACAAACGGAAGA**
TGGTCAGGCGATACTTACTATGCCATCACTGGGGCCAGGGGACCCAG
 (SEQ ID NO: 73 - clone 2D10), or

(vi) CAGGTGCAGCTGGTGGAGTCTGGGGGAGGATTGGTGCAGGCTGGGGGC
 5 TCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCATCTTCAGTGGCAATGA**
CATGTCCTGGTACCGCCAGGCTCCAGGGAAGGGACTCGAGTTGGTCGCG
 GTT**ATTACTAGCGGTGGTAGTAC**ATACTATGCAGACTCCGTGAAGGGCCG
 ATTCACCATCTCCAGAGACAATTCCaAGAAcACCcTATATCTGCAAATGAAC
 AGCCTGAGAGCTGAGGACACGGCCGTGTATTACTGC**ACAAACGGAAGAT**
 10 **GGTCAGGCGATACTTACTATGCCATCACT**GGGGCCAGGGGACCCCTG
 (SEQ ID NO: 74 - clone hu2C4-VHH)

or a sequence at least 60% identical thereto and/or having 10-20 nucleic acid substitutions.

In one embodiment, the antigen binding protein, variant or fragment thereof
 15 comprises a light chain constant domain encoded by the nucleotide sequence comprising

CGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCA
 GTTGAAATCTGGAAGTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCC
 CAGAGAGGCCAAAGTACAGTGAAGGTGGATAACGCCCTCCAATCGGGTA
 20 ACTCCAGGAGAGTGTGCGAGAGCAGGACAGCAAGGACAGCACCTACAG
 CCTCAGCAGCACCCCTGACGCTGAGCAAAGCAGACTACGAGAAACACAAAC
 TCTACGCCTGCGAAGTCACCCATCAGGGCCTGAGCTCGCCCGTCACAAAG
 AGCTTCAACAGGGGAGAGTGT (SEQ ID NO: 50 - clone 1B6 – light chain
 constant domain), or

CGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCA
 25 GTTGAAATCTGGAAGTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCC
 CAGAGAGGCCAAAGTACAGTGAAGGTGGATAACGCCCTCCAATCGGGTA
 ACTCCAGGAGAGTGTACAGAGCAGGACAGCAAGGACAGCACCTACAG
 CCTCAGCAGCACCCCTGACGCTGAGCAAAGCAGACTACGAGAAACACAAAG
 30 TCTACGCCTGCGAAGTCACCCATCAGGGCCTGAGCTCGCCCGTCACAAAG
 AGCTTCAGCAGGGGAGAGTGT (SEQ ID NO: 51 - clone 1C1 – light chain
 constant domain), or

a sequence at least 60% identical thereto and/or having 10-20 nucleic acid substitutions.

In some embodiments, an "antibody variant" refers to an antibody or antigen-binding fragment thereof comprising a VH and/or a VL wherein the non-CDR regions of the antibody or antigen-binding fragment thereof has at least about 85%, 90%, 95%, 96%, 97%, 98%, or 99% amino acid sequence identity, to an antibody described herein.

In one embodiment, the antigen binding protein is an antibody. In some embodiments, the antigen binding protein may include an isolated antibody.

In some embodiments, the antigen binding protein comprises a sequence that is at least 60% identical to any one of the sequences disclosed herein. For example, the antigen binding protein may comprise a sequence that is at least about 60%, at least about 61%, at least about 62%, at least about 63%, at least about 64%, at least about 65%, at least about 66%, at least about 67%, at least about 68%, at least about 69%, at least about 70%, at least about 71%, at least about 72%, at least about 73%, at least about 74%, at least about 75%, at least about 76%, at least about 77%, at least about 78%, at least about 79%, at least about 80%, at least about 81%, at least about 82%, at least about 83%, at least about 84%, at least about 85%, at least about 86%, at least about 87%, at least about 88%, at least about 89%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or 100% identical to any of the sequences disclosed herein (e.g. about 60%, or about 61%, or about 62%, or about 63%, or about 64%, or about 65%, or about 66%, or about 67%, or about 68%, or about 69%, or about 70%, or about 71%, or about 72%, or about 73%, or about 74%, or about 75%, or about 76%, or about 77%, or about 78%, or about 79%, or about 80%, or about 81%, or about 82%, or about 83%, or about 84%, or about 85%, or about 86%, or about 87%, or about 88%, or about 89%, or about 90%, or about 91%, or about 92%, or about 93%, or about 94%, or about 95%, or about 96%, or about 97%, or about 98%, about 99% or about 100% sequence identity to any one of the sequences disclosed herein. The percentage sequence identity can be determined using methods known in the art, for example using programs such as BLAST, NBLAST, XBLAST, and the like.

In some embodiments, the antigen binding protein comprises a sequence or an amino acid region or is encoded by a nucleotide region that differs by about one, about

two, about three, about four, about five, about six, about seven, about eight, about nine, about ten or more amino acids or nucleobase with the sequence as disclosed herein.

In some embodiments, the antigen binding protein comprises an amino acid sequence having one or more amino acid mutations with respect to any one of the sequences disclosed herein. In some examples, the antigen binding protein comprises an amino acid sequence having one, or two, or three, or four, or five, or six, or seven, or eight, or nine, or ten, or fifteen, or twenty amino acid mutations with respect to any one of the sequences disclosed herein. In some examples, the one or more amino acid mutations may be independently selected from substitutions, insertions, deletions, and truncations.

In some embodiments, the amino acid mutations are amino acid substitutions, and may include conservative and/or non-conservative substitutions.

In various embodiments, the substitution replaces the amino acids with natural amino acids, such as, but is not limited to, alanine, arginine, asparagine, aspartate, cysteine, glutamine, glutamate, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, and valine. In various embodiments, the substitution does not include replacement with cysteine. Thus, in various embodiments, the substitution replaces the amino acids with natural amino acids, such as, but is not limited to, alanine, arginine, asparagine, aspartate, glutamine, glutamate, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine, and valine.

In various embodiments, the substitution may be a conservative substitution that substitute one for another of similar properties. For example, substitution of one amino acid with another from the same group. In various embodiments, the substitution may include substitution with an amino acid with different properties. In various embodiments, the substitution is made without affecting the biological activity of the antigen binding protein, variant or fragment thereof as described herein. In various embodiments, the substitution increases binding affinity to EpCAM.

In some embodiments, the substitutions may also include non-classical amino acids. Illustrative non-classical amino acids include, but are not limited to, selenocysteine, pyrrolysine, N-formylmethionine β -alanine, GABA and δ -Aminolevulinic acid, 4-aminobenzoic acid (PABA), D-isomers of the common amino acids, 2,4-diaminobutyric acid, α -amino isobutyric acid, 4-aminobutyric acid, Abu, 2-amino butyric

acid, γ -Abu, ϵ -Ahx, 6-amino hexanoic acid, Aib, 2-amino isobutyric acid, 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, homocitrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine, β -alanine, fluoro-amino acids, designer amino acids such as β methyl amino acids, C α -methyl amino acids, N α -methyl amino acids, and amino acid analogs in general.

In some embodiments, the amino acid mutation may be in the CDRs of the antigen binding protein (e.g., the CDR1, CDR2 or CDR3 regions). In another example, amino acid alteration may be in the framework regions (FRs) of the antigen binding protein (e.g., the FR1, FR2, FR3, or FR4 regions).

Modification of the amino acid sequences may be achieved using any known technique in the art e.g., site-directed mutagenesis or PCR based mutagenesis.

In some embodiments, the mutations do not substantially reduce the antigen binding protein's capability to specifically bind to a target. In some examples, the mutations do not substantially reduce the antigen binding protein's capability to specifically bind to a target and without functionally modulating (e.g., partially or fully neutralizing) the target.

In one embodiment, the antigen binding protein, variant or fragment thereof binds to EpCAM with low affinity, moderate affinity or high affinity.

In one embodiment, the antigen binding protein, variant or fragment thereof binds to EpCAM with low to moderate affinity (K_D ranging from 1-100nM).

In one embodiment, the antigen binding protein, variant or fragment thereof binds with moderate affinity K_D ranging from 10-100nM.

In one embodiment, the antigen binding protein, variant or fragment thereof binds to EpCAM with low affinity.

In some embodiments, the binding affinity of the antigen binding protein of the disclosure for the full-length and/or mature forms and/or isoforms and/or splice variants and/or fragments and/or monomeric and/or dimeric forms and/or any other naturally occurring or synthetic analogs, variants, or mutants (including monomeric and/or dimer forms) of antigen binding protein may be described by the equilibrium dissociation constant (K_D). In some examples, the antigen binding protein binds to the full-length and/or mature forms and/or isoforms and/or splice variants and/or fragments and/or any other naturally occurring or synthetic analogs, variants, or mutants (including

monomeric and/or dimeric forms) of antigen binding protein with a K_D of less than about 1 μ M, about 900 nM, about 800 nM, about 700 nM, about 600 nM, about 500 nM, about 400 nM, about 300 nM, about 200 nM, about 100 nM, about 90 nM, about 80 nM, about 70 nM, about 60 nM, about 50 nM, about 40 nM, about 30 nM, about 20 nM, about 10 nM, or about 5 nM, or about 1 nM, about 50 pM, about 40 pM, about 30 pM, about 20 pM, about 10 pM, about 5 pM or about 1 pM.

The binding affinities of the presently disclosed antigen binding proteins/fragments thereof were measured via Bio-layer Interferometry (BLI) analysis (results shown in Example 4). Advantageously, all the clones tested exhibited binding affinities to EpCAM at nanomolar scales. Surprisingly, compared to the 2 high affinity clinically tested prior art anti-EpCAM antibodies 3622W94 and ING-1, the K_{off} rates of which were 1.8×10^{-5} , and 3.2×10^{-5} respectively, the presently disclosed anti-EpCAM IgG1 antibodies had a much slower off-rates (6.134×10^{-4} to 3.278×10^{-3}) and resulted in much lower binding affinities than 3622W94 and ING-1.

In some embodiments, the antigen binding protein may be a nanobody.

In some examples, the multi-specific antigen binding protein, variant or fragment thereof is a bispecific antibody. In some examples, the multi-specific antigen binding protein may be provided as a nanobody. In some examples, the multi-specific antigen binding protein may be provided as a Fc region.

In some examples, the second antigen binding protein, variant or binding fragment thereof of the immune cell engager binds to the immune marker selected from the group consisting of CD3, NKG2D, CD4, CD8, CD16, and CD64.

In some examples, the multi-specific antigen binding protein, variant or fragment thereof is an inducible bispecific T cell engager. In some examples, the multi-specific antigen binding protein is secreted by the immune cell.

In some examples, the multispecific antigen binding protein is an inducible bispecific T cell engager comprising a Heavy chain antibody variable region (i.e. VHH) and/or a single chain variable fragment (scFv).

In some embodiments, the bispecific T cell engager (BiTE) is a nanobody with a heavy chain only (VHH).

In one embodiment, the bispecific T cell engager (BiTE) binds to two antigens; wherein the first antigen is EpCAM; and wherein the second antigen is an immune cell marker.

In some embodiments, the second antigen is an immune cell marker that is involved in the activation of the immune cell.

In some embodiments, the antigen binding protein binds to EpCAM in solution and / or on the surface of the host cells.

5 In one embodiment, the second antigen targeted by the bispecific T cell engager (BiTE) is selected from the group comprising CD3, NKG2D, CD28, CD16, and the like.

In one embodiment, the antigen binding protein, variant or fragment thereof binds the bispecific T cell engager (BiTE) targets one or more antigens;

wherein the first antigen is EpCAM; and

10 wherein the second antigen is CD3.

In some embodiments, the CD3 may include CD3 ζ , CD3 ϵ , CD3 γ , CD3 δ , and the like.

As shown in the experimental data in FIG. 3A, the Fab region of one arm from each of the six anti-EpCAM human antibody clones (clones 1B6, 1C1, 1C11, 1D4, 1E4, 15 1H6 disclosed herein) was replaced by an anti-CD3 scFv fragment, and knob-in-holes mutations were introduced to facilitate correct heavy-chain pairing and LALA mutations to remove Fc γ receptor binding.

In one embodiment, the antigen binding protein, variant or fragment thereof is a human anti-EpCAM BiTE (for example clones 1B6, 1C1, 1C11, 1D4, 1E4, 1H6 as 20 disclosed herein).

Advantageously, as shown in the experimental data of the present disclosure, compared to the benchmark BiTE (anti-EpCAM arm of "MT110" with scFv fragment of anti-CD3 clone Okt3) that show binding EC50 of 0.666 nM, BiTE antibodies using the six presently disclosed anti-EpCAM clones (clone 1B6, 1C1, 1C11, 1D4, 1E4 and 1H6) 25 showed much lower absorbance reading at 450nm and higher EC50 values, suggesting lower binding affinities to the antigen (FIG. 3B). All six antibodies bound to EpCAM expressing cell HT-29 (FIG. 3C) and human T cells expressing CD3 (FIG. 3E), with binding EC50 to HT-29 cells being 5-20 times lower than that of the benchmark BiTE (MT110, Solitomab) which is 4.106 pM (FIG. 3G). Unlike the benchmark BiTE which 30 showed 10-20% of non-specific binding to EpCAM negative HeyA8 cells at high concentrations of 0.5, 2, 10 nM (FIG. 3F), the six antibodies did not bind to HeyA8 cells at all concentrations (FIG. 3D) suggesting that these anti-EpCAM BiTE could bind to cell surface antigen in a highly specific manner. All anti-EpCAM BiTE antibodies (clone 1B6,

1C1, 1C11, 1D4 and 1H6) were able to kill 70-100% of HT-29 cells within 72 hours at a concentration as low as 0.1 nM (FIG. 4B). 1C1 and 1H6 BiTE antibodies showed the lowest EC50 values of 18.5 pM and 12.02 pM, respectively; while 1B6, 1C11, 1D4, 1E4 showed low to moderate cell killing efficacy (FIG. 4C).

5 In some embodiments, the anti-EpCAM antibody is a nanobody with a heavy chain only (VHH). In some embodiments, the nanobody with heavy chain only includes an anti-EpCAM BiTE (such as 1A5-VHH, 1B8-VHH, 2B7-VHH, 2C4-VHH, 2D10-VHH).

10 In some embodiments, the anti-EpCAM BiTE may comprise single domain antibody fragments (VHH) against EpCAM that may be fused to another antigen binding protein (such as anti-CD3 T cell agonist antibodies) via a flexible linker.

As shown in the experimental data of the present disclosure, the five disclosed llama anti-EpCAM BiTE clones (1A5-VHH, 1B8-VHH, 2B7-VHH, 2C4-VHH and 2D10-VHH) show comparable or even lower EC50 values as the benchmark known in the art anti-EpCAM BiTE "MT110" in binding to both mobilized biotinylated EpCAM antigen (FIG. 6B) and EpCAM expressing cells (FIG. 6F), suggesting high binding affinities to the antigen. The four clones (1A5-VHH, 2B7-VHH, 2C4-VHH and 2D10-VHH) all bound strongly to EpCAM expressing cell HT-29 as well as human T cells (CD3 positive) (FIG. 6E). The binding EC50 to HT-29 cells range from 10-20 pM and binding specificity was superior as none of them showed binding to EpCAM negative, CD negative HeyA8 cells even at the highest concentration of 10nM (FIG. 6D). The four anti-EpCAM VHH BiTE antibodies (1A5-VHH, 2B7-VHH, 2C4-VHH and 2D10-VHH) were able to kill 90-100% of EpCAM positive HT-29 (FIG. 7A), HepG2 (FIG. 7B), and Hep3B (FIG. 7C) within 96 hours at concentrations as low as 10 pM but spared EpCAM negative cell HeyA8.

25 Without wishing to be bound by theory, the antigen binding fragments of VHH antibodies are advantageously much smaller and have a higher tissue penetration than conventional antibodies, and possess superior properties including high solubility, stability and resistance to heat-denaturation.

In one embodiment, the antigen binding protein, variant or fragment thereof is a humanized anti-EpCAM BiTE (e.g. hu2C4-VHH as disclosed herein).

30 As shown in the experimental results of the present disclosure, humanized clone hu2C4-VHH was constructed by pairing EpCAM BiTE antibodies with well-characterized anti-CD3 Okt3. Humanization resulted in 10 times reduction in binding affinity to the antigen protein (FIG. 8A) and about 2.5 times weaker cell binding (FIG. 8B and FIG. 8E).

The binding to the CD3 arm was not affected by the VHH arm humanization (FIG. 8D), while hu2C4-VHH/Okt3 showed high binding specificity as it didn't bind to EpCAM negative HeyA8 cells, similar to its parental clone (FIG. 8C). hu2C4-VHH/Okt3 EC50 still falls within the range of pM (FIG. 9A, FIG. 9C). Similar to the 2C4-VHH/Okt3 BiTE, hu2C4-VHH/Okt3 also did not kill EpCAM negative HeyA8 cells (FIG. 9B) nor did it induce interferon γ (FIG. 9D) or IL-2 (FIG. 9E) secretion upon co-culture as otherwise shown in the case of the benchmark BiTE (MT110).

In one embodiment, the antigen binding protein, variant or fragment thereof further comprises an immunoglobulin Fc fragment (region/domain), a protein capable of extending the half-life of the recombinant/fusion polypeptide (such as albumin, human albumin, and the like), linkers capable of enhancing binding valency (for example a rigid linker or a flexible linker, such as a GGGGS repeat), or combinations thereof, for example wherein the immunoglobulin Fc fragment is an IgG Fc fragment, such as a human IgG Fc fragment.

In various embodiments, the immunoglobulin fragment may include but is not limited to a Fc region/domain, one or more CH regions/domains, a Fab, a Fab', a F(ab')₂, a single chain Fv (ScFv) and/or Fv fragments, hinge regions/domain, as well as fragments or portions thereof.

In some embodiments, the antigen binding protein, variant or fragment thereof may also contain portions of immunoglobulin molecules or antibodies, such as including, but not limited to, all or portions of a constant heavy chain, a variable heavy chain, a constant light chain, a variable light chain, a hinge region, and/or an Fc domain of a Ig, as well as variants thereof. For example, a recombinant/fusion polypeptide as described herein may be combined with portions of a human IgG. In various embodiments, the type of immunoglobulin may include one or more type such as, but is not limited to, IgG, IgE, IgM, IgD, IgA, and IgY. In various embodiments, the type of immunoglobulin may be an immunoglobulin of class IgG1, IgG2, IgG3, IgG4, IgA1, IgA2, or subclass thereof.

In one embodiment, the immunoglobulin fragment is an IgG Fc fragment.

In another aspect, there is provided a polynucleotide encoding the immune cell and/or the polypeptide and/or multispecific antigen binding protein as described herein.

In some examples, the sequence encoding the immune cell engager encodes for a multi-specific antigen binding protein. In some examples, the multi-specific antigen binding protein is a bi-specific antibody. In some examples, the sequence encoding the

immune cell engager encodes for an antigen binding protein capable of binding EpCAM (anti-EpCAM antigen binding protein), or fragment, or variant thereof, and an anti-immune cell antigen binding protein.

In some examples, the sequence encoding the immune cell engager encodes for a single chain variable fragment (scFv) or a single variable domain located on a heavy chain (VHH).

In some examples, the sequence encoding the immune cell engager encodes for an anti-EpCAM scFv or an anti-EpCAM VHH.

In some examples, the sequence encoding the immune cell engager encodes for an anti-immune cell antigen binding protein that binds to an immune cell activation marker. In some examples, the immune cell activation marker may include but is not limited to CD3, NKG2D, CD4, CD8, CD16, CD64, and the like.

In some examples, the immune activation marker may be in VHH form or scFv form. In some examples, the sequence encoding the immune cell engager encodes for an antigen binding protein capable of binding CD3 (anti-CD3 antigen binding protein), or fragment, or variant thereof. In some examples, the sequence encoding the immune cell engager encodes for a single chain variable fragment (scFv) or a VHH form. In some examples, the sequence encoding the immune cell engager encodes for an anti-CD3 scFv or an anti-CD3 single domain VHH.

In some examples, the immune cell engager may include a His-tag.

In some examples, the immune cell engager comprises an anti-EpCAM antigen binding protein, a linker, an anti-CD3 scFv or an anti-CD3 single domain VHH, and a His-tag. In some examples, the linker is a cleavable linker, which may include but is not limited to, P2A, T2A, F2A, and the like.

In some examples, the polynucleotide comprises the sequence encoding an immune cell engager is a sequence encoding a bispecific T cell engager (BiTE). In some examples, the polynucleotide comprises the sequence encoding BiTE that bi-specifically binds to EpCAM and a T cell. In some examples, the polynucleotide comprises the sequence encoding for an anti-GPC3 scFv CAR with CD3 intracellular domain, and the sequence encoding for a BiTE that binds to EpCAM and a T cell.

In one embodiment, the polynucleotide comprises one or more nucleic acid sequences selected from the group consisting of SEQ ID NOs: 38 to 74.

In some embodiments, the polynucleotide sequence may be at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical to the sequences as disclosed herein.

In one embodiment, the vector is an expression vector.

5 In some embodiments, the vector is selected from the group consisting of a plasmid, a viral particle, a phage, a baculovirus, a yeast plasmid, a lipid based vehicle, a polymer microsphere, a liposome, and a cell based vehicle, a colloidal gold particle, lipopolysaccharide, polypeptide, polysaccharide, a viral vehicle, an adenovirus, a retrovirus, a lentivirus, an adeno-associated viruses, a herpesvirus, a vaccinia virus, a
10 foamy virus, a cytomegalovirus, a Semliki forest virus, a poxvirus, a pseudorabies virus, an RNA virus vector, a DNA virus vector and a vector derived from a combination of a plasmid and a phage DNA, further optionally wherein said polynucleotide is operatively linked to an expression control sequence(s) to direct peptide synthesis, even further optionally wherein the vector comprises one or more selectable marker genes to provide
15 a phenotypic trait for selection of transformed host cells.

In one embodiment, the vector is selected from the group comprising DNA vector, an RNA vector, a plasmid, a lentiviral vector, an adenoviral vector, an adeno-associated viral vector, and a retroviral vector.

In some embodiments, the vector is a lentiviral vector.

20 In some embodiments, the host cell comprises cloning or expression vectors as described above and/or nucleic acid sequences encoding for the antigen binding protein, antibodies and binding fragments thereof as described above.

The host cell can be any type of cell capable of being transformed or transfected with the nucleic acid or vector so as to produce an antigen binding protein or binding
25 fragment/protein thereof encoded thereby. The host cell comprising the nucleic acid or vector can be used to produce the antigen binding protein or binding fragment/protein thereof, or a portion thereof (e.g., a heavy chain sequence, or a light chain sequence encoded by the nucleic acid or vector). After introducing the nucleic acid or vector into the cell, the cell is cultured under conditions suitable for expression of the encoded
30 sequence. The antibody, antigen binding protein, or fragment, or portion of the antibody then can be isolated from the cell.

The host cells may be prokaryotic host cells (such as *E. coli*) or eukaryotic host cells (such as a yeast cell, an insect cell, or a vertebrate cell). The host cell, when

cultured under appropriate conditions, expresses an antibody or binding fragment thereof which can subsequently be collected from the culture medium (if the host cell secretes it into the medium) or directly from the host cell producing it (if it is not secreted). Selection of an appropriate host cell will depend upon various factors, such as desired expression levels, polypeptide modifications that are desirable or necessary for activity, such as glycosylation or phosphorylation, and ease of folding into a biologically active molecule. Selection of the host cell will depend in part on whether the antibody or binding fragment thereof is to be post-transcriptionally modified (e.g., glycosylated and/or phosphorylated). The host cell may comprise a bacterial cell, a yeast cell, an animal cell e.g., a mammalian cell and/or a plant cell.

Suitable mammalian host cells include CHO, myeloma or hybridoma cells. Many are available from the American Type Culture Collection (ATCC), Manassas, Va. Examples include mammalian cells, such as Chinese hamster ovary cells (CHO) (ATCC No. CCL61), human embryonic kidney (HEK) 293 or 293T cells (ATCC No. CRL1573), 3T3 cells (ATCC No. CCL92), or PER.C6 cells. Other cell types of use in expressing antibodies include lymphocytic cell lines, e.g., NSO myeloma cells and SP2 cells, COS cells.

In one embodiment, the host cell expresses/secretates the antigen binding protein, variant or fragment thereof as disclosed herein.

In one aspect, there is provided a cell that expresses/secretates an immune cell engager that is specific to EpCAM.

In one embodiment, the cell is a stem cell, for example selected from the group comprising a mesenchymal stem cell, neural stem cell and a pluripotent stem cell, such as an induced pluripotent stem cell (iPSC). Thus, in one embodiment, the stem cell is a mesenchymal stem cell. In one embodiment, the stem cell is a neural stem cell. In one embodiment, the stem cell is a pluripotent stem cell, such as an iPSC.

In one embodiment, the cell is an immune cell.

In one embodiment, the immune cell expresses/secretates an immune cell engager which is specific to EpCAM. In one embodiment, the immune cell engager is selected from the group comprising a T cell engager, an NK cell engager, a monocyte engager and a macrophage engager.

In one embodiment, the immune cell expresses/secretates a bispecific T cell engager (BiTE) which is specific to EpCAM.

In one embodiment, the immune cell expresses/secretates a bispecific T cell engager (BiTE), such as an inducible, non-inducible or constitutive expression BiTE comprising the antigen binding protein, variant or fragment thereof as disclosed herein.

Surprisingly, the present inventors have established that immune cells, such as
5 CAR T-cells, are able to secrete EpCAM BiTEs. See Examples 5 to 7. It was not previously known that this was possible. Advantageously, by engineering immune cells to express EpCAM immune engagers, this allows the EpCAM immune engagers to be secreted at the target site (for example at the site of a solid tumour), thereby minimising toxicity and/or side effects.

10 In some examples, the immune cell may include but is not limited to a macrophage, a dendritic cell, a T cell, a B cell, an eosinophil, a basophil, a neutrophil, a mast cell, a natural killer T cell (NKT cell), natural killer cell (NK cell), a macrophage, a monocyte, and the like. In one embodiment, the immune cell is a NK cell. In one embodiment, the immune cell is a macrophage. In one embodiment, the immune cell is
15 a dendritic cell. In one embodiment, the immune cell is a monocyte.

In one embodiment, the immune cell is a T-cell. In one embodiment, the immune cell is a CAR T-cell, such as an anti-GPC3, anti-HER2 or anti-CD19 CAR T-cell. Thus, in one embodiment, the CAR T-cell is an anti-GPC3 CAR T-cell. In one embodiment, the CAR T-cell is an anti-HER2 CAR T-cell. In one embodiment, the CAR T-cell is an anti-
20 CD19 CAR T-cell. In one embodiment, the immune cell is a CAR T-, CAR NK-, CAR macrophage-, or CAR monocyte-cell,

In some examples, the immune cell may bind to more than one host cell antigen. Therefore, in some examples, the immune cell may further bind to one host cell antigen, two host cell antigens, three host cell antigens, four host cell antigens, and the like.

25 Without wishing to be bound by theory, while the CAR T cells kill target cells directly, the anti-EpCAM BiTEs exert their cytotoxicity by recruiting the immune cells (such as T cells) nearby. By this approach, bystander T cells are physically directed to the close proximity of tumors and at the same time being activated and help the clearance of tumor cells. By applying local secretion of anti-EpCAM BiTE by CAR T cells,
30 it converts a non-druggable target into a druggable target as it diminishes the on-target, off-tumor toxicity that might be brought by systematic delivery.

As EpCAM is also defined as a cancer stem cell marker expressed on cancer progenitor cells and cancer stem cells, secretion of anti-EpCAM BiTE by CAR T cells will

concomitantly contribute to the prevention of cancer relapse and recurrence by eliminating cancer stem cells and progenitor cells.

In one embodiment, the immune cell expresses an inducible bispecific T cell engager (BiTE) comprising a heavy chain variable region that comprises the heavy chain

5 complementarity determining region:

(i) (for clones 1B6, 1C1, 1C11, 1D4, 1H6)

a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**),

a CDR-H2 comprising IIPFGTA (**SEQ ID NO: 2**), and

a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**), or

10 (ii) (for clone 1E4)

a CDR-H1 comprising GDISNSVA (**SEQ ID NO: 5**),

a CDR-H2 comprising TYRSKWYS (**SEQ ID NO: 6**), and

a CDR-H3 comprising AREVEGSSYDAFDI (**SEQ ID NO: 7**), and

(iii) (for clone 1A5-VHH)

15 a CDR-H1 comprising GSIFSGND (**SEQ ID NO: 25**),

a CDR-H2 comprising ITSGGST (**SEQ ID NO: 26**), and

a CDR-H3 comprising TNGRWSGDTYYAHH (**SEQ ID NO: 27**), or

(iv) (for clones 1B8-VHH, 2C4-VHH, 2D10-VHH and hu2C4-VHH)

a CDR-H1 comprising GSSEFTS (**SEQ ID NO: 29**),

20 a CDR-H2 comprising ITNGGST (**SEQ ID NO: 30**), and

a CDR-H3 comprising MAGTS (**SEQ ID NO: 31**), or

(v) (for clone 2B7-VHH)

a CDR-H1 comprising GSIFSGND (**SEQ ID NO: 25**),

a CDR-H2 comprising ITSGGST (**SEQ ID NO: 26**), and

25 a CDR-H3 comprising TNGRWSGDTYYAHL (**SEQ ID NO: 33**), or

or fragment or variation or sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity, and a light chain variable region that comprises the light chain complementarity determining region:

(i) (for clones 1B6 and 1C1)

30 a CDR-L1 comprising QSLHSNGYNY (**SEQ ID NO: 9**),

a CDR-L2 comprising LGS (**SEQ ID NO: 10**), and

a CDR-L3 comprising MQALQTPYT (**SEQ ID NO: 11**), or

(iii) (for clone 1C11)

- a CDR-L1 comprising QSLHNSGYNY (**SEQ ID NO: 9**),
 a CDR-L2 comprising LGS (**SEQ ID NO: 10**), and
 a CDR-L3 comprising MQGLQSPWT (**SEQ ID NO: 15**), or
 (iii) (for clone 1D4)
- 5 a CDR-L1 comprising QSLHNSRYNY (**SEQ ID NO: 17**),
 a CDR-L2 comprising LGS (**SEQ ID NO: 10**), and
 a CDR-L3 comprising MQALQTPYT (**SEQ ID NO: 11**), or
 (iv) (for clone 1E4)
- 10 a CDR-L1 comprising QSISDF (**SEQ ID NO: 19**),
 a CDR-L2 comprising AAS (**SEQ ID NO: 20**), and
 a CDR-L3 comprising QQSYIMPDT (**SEQ ID NO: 21**), or
 (v) (for clone 1H6)
- 15 a CDR-L1 comprising QSLHNSGYNY (**SEQ ID NO: 9**), (1H6)
 a CDR-L2 comprising LGS (**SEQ ID NO: 10**), and
 a CDR-L3 comprising MQGLQTPYT (**SEQ ID NO: 23**),
 or fragment or variation or sequence at least 80%, 85%, 90%, 95%, 96%, 97%,
 98%, 99%, or 100% identity.
- In one embodiment, the immune cell expresses an inducible bispecific T cell
 engager (BiTE) that further binds to an immune cell activator.
- 20 In one embodiment, the immune cell expresses an inducible bispecific T cell
 engager (BiTE) that further binds to CD3.
- Also disclosed is a polypeptide comprising a multispecific antigen binding protein.
 In some examples, the multispecific antigen binding protein is a bi-specific antibody. In
 some examples, the polypeptide comprises a multispecific antigen binding protein that
 25 binds to EpCAM (epithelial cell adhesion molecule) and an immune cell.
- In some examples, the multispecific antigen binding protein is a bispecific
 immune cell engager that is capable of engaging both an antigen and an immune cell.
- In some examples, the polypeptide comprises an anti-EpCAM antigen binding
 protein. In some examples, the polypeptide comprises a single domain anti-EpCAM
 30 antibody, optionally an anti-EpCAM H-chain antibody variable region (i.e., VHH).
- In some examples, the polypeptide comprises an anti-immune cell antigen
 binding protein. In some examples, the polypeptide comprises an anti-immune cell
 antigen binding protein that binds to an immune cell activation marker. In some

examples, the immune cell activation marker is CD3, NKG2D, CD4, CD8, CD16, CD64, and the like. In some examples, the polypeptide binds to CD3. In some examples, the CD3 may include CD3 ζ , CD3 ϵ , CD3 γ , CD3 δ , and the like.

5 In some examples, the polypeptide comprises an anti-CD3 antigen binding protein. In some examples, the polypeptide comprises a single-chain variable fragment of an anti-CD3 antibody (anti-CD3 scFv).

10 In some examples, the polypeptide is a bispecific antibody/antigen binding protein. In some examples, the bispecific antibody/antigen binding protein is a bispecific T cell engager (BiTE). In some examples, the bispecific T cell engager (BiTE) binds to two antigens, wherein the first antigen is EpCAM and wherein the second antigen is an immune cell marker. In some examples, the second antigen is an immune cell marker that is involved in the activation of the immune cell. In some examples, the second antigen targeted by the bispecific T cell engager (BiTE) may include but is not limited to CD3, NKG2D, CD28, CD16, CD64, and the like. In some examples, the polypeptide
15 comprises a BiTE that bi-specifically binds to EpCAM and a T cell. In some examples, the polypeptide comprises a BiTE that bi-specifically binds to EpCAM and CD3. In some examples, the EpCAM targeted by the BiTE is modified with an anti-EpCAM VHH paired with an anti-CD3 scFv.

20 In some examples, the anti-EpCAM VHH pairs with an anti-CD3 scFv that may include a clone Okt3 (Nb01-013A), or an anti-CD3 clone used by another anti-EpCAM in the art (MT110)(Nb01-013B).

In one aspect, there is provided a method of producing / generating the antigen binding protein, variant or fragment thereof as disclosed herein, comprising

- 25 (a) using a library to screen for antigen binding proteins; and
(b) cloning of the antigen binding proteins into an antibody format.

30 As shown in the experimental data of the present disclosure, 21 clones showing the positive Fab supernatant binding signals to biotinylated hEpCAM-His protein and specific binding to EpCAM positive HT-29 cells were sequence and 6 unique sequences were identified: 1B6, 1C1, 1C11, 1D4, 1E4 and 1H6. Clones 1B6, 1C1, 1C11, 1D4 and 1H6 share the same sequence in the heavy chain variable region but differ in the light chain sequences of the variable region, except for 1B6 and 1C1 that share the same sequences in the light chain variable region but with three amino acid differences in the kappa light chain constant region.

As shown in the experimental data of the present disclosure, 5 unique clones were identified (1A5, 1B8, 2B7, 2C4 and 2D10) with specific binding to gastric cancer cells (such as AGS cell line).

In one embodiment, the disclosed anti-EpCAM human Fab / VHH clones are used to construct anti-EpCAM BiTE antibodies.

In one aspect, there is provided a method of producing / generating the antigen binding protein, variant or fragment thereof as disclosed herein, comprising expressing the polynucleotide as described herein in a host cell.

In one aspect, there is provided a pharmaceutical composition comprising the antigen binding protein, variant or fragment thereof as disclosed herein.

In one embodiment, the composition is a prophylactic and/or therapeutic composition

In some embodiments, the composition comprises one or more pharmaceutically acceptable agents. Pharmaceutically acceptable agents for use in the present pharmaceutical compositions include carriers, excipients, diluents, antioxidants, preservatives, colouring, flavouring and diluting agents, emulsifying agents, suspending agents, solvents, fillers, bulking agents, buffers, delivery vehicles, tonicity agents, cosolvents, wetting agents, complexing agents, buffering agents, antimicrobials, and surfactants.

Pharmaceutically acceptable salts can be used, for example mineral acid salts, such as hydrochlorides, hydrobromides, phosphates and sulphates, or salts of organic acids, such as acetates, propionates, malonates and benzoates.

Pharmaceutically acceptable carriers in therapeutic compositions may additionally contain liquids such as water, saline, glycerol and ethanol. Additionally, auxiliary substances, such as wetting or emulsifying agents or pH buffering substances, may be present in such compositions. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries and suspensions, for ingestion by the patient. A thorough discussion of pharmaceutically acceptable carriers is available in Remington's Pharmaceutical Sciences (Mack Publishing Company, N.J. 1991).

In one embodiment, the composition further comprises a pharmaceutically acceptable excipient, a buffer or carrier.

In one embodiment, the composition is for use in therapy/medicine/vaccine.

In one aspect, there is provided a method of preventing and/or reducing the severity of symptoms caused by a disease in a subject in need thereof, the method comprises administering to the subject an antigen binding protein, variant or fragment thereof, or composition as disclosed herein.

5 In one embodiment, the antigen binding protein, variant or fragment thereof or composition thereof is administered to the subject by mode of administrations known in the art, including but not limited to intramuscularly, subcutaneously, intravenously, intraarterially, intraarticularly, intraperitoneally, intranasally, parenterally, and the like.

 In one embodiment, the subject is a mammal, such as a monkey, rabbit, mouse, 10 rat, pig or dog. In one embodiment, the subject is a human.

 In one aspect, there is provided a protein, composition, therapy, method of production or method as described herein.

SEQUENCES

15 CDRs are in bold (**CDR1**); in bold and in italics (***CDR2***); or in bold, in italics and underlined (**CDR3**). Highlighted residues/bases show differences between the 1B6 and 1C1 light chain constant domains.

 1. Amino-acid Sequences

20 1.1. Heavy chain variable domain of human IgG1

1B6, 1C1, 1C11, 1D4, 1H6 (SEQ ID NO: 4)

 EVQLVQSGAEVKKPGSSVKVSKASGGT**FSSY**AIWVRQAPGQGLEWMGG***IIPIFGT***
 ANYAQNFQGRVTMTADTSISTAYMELSSLRSEDTAVYYC**ARSLGGRFRY**WGQGTL

25 **1E4 (SEQ ID NO: 8)**

 QVQLQQSGPGLVKPSQTLTLCAIS**GDSISNS**VAWNWIRQSPSRGLEWLGR***TYYS***
 KWYSDAISVKGRLDINPDTSKNQFSLQLNSVTPEDTAVYYC**AREVEGSSYDAFDW**
 GQGTM

30 1.2. Light chain variable domain of human IgG1

1B6 (SEQ ID NO: 12)

 DVVMTQSPSLPVTGPGEPA**ISCRSSQSL**LHSNGYNYLDWYLQKPGQSPQLLIY**LGS**
 NRASGV**PDRFSGSG**SGTDFTLKISRVEAEDVGVYYC**MQALQTPYT**FGQGTK

1B6 (light chain constant domain) (SEQ ID NO: 13)

RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESV
EQDSKDYSLSTLTLSKADYEKHKYACEVTHQGLSSPVTKSFNRGEC

1C1 (SEQ ID NO: 12)

5 DVVMTQSPSLPVTGPGEPAISCRSSQSLHSHNGYNYLDWYLQKPGQSPQLLIYLG
NRASGVPRDFSGSGSGTDFTLKISRVEAEDVGVYYCMQALQTPYTFGQGTK

1C1 (light chain constant domain) (SEQ ID NO: 14)

RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESV
EQDSKDYSLSTLTLSKADYEKHKYACEVTHQGLSSPVTKSFNRGEC

10 1C11 (SEQ ID NO: 16)

EIVLTQSPSLPVTGPGEPAISCRSSQSLHSHNGYNYLDWYLQKPGQSPQLLIYLG
RASGVPRDFSGSGSGTDFTLKISRVEAEDVGVYYCMQGLQSPWTFGQGTK

1D4 (SEQ ID NO: 18)

15 DVVMTQSPSLPVTGPGEPAISCRSSQSLHSHNRYNYLDWYLQKPGQSPQLLIYLG
NRASGVPRDFSGSGSGTDFTLKISRVEAEDVGVYYCMQALQTPYTFGQGTK

1E4 (SEQ ID NO: 22)

DQLTQSPSSLSASVGRVTITCRASQSIDFLNWFYQQKPGKAPKLLIYAASLQTGVP
SRFGGSGSGTEFTLTISSLQPEDLGTYYCQQSYMPDTFGQGTK

1H6 (SEQ ID NO: 24)

20 DVVMTQSPSLPVTGPGEPAISCRSSQSLHSHNGYNYLDWYLQKPGQSPQLLIYLG
NRASGVPRDFSGSGSGTDFTLKISRVEAEDAGVYYCMQGLQTPYTFGQGTK

1.3. Single domain heavy chain variable domain of llama VHH

1A5-VHH (SEQ ID NO: 28)

25 QVQLQESGGGLVQPGGSLRLSCADSGSIFSGNDMAWYRRAPGVERELVAVITSGGS
THYADSVKGRFTISRDNAAKTQVYLQTNLKPEDTAVYYCTNGRWSGDTYYAHHWGQ
GTQ

1B8-VHH (SEQ ID NO: 32)

30 QVQLQESGGGLVQPGGSLRLSCAASGSSERFTSVAWYRQAPGKERELVAFITNGGS
TRYTDPVKGRFTISRDNAAKNTVYLMNSLKAEDTAVYYCMAGTSWGQGTQ

2B7-VHH (SEQ ID NO: 34)

QVQLQESGGGLVQPGGSLRLSCADSGSIFSGNDMAWYRRAPGVERELVAVITSGGS
THYADSVKGRFTISRDNQKTVYLQTNLKPEDTAVYYC TNGRWSGDTYYAHLWGQ
GTQ

2C4-VHH (SEQ ID NO: 35)

5 QVQLQESGGGLVQAGGSLRLSCADSGSIFSGNDMAWYRRAPGVERELVAVITSGGS
THYADSVKGRFTISRDNQKTVYLQTNLKPEDTAVYYC TNGRWSGDTYYAHHWGQ
GTQ

2D10-VHH (SEQ ID NO: 36)

10 QVQLQESGGGLVQAGDSLRLSCADSGSIFSGNDMAWYRRAPGVERELVAVITSGGS
THYADSVKGRFTISRDNQKTVYLQTNLKPEDTAVYYC TNGRWSGDTYYAHHWGQ
GTQ

1.4.	Single domain heavy chain variable domain of humanized llama VHH
------	--

hu2C4-VHH (SEQ ID NO: 37)

15 QVQLVESGGGLVQAGGSLRLSCAASGSIFSGNDMSWYRQAPGKGLVAVITSGGS
TYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYC TNGRWSGDTYYAHHWGQ
GTL

2. Nucleotide Sequences

2.1.	Heavy chain variable domain of human IgG1
------	---

1B6, 1C1, 1C11, 1D4, 1H6 (SEQ ID NO: 41)

GAGGTCCAGCTGGTGCACTCTGGGGCTGAGGTGAAGAAGCCTGGGTCTCGGT
GAAGGTCTCTGCAAGGCTTCTGGAGGCACCTTCAGCAGCTATGCTATCAGCTG
25 GGTGCGACAGGCCCTGGACAAGGGCTTGAGTGGATGGGAGGGATCATCCCTAT
CTTTGGTACAGCAAACTACGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCA
GACACCTCCATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGAC
ACGGCCGTGTATTACTGT GCGAGATCGTTGGGTGGGAGATTTCGCTACTTGGGGC
CAGGGAACCCCTG

30 1E4 (SEQ ID NO: 45)

CAGGTACAGCTGCAGCAGTCAGGTCCAGGGCTGGTGAAGCCCTCGCAGACCCCTC
TCACTCACCTGTGCCATCTCCGGGGACAGTATCTCTAGTAACAGTGTTGCTTGGA
ACTGGATCAGGCAGTCCCATCGAGAGGCCTTGAGTGGCTGGGAAGGACATACT
ACAGGTCCAAGTGGTACAGTGATTATGCAATATCTGTGAAAGGTCGATTAGACAT
35 CAACCCAGACACATCCAAGAACCAGTTCTCCCTGCAGCTGAACTCTGTGACTCCC

GAGGACACGGCTGTGTATTATTGT GCAAGAGAAGTTGAGGGCAGCAGCTATGAT
GCTTTTGATATCTGGGGCCAAGGGACAATG

2.2. Light chain variable domain of human IgG1

5

1B6 (SEQ ID NO: 49)

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTGGAGAGCCGG
CCTCCATCTCCTGCAGGTCTAGT **CAGAGCCTCCTGCATAGTAATGGATACAAC**
TTTGGATTGGTACCTGCAGAAGCCAGGGCAGTCTCCACAGCTCCTGATCTAT **TTG**
10 **GGTTCT**AATCGGGCCTCCGGGGTCCCTGACAGGTTCAAGTGGCAGTGGATCAGGC
ACAGATTTTACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTATT
ACTGC **ATGCAAGCTCTACAAACTCCGTACACT**TTTGGCCAGGGGACCAAG

1B6 (light chain constant domain) (SEQ ID NO: 50)

CGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCAGTTGA
15 AATCTGGAAGTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGAGGC
CAAAGTACAGTGGAAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAG
TGTC **CAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCTGAC**
GCTGAGCAAAGCAGACTACGAGAAACACAAA **TCTACGCCTGCGAAGTCACCCAT**
CAGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCA **CAGGGGAGAGTGT**

1C1 (SEQ ID NO: 49)

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTGGAGAGCCGG
CCTCCATCTCCTGCAGGTCTAGT **CAGAGCCTCCTGCATAGTAATGGATACAAC**
TTTGGATTGGTACCTGCAGAAGCCAGGGCAGTCTCCACAGCTCCTGATCTAT **TTG**
GGTTCTAATCGGGCCTCCGGGGTCCCTGACAGGTTCAAGTGGCAGTGGATCAGGC
25 ACAGATTTTACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTATT
ACTGC **ATGCAAGCTCTACAAACTCCGTACACT**TTTGGCCAGGGGACCAAG

1C1 (light chain constant domain) (SEQ ID NO: 51)

CGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCAGTTGA
AATCTGGAAGTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGAGGC
30 CAAAGTACAGTGGAAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAG
TGTC **CAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCTGAC**
GCTGAGCAAAGCAGACTACGAGAAACACAAA **TCTACGCCTGCGAAGTCACCCAT**
CAGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCA **CAGGGGAGAGTGT**

1C11 (SEQ ID NO: 53)

GAAATTGTGCTGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTGGAGAGCCGG
CCTCCATCTCCTGCAGGTCTAGT **CAGAGCCTCCTGCATAGTAATGGATACAAC**
TTTGGATTGGTACCTGCAGAAGCCAGGGCAGTCTCCACAGCTCCTGATCTAT **TTG**

GGTTCTAATCGGGCCTCCGGGGTCCCTGACAGGTTCAAGTGGCAGTGGATCAGGT
ACAGATTTTACACTGAAAATAAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTATTA
CTGC**ATGCAAGGTCTACAAAGTCCCTGGACG**TTCCGGCCAAGGGACCAAG

1D4 (SEQ ID NO: 55)

5 GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTGGAGAGTCGG
CCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGCATAGTAATAGATACAACTA**
TTTGGATTGGTACCTGCAGAAGCCAGGGCAGTCTCCACAGCTCCTGATCTAT**TTG**
GGTTCTAATCGGGCCTCCGGGGTCCCTGACAGGTTCAAGTGGCAGTGGATCAGGC
ACAGATTTTACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTATT
10 ACTGC**ATGCAAGCTCTACAAACTCCGTACACT**TTTGGCCAGGGGACCAAG

1E4 (SEQ ID NO: 59)

GACATCCAGTTGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAGGAGACAGAG
TCACCATCACTTGCCGGGCAAGT**CAGAGTATTAGCGACTTTTTAAATTGGTACCA**
GCAGAAACCAGGTAAAGCCCCGAAGCTCCTGATCTAT**GCTGCATCGAGTTTACAA**
15 ACTGGGGTCCCCTCAAGATTCCGTGGCAGTGGATCTGGGACAGAATTCACTCTCA
CCATAAGCAGTCTACAACCTGAAGATTTGGGAACCTATTACTGT**CAACAGAGTTAC**
ATTATGCCCCGACACTTTTGGCCAGGGGACGAAA

1H6 (SEQ ID NO: 61)

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTGGAGAGCCGG
20 CCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGCATAGTAATGGATACAACTA**
TTTGGATTGGTACCTGCAGAAGCCAGGGCAGTCTCCACAGCTCCTGATCTAT**TTG**
GGTTCTAATCGGGCCTCCGGGGTCCCTGACAGGTTCAAGTGGCAGTGGATCAGGC
ACAGATTTTACACTGCAAATCAGCAGAGTGGAGGCTGAGGATGCTGGGGTTTATT
ACTGC**ATGCAAGGTCTACAGACTCCGTACACT**TTTGGCCAGGGGACCAAG

25

2.3.	Single domain heavy chain variable domain of llama VHH
------	--

1A5 (SEQ ID NO: 65)

CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGAGGGTCTCT
GAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCAATGACATGGCCTG**
30 GTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTTGGTCGCGGTT**ATTACTAGCG**
GTGGTAGTACACACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGA
CAATGCCCAGAAGACCGTATATCTGCAAACGAACGACTTGAAACCTGAGGACACG
GCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCAT**
CACTGGGGGCCAGGGGACCCAG

1B8 (SEQ ID NO: 69)

35 CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGGGGGTCTCT
GAGACTCTCCTGTGCAGCCTCT**GGAAGCTCCGAAAGATTACATCAGTGGCCTG**

GTACCGCCAGGCTCCAGGAAAGGAGCGCGAGTTGGTGCATTT**ATTACTAATGG**
TGGTAGCACAAAGATATACAGACCCCGTGAAGGGCCGATTACCATCTCCAGAGA
CAACGCCAAGAACACGGTGTATCTGCAAATGAACAGCCTGAAAGCTGAGGACACG
GCCGTCTATTATTGT**ATGGCGGGTACGTCC**TGGGGCCAGGGGACCCAG

5 **2B7 (SEQ ID NO: 71)**

CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGAGGGTCTCT
GAGACTCTCCTGTGCAGACTCT**TGGAAGCATCTTCAGTGGCAATGACAT**GGCCTG
GTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTTGGTCGCGGTT**ATTACTAGCG**
10 **GTGGTAGTACAC**ACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGA
CAATGCCCAGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACACG
GCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCAT**
CTCTTGGGGCCAGGGGACCCAG

2C4 (SEQ ID NO: 72)

CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGGGGGCTCTCT
15 GAGACTCTCCTGTGCAGACTCT**TGGAAGCATCTTCAGTGGCAATGACAT**GGCCTG
GTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTTGGTCGCGGTT**ATTACTAGCG**
GTGGTAGTACACACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGA
CAATGCCCAGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACACG
GCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCAT**
20 **CACT**TGGGGCCAGGGGACCCAG

2D10 (SEQ ID NO: 73)

CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGGGGACTCTCT
GAGACTCTCCTGTGCAGACTCT**TGGAAGCATCTTCAGTGGCAATGACAT**GGCCTG
GTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTTGGTCGCGGTT**ATTACTAGCG**
25 **GTGGTAGTACAC**ACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGA
CAATGCCCAGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACACG
GCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCAT**
CACTTGGGGCCAGGGGACCCAG

30

2.4. Single domain heavy chain variable domain of humanized llama VHH

hu2C4-VHH (SEQ ID NO: 74)

CAGGTGCAGCTGGTGGAGTCTGGGGGAGGATTGGTGCAGGCTGGGGGCTCTCT
GAGACTCTCCTGTGCAGCCTCT**TGGAAGCATCTTCAGTGGCAATGACAT**GTCCCTG
35 GTACCGCCAGGCTCCAGGGAAGGGA

CGTGTATTACTGCACAAACGGAAGATGGTCAGGCGATACTTACTATGCCATCA
CTGGGGCCAGGGGACCCTG

BRIEF DESCRIPTION OF FIGURES

- 5
- Figure 1** shows the flow cytometry analysis of the surface expression of human EpCAM protein on multiple cells lines with epithelial morphology (HT-29, AGS, HepG2, Hep3B, MCF-7) and non-epithelial morphology (HeyA8, A172, U-87) and human T cells isolated from healthy donor blood tissue. A commercially available mouse anti-human
- 10 EpCAM monoclonal antibody directly conjugated to BV421 fluorescent dye was used.
- Figure 2A** shows the binding ELISA assay results of 6 Fab clones to the antigen protein. The 6 Fab clones were tested using crude culture supernatants in a binding ELISA assay to assess antigen binding to biotinylated human EpCAM with a poly-His tag, detected by goat-anti-human Fab-HRP.
- 15 **Figure 2B** shows the flow cytometry analysis of the 6 Fab clones in their binding to cell surface expressed EpCAM. Using crude culture supernatants, the 6 Fab clones were tested on their binding to EpCAM positive cell HT-29 and EpCAM negative human T cells. In both assays, supernatant from a culture medium alone well ("Blank") was used as background control.
- 20 **Figure 2C** shows the binding affinity results of the IgG1 antibodies. Six anti-EpCAM Fab clones were cloned and expressed as IgG1 antibodies and their binding affinity to the biotinylated human EpCAM protein antigen were tested using affinity binding ELISA and compared to the benchmark IgG (the anti-EpCAM arm of "MT110"). Antigen binding EC50 values of anti-EpCAM IgG1 antibodies were calculated by PRISM.
- 25 **Figure 3A** shows the structural format of anti-EpCAM BiTE antibodies.
- Figure 3B** shows the binding kinetics of anti-EpCAM BiTE antibodies for biotinylated human EpCAM protein antigen using binding ELISA.
- Figure 3C** shows the results of the flow cytometry analysis at different concentrations for the anti-EpCAM BiTE antibodies to EpCAM positive HT-29 cells,
- 30 **Figure 3D** shows the results of the flow cytometry analysis at different concentrations for the anti-EpCAM BiTE antibodies to EpCAM negative HeyA8 cells, and **Figure 3E** shows the results of the flow cytometry analysis at different concentrations for the anti-

EpCAM BiTE antibodies to CD3 positive human T cells. The percentages of cells positively stained by anti-EpCAM BiTE antibodies are plotted.

Figure 3F shows the binding of Benchmark BiTE antibody ("MT110") to HT-29 and HeyA8 cells.

5 **Figure 3G** shows the binding EC50 values of anti-EpCAM BiTE antibodies to HT-29 cells as calculated by PRISM, with top and bottom constraints set at 100% and 0% respectively.

10 **Figure 4A** shows the xCelligence impedance assay results for the anti-EpCAM BiTE antibodies. Cytolysis of EpCAM positive HT-29 cells was mediated by activated human T cells with time-course measurement using xCelligence impedance assay, upon treatment with 6 different anti-EpCAM BiTE antibodies (clones 1B6, 1C1, 1C11, 1D4, 1E4 and 1H6).

Figure 4B shows the Mean % cytolysis $\pm$ SD in duplicate wells at different concentrations at 72 hours.

15 **Figure 4C** shows the EC50 values of anti-EpCAM BiTE antibodies in killing of HT-29 cells as calculated by PRISM, with top and bottom constraints set at 100% and 0% respectively.

20 **Figure 5A** shows the binding ELISA results of 5 llama VHH clones to the antigen protein. The 5 llama VHH clones were tested using crude culture supernatants in a binding ELISA assay to assess their antigen binding to biotinylated human EpCAM with a poly-His tag, and detected by HRP-conjugated anti-HA.11 epitope tag antibody.

25 **Figure 5B** shows the flow cytometry analysis of the 5 llama VHH clones in their binding to cell surface expressed EpCAM. Using crude culture supernatants, the 5 VHH clones were tested for binding to EpCAM positive cell AGS. In both assays, supernatant from a culture medium alone well ("Blank") was used as background control.

Figure 6A shows the structural format of anti-EpCAM VHH BiTE antibodies.

Figure 6B binding kinetics of anti-EpCAM VHH BiTE antibodies for biotinylated human EpCAM protein antigen using binding ELISA.

30 **Figure 6C** shows the results of the flow cytometry analysis at different concentrations for the anti-EpCAM VHH BiTE antibodies to EpCAM positive HT-29 cells, **Figure 6D** shows the results of the flow cytometry analysis at different concentrations for the anti-EpCAM VHH BiTE antibodies to EpCAM negative HeyA8 cells, and **Figure 6E** shows the results of the flow cytometry analysis at different concentrations for the

anti-EpCAM VHH BiTE antibodies to CD3 positive human T cells. The percentages of cells positively stained by anti-EpCAM VHH BiTE antibodies were plotted using PRISM software.

5 **Figure 6F** shows the binding EC50 values of anti-EpCAM VHH BiTE antibodies to HT-29 cells as calculated by PRISM, with top and bottom constraints set at 100% and 0% respectively. All anti-EpCAM VHH BiTE antibodies used the same anti-CD3 clone as used in the benchmark antibody ("MT110").

10 **Figure 7A** shows the cytotoxicity of EpCAM positive HT-29 cells, **Figure 7B** shows the cytotoxicity of EpCAM positive HepG2 cells, **Figure 7C** shows the cytotoxicity of EpCAM positive Hep3B cells and **Figure 7D** shows the cytotoxicity of EpCAM negative HeyA8 cells. Cytotoxicity of the cells was mediated by activated human T cells with time-course measurement using xCelligence impedance assay, upon treatment with 5 different anti-EpCAM VHH BiTE antibodies (clones 1A5-VHH, 1B8-VHH, 2B7-VHH, 2C4-VHH, and 2D10-VHH) with the E:T ratio of 4:1. Data was presented as % cytotoxicity at different
15 concentrations at either 96 hours (**Figures 7A -C**) or 72 hours (**Figure 7D**).

Figure 7E shows the EC50 values of anti-EpCAM VHH BiTE antibodies in killing of different cells at either 96 hours (96h) or 72 hours (72h) calculated by PRISM, with top and bottom constraints set at 100% and 0% respectively. All anti-EpCAM VHH BiTE antibodies used the same anti-CD3 clone as used in the benchmark antibody ("MT110").

20 **Figure 8A** shows the binding kinetics of llama derived anti-EpCAM bi-specific T-cell engagers using 2C4-VHH and its humanized clone hu2C4-VHH paired with anti-CD3 clone Okt3. Anti-EpCAM 2C4-VHH/Okt3, hu2C4-VHH/Okt3 BiTE antibodies were tested for their binding to biotinylated human EpCAM protein antigen using binding ELISA and the antigen binding EC50 values were tabulated.

25 **Figure 8B** shows the binding of the anti-EpCAM 2C4-VHH/Okt3, hu2C4-VHH/Okt3 BiTE antibodies tested by flow cytometry analysis at different concentrations to EpCAM positive HT-29 cells, **Figure 8C** shows the binding of the anti-EpCAM 2C4-VHH/Okt3, hu2C4-VHH/Okt3 BiTE antibodies to EpCAM negative, CD3 negative HeyA8 cells, and **Figure 8D** shows the binding of the anti-EpCAM 2C4-VHH/Okt3, hu2C4-VHH/Okt3 BiTE antibodies to CD3 positive human T cells. The percentage of cells
30 positively stained by anti-EpCAM BiTE antibodies were plotted.

Figure 8E shows the binding EC50 values of anti-EpCAM 2C4-VHH/Okt3 and hu2C4-VHH/Okt3 BiTE antibodies to HT-29 cells as calculated by PRISM, with top and bottom constraints set at 100% and 0% respectively.

Figure 9A shows the T cell mediated cytotoxicity of EpCAM positive cells by anti-EpCAM bi-specific T-cell engagers using 2C4-VHH and its humanized clone hu2C4-VHH paired with anti-CD3 clone Okt3, and **Figure 9B** shows the T cell mediated cytotoxicity of EpCAM negative cells by anti-EpCAM bi-specific T-cell engagers using 2C4-VHH and its humanized clone hu2C4-VHH paired with anti-CD3 clone Okt3. Cytolysis of EpCAM positive MCF-7 cells (**Figure 9A**) and EpCAM negative HeyA8 cells (**Figure 9B**) was mediated by activated human T cells with time-course measurement using xCelligence impedance assay, upon treatment with anti-EpCAM VHH BiTE antibodies 2C4-VHH/Okt3 and hu2C4-VHH/Okt3, with the E:T ratio of 4:1. Data was presented as % cytotoxicity at different concentrations at 48 hours.

Figure 9C shows the EC50 values of the benchmark BiTE, 2C4-VHH/Okt3 and hu2C4-VHH/Okt3 BiTE antibodies in killing MCF-7 cells at 48 hours as calculated by PRISM, with top and bottom constraints set at 100% and 0% respectively.

Figure 9D shows the ELISA measurement of interferon- γ secretion from human T cells activated by different anti-EpCAM BiTE antibodies after 48 hours of co-culture with either EpCAM positive MCF-7 cells or EpCAM negative HeyA8 cells, and **Figure 9E** shows the ELISA measurement of IL-2 secretion from human T cells activated by different anti-EpCAM BiTE antibodies after 48 hours of co-culture with either EpCAM positive MCF-7 cells or EpCAM negative HeyA8 cells.

Figure 10A shows structural formats of anti-GPC3 CAR T (5C4) or GE CAR-BiTE T cells (anti-GPC3 CAR T secreting anti-EpCAM BiTE, "MT110", "Nb01-013A" or "Nb01-013B" was used as BiTE) or 19E CAR-BiTE T cells (anti-CD19 CAR T secreting anti-EpCAM BiTE, "Nb01-013A" or "Nb01-013B" was used as BiTE).

Figure 10B shows the results of an *in vitro* assay of Hep3B (GPC3^{High}, EpCAM^{High}) cells by anti-GPC3 CAR T, GE CAR-BiTE T cells or 19E CAR-BiTE T cells. **Figure 10C** shows the results of an *in vitro* assay of HT-29 (GPC3^{Low}, EpCAM^{High}) cells by anti-GPC3 CAR T, GE CAR-BiTE T cells or 19E CAR-BiTE T cells. **Figure 10D** shows the results of an *in vitro* assay of HeyA8 (GPC3^{-ve}, EpCAM^{-ve}) cells by anti-GPC3 CAR T, GE CAR-BiTE T cells or 19E CAR-BiTE T cells. The assays were

performed using xCelligence cell impedance assay and presented over 72 hours post co-culture of target cells and effector cells (E:T ratio = 1:1).

Figure 11A shows the structure map of HE CAR-BiTE T (clone 4D5 or F5) using anti-EpCAM VHH BiTE (Nb01-013A).

5 **Figure 11B** shows the FACS analysis results of tumor markers (HER2 and EpCAM) expressed on AGS and MDA-MB468 cells.

Figure 11C shows the percentage of CAR expression on anti-HER2 CAR T (4D5 or F5) and HE CAR- BiTE T (4D5 or F5) cells detected by flow cytometry analysis.

10 **Figure 11D** shows the percentage of cytolysis of AGS cells mediated by anti-HER2 CAR T (4D5 or 5F) and HE CAR-BiTE T (4D5 or F5) cells with time-course measurement using xCelligence impedance assay (E:T ratio = 2:1).

Figure 11E shows the percentage of cytolysis of MDA-MB468 cells mediated by anti-HER2 CAR T (4D5 or 5F) and HE CAR-BiTE T (4D5 or F5) cells with time-course measurement using xCelligence impedance assay (E:T ratio = 1:1).

15 **Figure 11F** shows the ELISA measurement of Interferon- γ and IL-2 measured using culture supernatant collected at 36 hours post co-culture effector CAR T cells with AGS cells.

Figure 11G shows the ELISA measurement of Interferon- γ and IL-2 measured using culture supernatant collected at 36 hours post co-culture effector CAR T cells with
20 MDA-MB468 cells.

Figure 12 shows a graph demonstrating that the anti-EpCAM BiTE secreted from GE CAR-BiTE T cells can mediate strong T cell killing of cancer cells. The anti-EpCAM BiTE (The benchmark BiTE "MT110" was used) were secreted by human GE CAR-BiTE T (MT110) cells and naïve human T cells were isolated from PBMCs of a healthy donor.
25 The time-course measurement was performed by xCelligence impedance assay (E:T ratio = 6:1), and the percentages of cytolysis of Hep3B cells mediated by anti-EpCAM BiTE activated naïve T cells are shown.

Figure 13 shows a graph demonstrating the detection of anti-EpCAM BiTE molecule Nb01-013A secreted from GE CAR-BiTE T cells by ELISA. Anti-GPC3 CAR T (5C4) cells or anti-GPC3 CAR T cells secreting anti-EpCAM BiTE (Nb01-013A) (named as "GE CAR-BiTE T (Nb01-013A)") were cultured in T cell growth medium containing IL-7 (20 ng/ml) and IL-15 (5 ng/ml). The starting cell density was 0.5 million per ml and the percentages of CAR expression in both cultures were around 60%. Cell culture
30

supernatants were collected daily (24h, 48h, 72h, 96h) for an ELISA to detect the amount of secreted anti-EpCAM BiTE molecule Nb01-013A. Briefly, human EpCAM-Fc tag protein was used to coat the ELISA plate overnight, after blocking with Casein for 2 hours, the culture supernatants containing the anti-EpCAM BiTE molecule Nb01-013A were added to the plate. After 1-hour incubation, the plate was washed and the bound BiTE molecules were detected by an HRP conjugated secondary antibody against the His-tag.

Figure 14 shows a graph demonstrating the detection of anti-EpCAM BiTE molecule Nb01-013A secreted from HE CAR-BiTE T cells by ELISA. Anti-HER2 CAR T (4D5 or F5) cells or anti-HER2 CAR T cells secreting anti-EpCAM BiTE (Nb01-013A) (named as "HE CAR-BiTE T (4D5 or F5)") were cultured in T cell growth medium containing IL-7 (20 ng/ml) and IL-15 (5 ng/ml). The starting cell density was 1.5 million per ml and the percentages of CAR expression ranged from 36.8% to 57.2%. Cell culture supernatants were collected at 24h and 48h for an ELISA to detect the amount of secreted anti-EpCAM BiTE molecule Nb01-013A. Briefly, human EpCAM-Fc tag protein was used to coat the ELISA plate overnight, after blocking with Casein for 2 hours, the culture supernatants containing the anti-EpCAM BiTE molecule Nb01-013A were added to the plate. After 1-hour incubation, the plate was washed and the bound BiTE molecules were detected by an HRP conjugated secondary antibody against the His-tag.

Figure 15 shows a graph demonstrating that human GE CAR-BiTE T cells using anti-EpCAM VHH constructed BiTEs showed superior tumor killing in *in vivo* Hep3B xenografts. Two million Hep3B (GPC3^{High}, EpCAM^{High}) cells were subcutaneously injected into the right flank of 50 male NSG mice (Day -21). On Day 0, mice were re-grouped according to the measurable tumor size and eight million CAR T (5C4) cells, GE CAR-BiTE T cells (Anti-EpCAM BiTE secreting anti-GPC3 CAR T cells, "MT110", "Nb01-013A" or "Nb01-013B" was used as BiTE), 19E CAR-BiTE T cells (Anti-EpCAM BiTE secreting anti-CD19 CAR T cells, "Nb01-013A" or "Nb01-013B" was used as BiTE), or Mock T cells were intravenously injected into these mice via tail vein. Tumor sizes of each mouse were measured and recorded every 3 to 7 days. The curve shows tumor sizes up to Day 92.

Figure 16 shows a graph demonstrating that human HE CAR-BiTE T cells using anti-EpCAM VHH constructed BiTEs showed superior tumor killing in *in vivo* AGS xenografts. 1.3 million AGS (Her2^{High}, EpCAM^{High}) cells were subcutaneously injected

into the right flank of female NSG mice (Day -9). On Day 0, mice were re-grouped according to the measurable tumor size and five million CAR T (4D5) cells, CAR T (F5) cells, HE CAR-BiTE T (4D5) cells ("Nb01-013A" was used as BiTE), HE CAR-BiTE T (F5) cells ("Nb01-013A" was used as BiTE), or Mock T cells were intravenously injected
5 into these mice via tail vein. Tumor sizes of each mouse were measured and recorded every 3 to 7 days.

EXAMPLES

10

Example embodiments of the disclosure will be better understood and readily apparent to one of ordinary skill in the art from the following discussions and if applicable, in conjunction with the figures. It should be appreciated that other modifications may be made without deviating from the scope of the invention. Example embodiments are not
15 necessarily mutually exclusive as some may be combined with one or more embodiments to form new exemplary embodiments. The example embodiments should not be construed as limiting the scope of the disclosure.

20 **Example 1 – Discovery of fully human anti-EpCAM antibodies from naïve human Fab phage display library**

To identify cell lines that can be used for characterizing the properties of anti-EpCAM antibodies involved in this invention, the present inventors analysed the surface EpCAM antigen expression using a commercially available anti-EpCAM antibody
25 **(Figure 1)**. As expected, EpCAM had high level of expression on all cells tested with epithelial morphology such as a lung adenocarcinoma cell line HT-29, a gastric adenocarcinoma cell line AGS, a hepatoblastoma cell line HepG2, a hepatocellular carcinoma cell line Hep3B, and a breast cancer cell line MCF-7. However, EpCAM was not expressed on non-epithelial cells, such as human T cells, an ovarian cell line HeyA8,
30 and two glioblastoma cell lines A172 and U-87.

Next, biotinylated recombinant human EpCAM protein with a poly-His tag (hEpCAM-His) was used to isolate low-affinity EpCAM binders from a naïve library of Fab sequences constructed in SIgN, using phage display technology. Out of 190 clones

screened, 21 clones showing the positive Fab supernatant binding signals to biotinylated hEpCAM-His protein and specific binding to EpCAM positive HT-29 cells were sequenced and 6 unique sequences were identified: 1B6, 1C1, 1C11, 1D4, 1E4 and 1H6. Clone 1B6, 1C1, 1C11, 1D4 and 1H6 share the same sequence in the heavy chain variable regions, however the light chain sequences of the variable region are different, except for clone 1B6 and 1C1 which also share the same sequences in their light chain variable region but bearing three amino acid differences in the kappa light chain constant region (see Annex). The results of binding ELISA and flow cytometry analysis of cell binding using Fab supernatants for the 6 unique clones were shown in **Figure 2A** and **Figure 2B**, respectively. Next, these Fab antibodies were cloned into IgG1 format and their binding affinity to the antigen protein were determined by affinity binding ELISA (**Figure 2C**). The binding EC50 values ranged between 3 to 6 nM, which are higher than the binding EC50 of the benchmark IgG (1.161nM, the sequence of the anti-EpCAM arm of "MT110" was used), suggesting that these anti-EpCAM antibodies derived from the human naïve Fab library had a relatively low binding affinity to the EpCAM antigen.

Example 2 - Use of anti-EpCAM bi-specific T-cell engagers

The anti-EpCAM antibodies were used to construct anti-EpCAM bi-specific T-cell engagers (BiTE). In these constructs, the Fab region of one arm from each of the 6 anti-EpCAM human antibody clones (1B6, 1C1, 1C11, 1D4, 1E4 and 1H6) was replaced by an anti-CD3 scFv fragment, and knob-in-hole mutations were introduced to facilitate correct heavy-chain pairing and LALA mutations to remove Fcγ receptor binding (**Figure 3A**). These anti-EpCAM BiTE antibodies were first examined for their binding to the antigen protein by affinity binding ELISA. Compared to the benchmark BiTE (anti-EpCAM arm of "MT110" with scFv fragment of anti-CD3 clone Okt3, constructed in the same structure as shown in **Figure 3A**) showing a binding EC50 of 0.666 nM, BiTE antibodies using the 6 newly identified anti-EpCAM clones showed much lower absorbance reading at 450nm and higher EC50 values, suggesting lower binding affinities to the antigen (**Figure 3B**).

The abilities of these antibodies to specifically recognize and bind to both EpCAM and human CD3 molecules on the cell surfaces were measured using HT-29, HeyA8, and human T cells. HT-29 cells and human T cells constitutively express high level of EpCAM and CD3, respectively, while HeyA8 cells do not express EpCAM or CD3. In this

assay, all 6 anti-EpCAM BiTE antibodies bound to the EpCAM expressing cell HT-29 (Figure 3C) as well as human T cells expressing CD3 (Figure 3E). The binding EC50 to HT-29 cells is 5-20 times lower than that of the benchmark BiTE (MT110, Solitomab, constructed in the same structure as shown in Figure 3A), which is 4.106 pM (Figure 3G). However, unlike the benchmark BiTE, which showed 10-20% of non-specific binding to EpCAM negative HeyA8 cells at high concentrations of 0.5, 2, and 10 nM (Figure 3F), these antibodies did not bind to HeyA8 cells at all tested concentrations (Figure 3D). **This suggests that these anti-EpCAM bi-specific T-cell engagers could bind to cell surface antigen in a highly specific manner.**

These anti-EpCAM BiTE antibodies were then tested for their efficacy in antibody dependent T cell mediated cytotoxicity assay. Briefly, EpCAM positive HT-29 cells (T: Target cells) were seeded onto the xCelligence E-plate and cultured for overnight with cell index values measured. The next day, immediately after addition of anti-EpCAM BiTE antibodies at various concentrations, primary T cells isolated from human PBMCs (E: Effector cells) were added to the target cells at an Effector to Target ratio of 6 to 1 (E:T = 6:1). The assay was continued for a further 96 hours to measure the target cell index values using xCelligence RTCA system (Figure 4A). Except for clone 1E4, which only induced profound cytotoxicity of target cell at high concentration of 10 nM and 1 nM, all the other anti-EpCAM BiTE antibodies (clone 1B6, 1C1, 1C11, 1D4 and 1H6) were able to kill 70-100% of HT-29 cells within 72 hours at a concentration as low as 0.1 nM (Figure 4B). Among the 6 clones, clone 1C1 and 1H6 BiTE antibodies showed the lowest EC50 values of 18.5 pM and 12.02 pM, respectively; while 1B6, 1C11, 1D4, and 1E4 showed low to moderate cell killing efficacy (Figure 4C).

In summary, results from the above assays, depicted in Figures 1 – 4, showed that the disclosed highly specific human anti-EpCAM antibodies in the format of bi-specific T cell engagers could elicit a wide range of low to high potency in inducing cell death of EpCAM expressing cells, which have the potential to be used in treating solid tumors with enhanced safety profiles (based on in vitro binding and killing assay results).

Example 3 – Discovery of anti-EpCAM llama antibodies from immunized llama

To discover anti-EpCAM llama VHH antibodies, the present inventors immunized a male alpaca with recombinant human EpCAM protein with a poly-His tag (hEpCAM-His) for a total of six immunizations with time intervals of 1-2 weeks. The process of llama

immunization was done by VIB Nanobody Core at VUB (Vrije Universiteit Brussel) as a paid service. Subsequent humanization of the llama antibodies was performed at SlgN. Around 150 µg of proteins was used for the two initial immunizations and 100 µg of proteins were used for the last four booster immunizations. After six immunizations, peripheral blood lymphocytes were collected from the alpaca for RNA extraction, and cDNA was used as the PCR template to amplify VHH sequences for a VHH library to be constructed. Biotinylated recombinant human EpCAM protein with a poly-His tag (hEpCAM-His) was used to isolate EpCAM binders from the constructed llama VHH library, using phage display technology. Out of 190 clones screened, 8 clones showing the positive VHH supernatant binding signals to biotinylated hEpCAM-His protein and specific binding to EpCAM positive AGS cells were sequenced and 5 unique sequences were identified: 1A5, 1B8, 2B7, 2C4 and 2D10 (See sequences below). The results of binding ELISA and flow cytometry analysis of cell binding using crude supernatants for the 5 unique VHH clones were shown in **Figure 5**.

Similar to those anti-EpCAM human Fab clones, the 5 VHH clones of anti-EpCAM antibodies were also used to construct anti-EpCAM VHH BiTE antibodies (**Figure 6A**). These anti-EpCAM VHH BiTE antibodies were first examined for their binding to the antigen protein by affinity binding ELISA. Compared to the benchmark BiTE (**Figure 3B**), BiTE antibodies using these 5 anti-EpCAM VHH clones showed comparable or even lower binding EC50 values, suggesting high binding affinities to the antigen (**Figure 6B**).

The ability of these antibodies to specifically recognize and bind to both EpCAM and human CD3 molecules on the cell surfaces was also measured using HT-29 (EpCAM positive), HeyA8 (EpCAM negative, CD3 negative), and human T cells (CD3 positive) cells. Except for clone 1B8-VHH, the other four clones (1A5-VHH, 2B7-VHH, 2C4-VHH and 2D10-VHH) all bound strongly to EpCAM expressing cell HT-29 (**Figure 6C**) as well as human T cells (**Figure 6E**). The binding EC50 to HT-29 cells ranged from 10-20 pM and the binding specificity was superior as none of them showed binding to EpCAM negative HeyA8 cells even at the highest concentration of 10 nM (**Figure 6D**).

These anti-EpCAM VHH BiTE antibodies were also tested for their killing efficacy in antibody dependent T cell mediated cytotoxicity assay. The results showed that, except for clone 1B8-VHH, which only induced profound cytolysis of HT-29 cells at the highest concentration of 10nM, all the other four anti-EpCAM VHH BiTE antibodies (1A5-

VHH, 2B7-VHH, 2C4-VHH and 2D10-VHH) were able to kill 90-100% of EpCAM positive HT-29 (**Figure 7A**), HepG2 (**Figure 7B**), and Hep3B (**Figure 7C**) cells within 96 hours at concentrations as low as 10 pM, but spared EpCAM negative cell HeyA8 (**Figure 7D**). Among the 5 clones, clone 2B7-VHH, 2C4-VHH and 2D10-VHH all showed very potent cell killing efficacy (**Figure 7E**), while 2C4-VHH and 2D10-VHH exhibited better killing specificity than 2B7-VHH (**Figure 7D**).

As 2C4-VHH and 2D10-VHH shared same CDR sequences but the framework of 2C4-VHH was closer to human VH sequences, 2C4-VHH was chosen as the lead VHH clone for further development. A humanized clone hu2C4-VHH was constructed and its antigen binding ability was compared with the parental 2C4-VHH clone using BiTE antibodies paired with a well-characterized anti-CD3 agonist Otk3. Humanization resulted in 10 times reduction in the binding affinity to the antigen protein (**Figure 8A**) and about 2.5 times weaker cell binding (**Figure 8B** and **Figure 8E**). The binding to the CD3 arm was not affected by the VHH arm humanization (**Figure 8D**), while hu2C4-VHH/Otk3 also showed high binding specificity as it didn't bind to EpCAM negative HeyA8 cells, similar to its parental clone (**Figure 8C**).

Killing potency of the BiTE antibodies using hu2C4-VHH/Otk3 was next tested using EpCAM positive MCF-7 cells. Interestingly, use of clone Otk3 as the anti-CD3 arm made the 2C4-VHH BiTE antibody extremely potent to have shown a cell killing EC50 less than 0.1 pM. Although humanized version of hu2C4-VHH/Otk3 BiTE showed a 144.1 times reduction in killing potency as compared to the parental clone, the EC50 was still falling within the range of pico molar (**Figure 9A**, **Figure 9C**), making the **hu2C4-VHH/Otk3 a perfect lead** for the potential downstream applications. Similar to the 2C4-VHH/Otk3 BiTE, hu2C4-VHH/Otk3 also did not kill EpCAM negative HeyA8 cells (**Figure 9B**) nor did it induce interferon- γ (**Figure 9D**) or IL-2 (**Figure 9E**) secretion upon co-culture, as otherwise shown in the case of the benchmark BiTE (MT110, constructed in the same structure as shown in **Figure 3A**).

In summary, our BiTE antibodies derived from llama VHH (**Figures 5 - 7**) and a humanized counterpart hu2C4-VHH (**Figures 8 - 9**) all exerted highly potent and specific killing functions directed to EpCAM expressing cells, which can be used in treating solid tumors of epithelial origins if delivery is restricted locally at the tumor sites.

Example 4 – Experiments determine binding affinity of anti-EpCAM human IgG1 antibodies and anti-EpCAM llama VHH antibodies

Binding affinity of anti-EpCAM human IgG₁ antibodies

Binding affinities of 6 anti-EpCAM IgG₁ antibodies was measured via Bio-Layer Interferometry (BLI) analysis using Octet RED96 system and shown in **Table 1**. All 6 antibody clones exhibited binding affinities to the recombinant EpCAM protein at the nanomolar scales ranging from the lowest affinity at 32.39 nM (clone 1H6) to the highest affinity at 1.432 nM (clone 1D11). Compared to the 2 high affinity clinically tested anti-EpCAM antibodies 3622W94 and ING-1, the *K_{off}* rates of which were 1.8×10^{-5} , and 3.2×10^{-5} respectively, our anti-EpCAM IgG₁ antibodies had a much slower off-rates (6.134×10^{-4} to 3.278×10^{-3}) and resulted in much lower binding affinities than 3622W94 (0.19 nM) and ING-1 (0.16 nM) (Ref - Münz et al. Cancer Cell International 2010, 10:44).

Table 1. Binding affinity (*K_D*), association rate, and dissociation rate (*K_d*) of anti-EpCAM human IgG₁ antibodies to human EpCAM protein measured by Octet BLI analysis.

Clone	<i>K_D</i> (nM)	<i>K_a</i> (1/Ms)	<i>K_d</i> (1/s)
1B6 IgG1	1.778	2.223×10^5	3.953×10^{-4}
1C1 IgG1	1.691	2.592×10^5	4.384×10^{-4}
1C11 IgG1	1.432	1.988×10^5	2.848×10^{-4}
1D4 IgG1	3.818	1.607×10^5	6.134×10^{-4}
1E4 IgG1	3.296	4.841×10^5	1.595×10^{-3}
1H6 IgG1	32.39	1.012×10^5	3.278×10^{-3}

Binding affinity of anti-EpCAM llama VHH antibodies

Binding affinities of 5 anti-EpCAM VHH antibodies was measured via Bio-Layer Interferometry (BLI) analysis using Octet RED96 system and shown in **Table 2**. All 5 VHH antibody clones were expressed as bivalent Fc fusion protein for the assay and exhibited binding affinities to the recombinant EpCAM protein at the nanomolar scales ranging from the lowest affinity at 9.715 nM (clone 2C4-VHH) to the highest affinity at 1.975 nM (clone 1B8-VHH).

Table 2. Binding affinity (*K_D*), association rate, and dissociation rate (*K_d*) of anti-EpCAM llama VHH antibodies (expressed as bivalent Fc fusion) to human EpCAM protein measured by Octet BLI analysis.

Clone	<i>K_D</i> (nM)	<i>K_a</i> (1/Ms)	<i>K_d</i> (1/s)
-------	---------------------------	-----------------------------	----------------------------

1A5 VHH-Fc	9.691	8.019×10^4	7.771×10^{-4}
1B8 VHH-Fc	1.975	2.305×10^5	4.552×10^{-4}
2B7 VHH-Fc	4.399	8.784×10^4	3.864×10^{-4}
2C4 VHH-Fc	9.715	7.997×10^4	7.770×10^{-4}
2D10 VHH-Fc	8.871	8.511×10^4	7.551×10^{-4}

Example 5 – Experiments demonstrating that human CAR T cells secreting anti-EpCAM BiTE molecules showed *in vitro* cytotoxicity in killing EpCAM expressing cells.

5 The anti-EpCAM bi-specific T-cell engagers using 2C4-VHH paired with either the anti-CD3 clone Okt3 (named as “Nb01-013A”) or the anti-CD3 clone used in “MT110” (named as “Nb01-013B”) were used to construct secreted forms of anti-EpCAM BiTE with a His-tag. These anti-EpCAM BiTEs were inserted into different CAR constructs against different tumor antigens, for example, GPC3, CD19 or HER2.

10 To showcase that anti-EpCAM BiTE molecules can be secreted by human T cells and contribute to the cytotoxicity of EpCAM expressing target cell killing, the present inventors constructed anti-GPC3 CAR T cells secreting different anti-EpCAM BiTEs (“MT110”, “Nb01-013A”, and “Nb01-013B”) and named them as GE CAR-BiTE T cells. The present inventors also constructed anti-CD19 CAR T cells secreting different anti-
15 EpCAM BiTEs (“Nb01-013A” or “Nb01-013B”), which were named as “19E CAR-BiTE T” cells (**Figure 10A**). These GE CAR-BiTE T cells and 19E CAR-BiTE T cells were tested for their *in vitro* killing efficacies using Hep3B (GPC3^{High}, EpCAM^{High}), HT-29 (GPC3^{Low}, EpCAM^{High}) and HeyA8 (GPC3^{-ve}, EpCAM^{-ve}) cells.

20 An E:T ratio of 1:1 (adjusted to 10% CAR expression level) was applied in the *in vitro* killing assay with the xCelligence RTCA system. All GE CAR-BiTE T cells did not show killing of HeyA8 (GPC3-ve, EpCAM-ve) cells at the low E:T ratio (**Figure 10D**), however, a strong cytotoxicity was observed in killing EpCAM positive cells Hep3B (**Figure 10B**) and HT-29 (**Figure 10C**). The 19E CAR-BiTE T cells (“Nb01-013A” used as BiTE) showed partial killing of both EpCAM positive cells Hep3B (GPC3^{High}, EpCAM^{High})(**Figure 10B**) and HT-29 (GPC3^{Low}, EpCAM^{High})(**Figure 10C**), indicating that
25 the amount of anti-EpCAM BiTEs secreted by the 19E CAR-BiTE T (Nb01-013A) cells in the current *in vitro* settings exceeded the threshold for BiTE mediated cell killing and was independent of the CAR-specific antigen expression level. On the contrary, the 19E CAR-BiTE T cells (“Nb01-013B” used as BiTE) only showed a minimal killing of HT-29

(GPC3^{Low}, EpCAM^{High}) cells, suggesting that "Nb01-013B" BiTE has a much weaker potency in mediating T cell killing than "Nb01-013A" BiTE (**Figure 10C**).

The present inventors also constructed HE CAR-BiTE T cells by fusing the anti-EpCAM BiTE (sequence of "Nb01-013A" was used) to the anti-HER2 CAR (clone 4D5 from "Trastuzumab" or our in-house anti-HER2 clone F5) lentiviral construct via a P2A cleavable linker (**Figure 11A**). Following the lentiviral transduction and expansion (**Figure 11C**), both AGS (HER2^{High}, EpCAM^{High}) and MDA-MB468 (HER2^{ve}, EpCAM^{High}) cells (**Figure 11B**) were used to test the *in vitro* killing efficacy of anti-HER2 CAR T (4D5 or F5) and HE CAR-BiTE T (4D5 or F5) cells. The HE CAR-BiTE T cells were shown to be able to kill AGS cells at a much faster rate than the anti-HER2 CAR alone (**Figure 11D**), suggesting the participation of the BiTE in target cell killing. Surprisingly, clone 4D5 constructed anti-HER2 CAR T cells could not initiate efficient killing of AGS cells (**Figure 11D**) with low level of Interferon- γ secretion (**Figure 11F**), suggesting that the cytotoxic function of the HE CAR-BiTE T cells (clone 4D5) was mainly contributed by the secreted anti-EpCAM BiTE. More intriguingly, anti-HER2 CAR T cells (clone 4D5) showed comparable level of killing of HER2 negative MDA-MB468 cells, while the anti-HER2 CAR T cells constructed using our patented clone F5 showed high specificity by sparing the HER2 negative cells (**Figure 11E and 11G**).

Example 6 – Experiments showing that anti-EpCAM BiTE secreted by CAR T cells can elicit strong T cell mediated killing of cancer cells

In another experiment, the culture supernatants containing the anti-EpCAM BiTE molecules secreted by human GE CAR-BiTE T (sequence of "MT110" was used) cells were harvested and mixed with naïve human T cells isolated from PBMCs of a healthy donor, before adding to Hep3B (GPC3^{High}, EpCAM^{High}) target cells (E:T ratio = 6:1). The time-course measurement of BiTE elicited T cell-mediated cytotoxicity was performed by xCelligence impedance assay. The results showed that the anti-EpCAM BiTE molecules secreted by GE CAR-BiTE T (MT110) cells could elicit strong killing of Hep3B cells, while there is no significant killing of Hep3B target cells from the cultures with the addition of supernatants not containing the anti-EpCAM BiTE molecules (Mock T or CAR T) (**Figure 12**).

Example 7 – Experiments demonstrating that secretion of anti-EpCAM llama VHH BiTE from different CAR-BiTE T cells can be detected from the *in vitro* cultures

Both anti-GPC3 CAR T (5C4) and GE CAR-BiTE T (Nb01-013A) cells were cultured in T cell growth medium containing IL-7 and IL-15 at the starting cell density of 0.5 million per ml. Cell culture supernatants were collected daily for a consecutive four days, followed by an ELISA to detect the amount of secreted anti-EpCAM BiTE molecule Nb01-013A. The results clearly showed that the amount of secreted Nb01-013A molecules increased over time, suggesting a continuous secretion and accumulation of anti-EpCAM VHH BiTE from the GE CAR-BiTE T (Nb01-013A) cell cultures (**Figure 13**).

Similarly, the Nb01-013A BiTE molecules can also be detected from HE CAR-BiTE T (4D5) and HE CAR-BiTE T (F5) cells from the *in vitro* cultures. Anti-HER2 CAR T (4D5 or F5) cells or HE CAR-BiTE T (4D5 or F5) cells were cultured in T cell growth medium containing IL-7 and IL-15 at the starting cell density of 1.5 million per ml. Cell culture supernatants were collected at 24h and 48h followed by an ELISA to detect the amount of secreted anti-EpCAM BiTE molecule Nb01-013A. For both HE CAR-BiTE T (4D5) and HE CAR-BiTE T (F5), the amount of secreted Nb01-013A molecules increased overtime, suggesting a continuous secretion and accumulation of anti-EpCAM VHH BiTE from both cultures (**Figure 14**).

Example 8 – Experiments demonstrating that human CAR T cells secreting anti-EpCAM BiTE molecules showed superior *in vivo* cytotoxicity in killing EpCAM expressing tumors.

Besides the characterization of the *in vitro* cytotoxicity of human CAR T cells secreting anti-EpCAM BiTE molecules (**Figure 10 and Figure 11**), the present inventors also evaluated their *in vivo* efficacy in killing EpCAM expressing tumors using mouse xenograft models.

First, they used Hep3B (GPC3^{High}, EpCAM^{High}) xenograft models in NSG mice to test anti-GPC3 CAR T cells secreting different anti-EpCAM BiTE molecules (“MT110”, “Nb01-013A”, and “Nb01-013B”) as well as anti-CD19 CAR T cells secreting different anti-EpCAM BiTE molecules (“Nb01-013A” or “Nb01-013B”) (**Figure 15**). Interestingly, Hep3B xenografts from the two 19E CAR-BiTE T cell treated groups of mice showed a similar growth rate as the control groups (“Tumor only” group and “Mock T” group),

suggesting that the *in vivo* tumor control efficacy by CAR-BiTE T cells is dependent on the CAR-specific antigen expression. The anti-GPC3 CAR T (5C4) cells could effectively control the tumor growth in all tested mice, however, after 3 weeks of tumor shrinkage, the remaining tumors started to grow back, suggesting the tumor escape from the GPC3^{Negative} cell populations. In this experiment, the present inventors failed to observe the effect of GE CAR-BiTE T (MT110) cells in controlling the Hep3B xenografts. This was because they had to voluntarily cull the mice when the humane point was hit (tumor sizes exceeded 3000mm³), before the *in vivo* expansion of GE CAR-BiTE T (MT110) cells reached the threshold points in executing a significant tumor control effect. This may be due to the donor variations in the T cells used to generate the CAR T cells in different batches of experiments. Nevertheless, GE CAR-BiTE T ("Nb01-013A" or "Nb01-013B" used as BiTE) cells generated using the same T cell donor showed much potent tumor control efficacies, suggesting that replacing the anti-EpCAM arm with the presently disclosed anti-EpCAM 2C4-VHH significantly improved the *in vivo* expansion as well as the *in vivo* killing potency of the GE CAR-BiTE T cells. GE CAR-BiTE T (Nb01-013A) group of mice achieved efficient tumor control for all tested mice. In addition, three mice remained healthy for more than 90 days after the treatment.

The *in vivo* efficacy of human CAR T cells secreting anti-EpCAM BiTE molecules was also tested using a second xenograft model of AGS (Her2^{High}, EpCAM^{High}) cells treated with anti-HER2 CAR T cells secreting Nb01-013A (**Figure 16**). In line with the *in vitro* data (**Figure 11**), AGS appeared to be resistant to 4D5 CAR-T. While CAR T (4D5) failed to control the tumor growth, the other groups of CAR T (F5) or HE CAR-BiTE T cell treatments showed some *in vivo* tumor control efficacies. CAR T (F5) treatment showed effective control of tumor growth in all 5 mice. However, after 5-6 weeks of tumor shrinkage, the remaining tumors started to grow back and the tumor sizes reached more than 1000mm³, suggesting tumor escape from the Her2^{Negative} cell populations. Effective tumor control was observed in 3 and 2 mice, respectively from the HE CAR-BiTE T (4D5) and HE CAR-BiTE T (F5) group, and these mice remained tumor-free up to 181 days post T cell infusion, showing prolonged tumor control by clearing Her2^{Negative}EpCAM^{Positive} tumor cells.

In summary, this example indicates that human CAR T cells secreting anti-EpCAM BiTE molecules derived from in-house anti-EpCAM 2C4-VHH showed superior *in vivo* cytotoxicity in killing EpCAM expressing tumors.

5 APPLICATIONS

The presently disclosed antigen binding proteins, variants and fragments thereof may be useful for one or more of the following:

- As diagnostic antibodies for solid tumors that were originated from epithelium with high EpCAM expression;
- 10 • Anti-EpCAM BiTEs incorporating the disclosed human Fab clones can be used in treating EpCAM high solid tumors due to their low affinity to the EpCAM antigen protein; and
- Anti-EpCAM BiTEs incorporating the llama VHH clones or humanized VHH clone can be used as locally secreted molecules in treating EpCAM positive solid
15 tumors

CLAIMS

1. An antigen binding protein, an antigen binding variant, or an antigen binding
fragment thereof that binds specifically to Epithelial Cellular Adhesion Molecule
(EpCAM), wherein the antigen binding protein, variant or fragment thereof
comprises a heavy chain variable region and/or a light chain variable region
selected from the group consisting of:
- (i) a heavy chain variable region comprising:
- a CDR-H1 comprising GGTFSYA (**SEQ ID NO: 1**) or a sequence having
at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity
thereto,
- a CDR-H2 comprising IIPFGTA (**SEQ ID NO: 2**) or a sequence having at
least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity
thereto, and
- a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**) or a sequence
having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%
identity thereto
- (ii) a heavy chain variable region comprising:
- a CDR-H1 comprising GDSISSNSVA (**SEQ ID NO: 5**) or a sequence
having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%
identity thereto,
- a CDR-H2 comprising TYYRSKWYS (**SEQ ID NO: 6**) or a sequence
having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%
identity thereto, and
- a CDR-H3 comprising AREVEGSSYDAFDI (**SEQ ID NO: 7**) or a
sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%,
or 100% identity thereto
- (iii) a light chain variable region comprising:
- a CDR-L1 comprising:
- QSLHSNGYNY (**SEQ ID NO: 9**) or a sequence having at least
80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity
thereto,

- QSLHNSRNY (SEQ ID NO: 17) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
 - QSIDF (SEQ ID NO: 19) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto
- 5 a CDR-L2 comprising:
- LGS (SEQ ID NO: 10) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
 - AAS (SEQ ID NO: 20) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
- 10 a CDR-L3 comprising:
- MQALQTPYT (SEQ ID NO: 11) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
 - MQGLQSPWT (SEQ ID NO: 15) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
 - QQSYIMPDT (SEQ ID NO: 21) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
 - MQGLQTPYT (SEQ ID NO: 23) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and
- 15 (iv) a heavy chain variable region comprising:
- 20 a CDR-H1 comprising:
- GSIFSGND (SEQ ID NO: 25) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
 - GSSERFTS (SEQ ID NO: 29) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto
- 25 a CDR-H2 comprising:
- 30

- ITSGGST (**SEQ ID NO: 26**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
 - ITNGGST (**SEQ ID NO: 30**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto and
- a CDR-H3 comprising:
- TNGRWSGDTYYAHH (**SEQ ID NO: 27**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
 - MAGTS (**SEQ ID NO: 31**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
 - TNGRWSGDTYYAHL (**SEQ ID NO: 33**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.
2. The antigen binding protein, variant or fragment thereof of claim 1, wherein the antigen binding protein, variant or fragment thereof comprises a heavy chain variable region selected from the group consisting of:
- (i) a heavy chain variable region comprising:
- a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
 - a CDR-H2 comprising IIPFGTA (**SEQ ID NO: 2**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
 - a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto
- (ii) a heavy chain variable region comprising:

- a CDR-H1 comprising GDSISSNSVA (**SEQ ID NO: 5**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- 5 a CDR-H2 comprising TYYRSKWYS (**SEQ ID NO: 6**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
- a CDR-H3 comprising AREVEGSSYDAFDI (**SEQ ID NO: 7**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and
- 10 comprises a light chain variable region comprising:
- a CDR-L1 comprising:
- QSLHNSGYNY (**SEQ ID NO: 9**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
 - 15 • QSLHNSRYNY (**SEQ ID NO: 17**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
 - QSIDF (**SEQ ID NO: 19**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto
- 20 a CDR-L2 comprising:
- LGS (**SEQ ID NO: 10**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
 - AAS (**SEQ ID NO: 20**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
- 25 a CDR-L3 comprising:
- MQALQTPYT (**SEQ ID NO: 11**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
 - 30 • MQGLQSPWT (**SEQ ID NO: 15**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

- QQSYIMPDT (**SEQ ID NO: 21**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
 - MQGLQTPYT (**SEQ ID NO: 23**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.
3. The antigen binding protein, variant or fragment thereof of claim 1, wherein the antigen binding protein, variant or binding fragment thereof comprises a heavy chain variable region comprising:
- a CDR-H1 comprising:
- GSIFSGND (**SEQ ID NO: 25**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
 - GSSERFTS (**SEQ ID NO: 29**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto
- a CDR-H2 comprising:
- ITSGGST (**SEQ ID NO: 26**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
 - ITNGGST (**SEQ ID NO: 30**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and
- a CDR-H3 comprising:
- TNGRWSGDTYYAHH (**SEQ ID NO: 27**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
 - MAGTS (**SEQ ID NO: 31**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
 - TNGRWSGDTYYAHL (**SEQ ID NO: 33**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

4. The antigen binding protein, variant or fragment thereof of any one of the preceding claims, wherein the antigen binding protein, variant or fragment thereof comprises a heavy chain variable region and/or a light chain variable region selected from the group consisting of:
- 5 (i) a heavy chain variable region comprising:
- a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- 10 a CDR-H2 comprising IIPIFGTA (**SEQ ID NO: 2**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
- a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto
- 15 (ii) a heavy chain variable region comprising:
- a CDR-H1 comprising GDSISSNSVA (**SEQ ID NO: 5**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- 20 a CDR-H2 comprising TYYRSKWYS (**SEQ ID NO: 6**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
- a CDR-H3 comprising AREVEGSSYDAFDI (**SEQ ID NO: 7**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto
- 25 (iii) a light chain variable region comprising:
- a CDR-L1 comprising QSLHSNGYNY (**SEQ ID NO: 9**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- 30 a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-L3 comprising MQALQTPYT (**SEQ ID NO: 11**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto

(iv) a light chain variable region comprising:

5 a CDR-L1 comprising QSLHNSNGYNY (**SEQ ID NO: 9**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

10 a CDR-L3 comprising MQGLQSPWT (**SEQ ID NO: 15**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto

(v) a light chain variable region comprising:

15 a CDR-L1 comprising QSLHNSRNYNY (**SEQ ID NO: 17**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

20 a CDR-L3 comprising MQALQTPYT (**SEQ ID NO: 11**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto

(vi) a light chain variable region comprising:

25 a CDR-L1 comprising QSISDF (**SEQ ID NO: 19**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-L2 comprising AAS (**SEQ ID NO: 20**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

30 a CDR-L3 comprising QQSYIMPDT (**SEQ ID NO: 21**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto

- (vii) a light chain variable region comprising:
- a CDR-L1 comprising QSLHNSGYNY (**SEQ ID NO: 9**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
 - 5 a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
 - a CDR-L3 comprising MQGLQTPYT (**SEQ ID NO: 23**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto
- 10 (viii) a heavy chain variable region comprising:
- a CDR-H1 comprising GSIFSGND (**SEQ ID NO: 25**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
 - 15 a CDR-H2 comprising ITSGGST (**SEQ ID NO: 26**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
 - a CDR-H3 comprising TNGRWSGDTYYAHH (**SEQ ID NO: 27**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto
- 20 (ix) a heavy chain variable region comprising:
- a CDR-H1 comprising GSSERFTS (**SEQ ID NO: 29**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
 - 25 a CDR-H2 comprising ITNGGST (**SEQ ID NO: 30**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
 - a CDR-H3 comprising MAGTS (**SEQ ID NO: 31**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and
- 30 (x) a heavy chain variable region comprising:

- a CDR-H1 comprising GSIFSGND (**SEQ ID NO: 25**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- a CDR-H2 comprising ITSGGST (**SEQ ID NO: 26**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
- a CDR-H3 comprising TNGRWSGDTYYAHL (**SEQ ID NO: 33**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.
- 5
- 10
5. The antigen binding protein, variant or fragment thereof of any one of the preceding claims, wherein the antigen binding protein, variant or fragment thereof comprises a heavy chain variable region and a light chain variable region selected from the group consisting of:
- 15 (i) a heavy chain variable region comprising:
- a CDR-H1 comprising GGTFSYA (**SEQ ID NO: 1**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- a CDR-H2 comprising IPIFGTA (**SEQ ID NO: 2**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
- 20 a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and
- 25 a light chain variable region comprising:
- a CDR-L1 comprising QSLHSNGYNY (**SEQ ID NO: 9**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
- 30

a CDR-L3 comprising MQALQTPYT (**SEQ ID NO: 11**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto

(ii) a heavy chain variable region comprising:

5 a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-H2 comprising IIPFGTA (**SEQ ID NO: 2**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

10 a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and

a light chain variable region comprising:

15 a CDR-L1 comprising QSLHNSNGYNY (**SEQ ID NO: 9**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

20 a CDR-L3 comprising MQGLQSPWT (**SEQ ID NO: 15**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto

(iii) a heavy chain variable region comprising:

25 a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-H2 comprising IIPFGTA (**SEQ ID NO: 2**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

30 a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and

a light chain variable region comprising:

a CDR-L1 comprising QSLHNSRNYNY (**SEQ ID NO: 17**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

5 a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

10 a CDR-L3 comprising MQALQTPYT (**SEQ ID NO: 11**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto

(iv) a heavy chain variable region comprising:

a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

15 a CDR-H2 comprising IIPFGTA (**SEQ ID NO: 2**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

20 a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and

a light chain variable region comprising:

a CDR-L1 comprising QSLHNSGYNY (**SEQ ID NO: 9**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

25 a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

30 a CDR-L3 comprising MQGLQTPYT (**SEQ ID NO: 23**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and

(v) a heavy chain variable region comprising:

- a CDR-H1 comprising GDSISSNSVA (**SEQ ID NO: 5**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- 5 a CDR-H2 comprising TYYSRSKWYS (**SEQ ID NO: 6**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
- a CDR-H3 comprising AREVEGSSYDAFDI (**SEQ ID NO: 7**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and
- 10 a light chain variable region comprising:
- a CDR-L1 comprising QSISDF (**SEQ ID NO: 19**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- 15 a CDR-L2 comprising AAS (**SEQ ID NO: 20**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
- a CDR-L3 comprising QQSYIMPDT (**SEQ ID NO: 21**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.
- 20
6. The antigen binding protein, variant or fragment thereof of any one of the preceding claims, wherein the antigen binding protein, variant or fragment thereof comprises a heavy chain variable region selected from the group consisting of:
- (i) a heavy chain variable region comprising:
- 25 a CDR-H1 comprising GSIFSGND (**SEQ ID NO: 25**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- a CDR-H2 comprising ITSGGST (**SEQ ID NO: 26**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
- 30 a CDR-H3 comprising TNGRWSGDTYYAHH (**SEQ ID NO: 27**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto

- (ii) a heavy chain variable region comprising:
- a CDR-H1 comprising GSSEFRTS (**SEQ ID NO: 29**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
 - 5 a CDR-H2 comprising ITNGGST (**SEQ ID NO: 30**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
 - a CDR-H3 comprising MAGTS (**SEQ ID NO: 31**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and
- (iii) a heavy chain variable region comprising:
- a CDR-H1 comprising GSIFSGND (**SEQ ID NO: 25**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
 - 15 a CDR-H2 comprising ITSGGST (**SEQ ID NO: 26**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and
 - a CDR-H3 comprising TNGRWSGDTYYAHL (**SEQ ID NO: 33**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.
7. The antigen binding protein, variant or fragment thereof of any one of the preceding claims, wherein the antigen binding protein, variant or fragment thereof comprises a heavy chain variable domain and/or a light chain variable domain selected from the group consisting of:
- 25
- (i) a heavy chain variable domain comprising
 EVQLVQSGAEVKKPGSSVKVCKASGGTFSSYAISWVRQAPGQGLE
 WMGGIIPFGTANYAQNFGGRVTMTADTSISTAYMELSSLRSEDNAVY
 YCARSLGGRFRYWGGQTL (**SEQ ID NO: 4 - 1B6, 1C1, 1C11, 1D4**
 30 and 1H6) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions

- (ii) a heavy chain variable domain comprising
 QVQLQQSGPGLVKPSQTLSTCAIS**GDSISSNSVA**WNWIRQSPSRGL
 EWLGR**TYYSK**WYSDYAISVKGRLDINPDTSKNQFSLQLNSVTPEDT
 AVYYC**CAREVEGSSYDAFDI**WGQGT**M (SEQ ID NO: 8)** or a fragment,
 5 variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%,
 98%, 99%, or 100% identical thereto and/or having two or three amino
 acids substitutions,
- (iii) a light chain variable domain comprising
 DVVMTQSPLSLPVTPGEPASISCRSS**QSLH**SN**G**YNYLDWYLQKPGQ
 10 SPQLLIY**LGS**NRASGVDPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**M**
QALQTPYTFGQGTK (SEQ ID NO: 12) or a fragment, variant or
 sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%,
 or 100% identical thereto and/or having two or three amino acids
 substitutions
- (iv) a light chain variable domain comprising
 EIVLTQSPLSLPVTPGEPASISCRSS**QSLH**SN**G**YNYLDWYLQKPGQS
 15 PQLLIY**LGS**NRASGVDPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**MQ**
GLQSPWTFGQGTK (SEQ ID NO: 16) or a fragment, variant or
 sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%,
 20 or 100% identical thereto and/or having two or three amino acids
 substitutions
- (v) a light chain variable domain comprising
 DVVMTQSPLSLPVTPGESASISCRSS**QSLH**SN**R**YNYLDWYLQKPGQ
 25 SPQLLIY**LGS**NRASGVDPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**M**
QALQTPYTFGQGTK (SEQ ID NO: 18) or a fragment, variant or
 sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%,
 or 100% identical thereto and/or having two or three amino acids
 substitutions
- (vi) a light chain variable domain comprising
 30 DIQLTQSPSSLSASVGDRTITCRAS**QSISD**FLNWYQQKPGKAPKLLIY
AASSLQTVPSRFGSGSGTEFTLTISLQPEDLGTYYC**QQSY**IMPDT
FGQGTK (SEQ ID NO: 22) or a fragment, variant or sequence thereof

- at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions
- (vii) a light chain variable domain comprising
- 5 DVVMTQSPLSLPVTGPASISCRSS**QSL**LH**SNGY**NYLDWYLQKPGQ
 SPQLLIY**LGS**NRASGVPDRFSGSGSGTDFTLQISRVEAEDAGVYY**CM**
QGLQTPYTFGQGTK (**SEQ ID NO: 24**) or a fragment, variant or
 sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%,
 or 100% identical thereto and/or having two or three amino acids
 substitutions
- 10 (viii) a heavy chain variable domain comprising
 QVQLQESGGGLVQPGGSLRLSCAD**S**G**SIFSG**NDAWYRRAPGVERE
 LVAVIT**S**G**G**STHYADSVKGRFTISRDN**AQK**TVYLQTN**DLK**PEDTAVYY
 CT**N**GRW**S**GD**TYA**HHWGQGTQ (**SEQ ID NO: 28**) or a fragment,
 variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%,
 15 98%, 99%, or 100% identical thereto and/or having two or three amino
 acids substitutions
- (ix) a heavy chain variable domain comprising
 QVQLQESGGGLVQPGGSLRLSCA**A**S**G**S**SER**F**T**SVAWYRQAPGKERE
 LVAFIT**N**G**G**STRYTDPVKGRFTISRDN**A**KNTVYLQ**M**NSLKAEDTAVYY
 20 **CM**AGTSWGQGTQ (**SEQ ID NO: 32**) or a fragment, variant or
 sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%,
 or 100% identical thereto and/or having two or three amino acids
 substitutions
- (x) a heavy chain variable domain comprising
 25 QVQLQESGGGLVQPGGSLRLSCAD**S**G**SIFSG**NDAWYRRAPGVERE
 LVAVIT**S**G**G**STHYADSVKGRFTISRDN**AQK**TVYLQTN**DLK**PEDTAVYY
 CT**N**GRW**S**GD**TYA**HLWGQGTQ (**SEQ ID NO: 34**) or a fragment,
 variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%,
 98%, 99%, or 100% identical thereto and/or having two or three amino
 30 acids substitutions
- (xi) a heavy chain variable domain comprising
 QVQLQESGGGLVQAGGSLRLSCAD**S**G**SIFSG**NDAWYRRAPGVERE
 LVAVIT**S**G**G**STHYADSVKGRFTISRDN**AQK**TVYLQTN**DLK**PEDTAVYY

- CTNGRWSGDTYYAHHWGQGTQ (**SEQ ID NO: 35**) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions
- 5 (xii) a heavy chain variable domain comprising
 QVQLQESGGGLVQAGDSLRLSCADSGSIFSGNDMAWYRRAPGVERE
 LVAVITSGGSTHYADSVKGRFTISRDNAAQKTVYLQTNLKPEDTAVYY
 CTNGRWSGDTYYAHHWGQGTQ (**SEQ ID NO: 36**) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions; and
- 10 a heavy chain variable domain comprising
 QVQLVESGGGLVQAGGSLRLSCAASGSIFSGNDMSWYRQAPGKGLE
 LVAVITSGGSTYYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYY
 CTNGRWSGDTYYAHHWGQGTQ (**SEQ ID NO: 37**) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions.
- 15
- 20 8. The antigen binding protein, variant or fragment thereof of any one of the preceding claims, wherein the antigen binding protein, variant or fragment thereof comprises a heavy chain variable domain and a light chain variable domain selected from the group consisting of:
- (i) a heavy chain variable domain comprising:
- 25 EVQLVQSGAEVKKPGSSVKVCKASGGTFSSYAISWVRQAPGQGLEWM
 GGIPIFGTANYAQNFQGRVTMTADTSISTAYMELSSLRSEDTAVYYCARS
 LGGRFRYWQGTQ (**SEQ ID NO: 4**) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions, and
- 30 a light chain variable domain comprising
 DVVMTQSPLSLPVTPGEPASISCRSSQSLHNSGYNVLDWYLQKPGQSP
 QLLIYLGSNRASGVPRFSGSGSGTDFTLKISRVEAEDVGVYYCMQALQT
 PYTFGQGTK (**SEQ ID NO: 12**) or a fragment, variant or sequence thereof

at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions

(ii) a heavy chain variable domain comprising:

5 EVQLVQSGAEVKKPGSSVKVSKASGGTFSSYAISWVRQAPGQGLEWM
GGIIPFGTANYAQNFGGRVTMTADTSISTAYMELSSLRSEDTAVYYCARS
LGGRRFRYWGGTL (SEQ ID NO: 4) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions, and
10 a light chain variable domain comprising
EIVLTQSPSLPVTGPASISCRSSQSLLSHNGYNYLDWYLQKPGQSPQ
LLIYLGSNRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYCMQGLQSP
WTFGQGTK (SEQ ID NO: 16) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions

15 (iii) a heavy chain variable domain comprising:

EVQLVQSGAEVKKPGSSVKVSKASGGTFSSYAISWVRQAPGQGLEWM
GGIIPFGTANYAQNFGGRVTMTADTSISTAYMELSSLRSEDTAVYYCARS
LGGRRFRYWGGTL (SEQ ID NO: 4) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions, and
20 a light chain variable domain comprising
DVVMTQSPSLPVTGPASISCRSSQSLLSNRNYNYLDWYLQKPGQSP
QLLIYLGSNRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYCMQALQT
PYTFGQGTK (SEQ ID NO: 18) or a fragment, variant or sequence thereof
25 at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions

(iv) a heavy chain variable domain comprising:

30 QVQLQQSGPGLVKPSQTLSTCAISGDSISNSVAWNWIRQSPSRGL
EWLGRITYYRSKWYSDYAIKVGRLDINPDTSKNQFSLQLNSVTPEDT
AVYYCAREVEGSSYDAFDIWGGTM (SEQ ID NO: 8) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions, and

- a light chain variable domain comprising
 DIQLTQSPSSLSASVGDRTITCRAS**QSIDFLN**WYQQKPGKAPKLLIYAA
 SSLQTGVPSRFGGSGSGTEFTLTISLQPEDLGTY**YCQQSYIM**PDTFGQG
 TK (**SEQ ID NO: 22**) or a fragment, variant or sequence thereof at least
 5 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto
 and/or having two or three amino acids substitutions, or
- (v) a heavy chain variable domain comprising:
 EVQLVQSGAEVKKPGSSVKVCKAS**G**GTFSSYAISWVRQAPGGLEWM
 GGIIPIFGTANYAQNFGGRVTMTADTSISTAYMELSSLRSEDTAVYYC**CARS**
 10 **LGGRFRYWGQ**TL (**SEQ ID NO: 4**) or a fragment, variant or sequence
 thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%
 identical thereto and/or having two or three amino acids substitutions, and
 a light chain variable domain
 DVVMTQSPLSLPVTPGEPASISCRSS**Q**LLHSNGYNYLDWYLQKPGQSP
 15 **QLLIYLG**SNRASGVPDRFSGSGSGTDFTLQISRVEAEDAGVYYC**MQGLQT**
PYTFGQGTK (**SEQ ID NO: 24**) or a fragment, variant or sequence thereof
 at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical
 thereto and/or having two or three amino acids substitutions.
- 20 9. The antigen binding protein, variant or fragment thereof of any one of the
 preceding claims, wherein the antigen binding protein, variant or fragment thereof
 comprises a single domain heavy chain variable domain having a sequence:
- (i) QVQLQESGGGLVQPGGSLRLSCAD**SGSIFSGND**MAWYRRAPGVERE
 LVAVIT**SGGST**HYADSVKGRFTISRDN**AQKT**VYLTNDLKPEDTAVYY
 25 **CTNGRWSGDTYYA**HHWGQGTQ (**SEQ ID NO: 28**) or a fragment,
 variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%,
 98%, 99%, or 100% identical thereto and/or having two or three amino
 acids substitutions, or
- (ii) QVQLQESGGGLVQPGGSLRLSCAAS**GSSERFTS**VAWYRQAPGKERE
 30 LVAFIT**NGGSTRY**TDPVKGRFTISRDN**AKNT**VYLMNSLKAEDTAVYY
CMAGTSWGQGTQ (**SEQ ID NO: 32**) or a fragment, variant or sequence
 thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%
 identical thereto and/or having two or three amino acids substitutions, or

- (iii) QVQLQESGGGLVQPGGSLRLSCADSGSIFSGNDMAWYRRAPGVERE
 LVAVITSGGSTHYADSVKGRFTISRDNAAKTVYLQTNDLKPEDTAVYY
 CTNGRWSGDTYYAHLWGQGTQ (SEQ ID NO: 34) or a fragment,
 variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%,
 98%, 99%, or 100% identical thereto and/or having two or three amino
 acids substitutions, or
- (iv) QVQLQESGGGLVQAGGSLRLSCADSGSIFSGNDMAWYRRAPGVERE
 LVAVITSGGSTHYADSVKGRFTISRDNAAKTVYLQTNDLKPEDTAVYY
 CTNGRWSGDTYYAHHWGQGTQ (SEQ ID NO: 35) or a fragment,
 variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%,
 98%, 99%, or 100% identical thereto and/or having two or three amino
 acids substitutions, or
- (v) QVQLQESGGGLVQAGDSLRLSCADSGSIFSGNDMAWYRRAPGVERE
 LVAVITSGGSTHYADSVKGRFTISRDNAAKTVYLQTNDLKPEDTAVYY
 CTNGRWSGDTYYAHHWGQGTQ (SEQ ID NO: 36) or a fragment,
 variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%,
 98%, 99%, or 100% identical thereto and/or having two or three amino
 acids substitutions, or
- (vi) QVQLVESGGGLVQAGGSLRLSCAASGSIFSGNDMSWYRQAPGKGLE
 LVAVITSGGSTYYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYY
 CTNGRWSGDTYYAHHWGQGTQ (SEQ ID NO: 37) or a fragment,
 variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%,
 98%, 99%, or 100% identical thereto and/or having two or three amino
 acids substitutions.
10. The antigen binding protein, variant or fragment thereof of any one of the
 preceding claims, wherein the antigen binding protein, variant or fragment thereof
 comprises a light chain constant domain having a sequence:
- (i) RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSG
 NSQESVAEQDSKDSTYLSSTLTLSKADYEKHKLYACEVTHQGLSSPVTK
 SFNRGEC (SEQ ID NO: 13) or a fragment, variant or sequence thereof at
 least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto
 and/or having two or three amino acids substitutions, or

(ii) RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSG
NSQESVTEQDSKDSTYLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTK
SFSRGEC (SEQ ID NO: 14) or a fragment, variant or sequence thereof at
least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto
and/or having two or three amino acids substitutions.

5

11. The antigen binding protein, variant or fragment thereof according to any one of
the preceding claims, wherein the antigen binding protein, variant or fragment
thereof is an IgG antibody, in particular an IgG1 antibody.

10

12. The antigen binding protein, variant or fragment thereof according to any one of
the preceding claims, wherein the antigen binding protein, variant or fragment
thereof is a multi-specific antigen binding protein, variant or fragment thereof,
such as a bispecific antibody.

15

13. The antigen binding protein, variant or fragment thereof according to claim 12,
wherein the multi-specific antigen binding protein, variant or fragment thereof
binds to an immune marker selected from the group consisting of CD3, NKG2D,
CD4, CD8, CD16 and CD64.

20

14. The antigen binding protein, variant or fragment thereof according to any one of
the preceding claims, wherein the antigen binding protein, variant or fragment
thereof is a bispecific T cell engager (BiTE).

25

15. The antigen binding protein, variant or fragment thereof of claim 14, wherein the
bispecific T cell engager (BiTE) comprises an anti-EpCAM Heavy chain antibody
variable region (i.e. VHH) or a single chain variable fragment (scFv).

30

16. The antigen binding protein, variant or fragment thereof according to any one of
the preceding claims, wherein the antigen binding protein, variant or fragment
thereof comprises a Fc region.

17. A polynucleotide encoding the antigen binding protein, variant or fragment thereof of any one of the preceding claims.
18. The polynucleotide of claim 17, wherein the heavy chain variable domain is encoded by a nucleotide sequence comprising:
- 5
- (i) GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
 GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCTGGAGGCACCTTCAG
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCT
 TGAGTGGATGGGAGGGATCATCCCTATCTTTGGTACAGCAAACTA
 10 CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
 CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
 CACGGCCGTGTATTACTGTGCGAGATCGTTGGGTGGGAGATTTCG
CTACTGGGGCCAGGGAACCCTG (SEQ ID NO: 41) or a sequence at
 least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical
 thereto and/or having 10-20 nucleic acid substitutions, or
- 15
- (ii) CAGGTACAGCTGCAGCAGTCAGGTCCAGGGCTGGTGAAGCCCTC
 GCAGACCCTCTCACTCACCTGTGCCATCTCCGGGGACAGTATCTC
TAGTAACAGTGTTGCTTGGAACTGGATCAGGCAGTCCCCATCGAG
 AGGCCTTGAGTGGCTGGGAAGG**ACATACTACAGGTCCAAGTGGT**
 20 **ACAGT**GATTATGCAATATCTGTGAAAGGTGATTAGACATCAACCC
 AGACACATCCAAGAACCAGTTCTCCCTGCAGCTGAACTCTGTGACT
 CCGGAGGACACGGCTGTGTATTATTGT**GCAAGAGAAGTTGAGGGC**
AGCAGCTATGATGCTTTTGATATCTGGGGCCAAGGGACAATG
(SEQ ID NO: 45) or a sequence at least 80%, 85%, 90%, 95%, 96%,
 25 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic
 acid substitutions;
- (iii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
 AGGGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG**
TGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
 30 GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTAC**ACACTATG
 CAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATGCCC
 AGAAGACCGTATATCTGCAAACGAACGACTTGAAACCTGAGGACAC
 GGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTA**

CTATGCCCCATCACTGGGGCCAGGGGACCCAG (SEQ ID NO: 65) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions,

(iv) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
GGGGTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCTCCGAAAG**
ATTCACATCAGTGGCCTGGTACCGCCAGGCTCCAGGAAAGGAGC
GCGAGTTGGTTCGATTT**ATTACTAATGGTGGTAGCACA**AGATATAC
AGACCCCGTGAAGGGCCGATTACCATCTCCAGAGACAACGCCAA
GAACACGGTGTATCTGCAAATGAACAGCCTGAAAGCTGAGGACAC
GGCCGTCTATTATTGT**TATGGCGGGTACGTCCT**GGGGCCAGGGGAC
CCAG (**SEQ ID NO: 69**) or a sequence at least 80%, 85%, 90%, 95%,
96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20
nucleic acid substitutions,

(v) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
AGGGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG**
TGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTAC**ACTATG
CAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATGCCC
AGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACA
CGGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTT**
ACTATGCCCCATCTCTGGGGCCAGGGGACCCAG (**SEQ ID NO: 71**) or
a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or
100% identical thereto and/or having 10-20 nucleic acid substitutions,

(vi) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGG
GGGCTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG**
TGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTAC**ACTATG
CAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATGCCC
AGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACA
CGGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTT**
ACTATGCCCCATCACTGGGGCCAGGGGACCCAG (**SEQ ID NO: 72**) or
a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or
100% identical thereto and/or having 10-20 nucleic acid substitutions,

- (vii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGG
GGACTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGT**
GGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCG
CGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACAC**ACTATGC
5 AGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATGCCCA
GAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACAC
GGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTA**
CTATGCCCATCACTGGGGCCAGGGGACCCAG (**SEQ ID NO: 73**) or a
sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%
10 identical thereto and/or having 10-20 nucleic acid substitutions, or
- (viii) CAGGTGCAGCTGGTGGAGTCTGGGGGAGGATTGGTGCAGG
CTGGGGGCTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCATCT**
TCAGTGGCAATGACATGTCTGGTACCGCCAGGCTCCAGGGAAG
GGACTCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACATAC**
15 TATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATT
CCaAGAAcACCcTATATCTGCAAATGAACAGCCTGAGAGCTGAGGA
CACGGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATA**
CTTACTATGCCCATCACTGGGGCCAGGGGACCCTG (**SEQ ID NO:**
74) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%,
20 or 100% identical thereto and/or having 10-20 nucleic acid substitutions;
and/or
- the light chain variable domain is encoded by a nucleotide sequence comprising:
- (i) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
25 **ATAGTAATGGATACA**ACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
CGGGGTCCCTGACAGGTTCA GTGGCAGTGGATCAGGCACAGATTT
TACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGCATG**CAAGCTCTACAACTCCGTACACT**TTTGGCCAGGGG
30 ACCAAG (**SEQ ID NO: 49**) or a sequence at least 80%, 85%, 90%, 95%,
96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20
nucleic acid substitutions, or

- (ii) GAAATTGTGCTGACTCAGTCTCCACTCTCCCTGCCCCGTACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
5 CGGGGTCCCTGACAGGTTCA GTGGCAGTGGATCAGGTACAGATTT
TACACTGAAAATAAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGCAT**GCAAGGTCTACAAAGTCCCTGGACG**TTTCGGCCAAGGG
ACCAAG (**SEQ ID NO: 53**) or a sequence at least 80%, 85%, 90%, 95%,
96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20
10 nucleic acid substitutions, or
- (iii) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCCGTACCCCTG
GAGAGTCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATAGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
15 CGGGGTCCCTGACAGGTTCA GTGGCAGTGGATCAGGCACAGATTT
TACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGCAT**GCAAGCTCTACAACTCCGTACACT**TTTGGCCAGGGG
ACCAAG (**SEQ ID NO: 55**) or a sequence at least 80%, 85%, 90%, 95%,
96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20
20 nucleic acid substitutions, or
- (iv) GACATCCAGTTGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAG
GAGACAGAGTCACCATCACTTGCCGGGCAAGT**CAGAGTATTAGCG**
ACTTTTTAAATGGTACCAGCAGAAACCAGGTAAAGCCCCGAAGCT
CCTGATCTAT**GCTGCATCGAGTTT**ACAACTGGGGTCCCCCTCAAGA
25 TTCGGTGGCAGTGGATCTGGGACAGAATCACTCTCACCATAAGCA
GTCTACAACCTGAAGATTTGGGAATTATTACTGT**CAACAGAGTTA**
CATTATGCCCCGACACTTTTGGCCAGGGGACGAAA (**SEQ ID NO: 59**)
or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or
100% identical thereto and/or having 10-20 nucleic acid substitutions, or
30 GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCCGTACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC

CGGGGTCCCTGACAGGTTCACTGGCAGTGGATCAGGCACAGATTT
TACACTGCAAATCAGCAGAGTGGAGGCTGAGGATGCTGGGGTTTA
TTACTGCATGCAAGGTCTACAGACTCCGTACACTTTTGGCCAGGG
GACCAAG (**SEQ ID NO: 61**) or a sequence at least 80%, 85%, 90%,
5 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-
20 nucleic acid substitutions.

19. The polynucleotide according to any one of claims 17 or 18, wherein the heavy
chain and light chain variable domains are encoded by nucleotide sequences
10 selected from the group consisting of:

(i) a heavy chain variable domain encoded by the nucleotide sequence
comprising:

GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
GTCCCTCGGTGAAGGTCTCCTGCAAGGCTTCTGGAGGCACCTTCAG
15 **CAGCTATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCT**
TGAGTGGATGGGAGGGATCATCCCTATCTTTGGTACAGCAAATA
CGCACAGAACTCCAGGGCAGAGTACCATGACCGCAGACACCTC
CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
CACGGCCGTGTATTACTGTGCGAGATCGTTGGGTGGGAGATTTCG
20 **CTACTGGGGCCAGGGAACCCTG** (**SEQ ID NO: 41**) or a sequence at
least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical
thereto and/or having 10-20 nucleic acid substitutions, and
a light chain variable domain encoded by the nucleotide sequence
comprising:

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGTCAGAGCCTCCTGC
25 **ATAGTAATGGATACAACCTATTTGGATTGGTACCTGCAGAAGCCAG**
GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
CGGGGTCCCTGACAGGTTCACTGGCAGTGGATCAGGCACAGATTT
TACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
30 TACTGCATGCAAGCTCTACAACTCCGTACACTTTTGGCCAGGGG
ACCAAG (**SEQ ID NO: 49**) or a sequence at least 80%, 85%, 90%, 95%,

96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions

(ii) a heavy chain variable domain encoded by the nucleotide sequence comprising:

5 GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
 GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCT
 TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCA**AACTA
 CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
 10 CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
 CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
CTACTGGGGCCAGGGAACCCTG (SEQ ID NO: 41) or a sequence at

least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, and

15 a light chain variable domain encoded by the nucleotide sequence comprising:

GAAATTGTGCTGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
 GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTGGATTGGTACCTGCAGAAGCCAG
 20 GGCAGTCTCCACAGCTCCTGATCTATTTGGTTCTAATCGGGCCTC
 CGGGGTCCCTGACAGGTTCA GTGGCAGTGGATCAGGTACAGATTT
 TACTG**CAATGCAAGGTCTACAAAGTCCCTGGACG**TTCCGGCCAAGGG
 ACCAAG (SEQ ID NO: 53) or a sequence at least 80%, 85%, 90%, 95%,

25 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions

(iii) a heavy chain variable domain encoded by the nucleotide sequence comprising:

30 GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
 GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCT
 TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCA**AACTA
 CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC

- 5 CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
CTACTGGGGCCAGGGAACCCTG (SEQ ID NO: 41) or a sequence at
least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical
thereto and/or having 10-20 nucleic acid substitutions, and
a light chain variable domain encoded by the nucleotide sequence
comprising:
GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCCGTACCCCTG
GAGAGTCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
10 **ATAGTAATAGATACTAACTATTTGGATTGGTACCTGCAGAAGCCAG**
GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
CGGGGTCCCTGACAGGTTCACTGGCAGTGGATCAGGCACAGATT
TACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGC**ATGCAAGCTCTACAACTCCGTACACTTTTGGCCAGGGG**
15 **ACCAAG (SEQ ID NO: 55)** or a sequence at least 80%, 85%, 90%, 95%,
96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20
nucleic acid substitutions
(iv) a heavy chain variable domain encoded by the nucleotide sequence
comprising:
20 GAGGTCCAGCTGGTGCACTCTGGGGCTGAGGTGAAGAAGCCTGG
GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCT
TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCAACTA**
CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
25 CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
CTACTGGGGCCAGGGAACCCTG (SEQ ID NO: 41) or a sequence at
least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical
thereto and/or having 10-20 nucleic acid substitutions, and
30 a light chain variable domain encoded by the nucleotide sequence
comprising:
GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCCGTACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**

ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
 GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
 CGGGGTCCCTGACAGGTTCAAGTGGCAGTGGATCAGGCACAGATTT
 TAACTGCAAATCAGCAGAGTGGAGGCTGAGGATGCTGGGGTTTA
 5 TTAAGTGCAT**GCAAGGCTACAGACTCCGTACACT**TTTGGCCAGGG
 GACCAAG (**SEQ ID NO: 61**) or a sequence at least 80%, 85%, 90%,
 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-
 20 nucleic acid substitutions; and

(v) a heavy chain variable domain encoded by the nucleotide sequence
 10 comprising:

CAGGTACAGCTGCAGCAGTCAGGTCCAGGGCTGGTGAAGCCCTC
 GCAGACCCTCTCACTCACCTGTGCCATCTCCGGGGACAGTATCTC
TAGTAACAGTGTGCTTGGAAGTGGATCAGGCAGTCCCCATCGAG
 AGGCCTTGAGTGGCTGGGAAGG**ACATACTACAGGTCCAAGTGGT**
 15 **ACAGT**GATTATGCAATATCTGTGAAAGGTCGATTAGACATCAACCC
 AGACACATCCAAGAACCAGTTCTCCCTGCAGCTGAACTCTGTGACT
 CCCGAGGACACGGCTGTGTATTATTGT**GCAAGAGAAGTTGAGGGC**
AGCAGCTATGATGCTTTTGATATCTGGGGCCAAGGGACAATG
 (**SEQ ID NO: 45**) or a sequence at least 80%, 85%, 90%, 95%, 96%,
 20 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic
 acid substitutions, and

a light chain variable domain encoded by the nucleotide sequence
 comprising:

GACATCCAGTTGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAG
 25 GAGACAGAGTCACCATCACTTGCCGGGCAAGT**CAGAGTATTAGCG**
ACTTTTTAAATTGGTACCAGCAGAAACCAGGTAAGCCCCGAAGCT
 CCTGATCTAT**GCTGCATCGAGTTT**ACAACTGGGGTCCCCCAAGA
 TTCGGTGGCAGTGGATCTGGGACAGAATCACTCTCACCATAAGCA
 GTCTACAACCTGAAGATTTGGGAAGTTATTACTGT**CAACAGAGTTA**
 30 **CATTATGCCCCGACACTTTT**GGCCAGGGGACGAAA (**SEQ ID NO: 59**)
 or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or
 100% identical thereto and/or having 10-20 nucleic acid substitutions.

20. The polynucleotide according to any one of claims 17 or 18, wherein the single domain heavy chain variable domain is encoded by a nucleotide sequence comprising

- (i) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
 5 AGGGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG**
TGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
 GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTAC**ACACTATG
 CAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATGCCC
 AGAAGACCGTATATCTGCAAACGAACGACTTGAAACCTGAGGACAC
 10 GGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTA**
CTATGCCCATCACTGGGGCCAGGGGACCCAG (**SEQ ID NO: 65**) or a
 sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%
 identical thereto and/or having 10-20 nucleic acid substitutions,
- (ii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
 15 GGGGTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCTCCGAAAG**
ATTCACATCAGTGGCCTGGTACCGCCAGGCTCCAGGAAAGGAGC
 GCGAGTTGGTCGCATTT**ATTACTAATGGTGGTAGCACA**AGATATAC
 AGACCCCGTGAAGGGCCGATTACCATCTCCAGAGACAACGCCAA
 GAACACGGTGTATCTGCAAATGAACAGCCTGAAAGCTGAGGACAC
 20 GGCCGTCTATTATTGT**TATGGCGGGTACGTCCT**GGGGCCAGGGGAC
 CCAG (**SEQ ID NO: 69**) or a sequence at least 80%, 85%, 90%, 95%,
 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20
 nucleic acid substitutions,
- (iii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
 25 AGGGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG**
TGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
 GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTAC**ACACTATG
 CAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATGCCC
 AGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACA
 30 CGGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTT**
ACTATGCCCATCTCTGGGGCCAGGGGACCCAG (**SEQ ID NO: 71**) or
 a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or
 100% identical thereto and/or having 10-20 nucleic acid substitutions,

- (iv) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGG
GGGCTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG**
TGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACACACTATG**
5 CAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATGCCC
AGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACA
CGGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTT**
ACTATGCCCATCACTGGGGGCCAGGGGACCCAG (**SEQ ID NO: 72**) or
a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or
10 100% identical thereto and/or having 10-20 nucleic acid substitutions,
- (v) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGG
GGACTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGT**
GGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCG
CGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACACACTATGC**
15 AGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATGCCCA
GAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACAC
GGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTA**
CTATGCCCATCACTGGGGGCCAGGGGACCCAG (**SEQ ID NO: 73**) or a
sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%
20 identical thereto and/or having 10-20 nucleic acid substitutions, or
- (vi) CAGGTGCAGCTGGTGGAGTCTGGGGGAGGATTGGTGCAGGCTGG
GGGCTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCATCTTCAG**
TGGCAATGACATGTCTGGTACCGCCAGGCTCCAGGGAAGGGACT
CGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACATACTATGC**
25 AGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCaAG
AAcACCcTATATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGG
CCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTACT**
ATGCCCATCACTGGGGGCCAGGGGACCCTG (**SEQ ID NO: 74**) or a
sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%
30 identical thereto and/or having 10-20 nucleic acid substitutions.

21. The polynucleotide according to any one of claims 17 to 19, wherein the light chain constant domain is encoded by a nucleotide sequence comprising

- 5 (i) CGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATG
AGCAGTTGAAATCTGGAAGTGCCTCTGTTGTGTGCCTGCTGAATAA
CTTCTATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGC
CCTCCAATCGGGTAACTCCCAGGAGAGTGTGCGAGAGCAGGACAG
CAAGGACAGCACCTACAGCCTCAGCAGCACCTGACGCTGAGCAA
AGCAGACTACGAGAAACACAACTCTACGCCTGCGAAGTCACCCAT
CAGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCAACAGGGGAGA
GTGT (**SEQ ID NO: 50**) or a sequence at least 80%, 85%, 90%, 95%,
96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20
10 nucleic acid substitutions, or
- 15 (ii) CGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATG
AGCAGTTGAAATCTGGAAGTGCCTCTGTTGTGTGCCTGCTGAATAA
CTTCTATCCCAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGC
CCTCCAATCGGGTAACTCCCAGGAGAGTGTGCGAGAGCAGGACAG
CAAGGACAGCACCTACAGCCTCAGCAGCACCTGACGCTGAGCAA
AGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCA
TCAGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCAACAGGGGAGA
GTGT (**SEQ ID NO: 51**) or a sequence at least 80%, 85%, 90%, 95%,
96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20
20 nucleic acid substitutions.
22. A vector expressing the polynucleotide of any one of claims 17 to 21.
23. A host cell comprising the vector of claim 22.
- 25 24. A cell expressing/secreting the antigen binding protein, variant of fragment thereof according to any one of claims 1 to 16.
25. A cell expressing/secreting an immune cell engager which is specific to EpCAM.
- 30 26. The cell according to claim 25, wherein the immune cell engager is selected from the group comprising a T cell engager, an NK cell engager, a monocyte engager and a macrophage engager.

27. The cell of claims 25 or 26, wherein the immune cell engager is a bispecific T cell engager (BiTE), such as an inducible BiTE, a non-inducible BiTE or a constitutive expression BiTE, comprising the antigen binding protein, variant or fragment thereof of any one of claims 1 to 15.
- 5 28. The cell of any one of claims 24 to 27, wherein the cell is an immune cell, for example selected from the group comprising a T cell, a macrophage, a monocyte, and an NK cell.
- 10 29. The cell of any one of claims 24 to 28, wherein the immune cell is a T-cell, in particular a CAR T-cell.
30. The cell of any one of claims 24 to 27, wherein the cell is a stem cell, for example selected from the group comprising a mesenchymal stem cell, a neural stem cell and a pluripotent stem cell, such as an induced pluripotent stem cell (iPSC).
- 15 31. A composition comprising the antigen binding protein, variant or fragment thereof of any one of claims 1 to 16 and/or the cell of any one of claims 24 to 30.
- 20 32. An antigen binding protein, variant or fragment thereof of any one of claims 1 to 16, cell of any one of claims 24 to 30, or a composition of claim 31 for use in treatment of a disease.
- 25 33. A method of treating a disease in a subject in need thereof, comprising administering to the subject an antigen binding protein, variant or fragment thereof of any one of claims 1 to 16, a cell according to any one of claims 24 to 30 or a composition of claim 31.
- 30 34. Use of the antigen binding protein, variant or fragment thereof of any one of claims 1 to 16, a cell according to any one of claims 24 to 30 or the composition of 31 in the manufacture of a medicament for preventing and/or treating a disease.

35. The method, use, antigen binding protein, variant or fragment thereof for use, cell or composition for use according to any one of claims 32 to 34, wherein the disease is a proliferative disease such as a tumor or cancer.
- 5 36. A method of diagnosis / determining the prognosis or presence of a solid tumor originating from epithelium comprising detecting the expression of high EpCAM on the solid tumor using the antigen binding protein, variant or fragment thereof of any one of claims 1 to 16 or a composition of claim 31.