

TARGET-SPECIFIC EXTRACELLULAR VESICLES

FIELD OF THE INVENTION

The invention relates to the field of engineering target-specific extracellular vesicles. The invention further relates to methods of production of said target-specific extracellular vesicles.

BACKGROUND OF THE INVENTION

In the last decade, the research into exosomes has intensified as they have been recognized as significant mediators of cell-to-cell communication [1]. Exosomes are released from multivesicular bodies upon their fusion with the plasma membrane, and released vesicles function as delivery vehicles transferring functional RNAs, exosomal DNA and transmembrane proteins including receptors to cells in the surrounding environment [2]. Regarding their advantages as a potential therapeutic moiety, including favorable properties such as low immunogenicity and low cytotoxicity, the interest in their application has triggered further development, such as new methods for encapsulation, improved cytosolic release of exosomal contents, enhanced cellular uptake and more specific cellular targeting [3].

Vesicular uptake of exosomes is cell type-specific and can involve membrane fusion or endocytosis, and can even be induced by stimulation of oncogenic cancer receptors [4]. To achieve tissue specific delivery, targeting of exosomes can be optimized by engineering the source cells to overexpress exosomal membrane proteins, such as tetraspanins, harboring receptor-specific ligand peptides as recognition units [5]. Tetraspanins are known as molecular facilitators, associating in large cell-signaling complexes known as the tetraspan web, which also involves members of other protein families, such as integrins and coreceptor molecules. Furthermore, such large membrane protein assemblies may be associated with lipid rafts [6-7]. The function of tetraspanins CD9, CD63 and CD81 ligands for endocytosis of exosomes has been reported [8-10], however the mechanism of uptake has not yet been clarified.

In spite of its almost ubiquitous distribution, the tetraspan protein CD81 is the major protein enriched in the exosomal fraction of multivesicular bodies [11]. The large

extracellular loop (LEL) of CD81, topologically located between transmembrane domains 3 and 4, is characterized by five helical elements forming a mushroom-like structure [12-13], stabilized by two pairs of cysteines. This motif is conserved among the protein members of tetraspanin family [14] and the oxidation of cysteine bonds is a prerequisite for high-affinity binding of the E2 envelope protein of hepatitis C virus (HCV), the natural ligand of CD81 [15]. Correct pairing of the cysteines is described for the recognition by the antibody M38 [16], which does not bind to denatured or reduced proteins, but can react with the membrane-bound hCD81 as well as a native form of purified soluble hCD81.

The crystal structure of hCD81 LEL, solved at 1.6 Å, revealed a new type of protein fold [12], and a subsequent sequence analysis of 160 tetraspanin family members indicated that their fold and key structural features are conserved [17]. Apart from cysteine bridges, the hCD81 LEL is stabilized by the invariant residues Gly157 and Pro176, which are located to accommodate cysteine connections, as well as Tyr127, which is fully buried and contributes to the hydrogen bonding network together with His151 and Cys190. Soluble hCD81 LEL assembles into dimers around a 2-fold axis, and the contact between the protomers is a low-polarity region between helices of each interacting partner and between helix B and C-terminal residues of the opposite protomer. The N- and C-termini of the protomers fall in the central region on opposite faces of the assembled dimer, similar to the dimeric assembly at the cell surface, where transmembrane segments are also present. A second low-polarity region comprises the solvent-exposed surface of helices C and D, which is energetically unfavorable. According to the solution studies, helix D is fairly unstructured and attains helical conformation only upon binding with certain antigens [18]. The sequence alignments of the tetraspanin family members indeed show an increased variability in this region, including insertions and deletions [19]. It has been suggested that this surface area might be involved in a species- or tetraspanin-specific recognition process, [20] which could hint to the possibility of heterodimeric tetraspanin species assembly [21]. In particular, segment D of CD81 should be able to guide specific homomeric clustering [22].

WO2014/168548 discloses a therapeutic delivery vesicle, which can for example be an exosome or microvesicle, which has attached to its membrane a polypeptide construct comprising a carrier polypeptide fused to a decoy receptor which is signaling-incompetent.

WO2016/073864 discloses a B cell targeting agent comprising a CD19 or CD21 targeting antibody coupled to a nanoparticle, lipid-based carrier molecule or extracellular vesicle.

WO2018/075825 discloses bioengineered exosomes comprising a fusion protein which comprises a segment of an exosome protein fused to a cancer stem cell targeting peptide.

WO2018/015535A1 discloses EVs coated with proteins containing Fc binding domains. Exemplary EVs carry fusion constructs comprising exosomal proteins fused to Fc binders such as Protein A/G, the Z domain or ZZ domain of Protein A.

US2018/0015182A1 discloses exosomes delivering bioactive cargo by engineering tetraspanins comprising fusions to proteins, or including a terminal or loop peptide attachment site to attach proteins to the exosomes.

El Andaloussi et al. (Advanced Drug Delivery Reviews 2013, 65:391-397) describe exosomes for targeted siRNA delivery. Exemplary targeting exosomes comprise a Lamp2b-brain specific peptide (RVG, a 29-mer peptide) fusion protein.

Drummer et al. (Journal of Virology 2002, 76(21):11143-11147) describe a binding site on the LEL of CD81 binding to hepatitis C virus E2 glycoprotein.

To enhance their potential as the next-generation therapeutic carriers, exosome-mediated delivery systems need to be further developed, especially to improve their inherently low efficiency of cellular uptake. There is a specific need for exosome-mediated delivery systems comprising exosomal membrane proteins with increased stability and improved target-specificity.

SUMMARY OF THE INVENTION

It is the objective of the present invention, to provide target-specific extracellular vesicles with improved target binding characteristics. It is a further objective to provide target-specific EV surface proteins and binding domains thereof with improved target binding characteristics.

The problem is solved by the present invention.

According to the invention, there is provided a method of producing a protein comprising a target-specific extravesicular domain (TED) of an extracellular vesicle (EV) surface protein comprising modifying a polynucleotide comprising a nucleotide sequence encoding the extravesicular domain (ED) of an EV surface protein by a mutagenesis method within at least one modified region within the ED amino acid sequence with a length of 3-20 contiguous amino acids flanked by regions of the wild-type ED sequence at its N-terminus and C-terminus, to incorporate a target binding site within the ED, thereby producing a repertoire of polynucleotides encoding a variety of TEDs, each comprising a different target binding site, and selecting a TED specifically recognizing a predetermined target, and producing the protein comprising the selected TED.

It is specifically understood that all features of the proteins, in particular the target binding molecules ("binders", "target-specific molecules") described herein are features characterizing the method of the invention, and *vice versa*.

Specifically, the protein comprising the TED may consist of the TED, or a protein comprising the TED any one or more further regions *e.g.*, another ED and/or a transmembrane domain, or a recombinant fusion protein in particular comprising a heterologous sequence.

Specifically, the repertoire of polynucleotides is comprised in genetic packages displaying the variety of TEDs on the outer surface, preferably employing a display system selected from the group consisting of a yeast, phage, bacterium, ribosome, mRNA or mammalian cell display.

Specifically, the modified region has a lower target binding affinity when isolated from the TED.

Specifically, the target binding site comprises at least one further binding region which is within a further modified region distant at least 2 amino acids, or within the wild-type ED sequence. Said at least one further binding region is preferably of the same ED and/or positioned within the same TED at a certain distance further described herein.

Specifically, the TED comprises at least 70% sequence identity to the wild-type ED.

Specifically, the wild-type ED originates from (or is of) an EV surface protein selected from the group consisting of tetraspan-like proteins, proteins of the integrin family, proteoglycans, five-transmembrane domains proteins, type I transmembrane

proteins, Notch family proteins, enzymatic membrane proteins, immune regulatory surface proteins, surface markers of mesenchymal stem cells, glycoproteins, or channelling proteins.

Specifically, the tetraspan-like protein is a tetraspanin, preferably selected from the group consisting of:

- a) CD81 comprising or consisting of the amino acid sequence identified as SEQ ID NO:87;
- b) CD9 comprising or consisting of the amino acid sequence identified as SEQ ID NO:89;
- c) CD53 comprising or consisting of the amino acid sequence identified as SEQ ID NO:90;
- d) TSPAN32 comprising or consisting of the amino acid sequence identified as SEQ ID NO:91;
- e) CD82 comprising or consisting of the amino acid sequence identified as SEQ ID NO:92;
- f) CD63 comprising or consisting of the amino acid sequence identified as SEQ ID NO:93;
- g) CD151 comprising or consisting of the amino acid sequence identified as SEQ ID NO:94; and
- h) CD37 comprising or consisting of the amino acid sequence identified as SEQ ID NO:95

or wherein the tetraspan-like protein is a lysosome-associated membrane protein, preferably LAMP2 comprising or consisting of the amino acid sequence identified as SEQ ID NO:96.

Specifically, the wild-type ED is of any one of:

- a) CD81, wherein the ED comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:130 or SEQ ID NO:131;
- b) CD9, wherein the ED comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:132, SEQ ID NO:182 or SEQ ID NO:133;
- c) CD53, wherein the ED comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:134 or SEQ ID NO:135;
- d) TSPAN32, wherein the ED comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:136 or SEQ ID NO:137;

e) CD82, wherein the ED comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:138 or SEQ ID NO:139;

f) CD63, wherein the ED comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:140 or SEQ ID NO:141;

g) CD151, wherein the ED comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:142 or SEQ ID NO:143;

h) CD37, wherein the ED comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:144 or SEQ ID NO:145;

or

i) LAMP2, wherein the ED comprises or consists of the amino acid sequence identified as SEQ ID NO:146:

Specifically, the protein comprises a loop structure in the ED amino acid sequence which is stabilized by one or more cysteine(s) at position(s) to allow the formation of one or more disulfide bonds.

Specifically, the modified region is positioned within a loop region of the ED.

Specifically, the ED is

a) of CD81 and the amino acid sequence is modified to introduce cysteines to allow formation of one or more disulfide bonds not naturally-occurring in the wild-type ED sequence, preferably between positions 134 and 144 and/or 130 and 146 and/or 135 and 168, wherein numbering is of human CD81 identified as SEQ ID NO:87, or

b) of CD9 and the amino acid sequence is modified to introduce cysteines to allow formation of one or more disulfide bonds not naturally-occurring in the wild-type ED sequence, preferably between positions 20 and 28, wherein numbering of the positions is of the CD9 large extracellular loop (LEL, SEQ ID NO:118).

Specifically, the ED is of CD81, and the modified region is positioned within positions 160 and 172, wherein numbering is of human CD81 identified as SEQ ID NO:87.

Specifically, the TED comprises at least one further binding region positioned between positions 132 and 141, or between positions 180 and 189, wherein numbering is of human CD81 identified as SEQ ID NO:87.

Specifically, the ED is of CD9 and the modified region is positioned within any one of positions 155-166, positions 128-142, positions 130-140, or positions 169-180, wherein numbering is of human CD9 identified as SEQ ID NO:89.

Specifically, the target is selected from the group consisting of cellular targets, preferably mitogenic receptors, cytokine receptors, asialoglycoprotein receptors, membrane transporters, lipoproteins, liposaccharides, glycoproteins, proteoglycans, or acellular targets, preferably cytokines, artificial proteins or artificial surface structures.

Specifically, the protein comprising the TED is a target-specific EV surface protein (TSP) comprising said TED and at least one transmembrane domain.

Specifically, the transmembrane domain comprises at least 70% sequence identity to a transmembrane domain originating from a mammalian EV surface protein.

Specifically, the transmembrane domain is of any one of:

a) CD81, wherein the transmembrane domain comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:147, SEQ ID NO:148, SEQ ID NO:149 or SEQ ID NO:150;

b) CD9, wherein the transmembrane domain comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:151, SEQ ID NO:152, SEQ ID NO:153 or SEQ ID NO:154;

c) CD53, wherein the transmembrane domain comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:155, SEQ ID NO:156, SEQ ID NO:157 or SEQ ID NO:158;

d) TSPAN32, wherein the transmembrane domain comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:159, SEQ ID NO:160, SEQ ID NO:161 or SEQ ID NO:162;

e) CD82, wherein the transmembrane domain comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:163, SEQ ID NO:164, SEQ ID NO:165 or SEQ ID NO:166;

f) CD63, wherein the transmembrane domain comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:167, SEQ ID NO:168, SEQ ID NO:169 or SEQ ID NO:170;

g) CD151, wherein the transmembrane domain comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:171, SEQ ID NO:172, SEQ ID NO:173 or SEQ ID NO:174;

h) CD37, wherein the transmembrane domain comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:175, SEQ ID NO:176, SEQ ID NO:177 or SEQ ID NO:178;

or

i) LAMP2, wherein the transmembrane domain comprises or consists of the amino acid sequence identified as SEQ ID NO:179.

Specifically, both, the ED and the transmembrane domain, originate from (or are of) the same EV surface protein, preferably wherein the EV surface protein is selected from the group consisting of:

a) CD81 comprising or consisting of the amino acid sequence identified as SEQ ID NO:87;

b) CD9 comprising or consisting of the amino acid sequence identified as SEQ ID NO:89;

c) CD53 comprising or consisting of the amino acid sequence identified as SEQ ID NO:90;

d) TSPAN32 comprising or consisting of the amino acid sequence identified as SEQ ID NO:91;

e) CD82 comprising or consisting of the amino acid sequence identified as SEQ ID NO:92;

f) CD63 comprising or consisting of the amino acid sequence identified as SEQ ID NO:93;

g) CD151 comprising or consisting of the amino acid sequence identified as SEQ ID NO:94;

h) CD37 comprising or consisting of the amino acid sequence identified as SEQ ID NO:95; and

i) LAMP2 comprising or consisting of the amino acid sequence identified as SEQ ID NO:96.

The invention further provides for a method of producing a target-specific extracellular vesicle (TEV) preparation by

a) providing a polynucleotide encoding a protein comprising the TED obtainable by a method described herein;

b) introducing said polynucleotide into a source cell or source cell mixture;

b) culturing said cell(s) under conditions producing extracellular vesicles;

c) isolating a fraction comprising a TEV comprising the target binding site of the TED; and

d) producing a preparation of the TEV comprised in said fraction.

Specifically, the protein comprising the TED is a target-specific EV surface protein (TSP) comprising said TED and at least one transmembrane domain, and the TEV is displaying the target binding site on the outer surface of its membrane.

Specifically, the method described herein further comprises loading the TEV with an intravesicular load, wherein the load comprises any one or more of peptides, polypeptides, protein domains, proteins, lipids, genes, nucleic acids such as mRNAs, miRNAs, RNAi mediating molecules in particular locked nucleic acids or phosphorothioates, DNA, DNA fragments, plasmids such as minicircle DNA, drugs such as small molecules, in particular chemotherapeutics or senolytics.

Specifically, the source cell or source cell mixture is originating from (or is of) a eukaryotic or prokaryotic source, preferably of body tissue, body fluid, or a cell culture.

Specifically, the source cell or source cell mixture is obtained from a subject and the TEV preparation is formulated for autologous use.

According to a specific aspect, the invention provides for a TEV preparation obtainable or obtained by the method described herein.

According to a specific aspect, the invention provides for an autologous TEV preparation obtainable or obtained by the method described herein, wherein the source cell or source cell mixture is obtained from a subject and the TEV preparation is administered to the same subject.

The invention further provides for the medical use of such autologous TEV preparation, in particular for use in treating a subject in need thereof, wherein the source cell or source cell mixture is obtained from said subject.

Specifically, the subject is a patient, in particular a human patient suffering from a disorder or a disease.

According to a specific aspect, the invention provides for a protein comprising a target-specific extravesicular domain (TED) of an extracellular vesicle (EV) surface protein obtainable or obtained by a method described herein.

According to a specific aspect, the invention provides for a target-specific extravesicular domain (TED) of an extracellular vesicle (EV) surface protein comprising at least 70% sequence identity to a wild-type extravesicular domain (ED) of a mammalian EV surface protein sequence and at least one modified region with a length of 3-20 contiguous amino acids flanked by regions of the wild-type ED sequence at its N-terminus and C-terminus, which modified region is at least part of a target binding site not naturally-occurring in the wild-type ED, wherein the modified region has a lower

target binding affinity when isolated from the TED and/or wherein the target binding site comprises at least one further binding region within a further modified region distant at least 2 amino acids, or within the wild-type ED sequence.

According to a specific aspect, the invention provides for a target-specific EV surface protein (TSP) comprising at least one transmembrane domain and an extravesicular domain (ED) comprising at least 70% sequence identity to a wild-type ED of a mammalian EV surface protein sequence and at least one modified region with a length of 3-20 contiguous amino acids flanked by regions of the wild-type ED sequence at its N-terminus and C-terminus, which modified region is at least part of a target binding site not naturally-occurring in the wild-type ED, wherein the modified region has a lower target binding affinity when isolated from the TSP and/or wherein the target binding site comprises at least one further binding region within a further modified region distant at least 2 amino acids, or within the wild-type ED sequence.

According to a specific aspect, the invention provides for a polynucleotide encoding any of the target-specific molecules described herein, in particular any one of the protein comprising the TED described herein, the TED described herein, or the TSP described herein. Specifically, the polynucleotide is a cDNA molecule.

According to a specific aspect, the invention provides for a target-specific extracellular vesicle (TEV) comprising a membrane and displaying a target-specific molecule on the outer surface of the membrane, wherein the target-specific molecule is any of the target-specific molecules described herein, in particular any one of the protein comprising the TED described herein, the TED described herein, or the TSP described herein.

According to a specific aspect, the invention provides for a pharmaceutical preparation comprising any one of the target-specific molecules described herein, in particular any one of the protein comprising the TED described herein, the TED described herein, or the TSP described herein, or which comprises a TEV described herein, and a pharmaceutically acceptable carrier, preferably in a formulation for intradermal, subcutaneous, intravenous, topical, or oral use.

According to a specific aspect, the invention provides for a target-specific extracellular vesicle (TEV) library comprising a variety of at least 10^2 TEVs described herein, wherein the variety comprises or consists of TEVs with a different modified region flanked by the same regions of the same wild-type extravesicular domain (ED) at its N-terminus and C-terminus.

Specifically, the TEV library comprises a repertoire of TEVs that include target-specific molecules described herein, wherein the repertoire covers at least 10^2 different modified regions or target binding sites.

According to a specific aspect, the invention provides for a method of producing a library of target-specific extracellular vesicles (TEVs), comprising:

a) providing a repertoire of polynucleotides encoding a variety of target-specific EV surface proteins (TSPs), each comprising a different target binding site;

b) introducing said repertoire into said source cell(s); and

b) isolating a fraction comprising a repertoire of target-specific extracellular vesicles (TEVs), with a different target binding specificity, to produce a library of TEVs,

wherein the repertoire of polynucleotides is produced by mutagenizing a polynucleotide encoding an EV surface protein by a mutagenesis method to obtain mutations of the extravesicular domain (ED) of said EV surface protein within at least one modified region with a length of 3-20 contiguous amino acids flanked by regions of the wild-type ED sequence at its N-terminus and C-terminus, to incorporate a target binding site within the ED.

According to a specific aspect, the invention provides for a TEV library obtainable or obtained by the method described herein, preferably comprising at least 10^2 TEVs with different target specificity.

According to a specific aspect, the invention provides for a target-specific extravesicular domain (TED) of an extracellular vesicle (EV) surface protein comprising at least 70% sequence identity to a wild-type extravesicular domain (ED) of a mammalian EV surface protein sequence and at least one modified region with a length of 3-20 contiguous amino acids flanked by regions of the wild-type ED sequence at its N-terminus and C-terminus, which modified region is at least part of a target binding site not naturally-occurring in the wild-type ED, wherein the TED is characterized as further described herein.

Specifically, the modified region has a lower target binding affinity when isolated from the TED.

Specifically, the TED comprises at least any one of 70%, 80%, 85%, 90%, or 95% sequence identity to the respective wild-type ED sequence.

Specifically, the wild-type ED originates from an EV surface protein as further described herein.

Specific EV surface proteins are any one of (or combinations of):

- a) tetraspan-like proteins,
- b) proteins of the integrin family,
- c) proteoglycans,
- d) five-transmembrane domains proteins,
- e) type I transmembrane proteins,
- f) Notch family proteins,
- g) membrane proteins, in particular those with enzymatic activity (enzymatic TM proteins),
- h) immune regulatory surface proteins,
- i) surface markers of mesenchymal stem cells,
- j) glycoproteins,
- k) channelling proteins, or
- l) miscellaneous exosomal (vesicular) surface proteins.

According to a specific embodiment, the EV surface protein is a tetraspan-like protein, in particular a tetraspanin or a lysosome-associated membrane protein.

Specifically, the tetraspan-like protein is of the tetraspan junctional complex superfamily, such as a tetraspanin of class I or II, preferably selected from the group consisting of:

- a) CD81 comprising or consisting of the amino acid sequence identified as SEQ ID NO:87;
- b) CD9 comprising or consisting of the amino acid sequence identified as SEQ ID NO:89;
- c) CD53 comprising or consisting of the amino acid sequence identified as SEQ ID NO:90;
- d) TSPAN32 comprising or consisting of the amino acid sequence identified as SEQ ID NO:91;
- e) CD82 comprising or consisting of the amino acid sequence identified as SEQ ID NO:92;
- f) CD63 comprising or consisting of the amino acid sequence identified as SEQ ID NO:93;
- g) CD151 comprising or consisting of the amino acid sequence identified as SEQ ID NO:94; and
- h) CD37 comprising or consisting of the amino acid sequence identified as SEQ ID NO:9.

Specifically, the lysosome-associated membrane protein is LAMP2 comprising or consisting of the amino acid sequence identified as SEQ ID NO:96.

Specifically, the originating protein is a human wild-type EV surface protein or an artificial protein (or a wild-type protein of a non-human animal) comprising at least 90% sequence identity thereto.

Specifically, the originating ED is a wild-type ED which is originating from a human wild-type EV surface protein or an artificial protein (or a wild-type protein of a non-human animal) comprising at least 90% sequence identity thereto.

Specifically, the wild-type ED is of a mammalian EV surface protein, such as comprising or consisting of a human or non-human animal amino acid sequence.

Specifically, the wild-type ED comprises at least any one of 90, 95, 98, 99% sequence identity, or comprises 100% sequence identity, to any one of the ED amino acid sequences comprised in or composed of the wild-type EC sequences, in particular of a human EV surface protein, which is any one:

a) of CD81, wherein the ED comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:130 or SEQ ID NO:131;

b) of CD9, wherein the ED comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:132, SEQ ID NO:182 or SEQ ID NO:133;

c) of CD53, wherein the ED comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:134 or SEQ ID NO:135;

d) of TSPAN32, wherein the ED comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:136 or SEQ ID NO:137;

e) of CD82, wherein the ED comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:138 or SEQ ID NO:139;

f) of CD63, wherein the ED comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:140 or SEQ ID NO:141;

g) of CD151, wherein the ED comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:142 or SEQ ID NO:143;

h) of CD37, wherein the ED comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:144 or SEQ ID NO:145;

or

i) of LAMP2, wherein the ED comprises or consists of the amino acid sequence identified as SEQ ID NO:146:

Specifically, the sequence of an ED may vary at one or at both ends, such that the ED is prolonged or shortened by a number of amino acids, e.g., by 1, 2, 3, 4, or 5 amino acids, depending on the EV displaying the EV surface protein or the method of determining the regions of an ED.

Specifically, the ED may be of any non-human mammalian origin, such as comprising or consisting of the respective EC or extravascular loop sequences of the respective non-human homologs.

Specifically, the mammalian wild-type tetraspan-like protein comprises at least any one of 90, 95, 98, or 99% sequence identity, or comprises 100% sequence identity, to a mammalian wild-type tetraspan-like protein.

According to a specific example, the EV surface protein is CD81, and the modified region is positioned within positions 160 and 172, wherein numbering is of human CD81 identified as SEQ ID NO:87. Preferably, such TED comprising said modified region comprises at least one further binding region positioned between positions 132 and 141, or between positions 180 and 189, wherein numbering is of human CD81 identified as SEQ ID NO:87.

According to another specific example, the EV surface protein is CD9, and the modified region is positioned within any one of positions 155-166, positions 128-142, positions 130-140, or positions 169-180, wherein numbering is of human CD9 identified as SEQ ID NO:89.

According to a specific embodiment, said at least one modified region within the TED (in particular within at least two distant regions or EDs comprised in the target binding site) comprises solvent exposed residues, preferably wherein said at least one modified region is located in a loop and/or helical region of the ED. Solvent exposure of a position is typically indicative of favorable accessibility for target binding. Test methods to determine exposure to solvents are solvent accessible surface area and relative accessible surface area. Specifically, solvent exposed residues are those with relative accessible surface area of more than 20%.

According to a further specific embodiment, said at least one modified region is located in an alpha-helical region of the ED comprising solvent exposed residues. Specifically, an alpha-helical region comprises a series of coils e.g., coil amino acid sequences, in particular coil repeat sequences, comprising a repeated pattern of hydrophobic and charged amino acid residues, thereby forming a peptidic alpha-helix. Specifically, the helical region of the ED has a length ranging between 5 and 30,

preferably 7 to 18 amino acids. Specifically, at least one of the helical regions is part of the target binding site. Since helical regions typically tend to dimerize or multimerize, said target binding site suitably prevents dimerization and multimerization, respectively, if the target binding site is comprised in the ED monomer.

Specifically, the ED is of a tetraspanin such as CD81 or CD9, and the target binding site comprises or involves at least one helical and/or loop structure of the tetraspanin, in particular within the LEL of CD81 and CD9, respectively. Specifically, the EV surface protein is CD81 or CD9, in particular human CD81 or CD9. Specifically, the surface protein is monomeric CD81 or monomeric CD9.

Yet, in some cases, the ED or the EV surface protein incorporates the target binding site only when present as a dimer or multimer. In such cases, it is preferred to engineer the target binding site within non-helical regions, to ensure dimerization and multimerization, respectively, and effective target binding.

Specifically, the location of the target binding site varies in different types of EDs or EV surface proteins. Certain motifs are conserved, whereas others can vary within a protein family or in analogous sequences of different species. For example, helix D in the tetraspanin protein family is fairly unstructured and attains helical conformation only upon binding with certain antigens. Sequence alignment of the tetraspanin family members shows an increased natural variability in this region, including insertions and deletions. Specifically, such natural variability indicates good tolerability of site-directed mutagenesis.

Specifically, the modified region of the TED described herein is positioned within a loop region of the wild-type ED sequence, in particular a large extracellular loop region. EV surface proteins may have tertiary structure when attached to the EV including a loop structure, and exposing the loop to the vesicle surrounding. Such loop region is particularly suitable for engineering a target binding site within the ED.

For example, tetraspan-like molecules may have one or more small extravesicular loops and/or one or more large extravesicular loops (LEL). Exemplary LEL sequences of CD81 and CD9 are further disclosed herein. In the case of human CD81 the LEL sequence is identified as SEQ ID NO:7, and in the case of human CD9 the LEL sequence is identified as SEQ ID NO:118.

Specifically, the TED or EV surface protein as described herein comprises a loop structure in the amino acid sequence which is stabilized by one or more cysteine(s) at position(s) to allow the formation of one or more disulfide bonds. Specifically, the TED

or EV surface protein comprising said TED as described herein comprises at least one loop region which is stabilized by at least one intramolecular bond connecting at least two amino acid side chains *e.g.*, disulfide bonds.

The loop length of an ED or an EV surface protein can vary. Specifically, the ED or the EV surface protein comprises at least one large extravesicular loop, and a loop of a smaller size. The large extravesicular loop typically has a length ranging between 75 and 140 amino acids, preferably, between 78 and 132 amino acids. Specifically, a small loop has a length ranging between 25 and 35 amino acids, preferably between 26 and 32 amino acids.

Specifically, the ED or an EV surface protein comprises at least one helical region or domain, *e.g.*, 1, 2, 3, or 4 helical regions. Specifically, a helical region is within a loop region, or within the terminal region of the surface protein.

Specifically, the TED or an EV surface protein comprising a TED described herein is present on the surface of an EV as a monomer. Though extracellular or extravesicular surface proteins tend to dimerize or oligomerize for biological function, the surface protein is engineered to be target binding as a monomer.

Specifically, at least one of the loop and/or helical regions is mutagenized for producing the modified region within the ED to become at least part of the target-binding site.

Specifically, the TED described herein is comprised in a target-specific EV (TEV) surface protein which comprises at least one loop and/or helical structure anchored to the EV, typically a transmembrane domain.

Specifically, the modified region within the TED described herein has a length of 3-20 contiguous amino acids, preferably at least 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12, up to *e.g.*, 20, 19, 18, 17, 16, 15, 14, 13, or 12. Modifications typically result in at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or up to 20 point mutations within any such modified region.

Specifically, said modification introduces a number of point mutations, including substitution, insertion or deletion of one amino acid at one position, preferably an amino acid substitution, *e.g.*, at least (or not more than) any one of 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20.

Specifically preferred are a number of point mutations at contiguous positions within a first modified region *e.g.*, a number ranging from 3 to 20, in particular 3, 4, 5, 6, 7, 8, 9, or 10; and optionally a number of point mutations within a second modified region

at contiguous positions distant said first modified region e.g., a number ranging from 3 to 20, in particular 3, 4, 5, 6, 7, 8, 9, or 10, wherein both, said first and second modified regions are comprised in the target binding site. The distance between said first and second modified region is typically composed of flanking sequences originating from a wild-type ED.

Specifically, the modified region is flanked by regions of the wild-type ED sequence at its N-terminus and C-terminus, such that the flanking regions are adjacent to the respective terminus of the modified region. Typically, the flanking regions are characterized by the wild-type amino acid sequence of the wild-type ED spanning, wherein the wild-type flanking sequences have a length of at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 contiguous amino acids that are identical to the respective (unmodified) sequence of the wild-type ED.

Yet, according to a specific embodiment, one (but not both) of the flanking regions may be absent. For example, the modified region can be a terminal region including the C-terminus of the ED. In such case, the modified region is flanked by regions of the wild-type ED sequence only at its N-terminus, and the modified region is a C-terminal region of the ED.

Specifically, the modified region is binding to the target when positioned within the TED, thus, involving the surrounding secondary or tertiary structure of the TED to specifically recognize the target.

According to a specific embodiment, the target binding site comprises at least said modified region and a wild-type region of the TED, and/or another modified region within the TED, within the TSP, or on the surface of an EV comprising the TED and/or TSP.

When isolated from the TED, or when produced as a separate peptide consisting of the same amino acid sequence as the modified region, typically, such isolated modified region has a lower target binding affinity or even lacks specificity or selectivity to bind the target. Compared to the target binding of the modified region within the TED, the isolated modified region specifically, has less affinity or less selective binding, with a binding constant or binding dynamics which is at least 10 fold, or at least 100 fold, at least 1000 fold different, as determined with the same assay or in a comparable setting.

Suitable assays to compare the target binding property of a TED described herein and of the modified region isolated from said TED, may employ any of the following assays: ELISA, affinity determination e.g., using Biacore, biolayer interferometry,

fluorescence measurement of cells displaying the TED incubated with cognate antigen, isothermal titration microcalorimetry, fluorescence correlation spectroscopy.

Specifically, the target binding site comprises at least two regions (wherein at least one or two of them is the modified region as further described herein), which are within at least two different EDs of the same EV surface protein, such as EDs separated by at least one transmembrane domain of the EV surface protein.

Such target binding site incorporated within the TED and comprising or otherwise involving said at least one modified region within the same ED or within at least two EDs has the particular advantage of improved binding properties. Such improved binding properties are typically obtained when mutagenizing a predetermined region within the ED to become a modified region with binding properties embedded within the ED, and selecting the suitable binders according to their target binding specificity and/or affinity.

The binding properties are specifically improved over comparable fusions of specific (peptide) binders to an ED thereby producing fusion proteins. This is because upon fusion, the binding properties typically are changed such that comparable fusion proteins have less affinity or less selective binding upon fusion of a binder to the ED as compared to the isolated binders.

Specifically, the modified region is at least part of a target binding site or consists of the target binding site. According to a specific embodiment, the modified region includes all contact amino acid residues of the binding site.

According to a specific embodiment, the target binding site comprises more than one modified binding region, wherein at least a first modified region is positioned at a certain distance from a second modified region.

Specifically, the target binding site comprises at least one further binding region within a further modified region distant at least any one of 2, 3, 4, 5, 6, 7, 8, 9, or 10 contiguous amino acids, or within the wild-type ED sequence. Specifically, the region between said at least two distant modified regions includes one of the regions flanking the modified region described herein at its N-terminus or C-terminus, which is a region of the wild-type ED sequence.

Specifically, the target binding site comprises binding residues within at least two distant regions of said TED. Specifically, the TED is modified or mutagenized within said at least two distant regions to incorporate said target binding site.

According to a specific embodiment, the target binding site comprises contact amino acid residues within the ED, but outside the modified region. Such further contact

amino acids may be positioned in one or more further regions of the ED, which one or more further regions may comprise a modified (mutated) amino acid sequence e.g. comprising one or more point mutations, or may comprise the wild-type sequence of the ED.

According to a specific embodiment, the ED or the target binding site comprises said at least one modified region, and one or more further point mutations, or a series of at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 contiguous point mutations within the amino acid sequence, specifically at a certain distance e.g., distant at least 2, 3, 4, 5, 6, 7, 8, 9, or 10 amino acids.

According to a specific embodiment, the ED or the target binding site comprises at least two or three of said modified regions.

In specifically preferred embodiments, the target binding site is a conformational binding site comprising two or more non-contiguous regions within the same ED or at least two different EDs, such as involving a binding surface stretching over at least two regions, each at a certain distance e.g., distant at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20, typically up to 100 or 80 amino acids. Specifically, the distant binding regions are not adjacent to each other. Specifically, the distant binding regions are not contiguous in the sequence of the respective EV surface protein.

Specifically, the binding site comprises each of the distant binding regions and optionally one or more further contact points or regions of the TED. Specifically, the target binding site comprises contact points within at least two modified (e.g., synthetic or mutagenized) regions and optionally at least one further region which is neither synthetic nor mutagenized, but a native (wild-type) region of the ED.

Specifically, a target is bound by a conformational paratope of a binding site that is incorporated within the extravesicular part of an EV surface protein (in particular the part comprising the TED described herein), which paratope comprises contact areas in distant regions of the TED e.g., loop and/or helical region(s). Suitable TEDs and respective EV surface proteins are conveniently produced by mutagenizing selected areas within the distinct regions. For example, a TED as described herein is mutagenized within one, two, or three distant regions.

Specifically, the target binding site is a novel binding site recognizing a specific predetermined target, not naturally-occurring in the wild-type ED or wild-type EV surface protein, herein also referred to as "artificial" binding site.

According to a specific embodiment, the target binding site may be a new or additional binding site within one or more EDs, such that the ED(s) has/have a novel or an additional specificity to bind a target. Such new target binding site may have a specificity to bind the same target as any naturally-occurring binding site of the ED(s), or may have a different specificity e.g. recognizing a different target.

According to another specific embodiment, the target binding site may be a modified naturally-occurring binding site of the ED(s), such that the ED(s) has/have a specificity to bind the same target as the naturally-occurring binding site (though fine specificity or affinity may be changed).

According to a specific aspect, the TED provided as a separate molecule or a protein comprising such TED (such as an EV surface protein or a TSP) may comprise at least one or two different target binding sites specifically recognizing the same target or different targets.

The TED described herein is specifically used for medical purposes e.g. to deliver therapeutically effective amounts of proteins comprising a TED described herein, or target-specific EV surface proteins (TSPs) comprising such TED, in particular when attached to an EV, thereby providing a target-specific extracellular vesicle (TEV) as further described herein.

Specifically, the target is selected from the group consisting of cellular targets, preferably mitogenic receptors, cytokine receptors, asialoglycoprotein receptors, membrane transporters, lipoproteins, liposaccharides, glycoproteins, proteoglycans, or acellular targets, preferably cytokines, artificial proteins or artificial surface structures.

According to a specific aspect, the target is a human cell, e.g., a cell originating from a healthy or diseased subject, or the respective cell lysate. Human cells of the diseased phenotype are preferably used as a target.

According to a certain aspect, the target binding site specifically recognizes a novel target, *i.e.*, a target which would otherwise not be bound by the naturally-occurring ED or EV surface protein.

According to another aspect, the target binding site specifically recognizes a target which is the natural ligand of the ED or EV surface protein, yet with modified binding properties, such as selectivity, fine specificity, affinity and/or avidity e.g., for improved target binding. For example, natural ligands of tetraspanins are e.g., antigens of pathogens or pathogens such as cellular pathogens or viruses. By modifying the tertiary structure of a tetraspanin presented as a surface protein by an TED, TSP or TEV

as described herein, binding properties can be improved *e.g.*, to attain increased selectivity and/or affinity of binding to target the respective pathogens.

Specifically, the target consists of an antigen or antigenic structure of an antigen, in particular an epitope which is otherwise recognized by a target-specific antibody.

Specifically, the target is a cellular receptor. According to a specific example, an EV comprising the TED or TSP described herein, which is targeting a cellular receptor can directly fuse with the recipient cell membrane, thus incorporating its membrane proteins to the plasma membrane and delivering their cargo to the cytoplasm of the recipient cell.

Specifically, the target is an antigen *e.g.*, naturally-occurring antigens or synthetic antigens. In certain embodiments, the target antigens are present in a diseased patient's blood, which antigens are bound by the surface protein of the EV and thus removed from the patient's cardiovascular and/or lymphatic system. Specific examples are undesired natural agents such as pathogens, toxins, diseased or cancer cells, cytokines, or metabolites. In certain further embodiments, the target antigens are synthetic antigens *e.g.*, of solid surfaces such as grafts or implants, or soluble compounds such as chemicals or synthetic compounds, which can be removed upon effective binding by a specific binder (*e.g.*, a TED, TSP or TEV) described herein.

Specific targets may be natural targets which are typically recognized by wild-type EVs. However, the target binding site comprised in the binders described herein may specifically recognize novel targets, which are otherwise not recognized by a wild-type EV. Among natural targets there are pathogens, such as viral antigens. Specific binders (*e.g.*, a TED, TSP or TEV) described herein may bind such natural targets through a novel binding site, which may have improved binding characteristics, such as binding affinity, avidity or specificity.

According to a specific aspect, the invention provides for a target-specific EV surface protein (TSP) comprising at least one transmembrane domain and at least one extravesicular domain (ED) comprising at least 70% sequence identity to a wild-type ED of a mammalian EV surface protein sequence and at least one modified region with a length of 3-20 contiguous amino acids flanked by regions of the wild-type ED sequence at its N-terminus and C-terminus, which modified region is at least part of a target binding site not naturally-occurring in the wild-type ED, wherein the TSP is characterized as further described herein.

Specifically, the modified region has a lower target binding affinity when isolated from the TSP.

Specifically, the TSP described herein is characterized by one or more extravesicular domains, wherein at least one of them is target binding ("target-specific"). Specifically, the ED of the TSP described is a TED described herein.

Specifically, said at least one transmembrane domain is of a vesicular membrane protein, or an artificial transmembrane domain *e.g.*, produced by mutagenesis of wild-type transmembrane domains or by *de novo* synthesis of suitable amino acid sequences.

Specifically, the transmembrane domain (TM) comprises at least any one of 70%, 80%, 85%, 90%, 95%, 96%, 977%, 98%, 99% sequence identity, or 100% sequence identity to the respective wild-type transmembrane domain sequence originating from a mammalian EV surface protein, *e.g.* any of the EV surface proteins as further described herein.

Specifically, the wild-type TM is of a mammalian EV surface protein, such as comprising or consisting of a human or non-human animal amino acid sequence.

Specifically, the TM is of a vesicular membrane protein which is of exosomal, microvesicular, or apoptotic body origin, in particular a wild-type exosomal, microvesicular, apoptotic body protein, or such wild-type protein which includes modifications *e.g.*, in the extravesicular region.

Specifically, the wild-type TM comprises at least any one of 90, 95, 98, 99% sequence identity, or comprises 100% sequence identity, to any one of the TM amino acid sequences comprised in or composed of the wild-type sequence of an EV surface protein that is capable of being integrated within a cellular or vesicular membrane when binding the EV surface protein to a respective cell or vesicle, in particular of a human EV surface protein.

Specifically, the transmembrane domain is of any one of:

a) CD81, wherein the transmembrane domain comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:147, SEQ ID NO:148, SEQ ID NO:149 or SEQ ID NO:150;

b) CD9, wherein the transmembrane domain comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:151, SEQ ID NO:152, SEQ ID NO:153 or SEQ ID NO:154;

c) CD53, wherein the transmembrane domain comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:155, SEQ ID NO:156, SEQ ID NO:157 or SEQ ID NO:158;

d) TSPAN32, wherein the transmembrane domain comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:159, SEQ ID NO:160, SEQ ID NO:161 or SEQ ID NO:162;

e) CD82, wherein the transmembrane domain comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:163, SEQ ID NO:164, SEQ ID NO:165 or SEQ ID NO:166;

f) CD63, wherein the transmembrane domain comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:167, SEQ ID NO:168, SEQ ID NO:169 or SEQ ID NO:170;

g) CD151, wherein the transmembrane domain comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:171, SEQ ID NO:172, SEQ ID NO:173 or SEQ ID NO:174;

h) CD37, wherein the transmembrane domain comprises or consists of any one of the amino acid sequences identified as SEQ ID NO:175, SEQ ID NO:176, SEQ ID NO:177 or SEQ ID NO:178;

or

i) LAMP2, wherein the transmembrane domain comprises or consists of the amino acid sequence identified as SEQ ID NO:179.

Specifically, the sequence of a TM may vary at one or at both ends, such that the TM is prolonged or shortened by a number of amino acids, e.g., by 1, 2, 3, 4, or 5 amino acids, depending on the EV displaying the EV surface protein or the method of determining the regions of a transmembrane domain.

According to a specific embodiment, both, the wild-type ED and said at least one transmembrane domain, originate from the same mammalian EV surface protein, preferably selected from the group consisting of:

a) CD81 comprising or consisting of the amino acid sequence identified as SEQ ID NO:87;

b) CD9 comprising or consisting of the amino acid sequence identified as SEQ ID NO:89;

c) CD53 comprising or consisting of the amino acid sequence identified as SEQ ID NO:90;

d) TSPAN32 comprising or consisting of the amino acid sequence identified as SEQ ID NO:91;

e) CD82 comprising or consisting of or consisting of the amino acid sequence identified as SEQ ID NO:92;

f) CD63 comprising or consisting of the amino acid sequence identified as SEQ ID NO:93;

g) CD151 comprising or consisting of the amino acid sequence identified as SEQ ID NO:94;

h) CD37 comprising or consisting of the amino acid sequence identified as SEQ ID NO:95; and

i) LAMP2 comprising or consisting of the amino acid sequence identified as SEQ ID NO:96.

Specifically, the EV surface protein is a vesicular membrane protein which is of exosomal, microvesicular, or apoptotic body origin, in particular a wild-type exosomal, microvesicular, apoptotic body protein, or such wild-type protein which includes modifications *e.g.*, in the extravesicular region.

Specifically, the EV surface protein is a protein originating from a cellular exosome *i.e.*, an exosomal protein.

Specifically, the EV surface protein is a protein originating from a cellular microvesicle *i.e.*, a microvesicular protein.

Specifically, the EV surface protein is a protein originating from an apoptotic body *i.e.*, an apoptotic body protein.

It is well-understood that an exosomal protein, microvesicular, or apoptotic body protein as used herein for modification purposes is of cellular origin, *i.e.*, which can be produced by a respective cell or which can be artificial and produced by *de novo* synthesis.

Specifically, said EV surface protein is a polypeptide, protein domain or protein which has a structure to incorporate an artificial target binding site. Such polypeptide, protein domain or protein may be naturally-occurring, or partially or fully synthetic.

Specifically, one or two of the transmembrane domains is/are fused to the ED of the surface protein, and capable of attaching the ED to the membrane of an EV.

According to a specific aspect, the invention provides for a target-specific extracellular vesicle (TEV) comprising a lipid bilayer membrane and the TED described

herein, or the TSP described herein, displaying the target binding site on the outer surface of the lipid bilayer membrane.

Specifically, the TED described herein can be bound to an EV through a TSP described herein or any other suitable means, *e.g.*, by conjugation, fusion, or affinity binding.

Specifically, the TSP described herein can be bound to an EV through said at least one transmembrane domain comprised in the TSP, in particular wherein said transmembrane domain(s) is/are within the vesicle membrane, such that the TSP is presenting the TED and optionally at least one further ED (which may be target binding or not) to the outer surface of the EV.

Specifically, said at least one transmembrane domain serves as an anchor which can be fused to any kind of target binding molecule incorporating the TED described herein *e.g.*, a TSP described herein.

Specifically, the TSP comprises at least two, three, or four transmembrane domains, preferably all transmembrane membranes of the same wild-type EV surface protein.

Specifically, a series of more than one transmembrane domains of one or more (*e.g.*, the same type or different) proteins or artificial transmembrane domains can be used, thereby creating extravesicular loop regions *i.e.*, loop structures outside the EV. Specifically, the number of transmembrane domains used to anchor at least one loop structure is 2, 3, or 4. Specifically, the surface protein comprises one or more loops *e.g.*, at least 1, 2, 3, or 4 loops. Yet, in certain case, the surface protein comprises a tertiary structure without loops.

Specifically, two transmembrane domains are used to present one extravesicular loop structure on the surface of an EV comprising the TSP described herein. According to specific embodiments, three or four transmembrane domains are used to present two extravesicular loop structures, and four or five transmembrane domains are used to present three extravesicular loop structures. Specifically, within one TSP described herein, one transmembrane domain is followed by a loop structure, such as an amino acid sequence spanning said transmembrane domain to another transmembrane domain, or connecting a loop structure including one or more intramolecular bonds, such as those stabilizing the loop or stem structure, to the vesicular surface. It is understood that the loop structure can be flanked by at least one transmembrane domain, in order to be bound to a vesicular membrane *e.g.*, by either N-terminal or C-terminal fusion of a

peptide loop sequence. If flanked by only one transmembrane domain, the opposite end (the N- or C-terminus of the loop sequence which is opposite to the fusion with the transmembrane domain) is typically not anchored to the membrane, e.g., a loose end. If flanked by two transmembrane domains e.g., by both, N-terminal and C-terminal fusion of a peptide loop sequence to a transmembrane domain, the loop structure is anchored to the EV membrane on both sides.

Specifically, the TSP is a tetraspan-like protein comprising a number of transmembrane domains spanning one or more loop sequences, in particular four transmembrane domains capable of presenting the wild-type loop structure of a tetraspanin protein on the surface of an EV, in particular two extracellular loops anchored by the four transmembrane domains.

Specifically, the TEV is originating from eukaryotic or prokaryotic source cells of body tissue, body fluid, or a cell culture.

According to a specific embodiment, the TEV is carrying an intravesicular load, e.g., wherein the load comprises any one or more of peptides, polypeptides, protein domains, proteins, lipids, genes, nucleic acids such as mRNAs, miRNAs, RNAi mediating molecules in particular locked nucleic acids or phosphorothioates, DNA, DNA fragments, plasmids such as minicircle DNA, drugs such as small molecules, in particular chemotherapeutics or senolytics.

The EV can be a membrane vesicle, which is a submicroscopical vesicle comprising a lipid bilayer membrane, such as the membrane of an exosome, microvesicle or apoptotic body. Specifically, the EV is produced by cells or synthetically produced. When produced by a cell, the EV can be released by the cell into the extracellular space by biological processes, and the surface protein is engineered prior to or after the cell release. The EV may be surface decorated e.g., with carbohydrate structures and/or by fusion with amino acids or amino acid sequences and/or by coupling with chemical compounds, such as drugs, labels or tags.

Specifically, the lipid bilayer membrane is anchoring the TED or the EV surface protein (the TSP) through said at least one transmembrane domain of a membrane surface protein, fusion of a respective coding sequence, or alternatively by chemical and/or affinity binding. In certain cases, a surface protein is anchored to the lipid bilayer membrane of the extracellular vesicle by means other as a transmembrane domain. Such anchors may involve an affinity-based interaction such as e.g. streptavidin-biotin,

or protein A-Fc interaction, or covalent-based interaction such as e.g. maleimide-free cysteine. The anchorage may be achieved e.g. over click-chemistry coupling.

Exemplary techniques for binding a surface protein to the membrane comprises linking atoms of amino acid side chains, such as S-S linkage (disulfide bridges), binding by bioconjugation e.g., by click-chemistry, or otherwise covalently bonding. Alternative techniques employ affinity binders such as biotin and avidin to link the surface protein to the membrane.

In certain cases, one or more (artificial) target binding sites are comprised in one or two surface proteins (the same proteins or different) of the TEV described herein. Specifically, two different surface proteins may be used to produce at least two different binding sites e.g., to recognize different epitopes of the same or different antigens. Specifically, TEVs described herein can be mono- or bispecific, or even oligospecific.

In certain cases, additional target binding sites of a TEV described herein may originate from any binding structure, such as derived from proteins, polypeptides or peptides, including antibodies and antibody fragments or composite molecules with a binding part. Specifically, the binding site may be of an antigen-binding portion of an antibody, or the binding site of any one of an enzyme, an adhesion protein, a ligand or a ligand binding portion of a receptor, which binding site is capable of binding a cognate structure of a binding partner. The EV surface protein particularly may comprise one or more binding sites of protein domains of antibodies or antibody fragments, or the respective antibody domains or fragments, such as those comprising one, two or more variable antibody domains e.g., Fab, Fv, VH/VL dimer, scFv, dAb, F(ab)₂, or other biological binders, such as soluble T-cell receptor, Darpins, etc.

Specifically described herein are modified tetraspanin proteins comprising a novel target binding site characterized by at least any one of 75, 80, 85, or 90% sequence identity, preferably ranging between 80 and 90% sequence identity, to the amino acid sequence of a wild-type human tetraspanin protein LEL, such as in the case of CD81 the LEL identified as SEQ ID NO:7, or in the case of CD9 the LEL identified as SEQ ID NO:118.

Specific examples of a modified tetraspanin protein comprising a novel target binding site are characterized by at least any one of 75, 80, 85, 90, 95, 96, 97, or 98% sequence identity, preferably ranging between 90 and 98% sequence identity, to the amino acid sequence of a full-length wild-type human tetraspanin protein, such as in the

case of CD81 the amino acid sequence identified as SEQ ID NO:87, or in the case of CD9 the amino acid sequence identified as SEQ ID NO:89.

According to a specific example, the novel target binding site comprises a modified region (a binding region) within the LEL of CD81, identified as SEQ ID NO:7, in particular within the region aa113-201 of the full-length protein identified as SEQ ID NO:87. Specific embodiments refer to a novel target binding site within aa130-201.

Specific examples refer to modifications in the amino acid sequence of CD81 (SEQ ID NO:87) to substitute at least one amino acid at positions 160-162, and 181-189.

According to a further specific embodiment, the EV surface protein is CD9 and the binding residues are located within at least one or two different regions selected from the following regions: 155-166; 128-142; 130-140; 169-180. Specifically, the CD9 is human CD9 identified by SEQ ID NO:89. Preferred Cys residues for disulfide bridging are 132 and 140, wherein numbering is according to SEQ ID NO: 89.

According to a further specific embodiment, the EV surface protein is CD53 and the binding residues are located within at least two different regions selected from the following regions: 147-160; 121-134; 123-129; 164-169. Specifically, the CD53 is human CD53 identified by SEQ ID NO:90.

According to a further specific embodiment, the EV surface protein is TSPAN32 and the binding residues are located within at least two different regions selected from the following regions: 158-171; 130-144; 132-139; 174-183. Specifically, the TSPAN32 is human TSPAN32 identified by SEQ ID NO:91.

According to a further specific embodiment, the EV surface protein is CD82 and the binding residues are located within at least two different regions selected from the following regions: 153-172; 178-188; 128-137; 194-215. Specifically, the CD82 is human CD82 identified by SEQ ID NO:92.

According to a further specific embodiment, the EV surface protein is CD63 and the binding residues are located within at least two different regions selected from the following regions: 149-165; 171-176; 127-135; 179-189. Specifically, the CD63 is human CD63 identified by SEQ ID NO:93.

According to a further specific embodiment, the EV surface protein is CD151 and the binding residues are located within at least two different regions selected from the following regions: 159-177; 186-191; 135-145; 194-206. Specifically, the CD151 is human CD151 identified by SEQ ID NO:94.

According to a further specific embodiment, the EV surface protein is CD37 and the binding residues are located within at least two different regions selected from the following regions: 156-177; 183-206; 131-141; 218-229. Specifically, the CD37 is human CD37 identified by SEQ ID NO:95.

The EV surface proteins or the respective EDs (or any domain other than the transmembrane domain comprised in the EV surface protein) can be used for mutagenesis in the predetermined regions to introduce one or more novel target binding sites as described herein.

Further EV surface proteins comprising a novel target binding site may be produced by mutagenizing respective regions. For identifying the suitable regions, sequences of large extracellular loop of tetraspanins Class I (e.g., CD9, CD53, and TSPAN32) have been analysed using Swissmodel modeling server with a template such as CD81. The sequence of large extracellular loop of CD82, a representative of tetraspanin Class II, has been modeled using a program for protein structure and function prediction (i-Tasser modeling server). Sequences of tetraspanins Class II (CD63, CD151, and CD37) have been modeled using the CD82 model coordinates.

According to a specific embodiment, the EV surface protein (e.g., CD81 or any other tetraspanin) is stabilized or thermostabilized by introducing one or more cysteine(s) at position(s) to allow the formation of one or more additional disulfide bonds stabilizing the extravesicular loop structure of the protein. Specifically, at least one additional cysteine is introduced in at least one region of an extravesicular domain of the surface protein. Specifically, a cysteine is introduced in at least two different regions of the surface protein. Introducing novel cysteine(s) allows connection of the regions via a novel (artificial) disulfide bond, which are usually formed from the oxidation of sulfhydryl (–SH) groups of the cysteines.

In particular, the tertiary structure of the surface protein can be modified by one or more additional small loop(s) resulting from one or more additional intramolecular bonds. Such additional small loop(s) can be used to create further contact points of the artificial binding site.

One or more additional intramolecular disulfide bond(s) successfully increased the stability of a surface protein, as measured by methods determining thermostability of the protein by standard methods. A stabilized EV surface protein can be conveniently used as a scaffold to produce a repertoire of EV surface proteins by mutagenesis of pre-determined regions within the EV surface protein each with different target binding or

different target binding properties. Such repertoire is suitably used as a library to select binders to a target of interest.

Specifically, as the number of endorsed residues is theoretically decisive of the free energy change in respect to an unmutated protein, candidate positions for mutagenesis into pairs of cysteines and subsequent formation of a cysteine bond can be pre-selected in a first step by visual inspection of the crystal structure. Specifically, at least one novel cysteine bond is introduced in pre-selected candidates, preferably, new cysteine bonds are created by introduction of at least one or one pair of cysteines via mutagenesis.

Specifically preferred examples refer to a modified tetraspanin comprising novel disulfide bonds that connect Cys residues when reduced, which Cys residues are introduced by mutating the tetraspanin sequence (e.g., one of the ECs, in particular the LEL of the tetraspanin) at two distant sites which are e.g., at the N- and C-terminus of a loop, thereby stabilizing the loop structure. Any such loop structure stabilized by at least one disulfide bond is preferably used to mutagenize the tetraspanin for engineering a novel target binding site within such loop structure.

Specifically, at least two cysteines are introduced by any of insertion or substitution. According to a specific example, additional cysteines are introduced in CD81 to stabilize and/or modify the tertiary structure of the LEL.

According to a specific embodiment, the EV surface protein is CD81 and the amino acid sequence is modified to introduce cysteines to allow formation of one or more disulfide bonds not naturally-occurring in the wild-type ED sequence, preferably between positions 134 and 144 and/or between positions 130 and 146 and/or between positions 135 and 168.

Specifically, the CD81 is human CD81 and a first cysteine is introduced at a position within amino acid 120 and 200 and at least a second cysteine is introduced at a position within amino acid 143 and 201, wherein numbering is of human CD81 identified as SEQ ID NO:87.

Specifically, a first cysteine is introduced at a position within amino acid 130 and 140 and a second cysteine is introduced at a position within amino acid 144 and 170.

Specifically, cysteines are introduced into the human CD81 sequence (in particular an ED of CD81, such as the LEL) at positions 134 and 144, substituting A134C and L144C, respectively, thereby connecting helix A and helix B of the LEL of CD81 by an additional intramolecular disulfide bond. Alternatively, additional cysteines are

introduced at positions 135 and 168, respectively, substituting V134C and S144C, respectively, thereby connecting helix A and helix C of the large extracellular loop of CD81 by an additional intramolecular disulfide bond.

According to a specific embodiment, Cys residues are introduced in the CD81 sequence identified as SEQ ID NO:87 (in particular an ED of CD81, such as the LEL), e.g., by the mutation Ala134Cys and Lys144Cys, thereby introducing a novel disulfide bond spanning the cysteines at positions 134 and 144,

According to a further embodiment (additional or alternative) Cys residues are introduced in the CD81 sequence identified as SEQ ID NO:87, by the mutation Val135Cys and Ser168Cys, thereby introducing a novel disulfide bond spanning the cysteines at positions 135 and 168. According to a further embodiment (additional or alternative) Cys residues are introduced in the CD81 sequence identified as SEQ ID NO:87, by the mutation Ala130Cys and Ala146Cys, thereby introducing a novel disulfide bond spanning the cysteines at positions 130 and 146.

According to a further embodiment (additional or alternative) Cys residues are introduced in the CD81 sequence identified as SEQ ID NO:87, by the mutation Val135Cys and Ser168Cys, thereby introducing a novel disulfide bond spanning the cysteines at positions 135 and 168.

Specifically, at least one of the EDs of human CD81, in particular the LEL, is modified to introduce additional cysteines at positions 134 and 144, and at positions 135 and 168, thereby connecting helix A to helix B, and helix A to helix C. Such CD81 mutant with a combination of potentially stabilizing novel disulfide bonds Ala134Cys/Lys144Cys and Val135Cys/Ser168Cys in its LEL shows an increased positive shift in melting temperature of at least 20°C.

Specifically, at least one of the EDs of human CD9, in particular the LEL, is modified to introduce additional cysteines, thereby obtaining one or more new (additional) disulfide bridges. Preferably, a disulfide bridge is linking positions 20 and 28, such as obtainable by mutating Lys20Cys and Arg28Cys, wherein numbering of the positions is of the CD9 LEL (SEQ ID NO:118). The resulting sequence of the stabilized variant specifically comprises or consists of SEQ ID NO:125.

Specifically, an increase in thermostability of mutated pre-selected candidates is indicated by an increase in the temperature at which thermal unfolding occurs by at least 5, 10, 15, 20, 25, 30, 35, 40, 45, 50 or 55°C compared to wild-type protein.

Specifically, any one of the TED, TSP or TEV described herein has an affinity to bind said target with a K_D of less than 10^{-5} M, preferably less than 10^{-6} M, 10^{-7} M, or 10^{-8} M, or even less than 10^{-9} M.

Usually, a binder is considered a high affinity binder with a $K_D < 10$ nM, in some cases, *e.g.*, for therapeutic purposes EVs with higher affinities are provided, *e.g.*, with a $K_D < 1$ nM, or a $K_D < 0.1$ nM, or a $K_D < 0.01$ nM or a $K_D < \text{pM}$ (picomolar = 10^{-12} M).

Once a selected TED, TSP or TEV has proven to bind a target of interest, the selected binding ED or EV surface protein may undergo affinity maturation by standard methods of affinity maturation, *e.g.*, those methods typically used for producing affinity matured antibodies. For this purpose, only a few point mutations *e.g.*, 1, 2, 3, 4, or 5, up to 10 point mutations may be introduced within one region of the molecule, or within the whole molecule, to produce a new repertoire of target binders that can be selected to isolate a binder with an increased binding affinity. Such affinity matured binders may exhibit an increased binding affinity with a K_D difference of at least 1 or 2 logs.

The specific binding may be determined in a suitable binding assay, such as conventional immunoassays. There are numerous methods known in the art for detecting binding in an immunoassay. Various immunoassays known in the art can be used including competitive and non-competitive assay systems using techniques such as radioimmunoassays, ELISA (enzyme linked immunosorbent assay), immunoradiometric assays, gel diffusion precipitation reactions, immunodiffusion assays, western blot, BIAcore etc.

According to a preferred embodiment, the TEV described herein is originating from eukaryotic or prokaryotic source cells. Source cells are herein understood as donor cells capable of producing the EVs described herein. However, source cells may also just serve as a template to produce the respective synthetic EVs *in vitro*, *e.g.*, using one or more components (such as a transmembrane domain and/or a surface protein) that would otherwise be produced by the source cell, and constructing the TEV with features described herein without relying on cellular transport mechanisms.

Exemplary eukaryotes are mammalian, plant, insect, fungi or yeast. Specific examples are cells of human or non-human animal origin, in particular mammals including Chinese hamster derived cells such as CHO cells, plants, in particular *Arabidopsis thaliana* or *Zea mays*, or fungi, in particular *Saccharomyces cerevisiae* or *Pichia pastoris*.

Exemplary prokaryotes are bacteria. EVs from Gram-negative bacteria are known as outer-membrane vesicles (OMVs). Specific examples are EVs of *Lactobacillus* or *Mycobacteria* pathogens, or *Salmonella enterica*, *M. tuberculosis*, *Moraxella catarrhalis* or *Haemophilus influenzae*.

Specifically, the source cell is of a body tissue, body fluid, or a cell culture, preferably of animal or plant cells; or of a mammalian body fluid or tissue, preferably of blood, urine, amniotic fluid, ascites, cerebrospinal fluid, saliva, synovial fluid, or bone marrow. Specifically, the EV described herein is produced by a cell culture of a source cell.

Specifically, the tissue is of an organ, such as a kidney, brain, or of placenta. Specifically, the tissue is a tumor or metastasis tissue, or a benign tissue.

Specifically, the source cell is a stem cell, such as a mesenchymal stem cell (MSC), amniotic stem cell, or induced pluripotent stem (iPS) cell, a dendritic cell, a hematopoietic cell, epithelial cell, endothelial cell, nerve cell, blood cell or immune cell. Specifically, the source cells of EVs can be amnion-derived multipotent progenitor cell, chorion derived mesenchymal stem cell, induced pluripotent stem cell, keratinocyte, fibroblast, embryonic stem cell, ectodermal stromal cell, endodermal stromal cell, olfactory ensheathing cell, dental pulp stem cell, or immortalized mesenchymal stem cell.

Specifically, the source cell is selected from the group consisting of normal or immortalized human cells, such as induced pluripotent or adult stem cells, epithelial cells and cancer cells.

Specifically, the source cell is a cell line of a recombinant host cell, such as mammalian host cells *e.g.*, cell lines used as cell factories such as for example human primary cells, telomerase immortalized cell lines, or cell lines immortalized by viral oncogenes including adenoviral E1A, HPV derived E6, EBV derived oncogenes, SV40, or combinations of transcription factors. Specifically, cell lines including telomerase immortalized endothelial or mesenchymal stem cells, HEK293, CHO, Vero, HEK, or CAP.

Cells suitably employed in large-scale EV production include mesenchymal stem cells, dendritic cells, and HEK cells or 293T cells.

Specifically, the source cell is a mammalian stem cell or dendritic cell, preferably of human origin.

According to a specific embodiment, the EV surface protein is endogenous to the source cell. Yet, the endogenous EV surface protein is typically presented as a modified surface protein by the TEV described herein.

According to a further specific embodiment, the EV surface protein is heterologous to the source cell. The heterologous EV surface protein can originate from a source cell of a different type, or be a synthetic surface protein which is not naturally occurring. When using a synthetic surface protein, the novel target binding site can be synthesized within the molecule without any further modification. Unlike modified native surface proteins, synthetic surface proteins typically do not have a sequence identity (e.g., less than 50% sequence identity) to a native (wild-type) surface protein.

Specifically, the TEV described herein has a size ranging from 10 to 1000nm, preferably from 30 to 150nm.

Specifically, the TEV described herein has a buoyant density ranging from 1.0 to 1.4, preferably between 1.1 and 1.2 (g/cm³), such as measured by density-gradient ultracentrifugation.

According to a specific embodiment, the TEV described herein is carrying an intravesicular load. Specifically, the load is within a volume of 10⁻¹⁴ to 10⁻¹⁰ µl (volume per one EV). Specifically, the load comprises an active substance or a mixture of active substances e.g., at loading efficiencies between 5 and 90%.

According to a specific embodiment, the binders described herein (in particular the TED, TSP, or TEV described herein) are provided for medical use in treating a subject in need thereof. Medical use encompasses treatment for therapy of a disease condition, either by administering a TEV described herein, or by *ex vivo* use e.g., as reagent or affinity matrix such as for purging undesired substances from body fluids. Further medical uses are for triggering an immune response e.g., to present an antigen to a subject's immune system, such as for active immunotherapy.

EVs can be used as a therapeutic agent by themselves or as delivery systems to deliver a specific load. According to a specific example, the intravesicular load is an active substance or a drug encapsulated within the vesicular membrane. Specifically, an active substance is used which acts in collaboration with elements naturally present in the EVs. According to another specific example, EVs serve only as vehicles to reach a specific target, sometimes highly protected from conventional administration routes.

Specifically, the binders described herein are provided for any cosmetic, food or industrial purpose. Specific embodiments refer to such TEVs which are provided in a

lipid or oily composition, or encapsulated e.g., for cosmetic or food purposes. Industrial purposes encompass analytical or preparatory purposes, such as to analyse or prepare specific binders (on an industrial scale), respectively.

Specifically, the TEV is carrying a load comprising an autologous or heterologous active substance, in particular a heterologous compound, for medical use as a targeting vector in treating a subject in need of targeted therapy with said compound.

According to a specific aspect, the invention provides for a method of treating a subject in need thereof, by administering an effective amount of a binder described herein, in particular a TEV or a TEV preparation described herein to said subject for therapeutic or diagnostic purposes, such as to improve or detect a certain condition, in particular a disease condition.

Specifically, an effective amount of the TEV is used, wherein dosing is measured according to the administered intravesicular load.

Specifically, the load comprises at least one autologous or heterologous compound. Specifically, the load comprises any one or more of peptides, polypeptides, protein domains, proteins, lipids, genes, nucleic acids such as mRNAs, miRNAs, RNAi mediating molecules in particular locked nucleic acids or phosphorothioates, DNA, DNA fragments, plasmids such as minicircle DNA, drugs such as small molecules, in particular chemotherapeutics or senolytics.

Small molecule drugs are herein understood to as low molecular weight (< 900 D) organic compounds that may regulate a biological process. Small molecules can have a variety of biological functions or applications, serving as cell signaling molecules, drugs in medicine, pesticides in farming, and in many other roles. These compounds can be natural (such as secondary metabolites) or artificial (such as antiviral or chemotherapeutic drugs); they may have a beneficial effect against a disease (such as drugs) or may be detrimental (such as teratogens and carcinogens).

According to a specific aspect, the TEV is originating from a source cell which is autologous to said subject. Specifically, a TEV of autologous origin is used, which is modified and/or loaded *in vitro* (outside the subject's body) for administration *in vivo*, according to the subject's needs.

According to a specific aspect, the invention provides for a TEV preparation comprising isolated TEVs. The TEV preparation specifically is characterized by a homogenous EV population, which consists of at least 50% EVs with the same target

specificity, preferably at least any one of 30%, 40%, 50%, 60%, 70%, 80%, 90%, or 95%.

According to a specific aspect, the TEV preparation is a homogeneous preparation with a median size between 100 and 150 nm, or between 120 and 140 nm. The specific yield is preferably at least 1000 EVs per source cell, more preferably at least 1500, or at least 1600 EVs per source cell.

Specifically, the TEV preparation is provided in a storage-stable aqueous solution or as a lyophilized preparation.

The invention further provides for a pharmaceutical preparation comprising any of the TED, or the TSP, or the TEV described herein, and a pharmaceutically acceptable carrier, preferably in a formulation for intradermal, subcutaneous, intravenous, topical, or oral use.

According to a specific aspect, the invention provides for a method of producing a protein comprising a target-specific extravesicular domain (TED) of an extracellular vesicle (EV) surface protein comprising modifying a polynucleotide comprising a coding sequence encoding the extravesicular domain (ED) of an EV surface protein by a mutagenesis method to obtain mutations of the ED amino acid sequence within at least one modified region with a length of 3-20 contiguous amino acids flanked by regions of the wild-type ED sequence at its N-terminus and C-terminus, to incorporate a target binding site within the ED, thereby producing a repertoire of polynucleotides encoding a variety of TEDs, each comprising a different target binding site, and selecting a protein comprising the TED specifically recognizing a predetermined target, and producing the selected protein.

Specifically, the repertoire of polynucleotides is comprised in genetic packages displaying the variety of TEDs on the outer surface, preferably employing a display system selected from the group consisting of a yeast, phage, bacterium, ribosome, mRNA or mammalian cell display.

Specifically, the protein comprising the selected TED is the TSP described herein.

Specifically, the selected TED is characterized as further described herein.

The invention further provides for a method of producing the TEV preparation described herein, comprising:

- a) introducing a polynucleotide encoding the TSP described herein into a source cell or source cell mixture;
- b) culturing said cell(s) under conditions producing extracellular vesicles;

c) isolating a fraction comprising a TEV comprising the target binding site of the TSP; and

d) producing a preparation of the TEV comprised in said fraction.

Specifically, the source cell or source cell mixture is obtained from a biological sample of a subject.

Specifically, the biological sample of said subject is selected from the group consisting of blood, urine, amniotic fluid, ascites, cerebrospinal fluid, saliva, synovial fluid, or bone marrow.

Specifically, the source cell is isolated before culturing in a cell culture, or cultured within the biological sample.

Specifically, the subject is an animal, such as a mammal, including a human being or non-human animals.

According to a further specific embodiment, there is provided a method of producing a preparation of the TEV described herein, which is originating from a stem cell of a subject, comprising:

a) introducing a polynucleotide or gene encoding the TEV surface protein into a stem cell;

b) culturing said stem cell under conditions producing extracellular vesicles;

c) isolating a fraction comprising the TEV e.g. according to the target binding specificity; and

d) producing a TEV preparation.

Specifically, the stem cell is isolated before culturing in a cell culture, or cultured within the biological sample.

According to a specific embodiment, the EVs are obtained from mesenchymal stem cells (MSCs). MSCs can be prepared by an in vitro proliferation of cell culture, for example, by dispersing an embryonic stem cell colony. Isolation of the EVs, in particular exosomes, from MSCs may be done in a mesenchymal stem cell conditioned medium. The medium may be obtained by culturing MSCs, descendent thereof or a cell line derived therefrom in a cell culture medium and isolating the cell culture medium.

Specifically, the source cell or source cell mixture is obtained from a subject, and the TEV preparation is formulated for autologous use.

Specifically provided herein is an autologous TEV preparation produced by the method described herein, wherein the source cell or source cell mixture is obtained from a subject and the TEV preparation is administered to the same subject.

According to a specific embodiment, TEVs can be targeted to a tumor, and loaded TEVs can be produced upon loading with antigens directly obtained from said tumor.

Specifically, the polynucleotide or gene, which is introduced into a source cell or source cell mixture, encodes a TSP, and the ED of the TSP is bound to the surface of the EV via at least one of the transmembrane domains of the TSP.

According to a specific embodiment, introduction of the polynucleotide or gene encoding the TSP into the source cell can be achieved via transfection. Specifically, the cell(s) is/are transfected with said gene prior to culturing. Specifically, the coding gene is introduced into the cell by any of commonly used transfection methods, such as electroporation, or transfection with apoptosis-inducing agents, such as siRNA, in particular liposome-based transfection.

According to a specific embodiment, the source cell or source cell mixture is cultured in a cell culture under conditions to produce membrane vesicles and to release the TEVs, thereby obtaining the TEVs in a culture supernatant.

Specifically, the cell culture conditions are adapted to the different source cells, or the different biological samples comprising the source cells. According to a specific aspect, the biological samples consist of a biological fluid from a subject (bone marrow, peripheral blood, etc.), a culture supernatant, a cell lysate, a pre-purified solution or any other composition comprising membrane vesicles. Specific cell culture methods for the production of TEVs may further involve inducing oxidative stress. The oxidative stress may be induced by an externally added cytokine or by an oxidant such as hydrogen peroxide.

Exosomes may also be synthesized or manufactured artificially, i.e., not isolated from a human or non-human cell. Instead of being isolated, exosomes could be synthesized by various lipid formation technologies.

Specifically, the source cell or source cell mixture is cultured in a cell culture comprising an active substance *e.g.*, a heterologous compound under conditions to produce extracellular vesicles carrying said compound, preferably by intravesicular loading said vesicles.

Specifically, said intravesicular loading is by incubation, optionally disrupting the membrane, or by binding the load to a component of the membrane. In particular the TEV is loaded by any suitable transfection technique, including reagent-based methods (Calcium phosphate, polyethylenimine, cationic polymers, DEAE-dextran, activated dendrimers, magnetic beads), or instrument-based methods (electroporation,

sonication, bolistic technology, microinjection, laserfection, optoinjection). Alternatively, TEVs can be loaded by binding or fusing a compound to the membrane *e.g.*, by binding to or fusion with membrane lipoproteins.

Loading the TEVs can be carried out *in vitro*, *in vivo* or *ex vivo*.

TEVs can be loaded prior to or after producing the extracellular vesicles.

Specifically, a source cell can be further surface decorated by any suitable method employing biological, enzymatic and/or chemical reactions, such as to produce TEVs comprising modifications *e.g.*, modifying the surface protein glycosylation (*e.g.*, by sialylation, fucosylation, or aglycosylation), post-translational modifications, and/or coupling chemical compounds, drugs, labels, tags, or enzymatic (*e.g.*, enzyme substrate) or chemical reaction groups.

For isolation of TEVs, in particular microvesicles or exosomes, medium of a cell culture of source cells is collected, pre-cleared of cells and debris, and subjected to a series of (ultra)centrifugation steps. Subsequently, the resulting TEV pellet is usually subjected to sucrose density gradient centrifugation, to separate a homogeneous EV population. Specifically, a culture supernatant is treated so as to be enriched with membrane vesicles. In particular, a pre-purified solution obtained from a culture supernatant of a population of membrane vesicle-producing cells or from a biological sample, is subjected to treatments such as centrifugation, clarification, ultrafiltration, nanofiltration and/or affinity chromatography.

A cell culture medium or supernatant can be filtered *e.g.*, through a membrane which has a particular porous size or a particular molecular weight cut-off, particularly employing tangential force filtration or ultrafiltration.

Specifically, the TEV containing fraction is isolated and optionally concentrated by any one or more of binding to affinity ligands, centrifugation, chromatography, clarification, ultrafiltration, or nanofiltration.

According to a specific aspect, the method of preparing the TEVs, particularly of purifying from a biological sample, includes at least one anion exchange chromatography step. Different types of anion exchange materials may be used to perform the anion exchange chromatography, such as including cellulose, poly(styrene-divinylbenzene), agarose, dextran, acrylamide, silica, ethylene glycol-methacrylate copolymer, or mixtures thereof, *e.g.*, agarose-dextran mixtures. The EVs retained on the column may be eluted in different ways, particularly using the passage of a saline solution gradient of increasing concentration. Typically, different fractions purified in this

way are detected by measuring their optical density at the column outlet using a continuous spectrophotometric reading.

As an alternative, or in addition to anion exchange chromatography step, gel permeation chromatography can be used. Typically, a material selected from silica, acrylamide, agarose, dextran, ethylene glycol-methacrylate co-polymer or mixtures thereof, *e.g.*, agarose-dextran mixtures, is used to perform a gel permeation chromatography step.

The invention further provides for a TED library comprising a variety of at least at least any of 10^2 , 10^3 , 10^4 , 10^5 , or 10^6 TEDs described herein, each with a different modified region flanked by the same regions of the same wild-type extravesicular domain (ED) at its N-terminus and C-terminus.

Preferably, the library of TEDs comprises at least any of 10^2 , 10^3 , 10^4 , 10^5 , or 10^6 TEDs, each with different modified region and/or target specificity. Specifically, the repertoire of a library of TEDs comprises at least 2×10^6 , 10^7 , 2×10^7 , 10^8 or 2×10^8 TEDs with different target specificity.

The invention further provides for a TSP library comprising a variety of at least at least any of 10^2 , 10^3 , 10^4 , 10^5 , or 10^6 TSDs described herein, each with a different modified region flanked by the same regions of the same wild-type extravesicular domain (ED) at its N-terminus and C-terminus.

Preferably, the library of TSDs comprises at least any of 10^2 , 10^3 , 10^4 , 10^5 , or 10^6 TSDs, each with different modified region and/or target specificity. Specifically, the repertoire of a library of TSDs comprises at least 2×10^6 , 10^7 , 2×10^7 , 10^8 or 2×10^8 TSDs with different target specificity.

The invention further provides for a TEV library comprising a variety of at least at least any of 10^2 , 10^3 , 10^4 , 10^5 , or 10^6 TEVs described herein, each with a different modified region flanked by the same regions of the same wild-type extravesicular domain (ED) at its N-terminus and C-terminus.

Preferably, the library of TEVs comprises at least any of 10^2 , 10^3 , 10^4 , 10^5 , or 10^6 TEVs, each with different modified region and/or target specificity. Specifically, the repertoire of a library of TEVs comprises at least 2×10^6 , 10^7 , 2×10^7 , 10^8 or 2×10^8 TEVs with different target specificity.

According to a specific aspect, the invention further provides for a library of binders, such as TEDs, TSPs, or TEVs described herein, produced by a method further described herein. Specifically, the method of producing any such library comprises

mutagenizing a nucleic acid sequence comprising a polynucleotide encoding an ED of a EV surface protein by a mutagenesis method to obtain mutations of said ED within at least one predetermined modified region with a length of 3-20 contiguous amino acids flanked by regions of the wild-type ED sequence at its N-terminus and C-terminus, to incorporate a novel target binding site within the ED or the EV surface protein.

The invention further provides for a method of producing a library of target-specific extracellular vesicles (TEVs), comprising:

a) providing a repertoire of polynucleotides encoding a variety of target-specific EV surface proteins (TSPs), each comprising a different target binding site;

b) introducing said repertoire into said source cell(s); and

b) isolating a fraction comprising a repertoire of target-specific extracellular vesicles (TEVs), with a different target binding specificity, to produce a library of TEVs,

wherein the repertoire of polynucleotides is produced by mutagenizing a polynucleotide encoding an EV surface protein by a mutagenesis method to obtain mutations of the extravesicular domain (ED) of said EV surface protein within at least one modified region with a length of 3-20 contiguous amino acids flanked by regions of the wild-type ED sequence at its N-terminus and C-terminus, to incorporate a target binding site within the ED.

Specifically, the repertoire can be produced by a mutagenesis method to modify said polynucleotide or gene within at least a predetermined region of the ED to produce the modified region described herein. Specifically, said mutagenesis method is applied to obtain a randomized amino acid sequence within the modified region. Specifically, said mutagenesis method employs mutagenesis of at least one or two distant regions of said ED or EV surface protein as described herein, thereby producing a repertoire of polynucleotides encoding a variety of EDs or EV surface proteins, each with a different binding specificity, and selecting a polynucleotide encoding the TED or TEV which specifically recognizes the target. Specifically, the repertoire of polynucleotides is comprised in genetic packages displaying the variety of surface proteins on the outer surface of extracellular vesicles.

Specifically, the display package further described herein is used *e.g.*, a replicable genetic package encoding the surface protein described herein, wherein the display package is displayed on extracellular vesicles of natural or artificial origin, preferably selected from the group consisting of exosomes, microvesicles, nanoparticles and liposomes.

Optionally, a polynucleotide encoding the TED or the TSP is selected for producing the respective TED or TSP in a cell culture and/or for producing the respective TEV comprising said TED or TSP.

Mutagenesis of the surface protein or the respective coding sequence may be performed with or without using an EV, such as to introduce one or more point mutation(s) within predetermined regions of the EV surface protein, in particular within those regions which can be presented on the outer surface of an EV or a cell (*i.e.*, the extravesicular or extracellular part of the surface protein).

Specifically, a repertoire of polynucleotides is produced which is comprised in genetic packages displaying the variety of surface proteins on the outer surface.

Specifically, the display package is a replicable genetic package encoding the surface protein described herein, such as selected from the group consisting of a bacteriophage, phagemid, and a cellular display package, preferably a bacterial, mammalian, or insect cell, or yeast, or an *in vitro* display system, such as a ribosome display system. Such *in vitro* display systems specifically translate nucleic acid information into the corresponding protein sequence that exhibits the associated binding specificity and affinity. Specifically preferred methods employ a display system selected from the group consisting of a phage, yeast, bacterium, ribosome, mRNA or mammalian cell display.

According to a specific example, the ED or EV surface protein described herein can be displayed by a bacteriophage or phagemid via an anchor protein selected from the group consisting of the protein product of gene III, gene VI, gene VII, gene VIII or gene IX, where gene III is often favored as it is located at the N-terminus of the phagemid or phage.

According to another specific example, the ED or EV surface protein described herein can be displayed by yeast which routinely displays proteins, *e.g.*, by fusion to the Aga2 yeast surface protein. This is routinely used for yeast display libraries. Anchoring to the yeast cells can also be achieved via coexpression of an anchoring protein such as for example, Protein A. Alternatively, the soluble surface proteins can be secreted from the yeast cells and captured following derivatization of the yeast cell surface by an external addition of protein A or antibodies specifically recognizing a desired target sequence.

A series of display packages displaying different surface proteins, *e.g.*, with different binding properties, may be provided. In particular, a series of display packages

displaying different modifications of the same surface protein, e.g., different modifications within the artificial binding site of the surface protein, may be provided to display those proteins with different binding properties.

In the method as described herein, the display library is specifically contacted with a target, such that a series of members of the display library are selected according to their target binding properties. Amongst the members of the display library recognizing the target, one or more library members may be identified as having high target specificity and/or affinity. Such identification may be specifically understood as a step of selecting binders of any kind, *i.e.* a selection according to binding specificity.

Specific methods allow the simultaneous selection according to both binding and functionality, using any of the following methods and means: Fluorescence activated cell sorting (FACS) is routinely utilized for these studies, with any appropriate target cell, coupled with any detecting antibody, binder protein, or dye such as Annexin V that registers binding or phenotypic changes such as apoptosis. Any similarly sensitive system can track and characterize binding events to target cells by advanced imaging or photodetection systems of microscopic sensitivity. FACS is also a primary technique for fractionating the binders and characterizing their corresponding phenotypic changes in cells.

Additionally, there is a spectrum of technologies that are also used in high throughput proteomics and cell biology, of which highly automated microscopic analysis is one of the most useful for the present purposes. Improvements may be achieved with automated high-information content digital microscopy that robotically analyzes individual cells in samples.

Once a suitable binder is selected its coding sequence can be used to produce the surface protein or to engineer fusion proteins comprising said TSP e.g., a fusion of a transmembrane domain and said TED, and to produce recombinant cells expressing such proteins on the surface of vesicular and/or cellular membranes, or to otherwise produce the TEVs described herein comprising the TSP and/or presenting the TED on its outer surface.

Further provided herein is a method of selecting a candidate binder specifically recognizing a predetermined target from a library described herein, by contacting the repertoire with the target under conditions to allow specific target binding, selecting and isolating a candidate binder with proven target binding specificity.

In the method as described herein, the library is specifically contacted with a target *e.g.*, a target antigen or a target cell, such that the target is bound by members of the library. Amongst the members of the library recognizing the target, one or more library members may be identified as having high target specificity and/or affinity. Such identification may be specifically understood as a step of selecting binders of any kind, *i.e.* a selection according to binding specificity.

FIGURES

Figure 1: Sequences used herein.

SEQ ID NO:7, amino acid sequence of wildtype human CD81 LEL
 SEQ ID NO:8, nucleotide sequence of wildtype human CD81 LEL
 SEQ ID NO:87, amino acid sequence of human CD81
 SEQ ID NO:88, nucleotide sequence of human CD81
 SEQ ID NO:89, amino acid sequence of human CD9
 SEQ ID NO:90, amino acid sequence of human CD53
 SEQ ID NO:91, amino acid sequence of human TSPAN32
 SEQ ID NO:92, amino acid sequence of human CD82
 SEQ ID NO:93, amino acid sequence of human CD63
 SEQ ID NO:94, amino acid sequence of human CD151
 SEQ ID NO:95, amino acid sequence of human CD37
 SEQ ID NO:96, amino acid sequence of human LAMP2
Extravascular domains of CD81, CD9, CD53, TSPAN32, CD82, CD63, CD151, CD37, LAMP2:
 SEQ ID NO:130, EC1 of human CD81
 SEQ ID NO:131, EC2 of human CD81
 SEQ ID NO:132, EC1 of human CD9
 SEQ ID NO:182, EC1 of human CD9
 SEQ ID NO:133, EC2 of human CD9
 SEQ ID NO:134, EC1 of human CD53
 SEQ ID NO:135, EC2 of human CD53
 SEQ ID NO:136, EC1 of human TSPAN32
 SEQ ID NO:137, EC2 of human TSPAN32
 SEQ ID NO:138, EC1 of human CD82
 SEQ ID NO:139, EC2 of human CD82
 SEQ ID NO:140, EC1 of human CD63
 SEQ ID NO:141, EC2 of human CD63
 SEQ ID NO:142, EC1 of human CD151
 SEQ ID NO:143, EC2 of human CD151
 SEQ ID NO:144, EC1 of human CD37
 SEQ ID NO:145, EC2 of human CD37
 SEQ ID NO:146, ED (EC) of human LAMP2
Four transmembrane domains (TM1-4) of CD81, CD9, CD53, TSPAN32, CD82, CD63, CD151, CD37; one TM of LAMP2:
 SEQ ID NO:147, TM1 of human CD81
 SEQ ID NO:148, TM2 of human CD81
 SEQ ID NO:149, TM3 of human CD81
 SEQ ID NO:150, TM4 of human CD81
 SEQ ID NO:151, TM1 of human CD9
 SEQ ID NO:152, TM2 of human CD9
 SEQ ID NO:153, TM3 of human CD9
 SEQ ID NO:154, TM4 of human CD9
 SEQ ID NO:155, TM1 of human CD53
 SEQ ID NO:156, TM2 of human CD53

SEQ ID NO:157: TM3 of human CD53
SEQ ID NO:158: TM4 of human CD53
SEQ ID NO:159: TM1 of human TSPAN32
SEQ ID NO:160: TM2 of human TSPAN32
SEQ ID NO:161: TM3 of human TSPAN32
SEQ ID NO:162: TM4 of human TSPAN32
SEQ ID NO:163: TM1 of human CD82
SEQ ID NO:164: TM2 of human CD82
SEQ ID NO:165: TM3 of human CD82
SEQ ID NO:166: TM4 of human CD82
SEQ ID NO:167: TM1 of human CD63
SEQ ID NO:168: TM2 of human CD63
SEQ ID NO:169: TM3 of human CD63
SEQ ID NO:170: TM4 of human CD63
SEQ ID NO:171: TM1 of human CD151
SEQ ID NO:172: TM2 of human CD151
SEQ ID NO:173: TM3 of human CD151
SEQ ID NO:174: TM4 of human CD151
SEQ ID NO:175: TM1 of human CD37
SEQ ID NO:176: TM2 of human CD37
SEQ ID NO:177: TM3 of human CD37
SEQ ID NO:178: TM4 of human CD37
SEQ ID NO:179: TM of human LAMP2
Further exemplary tetraspanins:
TSPAN8, NP_001356689, SEQ ID NO:184
TSPAN14, NP_001121781, SEQ ID NO:185
CD231 (TSPAN7), NP_004606, SEQ ID NO:186
Integrin family of proteins
CD49d, NP_000876, SEQ ID NO:187
ITGB5, NP_001341694.1, SEQ ID NO:188
ITGB6, NP_000879.2, SEQ ID NO:189
ITGB7, NP_000880.1, SEQ ID NO:190
CD71, NP_003225, SEQ ID NO:191
Proteoglycans
CD138 (syndecan-1), NP_001006947, SEQ ID NO:192
syndecan-2, NP_002989, SEQ ID NO:193
syndecan-3, NP_055469, SEQ ID NO:194
syndecan-4, NP_002990, SEQ ID NO:195
HSPG2, NP_001278789.1, SEQ ID NO:196
5 transmembrane domains protein family
CD133, XP_011512195, SEQ ID NO:195
Type I transmembrane proteins
CD50, NP_001307534, SEQ ID NO:196
CD102, NP_001093259, SEQ ID NO:197
Notch family
NOTCH1, NP_060087, SEQ ID NO:198
NOTCH2, NP_001186930, SEQ ID NO:199
NOTCH3, NP_000426, SEQ ID NO:200
NOTCH4, NP_004548, SEQ ID NO:201
DLL 1, NP_005609, SEQ ID NO:202
DLL4, NP_061947.1, SEQ ID NO:203
JAG1, NP_000205.1, SEQ ID NO:204
JAG2, NP_002217.3, SEQ ID NO:205
CD11a, XP_011544151.1, SEQ ID NO:206
CD11b, NP_001139280.1, SEQ ID NO:207
CD11c, NP_001139280.1, SEQ ID NO:208
CD18/ITGB2, NP_001120963.2, SEQ ID NO:209
CD41, NP_000410.2, SEQ ID NO:210
CD51, NP_001138472.1, SEQ ID NO:211
CD61, NP_000203.2, SEQ ID NO:212

CD104, NP_001308052.1, SEQ ID NO:213

Membrane proteins with enzymatic activity (enzymatic TM proteins)

CD13, NP_001141.2, SEQ ID NO:214

Immune regulatory surface proteins including Fc receptors, T-cell receptor, complement receptors, interleukin receptors, immunoglobulins, MHC-I or MHC-II components

CD2, NP_001758.2, SEQ ID NO:215

CD3 epsilon, NP_000724.1, SEQ ID NO:216

CD3 zeta, NP_932170.1, SEQ ID NO:217

CD18, NP_001120963.2, SEQ ID NO:218

CD19, NP_001761.3, SEQ ID NO:219

CD30, NP_001268359.2, SEQ ID NO:220

CD34, NP_001764.1, SEQ ID NO:221

CD36, NP_001120915.1, SEQ ID NO:222

CD40, NP_001289682.1, SEQ ID NO:223

CD40L, NP_000065.1, SEQ ID NO:224

CD44, NP_001001391.1, SEQ ID NO:225

CD45, NP_001254727.1, SEQ ID NO:226

CD47, NP_001768.1, SEQ ID NO:227

CD86, NP_787058, SEQ ID NO:228

CD110, NP_005364.1, SEQ ID NO:229

CD111, NP_976031.1, SEQ ID NO:230

CD115, NP_001336665.1, SEQ ID NO:231

CD117, XP_016863667.1, SEQ ID NO:232

CD125, XP_011531979.1, SEQ ID NO:233

CD135, XP_011533319.1, SEQ ID NO:234

CD184, NP_001334985.1, SEQ ID NO:235

CD200, NP_001352780.1, SEQ ID NO:236

CD279, NP_005009.2, SEQ ID NO:237

CD273, NP_079515.2, SEQ ID NO:238

CD274, NP_054862.1, SEQ ID NO:239

CD362 = syndecan-2 (see SEQ ID NO:193)

EGFR, NP_958441.1, SEQ ID NO:240

L1CAM, NP_001137435.1, SEQ ID NO:241

LFA-1, NP_001120963.2, SEQ ID NO:242

LGALS3BP, NP_005558.1, SEQ ID NO:243

MFGE8, NP_001297248.1, SEQ ID NO:244

SLIT2, NP_001276064.1, SEQ ID NO:245

STX3, NP_004168.1, SEQ ID NO:246

Surface markers of mesenchymal stem cells

CD44 (SEQ ID NO:225, as above)

CD45 (SEQ ID NO:226, as above)

CD71 (SEQ ID NO:191, as above)

CD73 (also included in the group of enzymatic TM proteins), NP_001191742.1, SEQ ID NO:247

CD90, NP_001298091.1, SEQ ID NO:248

CD29 (also included in the group of the integrin family), NP_596867.1, SEQ ID NO:249

CD105, NP_001265067.1, SEQ ID NO:250

CD106 (also included in the group of the sialoglycoproteins), NP_001069.1, SEQ ID NO:251

CD146, NP_006491.2, SEQ ID NO:252

CD164, NP_001135874.1, SEQ ID NO:253

CD166, NP_001618.2, SEQ ID NO:254

STRO-1, NP_004958.2, SEQ ID NO:255

Glycoproteins

CD54, NP_000192.2, SEQ ID NO:256

Sialoglycoprotein CD235a, NP_001295116, SEQ ID NO:257

Channelling proteins, including Ca-channel proteins

GLUR2, NP_000817.3, SEQ ID NO:258

GLUR3, NP_000819.3, SEQ ID NO:259

HLA-DM, NP_006111.2, SEQ ID NO:260

Miscellaneous:

FLOT2, NP_004466, SEQ ID NO:261

EV surface protein sequences are herein provided as amino acid sequences which may or may not include a signal sequence. It is herein understood that the EV protein sequences as used for the purpose of engineering binders such as TED, TSP, or TEV described herein, are those without the signal sequence of the respective sequence information, if any. The skilled person can easily identify which is a signal sequence included as the N-terminal part of a sequence identified herein.

Protein reference numbers referred to herein are respective NCBI References (National Center for Biotechnology Information, U.S. National Library of Medicine 8600 Rockville Pike, Bethesda MD, 20894 USA).

Figure 2: Schematic illustration of a two-dimensional model of human CD81 (SEQ ID NO:87). The extracellular domain contains two loops, the small extravesicular loop (SEL) from amino acid Trp34 to Tyr63 and the large extravesicular loop (LEL) from Phe113 to Lys201. It comprises four transmembrane domains (TM), TM1 spanning Val12 to Leu33, TM2 spanning Tyr64 to Gln85, TM3 spanning Leu90 to Gly112 and TM4 spanning Leu202 to Met228.

Figure 3: Two different views of the structure of the wildtype tetraspanin CD81. The transmembrane regions TM1-TM4, the small extravesicular loop (indicated by arrows), the structurally conserved domain of the large extravesicular loop (helices A, B and E), and the variable region of this domain (helices C and D) is shown. The disulfide bridges (DB) of the LEL are indicated, and the intravesicular cysteine residues that are probable sites of palmitate attachment are marked by asterisks. The variable region of the LEL contains two (as for wildtype CD81 as shown here), three or four segments separated by cysteine residues.

Figure 4: Localization of recombinant CD81 at HeLa cell membranes.

Figure 5: Recombinant Snorkel tag at the membrane. (A): snorkel-tagged CD81 (B): snorkel-tagged CD63 as visualized using anti HA antibodies in transiently transfected HeLa cells. (C) Immunogold-labelling in electron microscopy to confirm CD81 and HA positivity of recombinant EVs.

Figure 6: Elution of recombinant EVs carrying a Snorkel Tag using PreScission protease.

Figure 7: Characterization of recombinant EVs by size, amount and ratio of EVs carrying the recombinant CD81-GFP fusion protein.

Figure 8: EV uptake in CaCo-2 (A) Uptake visualized by fluorescence microscopy (B) quantitation of uptake positive cells by flow cytometry.

Figure 9: Treatment of Caco-2 cells with recombinant EVs containing toxic siRNA.

DETAILED DESCRIPTION

Unless indicated or defined otherwise, all terms used herein have their usual meaning in the art, which will be clear to the skilled person. Reference is for example made to the standard handbooks, such as Sambrook et al, "Molecular Cloning: A Laboratory Manual" (2nd Ed.), Vols. 1 -3, Cold Spring Harbor Laboratory Press (1989); Lewin, "Genes IV", Oxford University Press, New York, (1990), and Janeway et al, "Immunobiology" (5th Ed., or more recent editions), Garland Science, New York, 2001.

The subject matter of the claims specifically refers to artificial products or methods employing or producing such artificial products, which may be variants of native (wild-type) products. Though there can be a certain degree of sequence identity to the native structure, it is well understood that the materials, methods and uses of the invention, *e.g.*, specifically referring to isolated nucleic acid sequences, amino acid sequences, expression constructs, transformed host cells and recombinant proteins, are "man-made" or synthetic, and are therefore not considered as a result of "laws of nature".

The term "domain" with respect to a protein domain such as an ED or a transmembrane domain is herein understood as a polypeptide or protein of a contiguous amino acid sequence which is at least a certain part (or the full length) of a polypeptide or protein. The domain can be comprised in a larger protein. Yet, a protein domain is also called domain while being isolated from a protein that is larger than the domain.

The term "expression" is understood in the following way. Nucleic acid molecules containing a desired coding sequence of an expression product such as *e.g.*, an antibody as described herein, and control sequences such as *e.g.*, a promoter in operable linkage, may be used for expression purposes. Hosts transformed or transfected with these sequences are capable of producing the encoded proteins. In order to effect transformation, the expression system may be included in a vector; however, the relevant DNA may also be integrated into the host chromosome. Specifically the term refers to a host cell and compatible vector under suitable conditions, *e.g.*, for the expression of a protein coded for by foreign DNA carried by the vector and introduced to the host cell.

Coding DNA is a DNA sequence that encodes a particular amino acid sequence for a particular polypeptide or protein such as *e.g.*, an antibody. Promoter DNA is a DNA sequence which initiates, regulates, or otherwise mediates or controls the expression of the coding DNA. Promoter DNA and coding DNA may be from the same gene or from different genes, and may be from the same or different organisms. Recombinant cloning vectors often include one or more replication systems for cloning or expression, one or more markers for selection in the host, *e.g.*, antibiotic resistance, and one or more expression cassettes.

“Expression vectors” or “vectors” as used herein are defined as DNA sequences that are required for the transcription of cloned recombinant nucleotide sequences, *i.e.* of recombinant genes and the translation of their mRNA in a suitable host organism.

An “expression cassette” refers to a DNA coding sequence or segment of DNA that codes for an expression product that can be inserted into a vector at defined restriction sites. The cassette restriction sites are designed to ensure insertion of the cassette in the proper reading frame. Generally, foreign DNA is inserted at one or more restriction sites of the vector DNA, and then is carried by the vector into a host cell along with the transmissible vector DNA. A segment or sequence of DNA having inserted or added DNA, such as an expression vector, can also be called a “DNA construct”.

Expression vectors comprise the expression cassette and additionally usually comprise an origin for autonomous replication in the host cells or a genome integration site, one or more selectable markers (*e.g.*, an amino acid synthesis gene or a gene conferring resistance to antibiotics such as zeocin, kanamycin, G418 or hygromycin), a number of restriction enzyme cleavage sites, a suitable promoter sequence and a transcription terminator, which components are operably linked together. The term “vector” as used herein includes autonomously replicating nucleotide sequences as well as genome integrating nucleotide sequences. A common type of vector is a “plasmid”, which generally is a self-contained molecule of double-stranded DNA that can readily accept additional (foreign) DNA and which can readily be introduced into a suitable host cell. A plasmid vector often contains coding DNA and promoter DNA and has one or more restriction sites suitable for inserting foreign DNA. Specifically, the term “vector” or “plasmid” refers to a vehicle by which a DNA or RNA sequence (*e.g.*, a foreign gene) can be introduced into a host cell, so as to transform the host and promote expression (*e.g.*, transcription and translation) of the introduced sequence.

The term “extracellular vesicles”, abbreviated as “EVs”, including those EVs with target binding specificity such as TEVs described herein, is herein understood as cellular vesicles, or those vesicles originating from a cell, which are provided outside a cell. EVs can not only be produced *in vivo* or *ex vivo* by respective source or donor cells, but can also be artificially produced without using a cell, *e.g.*, by *in vitro* methods of engineering liposomes or nanoparticles to produce synthetic EVs comprising the EV features as described herein.

EVs are typically membrane-packed vesicles that are secreted by a variety of cell types, including T cells, B cells, dendritic cells, platelets, mast cells, epithelial cells, endothelial cells, neuronal cells, cancerous cells, oligodendrocytes, Schwann cells, embryonic cells, and MSCs. EVs are also naturally occurring in physiological fluids such as normal urine, blood, bronchial lavage fluid, breast milk, saliva, cerebrospinal fluid, amniotic fluid, synovial fluid, and malignant ascites. It has been demonstrated that EVs perform an important role in cell-to-cell communication. They mediate intercellular communication, enabling the transfer of functional nucleic acids from the cell of origin to the recipient cells. They have been implicated in processes such as immune responses, homeostasis maintenance, coagulation, inflammation, cancer progression, angiogenesis, and antigen presentation. Thus, EVs participate in many physiological and pathological conditions.

EVs may contain biomolecular or synthetic cargo herein referred to as “load”. Thus, they make an attractive delivery vehicle for targeted therapeutics or diagnostics owing to their stability in circulation, biocompatibility, low immunogenicity and toxicity profiles. EVs are specifically able to transport compounds between cells, including neurons. Advantageously, they are able to cross the blood brain barrier. This natural trafficking ability gives extracellular vesicles the potential to be used as delivery vehicles for various autologous or heterologous compounds.

Exemplary EVs described herein are exosomes or microvesicles.

Exosomes are a type of extracellular membrane-enclosed vesicle, which contains molecular constituents of the cell in which it was secreted from.

Exosomes constitute one of the main subclasses of EVs and have an endosomal origin. Exosomal EVs are nanometer-sized vesicles of endocytic origin that form by inward budding of the limiting membrane of multivesicular endosomes (MVEs). Thus, their size is equivalent to that of the intraluminal vesicle within MVEs, which generally ranges between 30nm and 120nm, preferably 50 to 100nm.

The biogenesis of exosomal EVs occurs via the endocytosis-exocytosis pathway when cells absorb small amounts of intracellular fluid in a specific membrane region and form early endosomes. The early endosome begins to mature and expands into a late endosome; then intraluminal vesicles or multivesicular bodies (MVBs) are formed by internal budding of the endosomal membrane. The MVBs then fuse to the cell membrane and are released into the extracellular environment. At this point the vesicles are named exosomes, which are released via exocytosis that is regulated by p53 and under the control of the cytoskeleton activation pathway but not affected by calcium. Exosomal EVs may have a diameter of 30–100 nm and a density of 1.13 to 1.19 g/mL in a sucrose gradient; they can be collected by centrifugation *e.g.*, at 100,000 g. After isolation, they can be stored as a lyophilisate, or in an aqueous solution *e.g.*, at room temperature, or at refrigerator temperature (2°-8°C), or frozen, *e.g.*, without any toxic cryoprotectant agents (at -80°C, or higher up to -18°C) for more than 6 months while maintaining their functions.

Exosomal EVs may contain large amounts of surface proteins such as annexins, tetraspanins *e.g.*, CD63, CD81, and CD9, and heat-shock proteins, including Hsp60, Hsp70, and Hsp90. They also express Alix, tumor susceptibility gene 101 (Tsg101), and clathrin. Exosomal EVs specifically comprise a lipid bilayer membrane that protects their contents and enables them to move long distances in tissues. The membrane typically possesses small amounts of phosphatidylserine and large amounts of cholesterol, ceramide, and sphingolipids.

Microvesicles are a type of membrane-enclosed vesicle, derived from fragments of plasma membrane. Microvesicular EVs typically bud from the cell surface and their size may vary between 50nm to 1000nm. Artificial microvesicular vesicles, such as semi-synthetic EVs and fully-synthetic EVs, can have a size ranging from 10 to 1000nm, preferably 10 to 500nm, even more preferably 10 to 100nm.

Microvesicular EVs are typically isolated by ultracentrifugation with a density of 1.04 to 1.07 g/mL in a sucrose gradient. Microvesicular EVs typically contain high amounts of phosphatidylserine-containing proteins associated with lipid rafts and are rich in the surface marker CD40 as well as cholesterol, sphingomyelin, and ceramide. Specifically, they are also encapsulated in a lipid bilayer membrane and comprise transmembrane proteins, such as tetraspanins.

Typically, apoptotic body EVs are released through outward blebbing and fragmentation of the cell membrane of apoptotic cells, and have a broad size range of 50–2,000 nm in diameter.

EVs typically interact with targets via surface ligand and adhesion molecules *e.g.*, with target cells. In some cases, they may enter cells via endocytic uptake or by direct fusion of the vesicles to the cell membrane. They may also transmit their contents through adhesion to the cell surface mediated by the interaction of a lipid–ligand receptor. These interactions indicate that EVs may possess pivotal roles in cell-to-cell communication and immune modulation in different physiologic and pathologic conditions.

Nano-sized EVs represent an excellent alternative for drug delivery. As the composition of EV's membrane is typically from a source or donor cell (*e.g.*, stem cells), these particles are non-immunogenic in nature allowing them to resist to fast clearance from circulation and thereby increasing the drug delivery efficiency to target tissues. They are known to naturally possess specific cell tropism or homing ability by cell type specific proteins (with their surface ligand and adhesion molecules), one of the key requirements for targeted drug delivery. However, natural targets are limited and problems with affinity and specificity are common. By engineering EVs comprising a surface protein as described herein, the target-specific EVs are provided which can be targeted to any desired target or cell type with great specificity and affinity.

EVs described herein are particularly useful for medical purposes *e.g.*, to diagnose or treat diseases or disease conditions, in particular those diseases where targeted therapy has proven to ameliorate disease conditions.

Mesenchymal stem cell-derived EVs, especially exosomal EVs, may be particularly useful with regard to their use as regenerative therapies. EVs originating from mesenchymal stem cells (MSCs) can carry biologically active molecules which can be transferred to target cells to exert their therapeutic effects like regenerating tissue injuries suppressing inflammatory responses modulating the immune system and many other beneficial effects. Accordingly, EVs can be an effective safe and cheap therapeutic approach in cell-free regenerative medicine.

EVs can be a suitable drug delivery system, in particular to cross biological barriers (*e.g.*, the blood brain barrier) and deliver their load to otherwise inaccessible sites.

EVs can be formulated to exhibit intended drug carrying activity through various approaches including biological, chemical and physical means. Encapsulation of active substances or drugs (e.g., chemicals, RNAs, DNA, proteins or lipids) into EVs can greatly increase their bioavailability by preserving their integrity and biological activity *in vivo*. Lipid membranes from donor cells are suited to avoid phagocytosis, degradation and modification in host circulation. In particular, autologous but also heterologous EVs typically avoid entrapment in the reticuloendothelial system (also known as mononuclear phagocytic system) and are non-immunogenic in most, if not all, parameters.

Various approaches can be utilized for loading active substances agents into EVs. These include (1) loading to purified EVs *ex vivo*, or (2) pre-loading to donor (source) cells prior to EV production, each followed by isolation and optionally purification.

Ex vivo loading strategies mostly utilize passive packaging of therapeutic molecules, ranging from simple incubation to more sophisticated chemical and/or physical methods. Hydrophobic (*i.e.*, lipophilic) molecules, such as anti-oxidants, anti-cancer drugs, lipophilic dyes, can be spontaneously packaged into EV under ambient conditions. Indeed, successful loading of curcumin, doxorubicin and paclitaxel into EVs has been demonstrated. Compared to standard liposomes composed of phosphatidylcholine and cholesterol, EVs exhibit higher loading efficiency and loading capacity to hydrophobic chemical drugs.

Many active substances cannot penetrate the membrane of EVs freely, therefore loading of EVs is typically effected e.g., by means of electroporation, sonication, permeabilization, fusogenic liposomes, polymeric carriers and/or other physical insults. Sonication and extrusion, or permeabilization with saponin, have been shown to result in stable EV reformation with high loading efficiency.

Electroporation specifically applies an electrical field to create pores in the membrane of EVs temporally, thereby allowing the movement of active substances into the lumen of EVs. Electroporation is also known to induce vesicular aggregation thereby affecting the integrity of the vesicles. The skilled person may choose several parameters including EV sources and concentrations, the cargo molecules (the load) and the applying voltage with time for optimal loading of the cargo.

EVs are conveniently loaded upon electroporation. Studies demonstrated the enhanced efficacy with decreased adverse effects typically associated with chemotherapeutic drugs when compared to either EV-free drugs or drug-loaded liposomes.

EVs are natural carriers of various nucleic acid molecules e.g., mRNA, miRNA and various noncoding RNAs, or DNA molecules, and thus represent suitable vehicles for nucleic acid transfer. Although nucleic acid molecules are effective means for the regulation of genes of interests, their low stability and transducibility in circulation dictates the necessity of vehicles that can protect and deliver these therapeutic molecules to target cells and tissues. Again, electroporation can be performed to load the material into EVs.

Sonication can be a suitable alternative for active loading of molecules with minimal aggregation and degradation.

Different methods of pre-loading of drugs to donor (source) cells prior to the release of EVs exist in the field. For example, active substances can be incorporated to EVs from host cells, in particular recombinant host cells overexpressing a protein of interest or a cellular metabolite. As described herein, EVs can be isolated from donor cells transfected with heterologous genes in addition to the gene encoding the surface protein incorporating the target binding site. Since the load may comprise proteins, loading of recombinant proteins expressed by host cells can be an attractive mode of protein delivery by loaded EVs. A number of model proteins, including ovalbumin, catalase, glial cell-line derived neurotrophic factor (GDNF) have already been successfully loaded into EVs from gene-modified host cells.

The use of target-specific EVs as described herein represents a next generation drug delivery system with ability to transverse complex biological barriers such as the blood brain barrier, while avoiding or overcoming a number of safety concerns related to drugs or vehicles, such as cytotoxicity, short biodistribution and low efficiency of targeted delivery. Chemical drugs and biological molecules with low stability in circulation and/or low transducibility to target cells can be efficiently transferred to cytoplasm of target cells without undergoing endosomal and lysosomal degradation.

As is described herein, EVs may be produced to target certain tissues, cells, artificial surfaces or soluble compounds by the artificial target binding site. EVs can be engineered for the purpose of expressing suitable surface proteins and structures incorporating the binding site. For example, surface proteins can be overexpressed in source cells to be expressed on the surface of the EVs employing suitable recombinant expression systems in the source cells.

The artificial target binding site is either engineered prior to or after vesicle formation. For example, the binding site may be incorporated within the surface protein

as described herein at pre-determined regions to comprise e.g., loop, helical, and/or linear (peptide) structures. Specific target contact surfaces or binding residues within said at least two distant regions of said surface protein may be produced *in situ*, i.e. when producing the EV e.g., by methods of recombining nucleic acid molecules such as employing methods of mutagenizing a source cells producing point mutations in the respective surface proteins, and/or by further modifications of EVs involving biologic, enzymatic and/or chemical reactions.

The EVs described herein are suitably provided in an EV preparation comprising isolated EVs. EVs may be characterized by certain features which can be determined by suitable quality control measures, such as determining the size, density, the amount and composition of the load, the target binding affinity and/or selectivity, purity, etc.

Exemplary methods of quality control are further described in the Examples section.

The term “extravesicular domain”, abbreviated ED, as used herein is understood to encompass a protein domain which is positioned at the outer surface (the extravesicular surface) of an EV when attached or bound to the EV. Yet, a protein domain is also called an ED while being isolated from an EV or isolated from an EV surface protein.

The term “flanking” or “flanked” as used herein with respect to elements of an amino acid sequence or nucleotide sequence is herein understood as follows: A first sequence element is said to be “flanked” by a second sequence element when the first sequence element is located immediately adjacent to the second sequence element, thereby providing a contiguous sequence of said first and second sequences. The linear first sequence element may be flanked by another element at only one of its termini or at both termini, i.e. on one or both sides.

In a TED described herein, the modified region is specifically flanked by two flanking sequences, one at its N-terminus, and one at its C-terminus. Thus, such modified region is positioned between the flanking sequences, also called “embedded”

The term “host cell” as used herein shall refer to primary subject cells transformed to produce a particular recombinant protein, such as a surface protein as described herein, or to produce the EV described herein, and any progeny thereof. It should be understood that not all progeny are exactly identical to the parental cell (due to deliberate or inadvertent mutations or differences in environment), however, such altered progeny are included in these terms, so long as the progeny retain the same functionality as that

of the originally transformed cell. The term "host cell line" refers to a cell line of host cells as used for expressing a recombinant gene to produce recombinant polypeptides or proteins. The term "cell line" as used herein refers to an established clone of a particular cell type that has acquired the ability to proliferate over a prolonged period of time. Such host cell or host cell line may be maintained in cell culture and/or cultured to produce a recombinant polypeptide.

The term "isolated" or "isolation" as used herein with respect to a modified region of an ED shall refer to a peptide consisting of the amino acid sequence of the modified region that has been sufficiently separated from the flanking sequences of the ED, so as to exist in "substantially pure" form. "Isolated" does not necessarily mean the exclusion of artificial or synthetic peptides, or mixtures of such peptides with other compounds or materials, or the presence of impurities that do not interfere with the fundamental binding activity, and that may be present, for example, due to incomplete purification. The term "isolated" is also meant to include those chemically synthesized.

In particular, an modified region can be isolated from the TED as described herein, or isolated from the TSP described herein, or provided as a respective isolated peptide for the purpose of comparing its binding properties, such as target binding affinity and/or specificity as compared to the same modified region that is comprised in (not isolated from) the TED and TSP, respectively.

The term "isolated" or "isolation" as used herein with respect to an EV described herein shall refer to such vesicle that has been sufficiently separated from the environment with which it would naturally be associated, so as to exist in "substantially pure" form. "Isolated" does not necessarily mean the exclusion of artificial or synthetic mixtures with other compounds or materials, or the presence of impurities that do not interfere with the fundamental activity, and that may be present, for example, due to incomplete purification. In particular, isolated EVs as described herein are also meant to include those chemically synthesized.

With reference to polypeptides or proteins, such as EV surface proteins described herein, the term "isolated" shall specifically refer to compounds that are free or substantially free of material with which they are naturally associated such as other compounds with which they are found in their natural environment, or the environment in which they are prepared (e.g. cell culture) when such preparation is by recombinant DNA technology practiced *in vitro* or *in vivo*.

A “library” as described herein with respect to a binder described herein, in particular a TED, TSP, or TEV described herein, is understood to comprise a repertoire of binders that includes a number of different target binding species (library members) that covers a certain variety of binders. The library typically comprises library members which can be distinguished by their functional binding and thus be selected according to the desired binding properties.

A TED library described herein specifically includes a set or a collection of TEDs, in particular TEVs described herein, each with a different modified region embedded in the same wild-type sequences of the same wild-type ED.

A TSP library described herein specifically includes a set or a collection of TSPs, in particular TSPs described herein, each with a different TED embedded in the same wild-type sequences of the same wild-type EV surface protein. The TSP library may comprise a library of EDs, such as a TED library. A TED library is suitably produced by mutagenizing the ED or an EV surface protein thereby producing a repertoire of ED and EV surface protein mutants, respectively, either as soluble proteins or on the surface of an EV.

A TEV library described herein specifically includes a set or a collection of TEVs, in particular TEVs described herein, each with a different TED embedded in the same wild-type sequences of the same wild-type EV surface protein bound to the membrane of the vesicle. The TEV library may comprise a library of surface proteins, such as a TSP library.

Libraries can be constructed by well-known techniques, involving suitable methods of mutagenesis, *e.g.*, site-directed mutagenesis of extravesicular domains of a surface protein.

Libraries as described herein preferably comprise at least 10^2 library members, more preferred at least 10^3 , more preferred at least 10^4 , more preferred at least 10^5 , more preferred at least 10^6 library members, more preferred at least 10^7 , more preferred at least 10^8 , more preferred at least 10^9 , more preferred at least 10^{10} , more preferred at least 10^{11} , up to 10^{12} members of a library.

Specifically, the library comprises at least 10^2 , 10^3 , 10^4 , 10^5 , or 10^6 library members, wherein each library member differs in at least one nucleotide in the sequence of the modified surface protein. Specifically, the library comprises at least 10^6 library members, wherein each library member has a different target binding site or specificity.

Any protein or gene diversity library may be used for the purpose described herein, which, *e.g.*, includes a high number of individual library members, to create a diversity of sequences, or employing preselected libraries, which are *e.g.*, enriched in stabilized or functionally active library members.

For example, a display system can couple a given protein, herein the surface protein described herein, with its encoding nucleic acid, *e.g.*, its encoding mRNA, cDNA or genes. Thus, each member of a library comprises a nucleic acid encoding the modified surface protein which is displayed thereon. Display systems encompass, without being limited to, cells, virus such as phages, ribosomes, eukaryotic cells such as yeast, DNAs including plasmids, and mRNA display.

As is well-known in the art, there is a variety of display and selection technologies that may be used for the identification and isolation of proteins with certain binding characteristics and affinities, including, for example, display technologies such as cellular and non-cellular methods, in particular mobilized display systems. Among the cellular systems the phage display, virus display, yeast or other eukaryotic cell display, such as mammalian or insect cell display, may be used. Mobilized systems are relating to display systems in the soluble form, such as *in vitro* display systems, among them ribosome display, mRNA display or nucleic acid display.

Specific libraries are provided for herein to display a diversity of surface proteins, and/or a diversity of the surface proteins anchored to an EV or cell. Preferably, the libraries are phage display or yeast libraries. Specifically, the yeast host cell exhibits surface protein described herein at the surface of the yeast cell. Phage and phagemid display systems are well-known for their versatility and potential ability to streamline the selection process. Yeast display offers a number of attractive features: The eukaryotic transcription and translation machinery is very well suited for expression of proteins, and the use of flow cytometry allows high-throughput quantitative analysis of individual clones in real-time, using scaffold ligands.

The yeast host cell is preferably selected from the genera *Saccharomyces*, *Pichia*, *Hansenula*, *Schizosaccharomyces*, *Kluyveromyces*, *Yarrowia* and *Candida*. Most preferred the host cell is *Saccharomyces cerevisiae*.

In certain cases, a repertoire of surface proteins described herein is displayed such that entity comprising DNA, RNA or cDNA encoding a surface protein described herein may be directly connected to the surface protein that it encodes, as in RNA or

DNA display libraries. In such case, the surface protein variants are generated by modifications of the methods of cell-free protein synthesis.

Screening for binding activity (or any other desired activity) in the library is conducted according to methods well-known in the art, for instance from phage display technology. For example, targets immobilized to a solid phase can be used to identify and isolate binding members of a repertoire. Screening allows selection of members of a repertoire according to desired characteristics.

In a method of selecting suitable binders of a target, it is advantageous to provide a large multiplicity of each binder in the repertoire of library members, *e.g.*, of at least 10 copies, to increase the chance of selecting one or more candidate binding sequences, which can be further characterized for the suitability to engineer a target-specific extracellular vesicle construct.

Screening the library for library members comprising a target-binding structure may be done by any suitable selection method. The screening step may comprise one or several rounds of selection (also referred to as panning).

One or several rounds of selection encompass *e.g.*, 1, 2, or preferably 3, and may encompass 4, 5, 6, 7, 8, 9, or 10 rounds of selection. In particular, the rounds of selection may comprise incubating the library in the presence of said target, so as to select the proteins which bind said target, or an epitope thereof.

The term “mutagenesis” as used herein refers to any art recognized technique for altering a polynucleotide or polypeptide sequence. Preferred types of mutagenesis include error prone PCR mutagenesis, saturation mutagenesis, or other site directed mutagenesis. Any of the known mutagenesis methods may be employed, to introduce point mutations at desired positions, *e.g.*, by randomisation techniques. In some cases positions are chosen randomly, *e.g.*, with either any of the possible amino acids or a selection of preferred amino acids to randomise the antibody sequences.

The term “recombinant” as used herein shall mean “being prepared by or the result of genetic engineering”. Alternatively, the term “engineered” is used. For example, a surface protein may be mutated to produce a variant by engineering the respective parent sequence to produce a variant of the parent sequence. A recombinant host specifically comprises a recombinant expression vector or cloning vector, or it has been genetically engineered to contain a recombinant nucleic acid sequence, in particular employing nucleotide sequence foreign to the host. A recombinant protein is produced by expressing a respective recombinant nucleic acid in a host.

A “point mutation” is herein particularly understood as the engineering of a polynucleotide that results in the expression of an amino acid sequence that differs from the non-engineered amino acid sequence in the substitution or exchange, deletion or insertion of one or more amino acids for different amino acids, at a certain position (at one position) of the amino acid sequence. Specifically, one or more single (non-consecutive) or doublets of amino acid residues may be subject to a point mutation. Specifically preferred methods of mutagenesis provide of point mutations at selected positions, preferably the substitution of one amino acid by another one at one (predetermined) amino acid position or more amino acids at only a predetermined amino acid position. One or more point mutations can be within a modified region, in particular point mutations at non-consecutive or consecutive positions within the region.

The term “repertoire” as used herein shall refer to a collection of variants, such as variants of the modified surface protein with a variety of specificities to bind the target with high affinity. Typically, the structure of the surface protein, which comprises the extracellular domain with helix and loop regions, is the same in such repertoire. The variety will specifically reflect the diversity of the binding site comprising binding residues within one or more predetermined positions or regions to be modified e.g., at least two distant regions which are part of the same target binding site to specifically recognize and bind the target.

The repertoire as described herein is specifically provided within a library, which is a mixture of heterogeneous surface proteins, target-specific extracellular vesicle constructs or targets. The library may take the form of a simple mixture of the proteins or EVs, or may be in the form of respective binding regions such modified surface proteins, either as isolated polypeptides or proteins, or as nucleic acids encoding such polypeptides or proteins, or even organisms or cells expressing the nucleic acids, e.g., for example bacteria, viruses, animal or plant cells and the like, transformed with a library of nucleic acids, reflecting the variety of target-specific binders of said repertoire.

The term “subject” as used herein shall refer to a warm-blooded mammalian, particularly a human being or a non-human animal. Thus, the term “subject” may also particularly refer to animals including dogs, cats, rabbits, horses, cattle, pigs and poultry. In particular, the antibody as described herein is provided for medical use to treat a subject or patient in need of prophylaxis or treatment of a disease condition. The term “patient” includes human and other mammalian subjects that receive either prophylactic

or therapeutic treatment. The term “treatment” is thus meant to include both prophylactic and therapeutic treatment.

The term “surface protein” including “EV surface proteins” as used herein refers to a protein located on or at the surface of an EV, which is anchored within in the lipid bilayer membrane of the EV. To this end “anchoring” is herein understood to bind to the cellular surface by fusion to a protein domain which is located within the membrane. A surface protein is herein understood to comprise at least one ED and at least one transmembrane domain. Such surface protein can be bound to an EV through integration of said at least one transmembrane protein domain to the membrane of an EV. The transmembrane domain(s) may be part of the surface protein, or fused to a surface protein, e.g., for the purpose of anchoring the surface protein to the EV.

The term “EV surface protein” specifically includes “exosomal proteins” and those proteins of an EV that can be utilized to transport a polypeptide or protein construct to a suitable vesicular structure or EV. Specifically, the term includes any protein that enables transporting, trafficking or shuttling of a polypeptide or protein construct to a vesicular structure, such as an EV. Examples of such EV surface protein are for instance isolated, synthetic and or recombinant amino acid sequence, comprising in whole or in part (as fragments) one or more types of the EV surface proteins described herein, such as identified by the sequences provided herein (in particular in Figure 1) (excluding signal sequences, if any), or isoforms thereof, including cloned isoforms.

Exemplary EV surface proteins are

a) tetraspan-like proteins including

- i) tetraspanins e.g., any one of the group consisting of CD81 (such as human CD81, SEQ ID NO:87), CD9 (such as human CD9, SEQ ID NO:89), CD53 (such as human CD53, SEQ ID NO:90), TSPAN32 (such as human TSPAN32, SEQ ID NO:91), CD82 (such as human CD82, SEQ ID NO:92), CD63 (such as human CD63, SEQ ID NO:93), CD151 (such as human CD151, SEQ ID NO:94), CD37 (such as human CD37, SEQ ID NO:95), TSPAN8 (such as human TSPAN8, SEQ ID NO:184), TSPAN14 (such as human TSPAN14, SEQ ID NO:185), or CD231 (TSPAN7) (such as human CD231 (TSPAN7), SEQ ID NO:186; or
- ii) lysosome-associated membrane proteins, e.g., LAMP2 (such as human LAMP2, SEQ ID NO:96);

b) proteins of the integrin family *e.g.*, CD49d (such as human CD49d, SEQ ID NO:187), ITGB5 (such as human ITGB5, SEQ ID NO:188), ITGB6 (such as human ITGB6, SEQ ID NO:189), ITGB7 (such as human ITGB7, SEQ ID NO:190), CD71 (such as human CD71, SEQ ID NO:191), CD29 (such as human CD29, SEQ ID NO:249);

c) proteoglycans *e.g.*, CD138 (syndecan-1) (such as human CD138 (syndecan-1), SEQ ID NO:192), syndecan-2 (such as human syndecan-2, SEQ ID NO:193), syndecan-3 (such as human syndecan-3, SEQ ID NO:194), syndecan-4 (such as human syndecan-4, SEQ ID NO:195), HSPG2 (such as human HSPG2, SEQ ID NO:196);

d) family of five-transmembrane domains proteins *e.g.*, CD133 (such as human CD133, SEQ ID NO:195);

e) type I transmembrane proteins *e.g.*, (such as human CD50, SEQ ID NO:196), CD102 (such as human CD102, SEQ ID NO:197);

f) Notch family *e.g.*, NOTCH1 (such as human NOTCH1, SEQ ID NO:198), NOTCH2 (such as human NOTCH2, SEQ ID NO:199), NOTCH3 (such as human NOTCH3, SEQ ID NO:200), NOTCH4 (such as human NOTCH4, SEQ ID NO:201), DLL1 (such as human DLL1, SEQ ID NO:202), DLL4 (such as human DLL4, SEQ ID NO:203), JAG1 (such as human JAG1, SEQ ID NO:204), JAG2 (such as human JAG2, SEQ ID NO:205), CD11a (such as human CD11a, SEQ ID NO:206), CD11b (such as human CD11b, SEQ ID NO:207), CD11c (such as human CD11c, SEQ ID NO:208), CD18/ITGB2 (such as human CD18/ITGB2, SEQ ID NO:209), CD41 (such as human CD41, SEQ ID NO:210), CD51 (such as human CD51, SEQ ID NO:211), CD61 (such as human CD61, SEQ ID NO:212), CD104 (such as human CD104, SEQ ID NO:213);

g) membrane proteins with enzymatic activity (enzymatic TM proteins) *e.g.*, CD13 (such as human CD13, SEQ ID NO:214), CD73 (such as human CD73, SEQ ID NO:247);

h) immune regulatory surface proteins, including *e.g.*, Fc receptors, T-cell receptors, complement receptors, interleukin receptors, immunoglobulins, MHC-I or MHC-II components; exemplary proteins are CD2 (such as human CD2, SEQ ID NO:215), CD3 epsilon (such as human CD3 epsilon, SEQ ID NO:216), CD3 zeta (such as human CD3 zeta, SEQ ID NO:217), CD18 (such as human CD18, SEQ ID NO:218), CD19 (such as human CD19, SEQ ID NO:219), CD30 (such as human CD30, SEQ ID NO:220), CD34 (such as human CD34, SEQ ID NO:221), CD36 (such as human CD36, SEQ ID NO:222), CD40 (such as human CD40, SEQ ID NO:223), CD40L (such as human CD40L, SEQ ID NO:224), CD44 (such as human CD44, SEQ ID NO:225), CD45

(such as human CD45, SEQ ID NO:226), CD47 (such as human CD47, SEQ ID NO:227), CD86 (such as human CD86, SEQ ID NO:228), CD110 (such as human CD110, SEQ ID NO:229), CD111 (such as human CD111, SEQ ID NO:230), CD115 (such as human CD115, SEQ ID NO:231), CD117 (such as human CD117, SEQ ID NO:232), CD125 (such as human CD125, SEQ ID NO:233), CD135 (such as human CD135, SEQ ID NO:234), CD184 (such as human CD184, SEQ ID NO:235), CD200 (such as human CD200, SEQ ID NO:236), CD279 (such as human CD279, SEQ ID NO:237), CD273 (such as human CD273, SEQ ID NO:238), CD274 (such as human CD274, SEQ ID NO:239), CD362 = syndecan-2 (such as human, SEQ ID NO:193), EGFR (such as human EGFR, SEQ ID NO:240), L1CAM (such as human L1CAM, SEQ ID NO:241), LFA-1 (such as human LFA-1, SEQ ID NO:242), LGALS3BP (such as human LGALS3BP, SEQ ID NO:243), MFGE8 (such as human MFGE8, SEQ ID NO:244), SLIT2 (such as human SLIT2, SEQ ID NO:245), STX3 (such as human STX3, SEQ ID NO:246);

i) surface markers of mesenchymal stem cells *e.g.*, CD44 (such as human CD44, SEQ ID NO:225), CD45 (such as human CD45, SEQ ID NO:226), CD71 (such as human CD71, SEQ ID NO:191), CD73 (such as human CD73, SEQ ID NO:247), CD90 (such as human CD90, SEQ ID NO:248), CD29 (such as human CD29, SEQ ID NO:249), CD105 (such as human CD105, SEQ ID NO:250), CD106 (such as human CD106, SEQ ID NO:251), CD146 (such as human CD146, SEQ ID NO:252), CD164 (such as human CD164, SEQ ID NO:253), CD166 (such as human CD166, SEQ ID NO:254), STRO-1 (such as human STRO-1, SEQ ID NO:255);

j) glycoproteins *e.g.*, CD54 (such as human CD54, SEQ ID NO:256), CD235a (such as human CD235a, SEQ ID NO:257), CD106 (such as human CD106, SEQ ID NO:251);

k) channelling proteins, including Ca-channel proteins *e.g.*, GLUR2 (such as human GLUR2, SEQ ID NO:258), GLUR3 (such as human GLUR3, SEQ ID NO:259), HLA-DM (such as human HLA-DM, SEQ ID NO:260); or

l) miscellaneous exosomal (vesicular) surface proteins *e.g.*, FLOT2 (such as human FLOT2, SEQ ID NO:261).

Further miscellaneous EV surface proteins are selected from TCRA, TCRB, TCRD, TCRG, and T-cell receptors (T cell receptor loci), which have variable amino acid sequences. The skilled person may easily identify the suitable EV surface proteins based on the information provided herein, or from respective databases *e.g.*, databases

providing genomic loci and amino acid sequences of the human EV surface proteins (including fragments comprising at least one ED or TM, or isoforms of such EV surface proteins), or homologues or analogs from non-human animals.

Specifically, the EV surface protein comprises a tertiary structure comprising regions with loop, helical, and/or linear (peptide) structures, in particular a tetraspan-like tertiary structure such as described for CD81, which includes at least one large loop, and one or more helical regions. Such tertiary structure can be suitably engineered to incorporate an artificial binding site comprising contact points in distant regions of the tertiary structure.

In certain cases, the surface protein is anchored to the lipid bilayer membrane of the extracellular vesicle via a linker. Such linker may for example be an amino acid linker, linker based on hydrophilic and non-charged polymers, such as polyethylene glycol (PEG) and polysaccharide, or composed of zwitterionic polymers, containing both cationic and anionic groups.

Specific surface proteins are of mammalian origin, particularly those naturally occurring in species including human beings or non-human mammalian animals, such as warm-blooded animals, particularly dogs, cats, rabbits, horses, cattle, pigs and poultry.

The term “transmembrane domain” as used herein refers to a lipid membrane-spanning protein domain, which is typically hydrophobic. Specifically, the surface protein comprises at least two transmembrane domains anchoring an extravesicular loop to the EV. Tetraspanin proteins typically comprise four membrane-spanning domains. A transmembrane domain is typically positioned within the membrane of an EV when attached or bound to the EV. Yet, a protein domain is also called transmembrane domain while being isolated from an EV or isolated from an EV surface protein.

“Tetraspanins” also referred to as “tetraspans” or “tetraspan proteins” are herein understood as proteins of the transmembrane 4 superfamily, a protein superfamily that organize membrane microdomains termed tetraspanin-enriched microdomains by forming clusters and interacting with a large variety of transmembrane and cytosolic signaling proteins (also referred to as “tetraspanin superfamily”). Tetraspanins are typically cell-surface proteins that are characterized by the presence of four hydrophobic transmembrane domains. Naturally-occurring tetraspanin proteins mediate signal transduction events that play a role in the regulation of cell development, activation, growth and motility.

Tetraspanins as understood herein typically are comprised of extravascular domains (also referred to as extravesicular domains, ED), transmembrane domains and intravascular domains. For example, the N- and C-terminus of a tetraspanin is typically located within the EV, whereas the transmembrane domains are located within the lipid bilayer membrane, and the extravascular domains are placed on the outer surface of an EV. Specific examples of tetraspanins are glycosylated.

Extracellular (extravascular) domains, also referred to as extravesicular domains, ED, are the most variable regions in tetraspanins, which can be involved in binding a target. EC1 (first extracellular loop) is also referred to as small extracellular loop (SEL). EC2 ("large extracellular loop", LEL) of tetraspanins has been studied using the CD81LEL as a model protein for all tetraspanins. The LEL domain is divided into a constant region with conserved A, B, and E helices, suggested to mediate homodimerization through a hydrophobic surface, and a variable region with helices C and D flanking those sequences responsible for protein-protein interactions. Specifically, EC2 comprises at least two conserved cysteine residues forming disulfide bonds, for EC2 folding (the CCG motif), one cysteine residue proximal to transmembrane four present in all tetraspanins, and a Pro-Xaa-Xaa-Cys (PXXC, SEQ ID NO:183, wherein X can be any amino acid) motif in the majority of tetraspanins.

Among tetraspanins, CD9, CD63, CD81, CD82, and CD151 have a broad tissue distribution, while others are restricted to particular tissues, such as Tssc6, CD37, and CD53 in hematopoietic cells. Immunoelectron microscopy studies have shown that tetraspanins are abundant on various types of endocytic membranes and have been widely used as exosomal markers.

The tetraspan protein CD81 is the major protein enriched in the exosomal fraction of multivesicular bodies. Human CD81 comprises or consists of the amino acid sequence identified as SEQ ID NO:87 (coding sequence identified as SEQ ID NO:88).

The large extracellular loop of CD81, topologically located between transmembrane domains 3 and 4, is characterized by five helical elements forming a mushroom-like structure, stabilized by two pairs of cysteines. This motif is conserved among the protein members of tetraspanin family, and the oxidation of cysteine bonds is *e.g.*, involved in high-affinity binding of the E2 envelope protein of hepatitis C virus (HCV), a natural ligand of CD81.

Tetraspan proteins may be expressed as soluble proteins or be displayed on the surface of tetraspanin expressing cells or the respective EVs.

The crystal structure of hCD81 LEL, solved at 1.6 Å, revealed a new type of protein fold, and a subsequent sequence analysis of 160 tetraspanin family members indicated that their fold and key structural features are conserved. Apart from cysteine bridges, the hCD81 LEL can be stabilized by the invariant residues Gly157 and Pro176, which are located to accommodate cysteine connections, as well as Tyr127, which is fully buried and contributes to the hydrogen bonding network together with His151 and Cys190. Soluble hCD81 LEL can assemble into dimers around a 2-fold axis, and the contact between the protomers is a low-polarity region between helices of each interacting partner and between helix B and C-terminal residues of the opposite protomer. The N- and C-termini of the protomers fall in the central region on opposite faces of the assembled dimer, similar to the dimeric assembly at the cell surface, where transmembrane segments are also present. A second low-polarity region comprises the solvent-exposed surface of helices C and D. According to solution studies, helix D is fairly unstructured and attains helical conformation only upon binding with certain antigens. The sequence alignments of the tetraspanin family members indeed show an increased variability in this region, including insertions and deletions. It has been suggested that this surface area might be involved in a species- or tetraspanin-specific recognition process, which could hint to the possibility of heterodimeric tetraspanin species assembly. In particular, segment D of CD81 may guide specific homomeric clustering.

According to a specific embodiment, biophysical properties of tetraspanins like CD81 are improved by introduction of *de novo* pairs of cysteine residues to form novel disulfide bridges stabilizing the protein. Specifically, formation of novel disulfide bridges increases the stability of tetraspanins like CD81, which allows production of mutants with higher stability *e.g.*, mutants that incorporate one or more modified regions or a novel target-specific binding site comprising binding residues within said one or more modified regions. Specifically, the amino acid sequence can be modified by targeted or random mutagenesis to incorporate (in particular by substitution of) binding residues at predetermined positions or region(s) which make up a binding site, or to generate a library comprising a diversity of binding sites comprising binding residues.

Alike CD81, the tetraspanin CD9 is a cell-surface protein containing four hydrophobic transmembrane domains and two extracellular domains, EDs (such as comprising or consisting of EC1 and EC2).

Naturally-occurring CD9 consists of 228 amino acids and weighs 24-27 kDa. The four small and highly conserved hydrophobic transmembrane domains each comprise 24-27 amino acids. It has a small N-terminal (11 amino acids) and a C-terminal cytoplasmic (7 amino acids) tail as well as a very small intracellular domain (4 amino acids). The remaining part of the protein is composed of two extracellular domains (a small one, EC1, of 20 amino acids and a large one, EC2, of 83 amino acids). Two disulfide bonds, generated by four well-conserved cysteine residues (C), stabilize the large extracellular domain (EC2). CD9 also contains a tetraspanin signature (amino acids 65-89) and a CCG motif (amino acids 152 to 154). CD9 is one of the most ubiquitously expressed proteins on the surface of exosomes and therefore considered an exosome marker. Although there are variations in the amino acid sequence in the extracellular loops, the CD9 protein sequence is very well conserved between species (90% between human, mice and rat). CD9 share also some homologies with other tetraspanins, particularly in the transmembrane domains.

Wild-type CD9 can interact or form complexes with many other proteins, including other tetraspanins, integrins, EWI molecules, TGF- α , diphtheria toxin receptor, or tyrosine kinase, pregnancy specific glycoproteins, and proteins of the immune system such as MHC class II molecules and members of the Ig superfamily. Moreover, CD9 is involved in platelet activation and aggregation, as well as in cell adhesion, spreading, cell motility and tumor metastasis. CD9 also regulates paranodal junction formation, and is required for gamete fusion. Furthermore, CD9 promotes muscle cell fusion and supports myotube maintenance.

As described herein, tetraspanin CD63 is a highly glycosylated protein cell-surface protein containing four transmembrane domains and three putative N-glycosylation sites.

Wild-type CD63 typically resides in late endosomes, lysosomes, secretory vesicles, and at the plasma membrane, and it moves among these compartments. CD63 is extensively and variably glycosylated and its EC2 region contains three potential N-linked glycosylation sites (N130, N150, and N172). It is often used as a marker for multivesicular bodies, which are enriched with CD63. It has numerous natural interaction partners, including other tetraspanins, such as CD82; the MHC class II molecules HLA-DR, HLA-DM, and HLA-DO; several integrins; and phosphatidylinositol 4-kinase. CD63 also contains a tyrosine-based motif on its extreme C terminus. Tyrosine-based motifs in the cytoplasmic domains of membrane proteins are recognized by clathrin adaptor

complexes and play important roles in endocytosis, lysosomal targeting, and basolateral targeting. The tyrosine-based motif in CD63 mediates its interaction with the μ -subunits of adaptor protein complexes 2 and 3 (AP-2 and AP-3).

As described herein, tetraspanin CD151 has the characteristic structure of a tetraspanin. It is a 253-amino acid protein with a single N-glycosylation site in its LEL and is palmitoylated on several cysteine residues. Immunoblotting reveals a doublet of bands of apparent kDa 28 and 32, representing unglycosylated and glycosylated forms of CD151. Human CD151 comprises or consists of the amino acid sequence identified as SEQ ID NO:94.

Wild-type CD151 has a wide cell and tissue distribution, including epithelium, endothelium, muscle, renal glomeruli and proximal and distal tubules, Schwann cells, and dendritic cells, with a single RNA species observed in most human adult tissues. CD151 is expressed at high levels on platelets and megakaryocytes. As with other tetraspanins, CD151 is associated in cell membranes with several integrins.

While all immune cells express tetraspanins, most of these are present in a variety of other cell types. There are a few, such as CD37, CD53, TSPAN32 (Tssc6) and TSPAN33, which are found in hematopoietic lineages.

As described herein, tetraspanin CD37 is a cell surface glycoprotein with the typical structure of a tetraspanin that is known to complex with integrins and other transmembrane 4 superfamily proteins. Alternate splicing results in multiple transcript variants encoding different isoforms. Human CD37 comprises or consists of the amino acid sequence identified as SEQ ID NO:95.

CD37 is known to be expressed by cells of the immune system, with highest abundance on mature B cells, and lower expression is found on T cells and myeloid cells. Wild-type CD37 controls both humoral and cellular immune responses. CD37-deficiency in mice leads to spontaneous development on B cell lymphoma, and patients with CD37-negative lymphomas have a worse clinical outcome.

As described herein, tetraspanin CD53 is a pan-leukocyte surface glycoprotein which spans the plasma membrane four times and is a member of the transmembrane 4 superfamily. The protein sequence and gene structure of mouse CD53 (Cd53) were determined by isolation of both genomic and cDNA clones. CD53 is highly conserved in evolution, as mouse Cd53 was 91% identical to rat CD53 and 82% identical to human CD53. The tetraspanin CD53 has four transmembrane domains and it is glycosylated twice in its second extracellular loop. It has a length of 219 amino acids and is located

in the cell membrane, endosomes and in the lipid bilayer membrane of exosomes. Human CD53 comprises or consists of the amino acid sequence identified as SEQ ID NO:90.

Wild-type CD53 is expressed by virtually all immune cells, a subset of hematopoietic stem cells, and in a variety of haematological malignancies.

There are several tetraspanins present in platelets including CD9, CD151, Tssc6, and CD63. Recent studies in knockout mouse models have revealed that CD151 and Tssc6 are physically and functionally involved in regulation of the 'outside-in' signalling properties of the major platelet integrin, integrin $\alpha(\text{IIb})\beta(3)$ and thrombus stability in vivo.

As used herein, tetraspanin Tssc6, also called TSPAN32, is a member of the tetraspanin superfamily. The protein has a size of 320 amino acids and is expressed ubiquitously at low levels. High levels of TSPAN32 expression are typically confined to hematopoietic tissues including peripheral blood leukocytes, thymus and spleen.

Human TSPAN32 comprises or consists of the amino acid sequence identified as SEQ ID NO:91.

As used herein, lysosomal associated membrane protein 2 (LAMP2), is a lysosome-associated membrane glycoprotein. LAMP2 is an integral membrane protein with two conserved luminal domains (constituting 90% of the entire protein), a single transmembrane (TM) domain (about 20 amino acids), and a short (10–12-amino acid) C-terminal cytosolic tail. Glycosylation is found in its luminal domains. Human LAMP2 comprises or consists of the amino acid sequence identified as SEQ ID NO:96. LAMP2, as used herein, preferably comprises modification in the extravesicular loop regions to incorporate the artificial binding site.

Wild-type LAMP2 plays an important role in chaperone-mediated autophagy, a process that mediates lysosomal degradation of proteins in response to various stresses and as part of the normal turnover of proteins with a long biological half-life.

The term “tetraspan-like proteins” (sometimes called tetraspanin-like) as used herein shall refer to EV surface proteins which comprise at least two transmembrane domains and at least one ED positioned between said at least two transmembrane domains, preferably wherein the region between said at least two transmembrane domains comprises or consists of one ED.

Specific examples of tetraspan-like proteins are naturally-occurring or modified tetraspanin proteins and others, such as lysosome-associated membrane glycoproteins

(LAMPs), or recombinant or synthetic proteins e.g. chimeric proteins comprising one or more transmembrane domains of one protein and one or more EDs of another protein, thereby obtaining a recombinant tetraspan-like protein.

“Sequence identity” or “percent (%) amino acid sequence identity” with respect to protein sequences and mutants thereof is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in the specific polypeptide sequence to be compared (the “parent sequence”), after aligning the sequence and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared.

As used herein, the term “specificity”, “target-specific” or “specific binding” refers to a binding reaction which is determinative of the cognate ligand of interest in a heterogeneous population of molecules. Thus, under designated conditions (e.g., immunoassay conditions), the modified surface protein binds to its particular target and does not bind in a significant amount to other molecules present in a sample. The specific binding means that binding is selective in terms of target identity, high, medium or low binding affinity or avidity, as selected. Selective binding is usually achieved if the binding constant or binding dynamics is at least 10 fold different, preferably the difference is at least 100 fold, and more preferred at least 1000 fold.

A specific binding does not exclude certain cross-reactivity with similar antigens, or the same antigens of a different species (analogues). For example, a binding entity may also preferably cross-react with rodent targets analogous to human targets, to facilitate preclinical animal studies.

The term “target” as used herein shall in particular include all antigens and target molecules capable of being recognised by a binding site of an antibody. Surface proteins described herein are engineered to comprise an artificial target binding site, which is specifically recognizing antigenic structures or epitopes alike an antibody.

Specific targets are cellular targets or soluble targets. In many cases, targets are self-antigens such as receptors located on the surface of tumor cells or cytokines or growth factors that can be present in the circulation of subjects or patients. Further targets may be of pathogen origin, e.g., microbial or viral pathogens.

The target antigen is either recognized as a whole target molecule or as a fragment of such molecule, especially substructures, *e.g.*, a polypeptide or carbohydrate structure of targets, generally referred to as “epitopes”, *e.g.*, B-cell epitopes, T-cell epitope), which are immunologically relevant, *i.e.* are also recognisable by natural or monoclonal antibodies.

Specifically, a target antigen is selected from cell surface antigens, including receptors, in particular from the group consisting of erbB receptor tyrosine kinases (such as EGFR, HER2 including Her2neu, HER3 and HER4, in particular those epitopes of the extracellular domains of such receptors, *e.g.*, the 4D5 epitope). In addition further antigens may be targeted, *e.g.*, molecules of the TNF-receptor superfamily, such as Apo-1 receptor, TNFR1, TNFR2, nerve growth factor receptor NGFR, CD40, CD40-Ligand, OX40, TACI, BCMA, BAFF-receptor, T-cell surface molecules, T-cell receptors, T-cell antigen, Apo-3, DR4, DR5, DR6, decoy receptors, such as DcR1, DcR2, CAR1, HVEM, GITR, ZTNFR-5, NTR-1, TNFL1, IGFR-1, c-Met, but not limited to these molecules, B-cell surface antigens, such as CD10, CD19, CD20, CD21, CD22, , DC-SIGN, antigens or markers of solid tumors or hematologic cancer cells, cells of lymphoma or leukaemia, other blood cells including blood platelets, but not limited to these molecules.

The term “therapeutically effective amount”, used herein interchangeably with any of the terms “effective amount” or “sufficient amount” of a compound, *e.g.*, a binder as described herein, particularly a TEV as described herein, is a quantity or activity sufficient to, when administered to the subject effect beneficial or desired results, including clinical results, and, as such, an effective amount or synonym thereof depends upon the context in which it is being applied.

An effective amount is intended to mean that amount of a compound that is sufficient to treat, prevent or inhibit a disease or disorder. In the context of disease, therapeutically effective amounts of a binder or TEV as described herein are specifically used to treat, modulate, attenuate, reverse, or affect a disease or condition that benefits from the interaction of the binder or TEV with its target antigen, *e.g.* a tumor cell.

The amount of the compound that will correspond to such an effective amount will vary depending on various factors, such as the given drug or compound, the pharmaceutical formulation, the route of administration, the type of disease or disorder, the identity of the subject or host being treated, and the like, but can nevertheless be routinely determined by one skilled in the art.

A binder as described herein may specifically be used in a pharmaceutical composition. Therefore, a pharmaceutical composition is provided which comprise a binder as described herein and a pharmaceutically acceptable carrier or excipient. These pharmaceutical compositions can suitably be administered as a bolus injection or infusion or by continuous infusion. Besides parenteral administration, topical or oral administration may be preferred. Pharmaceutical carriers suitable for facilitating such means of administration are well-known in the art.

Pharmaceutically acceptable carriers generally include any and all suitable solvents, adjuvants, dispersion media, coatings, isotonic and absorption delaying agents, and the like that are physiologically compatible with a binder provided herein. Further examples of pharmaceutically acceptable carriers include sterile water, saline, phosphate buffered saline, dextrose, glycerol, ethanol, and the like, as well as combinations of any thereof.

Suitable pharmaceutically acceptable carriers or excipients specifically include one or more of any and all conventional solvents, dispersion media, fillers, solid carriers, aqueous solutions, coatings, vehicles suitable for topical administration, other antimicrobial agents, isotonic and absorption enhancing or delaying agents, or activity enhancing or delaying agents for pharmaceutically active substances. Common pharmaceutically acceptable additives are disclosed, by way of example, in Remington: the Science & Practice of Pharmacy by Alfonso Gennaro, 20th ed., Lippencott Williams & Wilkins, (2000).

In one embodiment, suitable pharmaceutically acceptable carriers include, but are not limited to, inert solid fillers or diluents and sterile aqueous or organic solutions (e.g., polyethylene glycol, propylene glycol, polyvinyl pyrrolidone, ethanol, benzyl alcohol, etc.). In certain such embodiments, suitable pharmaceutically acceptable excipients include, but are not limited to, water, salt solutions, alcohol, polyethylene glycols, gelatin, lactose, amylase, magnesium stearate, talc, silicic acid, viscous paraffin, hydroxymethylcellulose, polyvinylpyrrolidone, fillers, such as sugars (e.g., lactose, sucrose, mannitol, or sorbitol), and cellulose preparations (e.g., maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethylcellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone, PVP).

In one such aspect, a binder can be combined with one or more carriers appropriate for a desired route of administration, e.g., admixed with any of lactose,

sucrose, starch, cellulose esters of alkanolic acids, stearic acid, talc, magnesium stearate, magnesium oxide, sodium and calcium salts of phosphoric and sulphuric acids, acacia, gelatin, sodium alginate, polyvinylpyrrolidone, polyvinyl alcohol, and optionally further tableted or encapsulated for conventional administration. Other carriers, adjuvants, and modes of administration are well known in the pharmaceutical arts. A carrier may include a controlled release material or time delay material, such as glyceryl monostearate or glyceryl distearate alone or with a wax, or other materials well known in the art.

Additional pharmaceutically acceptable carriers are known in the art and described in, *e.g.*, REMINGTON'S PHARMACEUTICAL SCIENCES. Liquid formulations can be solutions, emulsions or suspensions and can include excipients such as suspending agents, solubilizers, surfactants, preservatives, and chelating agents.

Pharmaceutical compositions are contemplated wherein a binder as described herein and one or more therapeutically active agents are formulated. Stable formulations of the binder described herein are prepared for storage by mixing said construct having the desired degree of purity with optional pharmaceutically acceptable carriers, excipients or stabilizers, in the form of lyophilized formulations or aqueous solutions. Formulations for *in vivo* administration are preferably sterile, *e.g.*, in the form of a sterile aqueous solution. This is readily accomplished by filtration through sterile filtration membranes or other suitable sterilization methods.

Administration of the pharmaceutical composition comprising a binder described herein, may be done in a variety of ways, including orally, subcutaneously, intravenously, intranasally, intraotically, transdermally, mucosal, topically, *e.g.*, tablets, gels, ointments, salves, suppositories, patches, lotions, creams, etc., intraperitoneally, intramuscularly, intrapulmonary, vaginally, parenterally, rectally, or intraocularly.

In one embodiment, the pharmaceutical composition is administered orally, intravenously, or inhalationally. In a specific embodiment, the binder is administered in a dosage form selected from the group consisting of solid dosage form, a cream, an aqueous mixture, a lyophilized aqueous mixture and an aerosol.

Exemplary formulations as used for parenteral administration include those suitable for subcutaneous, intramuscular or intravenous injection as, for example, a sterile solution, emulsion or suspension.

The binder described herein may specifically be used in a diagnostic composition *e.g.*, for *in vitro* or *in vivo* use. Therefore, a diagnostic composition is provided which

comprises a binder as described herein, and optionally a diagnostic reagent in a composition or a kit of parts.

The diagnostic kit preferably comprises all essential components to qualitatively or quantitatively determine the target in a biological sample, optionally without common or unspecific substances or components, such as water, buffer or excipients. A storage stable kit can be provided with a shelf-life of preferably at least 6 months, more preferably at least 1 or 2 years. It may be composed of dry (e.g., lyophilized) components, and/or include preservatives.

The preferred diagnostic kit is provided as a packaged or prepackaged unit e.g., wherein the components are contained in only one package, which facilitates routine experiments. Such package may include the reagents necessary for one or more tests e.g., suitable to perform the tests of a series of biological samples. The kit may further suitably contain a standard or reference control.

The diagnostic composition may be a reagent ready-to-use in a reaction mixture with the biological sample, or a conserved form of such reagent e.g., a storage-stable form such as lyophilized; snap-frozen (e.g., in liquid nitrogen), ultra-low-temperature storage (-70°C and -80°C), cold-storage (-20°C and 5°C) and controlled room temperature (15°C-27°C); standard sample storage as e.g., glycerol-stocks, tissue paraffin-blocks, (buccal) swabs and other standard biological sample storage methods, which conserved form of a reagent can be reconstituted or prepared to obtain a ready-to-use reagent. Such ready-to-use reagent is typically in the form of an aqueous solution, specifically (physiological) buffer conditions (e.g., EDTA buffered, phosphate buffer, HBSS, citrate buffer etc.).

Specifically, the further diagnostic reagent is a reagent specifically reacting with the binder and/or a reaction product of the binder binding to its target. The appropriate diagnostic reagent can be a solvent, a buffer, a dye, an anticoagulant, a ligand that specifically binds to the binder described herein and/or the binder-target complex.

Specifically, the invention provides for a diagnostic preparation of a binder described herein, optionally containing the binder with a label and/or a further diagnostic reagent with a label, such as a reagent specifically recognizing the binder or a complex of the binder with the respective target, and/or a solid phase to immobilize at least one of the binder and the diagnostic reagent.

Specifically, the further diagnostic reagent is a diagnostic label or a reagent specifically reacting with the binder and/or the reaction product of the binder binding to its target.

The EV or the diagnostic reagent can be directly labeled or indirectly labeled. The indirect label may comprise a labeled binding agent that forms a complex with the binder or diagnostic reagent to the target.

The label is typically a molecule or part of a molecule that can be detected in an assay. Exemplary labels are chromophores, fluorochromes, or radioactive molecules. In some embodiments the EV or diagnostic reagent is conjugated to a detectable label which may include molecules that are themselves detectable (e.g., fluorescent moieties, electrochemical labels, metal chelates, etc.) as well as molecules that may be indirectly detected by production of a detectable reaction product (e.g., enzymes such as horseradish peroxidase, alkaline phosphatase, etc.) or by a specific binding molecule which itself may be detectable (e.g., biotin, digoxigenin, maltose, oligohistidine, 2,4-dinitrobenzene, phenylarsenate, ssDNA, dsDNA, etc.).

Preferred diagnostic preparations or assays comprise the EV described herein immobilized on a solid phase e.g., latex beads, gold particles, etc.

The following items are considered specific embodiments of the invention:

1. A target-specific extracellular vesicle (EV) comprising a lipid bilayer membrane anchoring a surface protein which comprises an artificial binding site specifically recognizing a target, wherein said binding site comprises binding residues within at least two distant regions of said surface protein.
2. The EV of item 1, wherein said surface protein comprises a transmembrane domain of a vesicular membrane protein, which transmembrane domain is anchoring the surface protein to the EV.
3. The EV of item 2, wherein the membrane protein is a tetraspan protein, preferably selected from the group consisting of CD81, CD9, CD37, CD53, CD63, and CD82, or LAMP2.
4. The EV of any of items 1 to 3, wherein the surface protein is CD81 and the binding residues are located within a first region between positions 160 and 172, and within at least one further region located between positions 132 and 139, or between positions 180 and 189, wherein numbering is of CD81 identified as SEQ ID NO:87.

5. The EV of item 4, wherein the surface protein is CD81 which is stabilized by introducing one or more cysteine(s) at position(s) to allow the formation of one or more additional disulfide bonds stabilizing the extravesicular loop structure of the protein.

6. The EV of any one items 1 to 5, which is originating from eukaryotic or prokaryotic source cells of body tissue, body fluid, or a cell culture.

7. The EV of any one of items 1 to 6, which is carrying an intravesicular load, wherein the load comprises any one or more of peptides, polypeptides, protein domains, proteins, lipids, genes, nucleic acids such as mRNAs, miRNAs, RNAi mediating molecules in particular locked nucleic acids or phosphorothioates, DNA, DNA fragments, plasmids such as minicircle DNA, drugs such as small molecules, in particular chemotherapeutics or senolytics.

8. The EV of any one of items 1 to 7, wherein the target is selected from the group consisting of cellular targets such as mitogenic receptors, cytokine receptors, asialoglycoprotein receptors, membrane transporters, lipoproteins, liposaccharides, glycoproteins, proteoglycans, or acellular targets such as cytokines, artificial proteins or artificial surface structures.

9. A pharmaceutical preparation comprising the EV of any one items 1 to 8, and a pharmaceutically acceptable carrier, preferably in a formulation for intradermal, subcutaneous, intravenous, topical, or oral use.

10. A method of producing a preparation of EV of any one of items 1 to 8, comprising:

- a) introducing a gene encoding the surface protein into a source cell or source cell mixture;
- b) culturing said cell(s) under conditions producing extracellular vesicles;
- c) isolating a fraction comprising the EV with the target binding specificity; and
- d) producing an EV preparation.

11. The method of item 10, wherein the gene is modified by a mutagenesis method to incorporate an artificial binding site specifically recognizing a target, wherein said mutagenesis method employs mutagenesis of at least two distant regions of said surface protein, thereby producing a repertoire of polynucleotides encoding a variety of surface proteins, each with a different binding specificity, and selecting a polynucleotide encoding the surface protein which specifically recognizes the target.

12. The method of item 11, wherein the repertoire of polynucleotides is comprised in genetic packages displaying the variety of surface proteins on the outer surface,

preferably employing a display system selected from the group consisting of a phage, yeast, bacterium, ribosome, mRNA or mammalian cell display.

13. The EV preparation produced according to any one of items 10 to 12, for autologous use, wherein the source cell or source cell mixture is obtained from a subject and the EV preparation is administered to the same subject.

14. The method of any one of items 10 to 12, to produce a library of EVs, wherein

a) a repertoire of a gene encoding a variety of surface proteins is introduced into said source cell(s); and

b) a repertoire of EVs is isolated, each with a different target binding specificity, to produce a library of EVs comprising a repertoire of target binding EVs with a variety of target binding specificity,

wherein the repertoire is produced by mutagenizing said gene within at least two predetermined distant regions of said gene.

15. A library of EVs produced by a method of item 14, preferably wherein the repertoire comprises at least 10^2 EVs with different target specificity.

The examples described herein are illustrative of the present invention and are not intended to be limitations thereon. Different embodiments of the present invention have been described according to the present invention. Many modifications and variations may be made to the techniques described and illustrated herein without departing from the spirit and scope of the invention. Accordingly, it should be understood that the examples are illustrative only and are not limiting upon the scope of the invention.

EXAMPLES

Example 1: Yeast Display of CD81 LEL

Wild-type human CD81 LEL sequence was cloned as a C-terminal fusion protein with Aga2, preceded by an Xpress-tag and with a C-terminally appended his-tag and V5-tag.

The amino acid sequence of CD81 LEL was

(SEQ ID NO:7):

FVNKDQIAKDVKQFYDQALQQAVVDDDDANNAKAVVKTFFHETLDCCGSSTLTALTTSVLKNNLCP
SGSNIISNLFKEDCHQKIDDLFSGK.

The coding nucleotide sequence was

(SEQ ID NO:8):

```
TTTGTCAACAAGGACCAGATCGCCAAGGATGTGAAGCAGTTCTATGACCAGGCCCTACAGCAGG
CCGTGGTGGATGATGACGCCAACACGCCAAGGCTGTGGTGAAGACCTTCCACGAGACGCTTGA
CTGCTGTGGCTCCAGCACACTGACTGCTTTGACCACCTCAGTGCTCAAGAACAATTTGTGTCCC
TCGGGCAGCAACATCATCAGCAACCTCTTCAAGGAGGACTGCCACCAGAAGATCGATGACCTCT
TCTCCGGGAAG.
```

Primers for amplification of human CD81 LEL were

(SEQ ID NO:9): ACGTGGATCCTTTGTCAACAAGGACCAGATC

(SEQ ID NO:10): ACGTGCGGCCGCGCCTTCCCGGAGAAGAGGTC.

PCR product was cut with *Bam*HI and *Nof*I and ligated with correspondingly cut vector pYD1 (Thermo Fisher Scientific). Ligation mixture was transformed into electrocompetent *E. coli* TOP10 and transformants were selected on ampicilline plates. Plasmid was isolated with minipreparation and transformed into *S. cerevisiae* EBY100 using chemical transformation. A starter culture of EBY100 (Thermo Fisher Scientific) in 20 ml YPD medium (2% peptone, 1% yeast extract, 2% glucose) (Merck) was incubated overnight at 30°C and 180 rpm. The culture was then diluted to an OD₆₀₀ of 0.4 and incubated for about 5 hours at 30°C and 180 rpm. Aliquots of 50 ml cell culture were then pelleted at 1000 g for 5 min at room temperature, then washed with 25 ml AD and pelleted again. The cells were resuspended in 3 ml 100 mM Li-acetate and incubated for 15 min at 30°C in a shaking incubator. 0.3 ml of cell suspension was then pelleted and supernatant removed. Components of the transformation mix were added as follows: 240 µl 50% PEG 3350, 36 µl 1.0 M Li-acetate, 50 µl 2 mg/ml ssDNA (salmon sperm carrier DNA, previously heated up to 95°C for 5 minutes and then placed on ice) (Sigma Aldrich) and 1 µg of pYD1-CD81 LEL plasmid. The cell pellets were resuspended in the transformation mix, incubated for 30 min at 30°C in a shaking incubator and then heat-shocked at 42°C for 45 min. Yeast cells were pelleted at 1000 g, 5 min at RT, resuspended in AD and transformants were selected on MDL medium at 30°C for three days.

Transformants were inoculated into SD-CAA (1% Casamino acids (Becton Dickinson), 100 mM K-phosphate buffer, pH 6.0 (Merck), 1xYNB (Becton Dickinson), 2% glucose (Merck)) and cultured overnight at 30°C. Induction was with SG/R-CAA medium (1% Casamino acids (Becton Dickinson), 100 mM K-phosphate buffer, pH 6.0 (Merck), 1xYNB (Becton Dickinson), 2% galactose (Merck), 1% raffinose (Merck))

overnight at 37°C or for 2 days at 20°C. Yeast cultures were then examined for the expression of the recombinant proteins. FACS analysis revealed display levels similar to yeast transformed with the unmodified expression vector encoding only the reporter tags. Further, the wild-type CD81 LEL displaying yeast culture was stained with structure-reporting antibodies M38 and 1.3.3.22 (Thermo Fisher Scientific), and the results have shown the same expression level as for the tags examined. FACS analysis revealed similar display levels for cultures induced at 20°C or under stress conditions of 37°C. The expression of tags was at a similar level as found for the yeast transformed with the vector encoding the tags only. Further, the yeast cells expressing wild-type CD81 were stained with the anti-CD81 antibodies similarly to the cells induced at 20°C.

Example 2: Phage Display of CD81 LEL

The sequence encoding wild-type CD81 LEL was cloned into the multiple cloning site of the expression vector fdmyc that allows the expression of phage particles with the recombinant protein N-terminally positioned from the c-myc tag and the g3p protein. The primers for amplifications were

(SEQ ID NO:11): ACGAGTGCACAGTTTGTCAACAAGGACCAGATC

(SEQ ID NO:12): ACGTGCGGCCCGCCTTCCCGGAGAAGAGGTC.

PCR product was cut with restriction enzymes ApaI and NotI and ligated into the correspondingly cut vector fdmyc. Ligation mixture was transformed into E. coli TG1 cells (Thermo Fisher Scientific) and selected on tetracycline-containing TYE-plates (1.5% agar, 1.6% peptone, 1% yeast extract (Merck)). After overnight cultivation in tetracycline-containing medium at 30°C, high titers of 10^{11} - 10^{12} phage particles /L culture could be obtained, indicating that the expressed proteins are not detrimental to phage multiplication. The expression level of the fusion protein was tested using analysis of phage particles with SDS-PAGE and Western blotting. Detection of displayed protein was performed with an anti-g3p antibody (New England Biolabs) and found to be 50% fused with wild-type CD81 LEL. Phage-borne CD81 LEL could be detected with both CD81-specific antibodies, M38 and 1.3.3.22, indicating the correct folding of the phage-displayed molecule.

Example 3: Expression of soluble CD81 LEL

CD81 LEL sequence was amplified with primers

(SEQ ID NO:13): ACGTGCTAGCTTTGTCAACAAGGACCAGATC

(SEQ ID NO:14): ACGTGGATCCTCATCACTTCCCGGAGAAGAGGTC.

PCR fragment was cut with restriction enzymes *NheI* and *BamHI* and ligated to vector pTT22SSP4 (CNRC) that was cut with the same enzymes. Ligation mixture was transformed into electrocompetent *E. coli* TOP10 (Thermo Fisher Scientific) and transformants were selected on ampicilline plates. Plasmid was isolated with miniprep and transfected into HEK-293-6E cells (CNRC), exactly according to manufacturer's instructions. Protein expression proceeded for 5 days with addition of TN-20 to an end concentration of 0.5%. hCD81 LEL was then purified via Ni-NTA chromatography using standard protocols. Supernatant was buffered with PBS and 20 mM imidazole and pH adjusted to 7.5. Excel Ni-NTA column (GE Healthcare) was equilibrated with PBS and 20 mM imidazole, pH 7.5, and loaded with the buffered supernatant. Elution was in 5 column volumes with a linear gradient from 20 to 500 mM imidazole in PBS, pH 7.5. Protein-containing fractions were pooled and dialysed against 100-times volume of PBS at 4°C overnight.

SEC analysis in native conditions revealed a monodisperse elution profile corresponding to a dimeric form of the protein, analogously to the soluble wild-type CD81 LEL.

Alternatively, protein was expressed in ExpiCHO expression system (Thermo Fisher Scientific) following MaxTiter protocol exactly according to manufacturer's instructions. hCD81 LEL was then purified via Ni-NTA chromatography using standard protocols. Supernatant was diluted with an equal volume of AD and buffered with PBS and 20 mM imidazole and pH adjusted to 7.5. Excel Ni-NTA column (GE Healthcare) was equilibrated with PBS and 20 mM imidazole, pH 7.5, and loaded with the buffered supernatant. Elution was in 5 column volumes with a linear gradient from 20 to 500 mM imidazole in PBS, pH 7.5. Protein-containing fractions were pooled and dialysed against 100-times volume of PBS at 4°C overnight.

Example 4: Construction of CD81 Library 1

A yeast-display library of CD81 LEL mutants randomized in the total of 12 amino acid residues in C- and D-segment of the CD81 LEL was constructed. Randomized were the amino acid residues: 160-162 and 181-189 (numbering as in 1G8Q). Yeast display library was prepared at a size of 7×10^7 independent members. PCR fragment for recombination was prepared using oligonucleotides

(SEQ ID NO:15):

CTTCCACGAGACGCTTGACTGCTGTGGATCCNNKNNKNNKACTGCTTTGACCACCTC,
wherein N is any one of A, C, G, or T

(SEQ ID NO:16):

CCGCGCCTTCCCGGAGAAGAGGTCATCGATCTTCTGGTGGCAAKMMNNMNNMNNMNNMNNM
NNMNNGTGCTGCCCCGAGGGAC, wherein N is any one of A, C, G, or T;
and M is any one of A or C.

using Q5 HiFi polymerase (New England Biolabs) and purified after gel electrophoresis. Recipient vector was modified to facilitate the recombination of PCR fragment. All mutagenesis steps were performed using QuikChange Lightning Mutagenesis kit (Agilent), exactly according to manufacturer's instructions.

A *Bam*HI site was introduced into pYD1_CD81 LEL with oligonucleotides

(SEQ ID NO:17): GCTTGACTGCTGTGGATCCAGCACACTGACTG

(SEQ ID NO:18): CAGTCAGTGTGCTGGATCCACAGCAGTCAAGC,

and then the naturally occurring *Bam*HI site was removed using oligonucleotides

(SEQ ID NO:19): CGATAAGGTACCAGGTTCTTTGTCAACAAG

(SEQ ID NO:20): CTTGTTGACAAAGGAACCTGGTACCTTATCG.

Vector was linearized with *Bam*HI and *Cla*I and vector backbone was purified from the agarose gel. Chemical transformation was used to introduce PCR fragment and vector backbone into *S. cerevisiae* EBY100. A starter culture of EBY100 in 20 ml YPD medium was incubated overnight at 30°C and 180 rpm. The culture was then diluted to an OD₆₀₀ = 0.4 and incubated for about 5 hours at 30°C and 180 rpm. Aliquots of 50 ml

cell culture were then pelleted at 1000 g for 5 min at room temperature, then washed with 25 ml AD and pelleted again. The cells were resuspended in 3 ml 100 mM Li-acetate and incubated for 15 min at 30°C in a shaking incubator. Cells were pelleted and supernatant removed. Components of the transformation mix were added as follows: 2400 µl 50% PEG 3350, 360 µl 1.0 M Li-acetate, 500 µl 2 mg/ml ssDNA (salmon sperm carrier DNA, previously heated up to 95°C for 5 minutes and then placed on ice), 10 µg linearized recipient vector and 7 µg DNA fragment. The cell pellets were resuspended in the transformation mix and incubated for 30 min at 30°C in a shaking incubator and heat-shocked at 42°C for 45 min.

After collecting the cells by centrifugation at 1000 g for 5 min at room temperature and removing the supernatant, the pellets were resuspended in 10 ml SD-CAA medium (1% Casamino acids (Becton Dickinson), 100 mM K-phosphate buffer, pH 6.0 (Merck), 1xYNB (Becton Dickinson), 2% glucose (Merck). Aliquots were removed for dilution plating to determine the library size: 10 µl of cell suspension were diluted in 990 µl SD-CAA medium of which 100 µl were plated on an MDL plate (1.5% agar (Merck), 1xYNB (Becton Dickinson), 2% glucose (Merck) and 0.01% leucin (Sigma-Aldrich) and incubated at 30°C for 3 days. Yeast cells were diluted in 50 ml SD-CAA medium and incubated at 30°C at 180 rpm for 24 h, passaged into fresh SD-CAA medium at 1:20 dilution and cultured for another 24 h under the same conditions. Cells were harvested at 1000 g, 5 min, at 4°C and pellets were resuspended in an equal volume of 30% glycerol before freezing at -80°C.

Example 5: Selection of CD81 LEL Library 1 with mouse EGFR-Fc

Mouse EGFR-Fc was purchased from Sino Biological. For biotinylation, EZ-Link™ Sulfo-NHS-LC-LC-Biotin reagent was used in a 3:1 molar ratio. The antigen was reconstituted exactly according to manufacturer's instructions to a concentration of 0.25 µg/µl. Incubation with biotinylation reagent proceeded for 1h at room temperature with shaking. Unbound biotin was removed with dialysis against 100x volume of PBS, using Snakeskin dialysis tubing (Thermo Fisher Scientific) with 10.000 Da MWCO, at 4°C overnight with stirring.

For selection, library was inoculated into SD-CAA and cultured overnight at 30°C. Induction was with SG/R-CAA medium overnight at 37°C. The induced cell suspensions were diluted to 10⁸ cells per 1 ml 10% BSA-PBS and centrifuged at 1000 g for 5 min at 20°C. To block the cells, the pellets were resuspended in 1 ml 10% BSA-PBS and

incubated for 30 min at 20°C on a rotating platform. The cells were centrifuged at 3000 rpm for 5 min at 20°C and resuspended in 250 µl 10% BSA-PBS containing 1 µM biotinylated EGFR-Fc. After incubation for 30 min at room temperature on a rotating platform the cells were centrifuged at 3000 rpm for 5 min at 4°C. To wash the cells, the pellets were resuspended in 1 ml ice-cold PBS and centrifuged at 1000 g for 5 min at 4°C. Then, the cells were resuspended in 250 µl 10% BSA-PBS with streptavidin-Alexafluor 647 (1:800) and anti-V5-FITC antibody (1:100) (Thermo Fisher Scientific) and incubated for 30 min on ice. The cells were centrifuged at 3000 rpm for 5 min at 4°C before they were resuspended in 1 ml ice-cold PBS and centrifuged again. Lastly, the cells were resuspended in 250 µl ice-cold PBS and were kept on ice until sorting with FACS Aria™. 1st sort covered 2.5x library size and 1% false positive yeast cells were collected and in the 2nd sort 20x output of 1st sort was processed and 0.1% false positive yeast cells were propagated. In the following two sorting rounds at least 100x output of previous sort was processed, again 0.1% of yeast cells were collected and after the 4th sort plated out to characterize single yeast display clones. 23 selected clones were screened using staining with mouse EGFR-Fc and all except clones 9, 10 and 23 stained significantly with the antigen, but not with the secondary reagent streptavidin-Alexafluor 647.

Table 1 showing median fluorescence of yeast clones displaying selected CD81 LEL variants with antigen mouse EGFR-Fc and secondary reagent streptavidin-Alexafluor 647 or only with secondary reagent streptavidin-Alexafluor 647.

Yeast clone	2nd only	Antigen + 2nd
1	15.69	33.12
2	12.81	22.36
3	12.49	23.39
4	12.42	16.48
5	16.96	34.12
6	11.39	24.6
7	8.84	18.65
8	16.14	51.32
9	14.59	11.76
10	13.96	16.18
11	19.67	36.52
12	25.87	56.93
13	21.36	47.54
14	19.85	26.23
15	16.93	43.68
16	14.61	26.19

17	18.56	21.22
18	16.67	57.94
19	16.72	20.57
20	15.67	13.24
21	10.08	33.69
22	14.03	18.58
23	20.27	15.97
wild-type CD81 LEL	28.27	9.1

Upon sequencing, 6 different sequences were identified (Table 2).

Table 2 showing identified sequences.

clone	Helix C 160-162) (AA	Segment D (AA 181-189)
1	VRR	RRPRKRTRS (SEQ ID NO:1)
2	ERL	RRSRRRVHD (SEQ ID NO:2)
3	RAN	IWRARRRRS (SEQ ID NO:3)
4	ERL	RRSRRRVHD (SEQ ID NO:2)
5	SFS	RRFRKRPGD (SEQ ID NO:4)
6	SSS	SRRWRHRIA (SEQ ID NO:5)
7	ERL	RRSRRRVHD (SEQ ID NO:2)
8	VRR	RRPRKRTRS (SEQ ID NO:1)
9	MGN	WWGRRFRVS (SEQ ID NO:6)
10	ERL	RRSRRRVHD (SEQ ID NO:2)
11	ERL	RRSRRRVHD (SEQ ID NO:2)
12	ERL	RRSRRRVHD (SEQ ID NO:2)
13	RAN	IWRARRRRS (SEQ ID NO:3)
14	ERL	RRSRRRVHD (SEQ ID NO:2)
15	VRR	RRPRKRTRS (SEQ ID NO:1)
16	ERL	RRSRRRVHD (SEQ ID NO:2)
17	ERL	RRSRRRVHD (SEQ ID NO:2)
18	SSS	SRRWRHRIA (SEQ ID NO:5)
19	SFS	RRFRKRPGD (SEQ ID NO:4)
20	ERL	RRSRRRVHD (SEQ ID NO:2)
21	RAN	IWRARRRRS (SEQ ID NO:3)
22	ERL	RRSRRRVHD (SEQ ID NO:2)
23	ERL	RRSRRRVHD (SEQ ID NO:2)

Example 6: Selections of CD81 Library 1 with human EGFR-Fc

Human EGFR-Fc was purchased from Sino Biological. For biotinylation, EZ-Link™ Sulfo-NHS-LC-LC-Biotin reagent was used in a 3:1 molar ratio. The antigen was reconstituted exactly according to manufacturer's instructions to a concentration of 0.25 µg/µl. Incubation with biotinylation reagent proceeded for 1h at room temperature with shaking. Unbound biotin was removed with dialysis against 100x volume of PBS, using Snakeskin dialysis tubing (Thermo Fisher Scientific) with 10.000 Da MWCO, at 4°C overnight with stirring.

For sorting, libraries were first cultured in SD-CAA medium supplemented with penicillin-streptomycin at 30°C with shaking overnight and then the expression of recombinant protein was induced by resuspending the yeast cells in SG/R-CAA medium with penicillin-streptomycin and incubating them with shaking overnight at 37°C.

For MACS, 10⁹ induced yeast cells were pelleted for 5 min at 2500 g and washed by resuspending the pellet in 50 ml wash buffer (PBS, pH 7.2, 0.25% BSA, 2mM EDTA) and pelleted again for 5 min at 2500 g. Washed cells were resuspended in 25 ml wash buffer containing 0.25 µM biotinylated antigen and incubated for 30 min at room temperature with gentle agitation and 10 min in an ice bath. Then, the cells and antigen solution was pelleted for 5 min at 2500 g and 4°C, washed twice with 50 ml wash buffer and pelleted after each washing step. The cells were resuspended in 5 ml wash buffer plus 25 µl streptavidin microbeads (µMACS streptavidin kit, MACS Miltenyi) and incubated on ice for 10 min. Finally, 15 ml wash buffer were added.

The LS column was placed in the separator and 3 ml of wash buffer were applied to precondition the column. Then, the cell solution was loaded in 7 ml batches. Once the column stopped dripping it was briefly removed from the magnet and placed back in the magnet to reorient the beads in the column to allow trapped cells to flow through. Before loading the next 7 ml cell suspension 1 ml of wash buffer was applied to the column. These steps were repeated until all cells were loaded. Then, the column was washed with 3 ml wash buffer, briefly removed from the magnet and washed again with 3 ml wash buffer. To elute the binding cells, the column was removed from the magnet and 5 ml wash buffer were added. Using the supplied plunger, the cells were pushed through the column into a new tube. The binding cell fraction was pelleted for 5 min at 2,500 g. The pellet was resuspended in 10 ml SD-CAA medium and incubated at 30°C and 180rpm overnight. 10 µl of eluted cells were diluted in 990 µl SD-CAA and 100 µl were

plated onto an MDL plate and incubated at 30°C for three days to estimate the output of MACS procedure.

In the FACS selection rounds that followed, the induced cell suspensions were diluted to 10^8 cells per 1 ml 10% BSA-PBS and centrifuged at 3000 rpm for 5 min at 20°C. To block the cells, the pellets were resuspended in 1 ml 10% BSA-PBS and incubated for 30 min at 20°C on a rotating platform. The cells were centrifuged at 3000 rpm for 5 min at 20°C and resuspended in 250 μ l 10% BSA-PBS containing 1 μ M biotinylated EGFR-Fc. After incubation for 30 min at 20°C on a rotating platform the cells were centrifuged at 3000 rpm for 5 min at 4°C. To wash the cells, the pellets were resuspended in 1 ml ice-cold PBS and centrifuged at 3000 rpm for 5 min at 4°C. Then, the cells were resuspended in 250 μ l 10% BSA-PBS with streptavidin-Alexafluor 647 (1:800) and anti-V5-FITC antibody (1:100) (Thermo Fisher Scientific) and incubated for 30 min on ice. The cells were centrifuged at 3000 rpm for 5 min at 4°C before they were resuspended in 1 ml ice-cold PBS and centrifuged again. Finally, the cells were resuspended in 250 μ l ice-cold PBS and were kept on ice until sorting with FACS Aria™. At least 20x output of the previous sorting round was processed and 0.1% top anti-V5-antibody positive yeast cells were collected. The 5th sort was plated out to characterize single yeast display clones.

Following sequences of binders were identified (Table 3):

Table 3 depicts the identified sequences of binders

clone	Helix C (AA 160-162)	Segment D (AA 181-189)
1	ILG	PHHWRFPA (SEQ ID NO:181)

Example 7: Construction of CD81 LEL Library 2

To expand the possible repertoire of antigen-specific CD81 LEL binders it was aimed to design libraries where amino acid residues different those randomized in CD81 LEL library 1 could form a potential antigen recognition site. This design involves randomization of 11 amino acid residues: in helix A amino acid residues 132-133, in AB loop 136-139 and in helix C 162-165, 167 and 171-172. The scaffold here is a CD81 LEL mutant that harbors two novel disulfide bonds: one connects helices A and B (C4) and one connects helices A and C (C9). This CD81LEL mutant with a combination of novel disulfide bonds Ala134Cys/Lys144Cys and Val135Cys/Ser168Cys was then

produced in HEK293-6E cells (CNRC) and tested for its thermostability. Thermal unfolding was recorded up to 130 °C and proceeded in a single event at a T_m of 109.40 ± 0.25 °C, higher from the wild-type hCD81LEL for 43°C. This protein termed C4C9 migrated in SEC in native conditions as a single sharp peak at a time characteristic for the wild-type protein. Importantly, the stabilized mutant was able to bind to the structure-reporter antibody M38 (Thermo Fisher Scientific) to the same extent as to the wild-type hCD81 LEL. Far-UV CD spectrum for this mutant was examined and found to be identical to one obtained for the wild-type hCD81 LEL, which was typical of a protein with high α -helix content with two characteristic minima at 208 and 222 nm, similarly to previously published results.

First, a series of mutagenesis steps the recipient vector for yeast display pyd1_CD81LEL was modified with deletion of inherent *Bam*HI and *Hind*III restriction sites. All mutagenesis steps were performed using Quickchange Lightning Mutagenesis Kit (Agilent) using oligonucleotides: *Bam*HI site was deleted using oligonucleotides

(SEQ ID NO:21): CGATAAGGTACCAGGTTCTTTGTCAACAAG

(SEQ ID NO:22): CTTGTTGACAAAGGAACCTGGTACCTTATCG,

and *Hind*III site was deleted using oligonucleotides

(SEQ ID NO:23): GTATGTTTTTAAGCTCCTGCAGGCTAGTGGTG

(SEQ ID NO:24): CACCACTAGCCTGCAGGAGCTTAAAAACATAC.

To facilitate the recombination of library fragment after linearization, novel *Bam*HI restriction site was introduced using oligonucleotides

(SEQ ID NO:25): TTTGTGTCCCTCGGGATCCAACATCATCAGCAA

(SEQ ID NO:26): TTGCTGATGATGTTGGATCCCGAGGGACACAAA,

and a novel *Hind*III restriction site was introduced using oligonucleotides

(SEQ ID NO:27): GCAGTTCTATGACCAAGCTTTACAGCAGGCCGTGG

(SEQ ID NO:28): CCACGGCCTGCTGTAAAGCTTGGTCATAGAACTGC.

Recipient vector DNA was isolated with midipreparation using Nucleobond (Macherey-Nagel) and the vector was linearized using restriction with *Bam*HI and *Hind*III. Vector backbone was isolated using purification after preparative gel electrophoresis.

Further, the template insert of CD81 with mutations Ala134Cys and Lys144Cys, which form the novel disulfide bond, and Val135Cys and Ser168Cys, which form another novel disulfide bond, was modified to introduce sequences with restriction sites to

facilitate the correct recombination of the library encoding fragment. Template for modifications was pTT22SSP4_CD81 LEL C4C9 mutant. *Hind*III site was introduced using oligonucleotides

(SEQ ID NO:29): GCAGTTCTATGACCAAGCTTTACAGCAGTGCTGTG

(SEQ ID NO:30): CACAGCACTGCTGTAAAGCTTGGTCATAGAACTGC

and *Bam*HI site was introduced using oligonucleotides

(SEQ ID NO:31): TTTGTGTCCCTCGGGATCCAACATCATCAGCAA

(SEQ ID NO:32): TTGCTGATGATGTTGGATCCCGAGGGACACAAA.

PCR fragment with randomized nucleotide sequences was produced with oligonucleotides

(SEQ ID NO:33):

AAGGATGTGAAGCAGTTCTATGACCAAGCTTTANNKNNKTGCTGTNNKNNKNNKNNKGCCAACA
ACGCCTGTGCTGTG, wherein N is any one of A, C, G, or T

and (SEQ ID NO:34):

CTTGAAGAGGTTGCTGATGATGTTGGATCCCGAGGGACACAAATTMNNMNNAGCACACAMNNG
GTMNNMNNMNNMNNMTGTGCTGGAGCCACAGCA, wherein N is any one of A, C,
G, or T; and M is any one of A or C.

(according to IUPAC nucleotide code).

In a PCR reaction using Q5 HiFi Polymerase (New England Biolabs) and library fragment was purified after gel electrophoresis.

For transformation, a starter culture of EBY100 (Thermo Fisher Scientific) in 20 ml YPD medium (2% peptone, 1% yeast extract, 2% glucose) (Merck) was incubated overnight at 30°C and 180 rpm. The culture was then diluted to an OD₆₀₀ of 0.4 and incubated for about 5 hours at 30°C and 180 rpm. Aliquots of 50 ml cell culture were then pelleted at 1000 g for 5 min at room temperature, then washed with 25 ml AD and pelleted again. The cells were resuspended in 3 ml 100 mM Li-acetate and incubated for 15 min at 30°C in a shaking incubator. Cells were pelleted and supernatant removed. Components of the transformation mix were added as follows: 2400 µl 50% PEG 3350, 360 µl 1.0 M Li-acetate, 500 µl 2 mg/ml ssDNA (salmon sperm carrier DNA, previously heated up to 95°C for 5 minutes and then placed on ice) (Sigma-Aldrich), 10 µg linearized recipient vector and 7 µg DNA fragment. The cell pellets were resuspended in the

transformation mix and incubated for 30 min at 30°C in a shaking incubator and heat-shocked at 42°C for 45 min.

After collecting the cells by centrifugation at 1000 g for 5 min at room temperature and removing the supernatant, the pellets were resuspended in 10 ml SD-CAA medium. Aliquots were removed for dilution plating to determine the library size: 10 µl of cell suspension were diluted in 990 µl SD-CAA medium of which 100 µl were plated on an MDL plate and incubated at 30°C for 3 days. Yeast cells were diluted in 50 ml SD-CAA medium and incubated at 30°C at 180 rpm for 24 h, passaged into fresh SD-CAA medium at 1:20 dilution and cultured for another 24h under the same conditions. Cells were harvested at 1000 g, 5 min, at 4°C and pellets were resuspended in an equal volume of 30% glycerol before freezing at -80°C.

Yeast display transformations of the CD81_LEL Library 2 resulted in:

Library 2_4: 9.5×10^6 independent members,

Library 2_5: 2.1×10^7 independent members, and

Library 2_6: 1.7×10^7 independent members.

Example 7: Selections of CD81 LEL Library 2 with human EGFR-Fc

For sorting, libraries were first cultured in SD-CAA medium supplemented with penicillin-streptomycin at 30°C with shaking overnight and then the expression of recombinant protein was induced by resuspending the yeast cells in SG/R-CAA medium with penicillin-streptomycin and incubating them with shaking overnight at 37°C.

Next, 10^9 cells were pelleted for 5 min at 2500 g, then they were washed by resuspending the pellet in 50 ml wash buffer (1x D-PBS (Thermo Fisher Scientific), 0.25% BSA (Sigma Aldrich), 2mM EDTA (Sigma Aldrich) and pelleted again for 5 min at 2500 g. The washed cells were resuspended in 25 ml wash buffer containing 0.25 µM biotinylated antigen and incubated for 30 min at room temperature with gentle agitation and 10 min in an ice bath. Then, the cells and antigen solution was pelleted for 5 min at 2,500 g and 4°C, washed twice with 50 ml wash buffer and pelleted after each washing step. The cells were resuspended in 5 ml wash buffer plus 25 µl streptavidin microbeads and incubated on ice for 10 min. Lastly, 15 ml wash buffer were added.

The LS column (Miltteny Biotec) was placed in the separator and 3 ml of wash buffer were applied to precondition the column. Then, the cell solution was loaded in 7 ml batches. Once the column stopped dripping it was briefly removed from the magnet and placed back in the magnet to reorient the beads in the column to allow trapped cells

to flow through. Before loading the next 7 ml cell suspension 1 ml of wash buffer was applied to the column. These steps were repeated until all cells were loaded. Then, the column was washed with 3 ml wash buffer, briefly removed from the magnet and washed again with 3 ml wash buffer. To elute the binding cells the column was removed from the magnet and 5 ml wash buffer were added. Using the supplied plunger, the cells were pushed through the column into a new tube. The binding cell fraction was pelleted for 5 min at 2500 g. The pellet was resuspended in 10 ml SD-CAA medium and incubated at 30°C and 180 rpm overnight. Additionally, 100 µl of a 10⁻² dilution were plated onto an MDL plate and incubated at 30°C.

The induced cell suspensions were diluted to 10⁸ cells per 1 ml 10% BSA-PBS and centrifuged at 3000 rpm for 5 min at 20°C. To block the cells, the pellets were resuspended in 1 ml 10% BSA-PBS and incubated for 30 min at 20°C on a rotating platform. The cells were centrifuged at 3000 rpm for 5 min at 20°C and resuspended in 250 µl 10% BSA-PBS containing 1 µM biotinylated EGFR-Fc (Sino Biological). After incubation for 30 min at 20°C on a rotating platform the cells were centrifuged at 3000 rpm for 5 min at 4°C. To wash the cells, the pellets were resuspended in 1 ml ice-cold PBS and centrifuged at 3000 rpm for 5 min at 4°C. Then, the cells were resuspended in 250 µl 10% BSA-PBS with streptavidin-Alexafluor 647 (1:800) and anti-V5-FITC antibody (1:100) (Thermo Fisher Scientific) and incubated for 30 min on ice. The cells were centrifuged at 3000 rpm for 5 min at 4°C before they were resuspended in 1 ml ice-cold PBS and centrifuged again. Lastly, the cells were resuspended in 250 µl ice-cold PBS and were kept on ice until sorting with FACS Aria™.

After sorting the cells were resuspended in an appropriate amount of SD-CAA and incubated at 30°C and 180 rpm before induction as described above.

Table 4

Library	AB1 (AA 132-133)	AB2 (AA 136-139)	CD1 (AA 162-165)	CD2 (AA167)	CD3 (AA171-172)
81_2_4	VI	GASA (SEQ ID NO:35)	DAGP (SEQ ID NO:36)	H	VE
81_2_4	KW	SFLV (SEQ ID NO:37)	PLTM (SEQ ID NO:38)	T	LK
81_2_6	WR	AKMY (SEQ ID NO:39)	SARF (SEQ ID NO:40)	E	TW

Example 8: Expression of modified CD81 LEL in soluble form

Sequences encoding CD81 were amplified from yeast display clones using oligonucleotides

(SEQ ID NO:41):

ACGTGCTAGCTTTGTCAACAAGGACCAGATC

(SEQ ID NO:42):

ACGTGGTGACCGGTGCTGGAACCCCTCCCGGAGAAGAGGTC.

PCR fragment was cut with restriction enzymes *NheI* and *BstEII* and ligated to vector pTT28 (CNRC) that was cut with the same enzymes. Ligation mixture was transformed into electrocompetent *E. coli* TOP10 (Thermo Fisher Scientific) and transformants were selected on ampicillin plates. Plasmid was isolated with miniprep and transfected into ExpiCHO cells (Thermo Fisher Scientific). Expression in ExpiCHO expression system (Thermo Fisher Scientific) was according to MaxTiter protocol exactly following the manufacturer's instructions. hCD81 LEL variants were then purified via Ni-NTA chromatography using standard protocols. Supernatant was diluted with an equal volume of AD and buffered with PBS and 20 mM imidazole and pH adjusted to 7.5. Excel Ni-NTA column (GE Healthcare) was equilibrated with PBS and 20 mM imidazole, pH 7.5, and loaded with the buffered supernatant. Elution was in 5 column volumes with a linear gradient from 20 to 500 mM imidazole in PBS, pH 7.5. Protein-containing fractions were pooled and dialysed against 100-times volume of PBS at 4°C overnight.

SEC analysis in native conditions revealed a monodisperse elution profile corresponding to a dimeric form of the protein, analogously to the soluble wild-type CD81 LEL.

Alternatively, protein was expressed in ExpiCHO expression system (Thermo Fisher Scientific) following MaxTiter protocol exactly according to manufacturer's instructions. hCD81 LEL was then purified via Ni-NTA chromatography using standard protocols. Supernatant was diluted with an equal volume of AD and buffered with PBS and 20 mM imidazole and pH adjusted to 7.5. Excel Ni-NTA column (GE Healthcare) was equilibrated with PBS and 20 mM imidazole, pH 7.5, and loaded with the buffered supernatant. Elution was in 5 column volumes with a linear gradient from 20 to 500 mM imidazole in PBS, pH 7.5. Protein-containing fractions were pooled and dialysed against 100-times volume of PBS at 4°C overnight.

Example 9: Thermal Stabilization of the CD81 LEL

The fragment encoding hCD81 LEL (fragment Phe113-Lys201) (numbering according to SEQ ID NO:87) was amplified from a synthetic construct with a full length CD81 sequence (Geneart). DSDBASE (Vinayagam et al. 2004, Nucleic Acids Research, Volume 32, Issue suppl_1, Pages D200–D202, <https://doi.org/10.1093/nar/gkh026>) was used as a prediction tool for the identification of positions with the potential to harbor cysteine residues suitable for the creation of intradomain disulfide bonds. The algorithm was used to analyze hCD81 LEL crystal structure 1G8Q for the distances between C α and C β atoms of neighboring amino acid residues as well as for torsion angles and resulting S-S bond lengths. Out of 36 predicted possible disulfide bonds 11 with the highest likelihood of success as judged by visual examination of the crystal structure were selected. Five of those were predicted by the DSDBASE program both in protomer A and protomer B of hCD81 LEL.

Mutagenesis of single chosen amino acid residues to cysteine was performed using QuikChange Lightning Mutagenesis kit (Agilent), exactly according to manufacturer's instructions with oligonucleotides listed in Table 5.

Table 5 showing oligonucleotides used for mutagenesis

hCD81 LEL variant	Substitution	Oligonucleotide sequence	SEQ ID NO:
C1	A134C	CAGGCCCTACAGCAGTGCGTGGTGGATGATGAC	43
		GTCATCATCCACCACGCACTGCTGTAGGGCCTG	44
	A143C	GATGATGCCAACAACTGCAAGGCTGTGGTGAAG	45
		CTTCACCACAGCCTTGCAAGTTGTTGGCATCATC	46
C2	A130C	CAGTTCTATGACCAGTGTCTACAGCAGGCCGTG	47
		CACGGCCTGCTGTAGACACTGGTCATAGAACTG	48
	V146C	AACAACGCCAAGGCTTGTGTGAAGACCTTCCAC	49
		GTGGAAGGTCTTCACACAAGCCTTGGCGTTGTT	50
C3	Q133C	GACCAGGCCCTACAGTGCGCCGTGGTGGATGATG	51
		CATCATCCACCACGGCGCACTGTAGGGCCTGGTC	52
	A143C	GATGATGCCAACAACTGCAAGGCTGTGGTGAAG	53
		CTTCACCACAGCCTTGCAAGTTGTTGGCATCATC	54
C4	A134C	CAGGCCCTACAGCAGTGCGTGGTGGATGATGAC	55
		GTCATCATCCACCACGCACTGCTGTAGGGCCTG	56
	K144C	GACGCCAACAAACGCCTGTGCTGTGGTGAAGACC	57
		GGTCTTCACCACAGCACAGGCGTTGTTGGCGTC	58
C5	L154C	GACCTTCCACGAGACGTGTGACTGCTGTGGCTC	59
		GAGCCACAGCAGTCACACGTCTCGTGGAAGGTC	60

	K193C	GAGGACTGCCACCAGTGTATCGATGACCTCTTC	61
		GAAGAGGTCATCGATACACTGGTGGCAGTCCTC	62
C6	A120C	CAACAAGGACCAGATCTGTAAGGATGTGAAGCAG	63
		CTGCTTCACATCCTTACAGATCTGGTCCTTGTTG	64
	F198C	AAGATCGATGACCTCTGTTCCGGGAAGTGATGAG	65
		CTCATCACTTCCCGGAACAGAGGTCATCGATCTT	66
C7	A130C	CAGTTCTATGACCAGTGTCTACAGCAGGCCGTG	67
		CACGGCCTGCTGTAGACACTGGTCATAGAACTG	68
	A143C	GATGATGCCAACAAGTCAAGGCTGTGGTGAAG	69
		CTTACCACAGCCTTGCAGTTGTTGGCATCATC	70
C8	L131C	TTCTATGACCAGGCCTGCCAGCAGGCCGTGGTG	71
		CACCACGGCCTGCTGGCAGGCCTGGTCATAGAA	72
	L165C	AGCACACTGACTGCTTGTACCACCTCAGTGCTC	73
		GAGCACTGAGGTGGTACAAGCAGTCAGTGTGCT	74
C9	V135C	GCCCTACAGCAGGCCTGTGTGGATGATGACGCC	75
		GGCGTCATCATCCACACAGGCCTGCTGTAGGGC	76
	S168C	GCTTTGACCACCTGTGTGCTCAAGAACAAAT	77
		ATTGTTCTTGAGCACACAGGTGGTCAAAGC	78
C10	V169C	TTGACCACCTCAGTGTGCAAGAACAATTTGTGTC	79
		GACACAAATTGTTCTTGACACTGAGGTGGTCAA	80
	L174C	GTGCTCAAGAACAATTGTTGTCCCTCGGGCAGC	81
		GCTGCCCCGAGGGACAACAATTGTTCTTGAGCAC	82
C11	S160C	GA CTGCTGTGGCTCCTGTACACTGACTGCTTTG	83
		CAAAGCAGTCAGTGTACAGGAGCCACAGCAGTC	84
	D189C	AACCTCTTCAAGGAGTGTGGCCACCAGAAGATC	85
		GATCTTCTGGTGGCAACACTCCTTGAAGAGGTT	86

hCD81 LEL variants were cloned into the pTT22SSP4 mammalian expression vector (CNRC) and expressed in two different expression systems. For prescreening, the constructs were expressed in HEK293-6E cells (CNRC) at a 2-mL-scale in F17 medium supplemented with 4 mM glutamine and 50 µg/mL G-418 (Thermo Fisher Scientific) on an orbital shaker at 180 rpm, at 37°C under 5% CO₂ for 4 days, with feeding of TN-20 to an end concentration of 0.8% on the second day after transfection. Mutant C5 did not express, C6 expressed poorly and C10 formed a conspicuous dimer, therefore they were omitted from further analysis. Mutants selected for further characterization (C1, C2, C3, C4, C7, C8, C9 and C11) were transfected into ExpiCHO cells (Thermo Fisher Scientific) exactly according to manufacturer's instructions. Cultivation of the cells proceeded according to MaxTiter protocol (Thermo Fisher Scientific). Supernatants were harvested after 14 days and purified using Ni-NTA affinity chromatography. After clarification, the samples were buffered with phosphate buffered saline (PBS) with 20 mM imidazole, pH 7.5, and passed over Excel Ni-NTA column (GE

Healthcare) equilibrated with the same buffer. His-tagged hCD81 LEL was eluted with a gradient from 20 to 500 mM imidazole in 5 column volumes. Fractions containing the target protein were pooled and dialyzed twice against 100-times volume of PBS overnight at 4 °C. The proteins were stored at -80 °C until use.

Shimadzu LC-20A Prominence system equipped with a diode array detector and a refractive index detector was used to perform SEC-HPLC with a *Superdex* 200 Increase 10/300 GL column (GE Healthcare). The mobile-phase buffer used was PBS with 200 mM NaCl. Chromatography was conducted with a constant flow rate of 0.75 mL/min. A total of 200 µg protein at about 2 mg/mL were loaded on the column for analysis. Column calibration was performed with a set of molecular weight standards ranging from 10 to 500 kDa (Bio-Rad). Mutants C1 and C8 did not show well resolved peaks and all other mutants were similar to the wild-type protein.

DSC experiments were performed using an automated MicroCal PEAQ-DSC Automated system (Malvern), using 80 µM protein solution, diluted in PBS at pH 7.4. The heating was performed from 20°C to 110°C at a rate of 1°C/min. Protein solution was then cooled *in situ* and an identical thermal scan was run to obtain the baseline for subtraction from the first scan. All measurements were taken in duplicates. Fitting was performed with MicroCal PEAQ-DSC Software using the non-2-state transition mechanism.

Measurement of thermal stability of wild-type CD81 LEL revealed a single melting point at 66.15 °C, however a low enthalpy of 2.4×10^4 kcal/mol.

Table 6 showing T_m of hCD81 LEL variants

hCD81 LEL	Mutated position 1		Mutated position 2		T _m (°C)
variant	Located on segment	Amino acid	Located on segment	Amino acid	
wild-type					66.15±0.25
C1	Helix A	Ala134	Helix B	Ala143	n.d. ¹
C2	Helix A	Ala130	Helix B	Val146	67.35±0.05
C3	Helix A	Gln133	Helix B	Ala143	82.15±0.05
C4	Helix A	Ala134	Helix B	Lys144	88.95±0.05
C5	Helix B	Leu154	Helix E	Lys193	n.d.
C6	Helix A	Ala120	Helix E	Phe198	n.d.
C7	Helix A	Ala130	Helix B	Ala143	76.90±0.00
C8	Helix A	Leu131	Helix C	Leu165	n.d.
C9	Helix A	Val135	Helix C	Ser168	90.45±0.15
C10	Helix C	Val169	Unstructured part of segment C	Leu174	n.d.
C11	Loop preceding helix C	Ser160	Start of helix D	Asp189	70.25±0.15

Differential thermal calorimetry scans were run with a re-scan of the denatured protein solution to be subtracted as background. Here it was discovered that the mutant C2 in contrast with wild-type CD81 LEL and other stabilized variants surprisingly exhibited reversible unfolding up to 110°C.

A hCD81LEL mutant with a combination of potentially stabilizing novel disulfide bonds Ala134Cys/Lys144Cys and Val135Cys/Ser168Cys was then produced and tested for its thermostability. Thermal unfolding was recorded up to 130 °C and proceeded in a single event at a T_m of 109.40±0.25 °C, higher from the wild-type CD81LEL for 43 °C. This protein termed C4C9 migrated in SEC in native conditions as a single sharp peak at a time characteristic for the wild-type protein.

Importantly, the stabilized mutant was able to bind to the structure-reporter antibody M38 to the same extent as to the wild-type CD81 LEL. ELISA plate (Maxisorp, NUNC) was coated with an anti-hCD81 M38 antibody (Thermo Fisher Scientific) at 5 µg/mL in PBS for 1h at room temperature. After blocking with 5% bovine serum albumin (BSA)-PBS for 1h at RT, supernatants of HEK293-6E cells transfected with hCD81 LEL variants or purified variants of hCD81 LEL diluted in 2.5 % BSA-PBS were allowed to bind for 1h at RT. After extensive washing, the binding of mutant proteins was detected with an anti-his-horseradish peroxidase (HRP) conjugated antibody (QIAgen), diluted 1:2000 in 2.5% BSA-PBS. Antibody binding was revealed with 3,3',5,5'-

tetramethylbenzidine (TMB) (Sigma Aldrich), the reaction was stopped by adding an equal volume of 30% H₂SO₄ and absorbance was read at 450/620 nm.

Table 7 showing absorbance values of wildtype and mutant CD81 LEL

CD variant concentration (μg/ml)	LEL	A 450/620 (wild-type LEL)	CD81	SE (wild-type LEL)	CD81	A 450/620 (C4C9)	SE (C4C9)
5		3.9684		0.0316		3.82795	0.01235
1.666667		3.694		0.077		3.82425	0.08275
0.555556		3.6268		0.1979		3.73145	0.02185
0.185185		3.67715		0.06925		3.69125	0.05935
0.061728		3.53355		0.09535		3.45675	0.04415
0.020576		1.8876		0.0352		1.97565	0.00285
0.006859		0.61595		0.01845		0.7503	0.0542
0.002286		0.17665		0.01115		0.187	0.0405
0 (2 nd only)		0.02185		0.00125			

Far-UV CD spectrum for the mutant C4C9 and wild-type CD81 was examined. The CD spectra were measured on a Chirascan spectropolarimeter (Applied Photophysics). A 1 mm quartz cuvette was used Protein preparations were diluted in PBS to 1 mg/mL. Spectra were found to be identical for the wild-type hCD81 LEL and C4C9. Spectrum observed was typical of a protein with high α -helix content with two characteristic minima at 208 and 222 nm, similarly to previously published results.

Table 8 showing Far-UV CD spectrum for wildtype and mutant CD81

Wavelength (nm)	Raw ellipticity (millideg) Wild-type CD81 LEL	Raw ellipticity (millideg) C4C9
260	0.206785	0.145295
259	0.19977	0.177317
258	0.265558	0.205747
257	0.206915	0.148188
256	0.221795	0.159402
255	0.257513	0.183345
254	0.223584	0.140538
253	0.231864	0.120346
252	0.157287	0.072734
251	0.0663846	-0.0527117
250	-0.157381	-0.263102
249	-0.325095	-0.453448
248	-0.617804	-0.716791

247	-1.118	-1.19926
246	-1.6976	-1.79587
245	-2.4624	-2.56849
244	-3.50611	-3.55678
243	-4.72366	-4.74579
242	-6.31973	-6.25803
241	-8.27745	-8.2164
240	-10.556	-10.3732
239	-13.4046	-13.1298
238	-16.736	-16.3624
237	-20.6866	-20.1611
236	-25.1074	-24.5011
235	-30.1779	-29.403
234	-35.8887	-35.0253
233	-42.0612	-40.9566
232	-48.5922	-47.4025
231	-55.3904	-53.9963
230	-61.9441	-60.4902
229	-68.2922	-66.5214
228	-74.0251	-72.2262
227	-79.1404	-77.2646
226	-83.4069	-81.4188
225	-86.7399	-84.7366
224	-88.9684	-86.9499
223	-90.4185	-88.2177
222	-90.8311	-88.7247
221	-90.4246	-88.2572
220	-89.1682	-87.186
219	-87.7019	-85.7441
218	-86.3482	-84.3822
217	-85.2084	-83.4309
216	-84.2489	-82.4846
215	-83.359	-81.8334
214	-82.945	-81.2163
213	-83.4676	-81.9112
212	-85.7387	-83.7029
211	-88.655	-87.0087
210	-92.3376	-90.7764
209	-95.7168	-93.9584
208	-97.4284	-95.5125
207	-93.7432	-93.6295

206	-84.8028	-86.8648
205	-63.0458	-70.4512
204	-45.8841	-52.7578
203	-17.4302	-33.2052
202	16.0949	13.3243
201	791.06	-95.0327
200	-113.547	-58.5575
199	-93.1264	-40.4351
198	-69.128	-23.4953
197	-58.7217	-17.9316
196	-65.7443	-35.5883
195	-50.8545	-23.3845

Example 10: Production of target-specific EVs in cell culture.

10.1 Transient transfection of target-specific EV donor cell line using HeLa cells

HeLa cells expressing C-terminal GFP-fused recombinant CD81 to secrete target-specific CD81-GFP-containing exosomes (CD81-GFP-exosomes) were transiently transfected using electroporation. HeLa cells (5×10^6 cells) were plated on a T175 cell culture flask (Sigma, Kremsmünster, Germany) 2-3 days prior to transfection. Complete RPMI-1640 medium (Merck, Darmstadt, Germany), containing 10% FCS, 4 mM L-Glutamine, was used for the regular culturing of donor HeLa cells. On the day of transfection (Day 0), HeLa cells (optimally at 70-85% of confluency) were harvested using 0.1% Trypsin solution (5 minutes treatment, neutralization with complete RPMI-1640). Detached HeLa cells in RPMI-1640 were quantified using the ViCell Cell Counter. 3×10^6 cells were transfected with CD81-GFP plasmid (pTT5, NRC Canada, Ottawa, Canada, 4 μ g) to fill one T175 flask. Electroporation was performed using an Amaxa Nucleofector I device (Lonza, Basel, Switzerland) according to the manufacturers brochure (program I-013 for HeLa cells) using an in-house prepared electroporation buffer (5mM KCl, 15mM MgCl₂, 120mM Na₂HPO₄/NaH₂PO₄ pH7.2, 50mM Mannitol, 0.005% Pluronic F-68). Electroporated cells were recovered in complete RPMI-1640 and cultured for a time period of 24 hours allowing the cells to attach to culture flask surface. Indeed, various forms of human transferrin receptor targeting recombinant CD81 localized to the plasma membrane suggesting functional integration of the recombinant tetraspanin (Figure 4).

10.2 Isolation and purification of target-specific EVs

On Day 1 (24 hours after transfection), the culturing medium was removed and replaced by a respective volume (45 mL for 1x T175) of EV-depleted secretion medium (RPMI-1640). EV-depleted RPMI-1640 was prepared by collecting the supernatant fraction from ultracentrifuged (14-18h, 4°C, 100,000 g) complete RPMI-1640 (containing 20% FCS) and diluting the medium to 10% FCS with RPMI-1640 and adding of L-Glutamine to a final concentration of 4 mM. HeLa cells were incubated with EV-depleted medium for 72 hours (37 °C, 5% CO₂ atmosphere) to secrete target-specific EVs.

Isolation of EVs was performed after 72 h of collection time by using ultracentrifugation. The conditioned cell culture medium was collected and centrifuged (3,100 × g) for 30 min at 4 °C to remove cells and cell debris. The supernatant was passed through a 450 nm syringe filter unit (Merck, Darmstadt, Germany) and transferred to 100 mL sealable ultracentrifuge tubes. The supernatant was then centrifuged (100,000 × g) for 90 min at 4 °C using an ultracentrifuge (Optima L-60, Beckman Coulter, Brea, USA). The supernatant was removed and the pellet was collected in 1 mL the remaining medium. A second step of ultracentrifugation was performed (100,000 × g, 90 min, 4 °C) to collect EVs in a smaller volume. The supernatant was again removed and the remaining EV pellet was resuspended in 200 to 400 µl of HEPES-based 1x Live Cell Imaging Solution (ThermoFisher Scientific, Waltham, USA). Using this method, around 10¹¹ EVs /ml were purified from HeLa supernatants.

Example 11: Production of EVs carrying a traceable and purifiable tag

11.1 Sequence and cloning of CD81-Snorkel tag in retroviral system

Wild-type human CD81 sequence amino acid sequence SEQ ID NO:87; nucleic acid sequence SEQ ID NO:88) was cloned into pBMN vector with C-terminally fused SNORKEL tag (Brown et al. 2013, PLoS ONE 8(9): e73255. doi:10.1371/journal.pone.0073255). SNORKEL tag enables tags displayed on the surface of vesicular membrane. The PCR product of CD81 was cut with NcoI and AgeI and PCR product of SNORKEL tag was cut with AgeI and NotI. pBMN plasmid was cut with NcoI and NotI. Digested PCR products were ligated with NcoI and NotI digested pBMN plasmid. The ligation mixture was transformed into competent *E.coli* cells and transformants were selected on ampicillin plates. Plasmids were isolated by endotoxin free plasmid preparation (Qiagen Maxiprep).

11.2 Stable expression of CD81-Snorkel tag in EV donor cells using retroviral transfection

Phoenix cells (5×10^6 cells) were plated in a T75 plate (Sigma, Kremsmünster, Germany) 1 day prior (day 0) to transfection. Cells were cultured in complete growth media with DMEM 4.5g glucose with 10% FCS and 4mM L-Glutamine. Once cells were 60-70% confluent, culture media was removed and replaced with DMEM 4.5g glucose with 4mM L-Glutamine without FCS. Cells are transfected with 10 μ g of pBMN-CD81-SNORKEL tag (SNORKEL tag fused to CD81 C-termini) plasmid using jetPRIME transfection reagent (Polyplus, France). 24hrs post transfection (day 1), media was changed to complete growth media of 7 ml. Target cells (HeLa) were plated in 6 well plate (Sigma, Kremsmünster, Germany) with 1×10^5 cells/well in RPMI-1640 medium containing 10% FCS and 4mM L-Glutamine. 24 hrs post seeding (day 2), HeLa cells (50-80% confluency) can be infected. 7mL supernatant from Phoenix cells (containing virus particles) was filtered using 0.45 μ m filter (Merck, Darmstadt, Germany) and polybrene (final conc 8 μ g/mL) to this filtrate and mixed. Fresh growth media was added onto Phoenix cells. Supernatant from target cells was removed and up to 2 mL of virus-containing supernatant were added. The entire plate was wrapped in parafilm and centrifuge at 800 x g for 60min at room temperature. Then, virus supernatant was discarded and fresh growth media was added onto Hela cells. Viral infection was carried out for next 3 more days until the cells reach 95-100% expressing CD81-SNORKEL tag. The tagged CD81 was localized to cell membranes indicating correct folding and localization (Figure 5A). Along with SNORKEL tag fused to C-termini of CD81 the method was also successfully performed with CD63 fusion protein in Hela cells (Figure 5B). Similar results were observed in human mesenchymal stem cells (ASCs). The EVs are simultaneously positive for the HA tag as well as the CD81 when subjected to electron microscopy after immunogold labelling by anti HA antibodies coupled to 10 nm gold particles and anti – CD81 antibodies coupled to 18 nm gold particles (Figure 5C).

11.3 Characterization of EVs for size, number and incorporation of CD81-SNORKEL tag

HeLa cells stably expressing CD81-SNORKEL tag were plated in 3 T75 plates (5×10^6 cells/ flask) in culture medium. 24hrs of post seeding (day 1) the culturing medium was removed and replaced by a respective volume (12 mL for 1x T75) of EV-depleted secretion medium (RPMI-1640). EV-depleted RPMI-1640 was prepared by collecting the supernatant fraction from ultracentrifuged (14-18h, 4°C, 100,000 g) complete RPMI-1640 (containing 20% FCS) and diluting the medium to 10% FCS with RPMI-1640 and adding of L-Glutamine to a final concentration of 4 mM. HeLa cells were incubated with EV-depleted medium for 48 hours (37 °C, 5% CO₂ atmosphere) to secrete EVs carrying SNORKEL tag.

Isolation of EVs carrying SNORKEL tag was performed after culturing cells in EV depleted media for 48hrs. The conditioned media was collected and centrifuged at 700g for 5 min at 4°C to remove cell debris. The supernatant was passed for 2nd round of centrifugation at 2000g for 10 min at 4°C to remove apoptotic bodies and larger particles. Next the supernatant is passed through 0.22 µm filter (Merck, Darmstadt, Germany) to remove larger particles like microvesicles. The processed supernatant was concentrated to 1/20th times of initial volume using tangential flow filtration columns with 300 kDa pore size hollow fibers (SpectrumLabs, Netherlands). The concentrated conditioned media used for further characterizing EVs. 10µl of concentrated conditioned media was diluted to 1000 times in filtered DPBS (Merck, Darmstadt, Germany) and used for nanoparticle tracking analysis (NTA) in a scatter mode with nanosight NS500. All of the acquisition parameters were identical to those of the size and concentration measurements. All experiments were measured in experimental triplicates. The size of the EVs from different isolations are around 110-140 nm.

The incorporation of CD81-SNORKEL tag into EVs was confirmed by western blot analysis. 1×10^{10} EVs quantified by nanosight were loaded per well along with 50 µg of cell lysate quantified by BCA kit (Thermo fisher scientific) in NuPAGE 4-12 % Bis-Tris protein gels and SDS-PAGE was performed. Proteins on gel were transferred to PVDF membrane using Trans-Blot turbo transfer system (Bio-Rad). Specific antibodies against EV specific markers along with SNORKEL tag for quantification. Indeed, EVs showed markers of typical EV proteins in Western blots including CD63, CD81, and TSG101 as well as absence of the intracellular protein calnexin.

11.4 Isolation of EVs exclusively carrying SNORKEL tag using anti-HA matrix and PreScission protease

The conditioned media from HeLa cells expressing CD81-SNORKEL tag was processed by differential centrifugation (700g & 2000g), filtered using 0.22 µm filter (Merck, Darmstadt, Germany) and concentrated to 1 ml using tangential flow filtration columns with 300 kDa pore size hollow fibers (SpectrumLabs, Netherlands). 350 µl of concentrated conditioned media was mixed with 200 µl of anti-HA antibody conjugated to agarose beads (Sigma) and incubated overnight at 4°C on test tube rotator. After 16-18 hrs of incubation, the agarose beads were spun down at 500g for 5 min at 4°C and flowthrough (unbound EVs) were collected and beads were washed with 1 ml of filtered DPBS by spinning down the agarose beads at 500g for 5 min at 4°C. Next the anti-HA antibody conjugated agarose beads bound to SNORKEL tag on EVs were subjected to PreScission protease treatment (sigma) 4 µl with 8 units in 50mM Tris (Sigma), 150mM NaCl (Sigma) with pH 7-7.4 for 16-18 hrs at 4°C on test tube rotator. After PreScission protease treatment for 16-18 hrs agarose beads were spun down at 500g and supernatant was collected. All the samples (input, flowthrough, wash and elute) were used for NTA analysis with 1 to 1000 times dilution to quantify the number of EVs carrying SNORKEL tag were purified. Purification using PreScission protease yielded around 80-90% of vesicles mildly eluted from the affinity matrix, as no bands are detectable in the elution fraction after PreScission digest using the cleaved off HA tag, while the remaining CLIP tag is well visible in the elution fraction. Comparing elution fraction to the fraction not eluted and remaining on the beads, a yield of about 80-90% is observed (Figure 6).

Insertion of antigen specific ligands such as Gas6 for AXL-specific targeting before the PreScission protease site in SNORKEL tag enables purification of specific EVs carrying SNORKEL tag and targeting EVs after purification is performed.

Example 12: Characterization of the target-specific extracellular EVs

12.1 Characterization of EV size and recombinant protein incorporation ratio

10µl of EV isolates were diluted with 6 mL of filtered DPBS (Merck, Darmstadt, Germany) and used for nanoparticle tracking analysis (NTA) in scattering and fluorescence mode with a ZetaView (Particle Metrix, Inning, Germany), followed by evaluation using the NTA software (Particle Metrix, Inning, Germany). All of the

acquisition parameters were identical to those of the size and concentration measurements. All experiments were measured in experimental triplicates.

The ratio of the particle numbers measured in scattering and fluorescence mode equals the ratio between recombinant EVs, harboring the target-specific CD81-LEL portion and the fluorescent GFP, and non-recombinant and non-fluorescent native EVs. Through different batches of recombinant EV production the ratio obtained from NTA measurements showed that approximately 30-50% of all isolated EVs are recombinant. The common median size of isolated EVs with the above-mentioned isolation method is around 120-140 nm and the typical yield of target-specific fluorescent EV amounts to 1600 to 5000 recombinant EVs per HeLa cell in a size range of 60-400 nm, peaking at 130 nm in diameter of particle size (Figure 7).

12.2 Qualification and quantification of specific EV uptake *in vitro*

(a) Qualification of EV uptake in target-antigen positive cells via fluorescence microscopy

Recipient cells (e.g. Caco-2 cells) were seeded in an 8-well ibidi glass bottom plate (2.5×10^4 cells per well) and allowed to attach for 24h in complete culture medium (37 °C, 5% CO₂ atmosphere). After the incubation time, the 10^9 recombinant EVs were added to each well. Cells were cultivated for another 24h allowing recombinant particles to be taken up by the recipient cells. Cells were eventually washed PBS and treated with acidic glycine buffer (100mM NaCl, 100 mM Glycine, pH 3.5) on ice to remove surface bound EVs. Cells were then imaged at 40× magnification using fluorescence microscopy (DMI3000B; Leica Microsystems, Wetzlar, Germany). Microscopic imaging parameters (exposure time, contrast, and gain) were the same for all experiments. Indeed, recombinant EVs were readily taken up by recipient cells (Figure 8A).

(b) Quantification of EV uptake in target-antigen positive cells via flow cytometry

Recipient cells (e.g. Caco-2 cells) were seeded in a 24-well plastic bottom plate (0.25×10^6 cells per well) and allowed to attach for 24h in complete culture medium (37 °C, 5% CO₂ atmosphere). After the incubation time, a fixed number of 5×10^9 recombinant EVs were added to each well. Cells were cultivated for another 24h allowing recombinant particles to be taken up by the recipient cells. Cells were eventually washed DPBS and treated with 0.1% Trypsin (5 minutes, 37°C) and neutralized with complete

medium containing 10% FBS. Cells were then centrifuged at 300 g and the pellet was resuspended in ice-cold DPBS.

Samples were kept on ice and measured with the Cytoflex S flow cytometer (Beckman Coulter, Brea, USA). Data was analyzed with the Cytoflex software. Mean fluorescence intensity was normalized over the control/untreated cell sample (Δ MFI). Here an increased uptake of recombinant EVs targeting the transferrin receptor by about 30-60% was observed depending on the recombinant CD81 construct (Tfr1EVsP1, Tfr2EVs P1, Tfr2CP4EVs P1) as compared to EVs of cells overexpressing non-Tfr1 targeting CD81 (CP4EVs) or wild type EVs (wtEVs) (Figure 8B).

Example 13: Production of therapeutic loaded target-specific EVs using apoptosis-inducing siRNA

13.1 Transfection of recombinant EVs with siRNAs

Purified EVs carrying either non-recombinant or recombinant CD81 targeting the human transferrin receptor were transfected either with apoptosis-inducing siRNA (TOX transfection control, Dharmacon) or with untargeted control siRNA (ON-TARGETplus Non-targeting siRNA, Dharmacon) by using the liposome-based transfection reagent Dharmafect (Dharmacon, Lafayette, USA). EVs were transfected with respective targeted or untargeted siRNA (to a final concentration of 25 nM siRNA) according to the manufacturer's brochure. Loading of siRNA was controlled for using qPCR.

13.2 Cell-based assay for the determination of target-specific cytotoxicity

Target-antigen expressing cells were seeded at a density of 7.5×10^3 cells/well in 96-well plates in complete medium followed by incubation with 5×10^8 recombinant/siRNA-transfected EVs per well for 72 h. Experiments were performed in six replicates per EV dose or EV control. Following medium removal and PBS washing, cells were incubated with 1x WST-1 (Sigma, St. Louis, USA) for 1h, 2h and 4h or 1x AlamarBlue (ThermoFisher Scientific, Waltham, USA) overnight in complete medium. Assays readouts were performed at the microplate reader Infinite 200 Pro (Tecan, Männedorf, Switzerland) according to the respective manufacturer's brochures. Indeed, an about 30% higher cytotoxicity of EVs targeting the transferrin receptor as compared to controls was observed, especially using the Tfr1CP4 recombinant CD81 construct (Figure 10).

Example 14: Design of the library CD81 LEL_L3

CD81 LEL_L3 library design was based on the reversibly refolding CD81 LEL stabilized mutant (SEQ ID NO:180)

SEQ ID NO:180:

FVNKDQIAKDVKQFYDQCLQQACVDDDDANNAKACVKTFHETLDCCGSSTLTALTTCVLKNNLCP
SGSNIIISNLFKEDCHQKIDDLFSGK

In this library, the stabilizing mutations Ala130Cys/Ala146Cys and Val135Cys/Ser168Cys work additively to increase the midpoint of thermal transition of CD81 LEL to 93.4 °C, and this combined mutant can reversibly refold when heated up to 110 °C. In the library, the amino acid residues at the positions 132-133, 136-141, 162-165, 167, and 171-172 were randomized to form a composite surface available for antigen binding (X in SEQ ID NO:1) (SEQ ID NO:97).

SEQ ID NO:97:

FVNKDQIAKDVKQFYDQCL**XX**AC**XXXXXX**NAKACVKTFHETLDCCGSST**XXXXT**XCVL**X**
XNLCPSGSNIIISNLFKEDCHQKIDDLFSGK

wherein X is any amino acid.

Example 15: Construction of yeast display libraries CD81 LEL_L2 and L3

For the yeast display library CD81 LEL_L2, PCR recombination products encoding the randomized inserts were produced in 100 µl aliquots using Q5 HiFi Polymerase MasterMix (New England Biolabs), 10 ng/µl template (pTT22SSP4_C4C9), and 50 pmol of oligonucleotides LIB2FWD (SEQ ID98) and EFrev (SEQ ID NO:99). The initial denaturation for 30 sec at 98°C was followed by 35 cycles of each 20 sec denaturation at 98°C, 20 sec annealing at 55°C and 20 sec extension at 72°C, and completed with an incubation step at 72°C for 5 min. PCR products were purified with Illustra GFX purification kit and eluted in AD.

SEQ ID NO:98:

AAGGATGTGAAGCAGTTCTATGACCAAGCTTTANNKNNKTGCTGTNNKNNKNNKNNKGCCAACA
ACGCCTGTGCTGTG

wherein N is any one of A, C, G, or T.

SEQ ID NO:99:

CTTGAAGAGGTTGCTGATGATGTTGGATCCCGAGGGACACAAATTMNNMNNAGCACACAMNNG
GTMNNMNNMNNMNNMTGTGCTGGAGCCACAGCA

wherein N is any one of A, C, G, or T; and M is any one of A or C.

For the yeast display library CD81 LEL_L3, PCR recombination products encoding the randomized inserts were produced in 100 µl aliquots using Q5 HiFi Polymerase MasterMix (New England Biolabs), 10 ng/µl template (pTT22SSP4_C2C9), and 50 pmol of oligonucleotides LIB3FWD (SEQ ID 100) and LIB2REV2 (SEQ ID 99). The initial denaturation for 30 sec at 98°C was followed by 35 cycles of each 20 sec denaturation at 98°C, 20 sec annealing at 55°C and 20 sec extension at 72°C, and completed with an incubation step at 72°C for 5 min. PCR products were purified with Illustra GFX purification kit (GE Healthcare) and eluted in AD.

SEQ ID NO:100:

AAGGATGTGAAGCAGTTCTATGACCAGTGTCTANNKNNKGCCTGTNNKNNKNNKNNKNNKNNKA
ACGCCAAGGCTTGTGTG

wherein N is any one of A, C, G, or T.

The recipient vector pYD1 delbamdelhind_bamhind was linearized using *Bam*HI and *Hind*III and purified using preparative agarose gel electrophoresis and Illustra GFX purification kit (GE Healthcare), and eluted in AD.

Yeast *Saccharomyces cerevisiae* EBY100 was transformed with linearized recipient vector and PCR recombination product using chemical PEG3350/Li-acetate/ssDNA transformation method. 2 libraries, L2A and L2B, and L3A and L3B, were produced with each of the recombination fragments. For each library, 250 ml YPD medium were inoculated with an overnight culture to OD₆₀₀ of 0.4. After 5 h incubation on shaking incubator at 30°C, the yeast cells were harvested in 50-ml-aliquots. Supernatant was removed and yeast cells were washed with 25 ml AD per aliquot at 3000 rpm, 5 min at RT. Cell pellet was resuspended in 3 ml per aliquot of 100 mM Li-acetate and shaken at 200 rpm, 15 min at 30°C. Cells were collected at 3000 rpm, 5 min

at RT. To each yeast cell aliquot a mix of 2750 μ l 50% PEG3350 solution, 360 μ l 1 M Li-acetate, 500 μ l heat-shocked ssDNA, and 10 μ g linearized recipient vector and 10 μ g of recombination PCR fragment was added. The incubation was performed at 200 rpm, 30 min at 30°C, and was followed by a heat shock for 45 min at 42°C. Yeast cells were collected at 3000 rpm, 5 min at RT and diluted into 250 ml SD-CAA medium with penicillin and streptomycin. Cultivation proceeded at 200 rpm for 24 h at 30°C and 12.5 ml of the culture were transferred into 250 ml of fresh SD-CAA medium with penicillin and streptomycin for 24 h at 30°C, pelleted at 3000 rpm, 10 min at 4°C and frozen after mixing with an equal volume of 30% glycerol. The size of the libraries was determined using dilution plating and was determined to be 1.1×10^8 independent members for L2 and in 1.2×10^8 independent members for L3.

To determine the level of correctness of the libraries, plasmid DNA was isolated from 10 μ l of pelleted yeast cells and transformed to *E.coli* TOP10 using electroporation. The expression cassette of CD81 LEL mutant was amplified using primers pydfwd (SEQ ID NO:101) and pydrev (SEQ ID NO:102) and sequenced using one of these primers. The correctness of the library L2A was found to be 62.5%, L2B 87.5%, L3A 62.5% and L3B 100%.

SEQ ID NO:101: AGTAACGTTTGTTCAGTAATTGC

SEQ ID NO:102: GTCGATTTTGTACATCTACAC

For quality control, the yeast cells were induced in SG/RCAA medium with penicillin and streptomycin either for 48h at 20°C or 24h at 37°C. For staining, the yeast cells were blocked in a 2% BSA-PBS solution for 30 min at RT at an OD₆₀₀ of 1. Then they were resuspended into 100 μ l-aliquots and stained with anti-Xpress antibody (Thermo Scientific) (1:1000) reactive with the N-terminally positioned Xpress tag, and M38 antibody (Thermo Scientific) (1 μ g/ml), which detects the properly folded CD81 LEL, in 2% BSA-PBS for 1h at RT. Cells were pelleted at 3000 rpm, for 5 min at 4°C and resuspended in 2% BSA-PBS with goat anti-mouse (Fab')₂-FITC conjugate (Sigma Aldrich), diluted 1:200 in 2% BSA-PBS. Other stainings were: anti-his-tag-Alexa Fluor 488 (QIAGEN), diluted 1:200 in 2% BSA-PBS and anti-V5-tag-FITC (Thermo Scientific), diluted 1:100 in 2% BSA-PBS, used to detect C-terminal his tag and C-terminal V-tag to determine the correct read-through of the clones. The incubation with fluorescent antibodies was for 30 min on ice. Finally, the cells were collected at 3000 rpm, 5 min at 4°C, and resuspended in 200 μ l ice-cold PBS. The fluorescence of stained samples and

unstained controls was determined using Guava EasyCyte flow cytometer. The percentage of positive cells was determined and is presented in Table 9.

	Sample	L2A	L2A	L2B	L2B	L3A	L3A	L3B	L3B
	Induction temperature (°C)	20	37	20	37	20	37	20	37
staining									
1 st step	2 nd step								
unstained	unstained	0.57	0.43	0.29	0.34	0.06	1.25	0.09	1.21
Anti-Xpress	goat anti-mouse (Fab') ₂ -FITC	75.4	68.65	77.13	65.79	78.04	77.89	80.15	70.84
M38	goat anti-mouse (Fab') ₂ -FITC	1.73	1.99	0.66	1.81	1.94	1.65	1.19	0.81
unstained	goat anti-mouse (Fab') ₂ -FITC	0.48	1.31	0.43	1.26	0.45	0.44	0.60	1.12
unstained	anti-his-tag-Alexa Fluor 488	57.73	49.16	61.29	50.21	59.52	48.44	65.13	46.45
unstained	anti-V5-tag-FITC	61.05	51.81	60.92	51.95	62.27	55.67	66.85	51.60

Table 9.

Example 16: CD81 LEL-based EGFR-specific binders

16.1 Selections of CD81 LEL libraries L2 and L3 with human EGFR-Fc

Human EGFR-Fc was purchased from Sino Biological. For biotinylation, EZ-Link™ Sulfo-NHS-LC-LC-Biotin reagent (Thermo Scientific) was used in a 3:1 biotin to protein molar ratio. The antigen was reconstituted exactly according to manufacturer's instructions to a concentration of 0.25 µg/µl. Incubation with biotinylation reagent proceeded for 1 h at room temperature with shaking. Unbound biotin was removed with dialysis against 100x volume of PBS, using Snakeskin dialysis tubing (Thermo Scientific) with 10.000 Da MWCO, at 4 °C overnight with stirring.

For sorting, libraries 2A, 2B, 3A and 3B were first cultured in SD-CAA medium supplemented with penicillin-streptomycin at 30 °C with shaking overnight and then the expression of recombinant protein was induced by resuspending the yeast cells in SG/R-CAA medium with penicillin-streptomycin and incubating them with shaking overnight at 37 °C.

For MACS, 10⁹ induced yeast cells were pelleted for 5 min at 1000g and washed by resuspending the pellet in 50 ml wash buffer (PBS, pH 7.2, 0.25% BSA, 2 mM EDTA) and pelleted again for 5 min at 2500 g. Washed cells were resuspended in 5 ml wash buffer containing 0.5 µM biotinylated antigen and incubated for 30 min at room temperature with gentle agitation. Antigen binding was quenched with the addition of 10 ml ice-cold wash buffer and the cells and antigen solution was pelleted for 5 min at 1000g and 4 °C. The cells were resuspended in 5 ml wash buffer plus 25 µl streptavidin microbeads (µMACS streptavidin kit, MACS Miltenyi) and incubated on ice for 10 min. Finally, 15 ml wash buffer were added and the cells were passed through a 40 µl-cell strainer.

The LS column (MACS Miltenyi) was placed in the separator and 3 ml of wash buffer were applied to precondition the column. Then, the cell solution was loaded in 7 ml batches. Once the column stopped dripping, it was briefly removed from the magnet and placed back in the magnet to reorient the beads in the column to allow trapped cells to flow through. Before loading the next 7 ml cell suspension, 1 ml of wash buffer was applied to the column. These steps were repeated until all cells were loaded. Then, the column was washed with 3 ml wash buffer, briefly removed from the magnet and washed again with 3 ml wash buffer. To elute the binding cells, the column was removed from the magnet and 5 ml wash buffer were added. Using the supplied plunger, the cells were pushed through the column into a new tube. The binding cell fraction was pelleted for 10 min at 2500g. The pellet was re-suspended in 10 ml SD-CAA medium and 10 μ l of eluted cells were diluted in 990 μ l SD-CAA and 100 μ l were plated onto an MDL plate and incubated at 30°C for three days to estimate the output of MACS procedure, and the rest of the cells were incubated at 30°C and 180 rpm overnight.

In the FACS selection rounds that followed, the induced cell suspensions were diluted to 10^8 cells per 1 ml 10% BSA-PBS and centrifuged at 3000 rpm for 5 min at 20°C. To block the cells, the pellets were resuspended in 1 ml 10% BSA-PBS and incubated for 30 min at 20°C on a rotating platform. The cells were centrifuged at 3000 rpm for 5 min at 20°C and resuspended in 250 μ l 10% BSA-PBS containing 0.5 μ M biotinylated EGFR-Fc. After incubation for 1 h at 20°C on a rotating platform the cells were centrifuged at 3000 rpm for 5 min at 4 °C. To wash the cells, the pellets were resuspended in 1 ml ice-cold PBS and centrifuged at 3000 rpm for 5 min at 4°C. Then, the cells were resuspended in 250 μ l 10% BSA-PBS with streptavidin-Alexa Fluor 647 (1:800) and anti-V5-FITC antibody (1:100) and incubated for 30 min on ice. The cells were centrifuged at 3000 rpm for 5 min at 4 °C, resuspended in 250 μ l ice-cold PBS and kept on ice until sorting with Sony cell sorter SH8000. At least 20x output of the previous sorting round was processed and 0.1% top anti-V5-antibody-positive yeast cells were collected. After visible enrichment, the sorts were plated out to characterize single yeast display clones. For some enriched sorts, an additional selection round using 100 nM antigen concentration was performed.

For screening, induced yeast cells were blocked with 2% BSA-PBS at OD₆₀₀ of 1, for 30 min at RT. Then they were incubated with a 2-fold serial dilution of biotinylated human EGFR-Fc, starting at 100 nM, in 2% BSA-PBS, for 30 min at RT. Binding was detected with streptavidin-Alexa Fluor 647 at 1:1000 dilution in 2% BSA-PBS, for 30 min

on ice. Finally, cells were resuspended in 200 μ l ice-cold PBS and fluorescence of the sample was recorded on a Guava EasyCyte flow cytometer. Percent of antigen-binding cells was evaluated (Table 10) and an EGFR-specific binder L2B_EU1_1 was identified with the amino acid sequence with SEQ ID 103. Residues different from the parental clone are in bold print.

Concentration (nM)	EGFR-Fc	% of antigen-binding cells
100		12.34
50		8.17
25		2.95
12.5		1.65
0		1.13

Table 10.

SEQ ID NO:103:

FVNKDQIAKDVKQFYDQAL**WWCCVLYK**ANNACAVVKTFHETLDCCGSST**GRSHTA**CVL**KGN**LCP
SGSNIIISNLFKEDCHQKIDDLFSGK

16.2 Mammalian display system for confirmation of antigen binding for the EGFR-specific clone

The sequences of wild-type CD81 LEL and EGFR-specific clone L2BEU1_1 were cloned between the *Sfi*I and *Sal*I restriction sites of the pDisplay vector (Thermo Scientific). This display system allows the expression of C-terminal anchored protein of interest between N-terminal HA-tag and C-terminal c-myc-tag. Constructs were transfected into HEK293-6E cells (CNRC). Cells were harvested after 48 or 72h, blocked for 30 min in 4% BSA-PBS on ice and stained with antibodies against CD81 (M38, Thermo Scientific) at 10 μ g/ml in 4% BSA-PBS and anti-c-myc (A-14, sc789, Santa Cruz) at 10 μ g/ml in 4% BSA- PBS, for 30 min on ice. Their binding was detected after incubation with secondary reagents anti-mouse F(ab)2-Alexa Fluor 555 (Thermo Scientific), diluted 1:1000 in 4% BSA-PBS, and anti-rabbit (H+L) antibody conjugated with Alexa Fluor 488 (Thermo Scientific), diluted 1:100 in 4% BSA-PBS, for 30 min on ice. Cells were then resuspended in PBS and kept on ice until analysis with Guava EasyCyte flow cytometer (Merck Millipore). Antigen reactivity was determined after incubation with biotinylated human EGFR-Fc at 300 nM and detection with streptavidin-Alexa Fluor 647 at 1:1000 in 4% BSA-PBS. Both wild-type CD81 LEL and EGFR-binding

mutant showed a good level of display on the mammalian cell surface judging from the reactivity with the anti-c-myc antibody (Table 11). Wild-type CD81 LEL reacted well with the anti-CD81 antibody while the mutant did not display any reactivity, presumably due to the modifications of the relevant epitope due to library mutagenesis (Table 11). Antigen binding in mammalian cell-display format could be confirmed for the antigen-binding CD81 LEL mutant (Table 11).

staining		MFI		
1 st step	2 nd step	construct	wild-type CD81 LEL	L2B_EU1_1
anti-c-myc antibody	anti-rabbit(H+L)-Alexa Fluor 488		69.60	55.57
unstained	anti-rabbit (H+L)-Alexa Fluor 488		13.50	13.57
M38	anti-mouse F(ab) ₂ -Alexa Fluor 555		83.81	50.55
unstained	anti-mouse F(ab) ₂ -Alexa Fluor 555		8.52	8.58
EGFR-Fc-biotin	streptavidin-Alexa Fluor 647		156.41	1825.91
unstained	streptavidin-Alexa Fluor 647		156.36	167.80

Table 11.

16.3 Characterization of specificity, species cross-reactivity and epitope mapping of EGFR binder

HEK293-6E cells have been transfected with pDisplay construct encoding wild-type CD81 LEL and anti-EGFR mutant CD81 LEL L2B_EU1_1. 1×10^5 cells have been blocked in 2% BSA-PBS for 30 min on ice and then stained with each 500 nM biotinylated human EGFR-Fc, biotinylated human Her2/neu-Fc and biotinylated mouse EGFR-Fc for 30 min on ice in 2% BSA-PBS. After the centrifugation at 300g, 5 min at 4°C, binding of the antigens has been detected with streptavidin-Alexa Fluor 647 (Thermo Scientific), diluted 1:1000 for 30 min on ice. Cells were then centrifuged at 300g, 5 min at 4°C and resuspended in 200 µl ice-cold PBS. The display was measured by staining the induced cultures with anti-c-myc antibody (A-14, sc789, Santa Cruz) at 10 µg/ml in 2% BSA-PBS and anti-rabbit (H+L) antibody conjugated with Alexa Fluor 488 (Thermo Scientific), diluted 1:1000 in 2% BSA-PBS, for 30 min on ice. The fluorescence has been determined with Guava EasyCyte flow cytometer. The anti-EGFR clone has shown binding only to its cognate antigen, but not to human Fc and Her2/neu Fc protein (Table 12). The EGFR-reactive clone was shown to be cross-reactive with mouse EGFR (Table 12).

staining		construct		
1 st step	2 nd step		wild-type CD81 LEL	L2BEU1_1
		MFI		
anti-c-myc antibody	anti-rabbit(H+L)-Alexa Fluor 488		176.45	117.88

unstained	anti-rabbit (H+L)-Alexa Fluor 488		16.10	17.09
		% antigen binding cells		
Human EGFR-Fc-biotin	streptavidin-Alexa Fluor 647		10.36	71.11
Human Her2/neu-Fc-biotin	streptavidin-Alexa Fluor 647		10.49	8.17
Mouse EGFR-Fc-biotin	streptavidin-Alexa Fluor 647		12.80	89.34
unstained	streptavidin-Alexa Fluor 647		14.11	5.82

Table 12.

In the epitope mapping experiment, 300 nM antigen solution was incubated with 3-fold molar excess of the validated anti-EGFR antibodies cetuximab (Li et al., Cancer Cell 7 (4), 301-311 (2005). DOI: 10.1016/j.ccr.2005.03.003) and matuzumab (Schmiedel et al., Cancer Cell 13 (4), 365-373 (2008). doi: 10.1016/j.ccr.2008.02.019.2005) or commercially available human IgG-kappa isotype antibody (Sigma-Aldrich) and then used for staining of the mutant CD81 LEL- displaying cells. The binding of validated antibodies to the antigen proceeded for 30 minutes in 2% BSA-PBS at room temperature and the solution was then used to stain CD81 LEL mutant-displaying cells. Detection proceeded with streptavidin-Alexa Fluor 647, diluted 1:1000 in 2% BSA-PBS for 30 min on ice. MFI values of displaying cells were recorded. The data indicates that the binding site of the L2B_EU1_1 clone overlaps with the binding site for the cetuximab antibody (Table 13). L2B_EU1_1 could still bind to the antigen in the presence of the excess of matuzumab antibody.

staining		MFI
1 st step	2 nd step	
Human EGFR-Fc-biotin	streptavidin-Alexa Fluor 647	5332
Human EGFR-Fc-biotin + human IgG-kappa isotype	streptavidin-Alexa Fluor 647	5546
Human EGFR-Fc-biotin + cetuximab	streptavidin-Alexa Fluor 647	1324
Human EGFR-Fc-biotin + matuzumab	streptavidin-Alexa Fluor 647	2141
unstained	streptavidin-Alexa Fluor 647	1080

Table 13.

16.4 Rerandomization of modified loops of CD81 LEL-based EGFR-binder

The aim of this experiment was to re-randomize the mutated stretches of residues in the EGFR-binding mutant and show that the binding of the clone is dependent on amino acid residues on both stretches of the polypeptide chain that were randomized in the parental clone to introduce an antigen binding site.

PCR fragments for recombination were produced by using once the mutagenic primer LIB2FWD (SEQ ID NO:98) and that randomizes the mutated residues in helix A and the AB-loop in Library L2BE1_A, and once the mutagenic primer LIB2REV2 (SEQ

ID NO:99) that randomized the mutated residues in helix C in Library L2BE1_B. PCR fragments were produced using Q5 High-Fidelity Polymerase (New England Biolabs). For recombination fragment for library L2BE1_A primer LIB2FWD was used together with primer AP2 (SEQ ID NO:104) and for library L2BE1_B primer LIB2REV2 (SEQ ID NO:99) was used together with primer AP1 (SEQ ID NO:105).

SEQ ID NO:104: CTTGAAGAGGTTGCTGAT

SEQ ID NO:105: AAGGATGTGAAGCAGTTC

Each recombination fragment was transformed together with *Bam*HI/*Hind*III linearized recipient vector pYD1 with cloned CD81 LEL_dellbamdelhind_bamhind sequence using chemical transformation into *S. cerevisiae* EBY100. 35 µg of linearized vector and 25 µg of the PCR fragment were used for library construction. The size of the libraries were 8.4×10^6 independent members for library L2BE1_A and 1.18×10^7 independent members for Library L2BE1_B. Quality control of the libraries included induction of library members as a stress temperature and subsequent measurement of yeast-surface displayed protein via the N-terminal tag, detected with an anti-Xpress tag antibody, followed by incubation with a goat anti-mouse (Fab')₂-FITC (Sigma Aldrich) and the determination of the correct reading frame of the displayed proteins via detection of the C-terminal V5-tag with an anti-V5-FITC antibody (Thermo Scientific). The percentage of antigen-binding cells for each library was determined after staining with human EGFR-Fc at 500 nM and the detection with goat anti-human gamma chain – PE conjugate (Sigma Aldrich). The percentages of positive cells for the libraries are presented along with the values characteristic for the parental clone in Table 14. Re-randomization of the mutated residues in both targeted regions caused a reduction in the number of antigen-positive clones, implicating that amino acid residues in both randomized stretches contribute to antigen binding.

	Sample	L2BE1_1	L2BE1_A	L2BE1_B
staining	Percentage of positive cells (%)			
1 st step	2 nd step			
Human EGFR-Fc	Goat anti-human gamma -PE	61.77	15.77	15.55
unstained	Goat anti-human gamma -PE	6.16	6.70	6.91
Anti-Xpress	goat anti-mouse (Fab') ₂ -FITC	75.17	65.84	60.67
unstained	goat anti-mouse (Fab') ₂ -FITC	1.53	0.84	0.48
unstained	anti-V5-tag-FITC	83.64	73.37	62.73
unstained	unstained	2.30	1.10	1.20

Table 14.

Example 17: CD81 LEL-based binders to human placental laminin**17.1 Selections of CD81 LEL libraries L2 and L3 with human laminin**

Human laminin was purchased from Sigma-Aldrich. For biotinylation, EZ-Link™ Sulfo-NHS-LC-LC-Biotin reagent (Thermo Scientific) was used in a 3:1 molar ratio, after dialysis of the antigen against 100-fold volume of PBS overnight at 4 °C. Incubation with biotinylation reagent proceeded for 1 h at room temperature with shaking. Unbound biotin was removed with dialysis against 100-fold volume of PBS, using Snakeskin dialysis tubing (Thermo Scientific) with 10.000 Da MWCO, at 4 °C overnight with stirring.

For sorting, libraries 2A, 2B, 3A and 3B were first cultured in SD-CAA medium supplemented with penicillin-streptomycin at 30 °C with shaking overnight and then the expression of recombinant protein was induced by resuspending the yeast cells in SG/R-CAA medium with penicillin-streptomycin and incubating them with shaking overnight at 37 °C.

For MACS, 10⁹ induced yeast cells were pelleted for 5 min at 1000g and washed by resuspending the pellet in 50 ml wash buffer (PBS, pH 7.2, 0.25% BSA, 2 mM EDTA) and pelleted again for 5 min at 1000 g. Washed cells were resuspended in 5 ml wash buffer containing 1 µM biotinylated antigen and incubated for 30 min at room temperature with gentle agitation. Antigen binding was quenched with the addition of 10 ml ice-cold wash buffer and the cells and antigen solution was pelleted for 5 min at 1000g and 4°C. The cells were resuspended in 5 ml wash buffer plus 25 µl streptavidin microbeads (µMACS streptavidin kit, MACS Miltenyi) and incubated on ice for 10 min. Finally, 15 ml wash buffer were added and cell suspension was filtered through a 40-µm strainer.

The LS column (MACS Miltenyi) was placed in the separator and 3 ml of wash buffer were applied to precondition the column. Then, the cell solution was loaded in 7 ml batches. Once the column stopped dripping it was briefly removed from the magnet and placed back in the magnet to reorient the beads in the column to allow trapped cells to flow through. Before loading the next 7 ml cell suspension, 1 ml of wash buffer was applied to the column. These steps were repeated until all cells were loaded. Then, the column was washed with 3 ml wash buffer, briefly removed from the magnet and washed again with 3 ml wash buffer. To elute the binding cells, the column was removed from the magnet and 5 ml wash buffer were added. Using the supplied plunger, the cells were pushed through the column into a new tube. The binding cell fraction was pelleted for 10 min at 2500g. The pellet was resuspended in 10 ml SD-CAA medium and 10 µl of

eluted cells were diluted in 990 μ l SD-CAA and 100 μ l were plated onto an MDL plate and incubated at 30 °C for three days to estimate the output of MACS procedure, and the rest of the cells were incubated at 30 °C and 180 rpm overnight.

In the FACS selection rounds that followed, the induced cell suspensions were diluted to 10^8 cells per 1 ml 10% BSA-PBS and centrifuged at 3000 rpm for 5 min at 20 °C. To block the cells, the pellets were resuspended in 1 ml 10% BSA-PBS and incubated for 30 min at 20 °C on a rotating platform. The cells were centrifuged at 3000 rpm for 5 min at 20 °C and resuspended in 250 μ l 10% BSA-PBS containing 1 μ M biotinylated laminin. After incubation for 1 h at 20 °C on a rotating platform the cells were centrifuged at 3000 rpm for 5 min at 4 °C. To wash the cells, the pellets were resuspended in 1 ml ice-cold PBS and centrifuged at 3000 rpm for 5 min at 4 °C. Then, the cells were resuspended in 250 μ l 10% BSA-PBS with streptavidin-Alexa Fluor 647 (1:800) and anti-V5-FITC antibody (1:100) and incubated for 30 min on ice. The cells were centrifuged at 3000 rpm for 5 min at 4 °C, resuspended in 250 μ l ice-cold PBS and kept on ice until sorting with Sony cell sorter. At least 20x output of the previous sorting round was processed and 0.1% top anti-V5-antibody positive yeast cells were collected. After visible enrichment, the sorts were plated out to characterize single yeast display clones. For some enriched sorts, an additional selection round using 100 nM antigen was performed.

17.2 Expression in mammalian cells

Unique binder sequences amplified using oligonucleotide sequences CD81hnhe1 (SEQ ID NO:106) and LELp28_bste2 (SEQ ID NO:107) and cloned between the *NheI* and *BstEII* sites of the mammalian expression vector pTT28 and the sequences of the constructs were verified using Sanger sequencing.

SEQ ID NO:106: ACGTGCTAGCTTTGTCAACAAGGACCAGATC

SEQ ID NO:107: ACGTGGTGACCGGTGCTGGAACCCCTCCCGGAGAAGAGGTC

Recombinant proteins were expressed in HEK293-6E exactly after manufacturer's instructions. The supernatant of 25-ml-cultures was harvested with centrifugation at 3500 rpm, 15 min at 4 °C, buffered with PBS and 20 mM imidazole and filtered through an 0.45-µm-filter before loading to a 1-ml-His Excel column (GE Healthcare), equilibrated with PBS/20 mM imidazole, pH 7.5. Column was then washed with the same buffer and his-tagged protein was eluted with a gradient from 20-500 mM imidazole in PBS, pH 7.5, in 5 column volumes. Fractions 4-6 were analyzed for the presence of the eluted protein with SDS-PAGE followed with Coomassie staining. Expression could be confirmed for following mutants: L2A_LU1 (SEQ ID 108), L2A_LU1_1 (SEQ ID 109), L2B_LU1 (SEQ ID 110), L3B_LU1_2 (SEQ ID 111). Residues different from the parental clone are in bold print.

SEQ ID NO:108: L2A_LU1:

FVNKDQIAKDVKQFYDQAL**RECCNTFD**ANNACAVVKTFFHETLDCCGSST**THSSTG**CVL**FF**NLCP
SGSNIIISNLFKEDCHQKIDDLFSGK

SEQ ID NO:109: L2A_LU1_1:

FVNKDQIAKDVKQFYDQAL**MICCH**RHYANNACAVVKTFFHETLDCCGSST**PSSTT**PCVLQ
NNLCPSGSNIISNLFKEDCHQKIDDLFSGK

SEQ ID NO:110: L2B_LU1:

FVNKDQIAKDVKQFYDQAL**VFCC****FGGH**ANNACAVVKTFFHETLDCCGSST**HPYKTK**CVLQ
RNLCPSGSNIISNLFKEDCHQKIDDLFSGK

SEQ ID NO:111: L3B_LU1_2:

FVNKDQIAKDVKQFYDQCL**KYAC****KQGDW**ENAKACVKTFFHETLDCCGSST**VANMTR**CVLP
ANLCPSGSNIISNLFKEDCHQKIDDLFSGK

17.3 Antigen binding in soluble format

Neat eluted proteins were tested for binding to their cognate antigen and BSA as a control antigen. Streptavidin-activated plates (Immobilizer, NUNC) were coated with biotinylated laminin at 5 µg/ml in PBS for 30 min at RT and blocked with 4% BSA-PBS for 1h at RT. Eluted proteins were added in 2% BSA-PBS and allowed to bind for 1h at RT. After three wash steps with PBS, binding was detected with a 1:3000 dilution of anti-pentaHis antibody-HRP conjugate (QIAGEN) in 2% BSA-PBS for 30 min and revealed upon addition of TMB (Sigma-Aldrich). The reaction was stopped with the addition of H₂SO₄ and the absorbance read at 450/620 nm (Table 15). Specific reactivity with the target protein could be confirmed for 3 laminin-specific clones: L2A_LU1, L2B_LU1 and L3B_LU1_2.

Clone	Antigen	A _{450/620} (average)	St.dev.
L2A_LU1	laminin	0.156	0.00745937
L2A_LU1	BSA	0.045	0.00511881
L3A_LU1	laminin	1.702	0.27844276
L3A_LU1	BSA	0.016	0.00171464
L3B_LU1_2	laminin	1.316	0.16702341
L3B_LU1_2	BSA	0.021	0.00283314

Table 15.

17.4 Expression of laminin binders in mammalian cell display system

The sequences of laminin binding clones were cloned between the *Sfi*I and *Sa*I restriction sites of the pDisplay vector (Thermo Scientific). This display system allows the expression of C-terminal anchored protein of interest between N-terminal HA-tag and C-terminal c-myc-tag. Constructs were transfected into HEK293-6E cells (CNRC). Cells were harvested after 48 or 72h, blocked for 30 min in 4% BSA-PBS on ice and stained with an anti-c-myc antibody (A-14, sc789, Santa Cruz) at 10 µg/ml in 4% BSA-PBS, for 30 min on ice. Their binding was detected after incubation with secondary anti-rabbit (H+L) antibody conjugated with Alexa Fluor 488 (Thermo Scientific A11034), diluted 1:1000 in 4% BSA-PBS, for 30 min on ice. Antigen reactivity was determined after incubation with biotinylated human laminin at 500 nM and detection with streptavidin-Alexa Fluor 647 at 1:1000 in 4% BSA-PBS. Cells were then resuspended in PBS and kept on ice until analysis with Guava EasyCyte flow cytometer (Merck Millipore). For all tested clones, display on mammalian cell surface could be detected

(Table 16). For nine clones, reactivity with the recombinant antigen could be established (Table 16).

clone	Staining	MFI		% antigen-positive cells	
	1 st step	Anti-c-myc	unstained	laminin-biotin	unstained
	2 nd step	anti-rabbit (H+L)-Alexa Fluor 488	anti-rabbit (H+L)-Alexa Fluor 488	streptavidin-Alexa Fluor 647	streptavidin-Alexa Fluor 647
L2A_LU1		37.68	16.77	28.98	5.92
L3B_LU1_2		71.30	17.90	29.35	2.75
L3A_LU1*		74.45	14.04	28.60	8.08
L2A_LU1_1		74.79	15.19	37.93	5.57
L2B_LU1_1		103.85	10.96	20.08	1.77
81_L1		41.04	15.68	54.17	6.09
81_L13		46.47	13.42	30.78	4.17
81_L17		80.52	12.05	28.11	4.11
81_L21		55.83	12.48	37.14	3.28

Table 16.

*For L3A_LU1 clone, neutravidin-PE (Thermo Scientific) was used as a secondary reagent.

The sequences of the discovered clones were determined for: L3A_LU1 (SEQ ID NO:112), L2B_LU1_1 (SEQ ID NO:113), 81_L1 (SEQ ID NO:110, same as L2B_LU1), 81L_13 (SEQ ID NO:114), 81L_17 (SEQ ID NO:115), 81L_21 (SEQ ID NO:116). Residues different from the parental clone are in bold print.

SEQ ID NO:112:

FVNKDQIAKDVKQFYDQCL**LWACSKRYKY**NAKACVKTFHETLDCCGSST**MGNLTECVLI**
ENLCPSGSNIIISNLFKEDCHQKIDDLFSGK

SEQ ID NO:113:

FVNKDQIAKDVKQFYDQAL**WKCCNLSQ**ANNACAVVKTFHETLDCCGSST**YLCATYCVLW**
PNLCPSGSNIIISNLFKEDCHQKIDDLFSGK

81L_1 (SEQ ID NO:110):

FVNKDQIAKDVKQFYDQAL**VFCCFGGH**ANNACAVVKTFHETLDCCGSST**HPYKTKCVLQ**
RNLCPSGSNIIISNLFKEDCHQKIDDLFSGK

SEQ ID NO:114:

FVNKDQIAKDVKQFYDQCL**SWACNRKYEP**NAKACVKTFHETLDCCGSST**KANYTHCVLE**
MNLCPSGSNIIISNLFKEDCHQKIDDLFSGK

SEQ ID NO:115:

FVNKDQIAKDVKQFYDQCL**RHAC****SGLSGR**NAKACVKTFHETLDCCGSST**IKHKTLCVLI**
KNLCPSGSNIIISNLFKEDCHQKIDDLFSGK

SEQ ID NO:116:

FVNKDQIAKDVKQFYDQAL**FICCFRRY**ANNACAVVKTFHETLDCCGSST**RNNETACVLA**
KNLCPSGSNIIISNLFKEDCHQKIDDLFSGK

Example 18 Anti-laminin CD81 mutants expressed on the surface of EVs

18.1 EV preparation

HeLa cells were transduced with wild-type CD81 or CD81 with modified LEL corresponding to the sequence of L2A_LU1 (SEQ ID NO:12) or L3B_LU1_2 (SEQ ID NO:15), cloned into pBMN expression vector in frame with eGFP. Extracellular vesicles (EVs) were prepared from transduced HeLa cells cultured to a confluence of 80% in RPMI (10% FCS, 4 mM L-Glutamine). Afterwards medium was changed and EV collection was performed for 72 - 96 hours in OptiPRO SFM. Cell culture supernatants were centrifuged at 3 000g for 30 min in order to remove cells and cell debris. Larger particles were excluded by filtering supernatants through a 0.45 µm cellulose acetate filter. Ultimately, EVs were pelleted by ultracentrifugation at 120000g for 90 min and resuspended in Live Cell imaging solution (Thermo Scientific). Recombinant EVs, harboring CD81-eGFP fusion, were verified using human CD9 capture beads for flow cytometry based detection (ImmunoStep).

18.2 Competition assay showing specific antigen binding of anti-laminin CD81 mutants expressed on the surface of EVs

5x 10⁹ recombinant EVs harboring wild-type CD81-eGFP and respective Laminin targeting variants L2A_LU1 and L3B_LU1_2 were incubated overnight with 6x 10³ CD9+ capture beads. 5 µg/mL of biotinylated human placental laminin were added to the beads, and in some parallel samples 15 µg/mL of unlabeled laminin for competition. Neutravidin-PE (1:800) was used to detect EV-bound laminin. Background (BG) fluorescence was determined by staining with neutravidin-PE only. Fluorescence of eGFP was measured at 488 nm and PE-fluorescence at 561 nm. For both laminin-binding CD81 mutants, specific binding to the target antigen was shown, while this was not observed for the wild-type CD81. In addition, outcompeting of labelled by unlabeled laminin resulted in reduction of signals indicating specific binding (Table 17).

	construct	Mean 488	Mean 561	Mean 561 - BG	ratio 488/561
without competition	Wild-type-eGFP-EVs	96203.4	20191.9	14762.2	0.209888
	L2A_LU1-eGFP-EVs	150790.6	31526.3	26096.6	0.209073
	L3B_LU1-2-eGFP-EVs	187387.7	36420.4	30990.7	0.194359
with competition of 3-fold molar excess of unlabelled laminin	Wild-type-eGFP-EVs	96391	23944.7	18665.1	0.248412
	L2A_LU1-eGFP-EVs	126038.4	23397.8	18118.2	0.18564
	L3B_LU1_2-eGFP-EVs	139326.7	16024.8	10745.2	0.115016

Table 17.

18.3 The uptake of laminin-targeting EVs into model hepatocarcinoma cell line Huh7

2.4x, 1.6x and 0.8x 10¹⁰ recombinant EVs harboring CD81-eGFP and its targeting variants were incubated with 0.4x 10⁶ Huh7 cells for 3 hours. Cells were trypsinized, neutralized and resuspended in 100 µl of PBS. Flow cytometry was used to measure uptake levels of cells with incorporated eGFP-positive EVs and the MFI values of the main population (gated through FSC/SSC) were determined. In Table 18, biological triplicates of results obtained of a single batch of EVs, normalized to MFI_{max} of CD81 wild-type-eGFP, are presented showing a 2 to 3-fold increased uptake of the laminin targeting EVs.

	construct	Mean 488	Mean 561	Mean 561 - BG	ratio 488/561
without competition	Wild-type-eGFP-EVs	96203.4	20191.9	14762.2	0.209888
	L2A_LU1-eGFP-EVs	150790.6	31526.3	26096.6	0.209073
	L3B_LU1-2-eGFP-EVs	187387.7	36420.4	30990.7	0.194359
with competition of 3-fold molar excess of unlabelled laminin	Wild-type-eGFP-EVs	96391	23944.7	18665.1	0.248412
	L2A_LU1-eGFP-EVs	126038.4	23397.8	18118.2	0.18564
	L3B_LU1_2-eGFP-EVs	139326.7	16024.8	10745.2	0.115016

Table 18.

In another experiment, replicates of uptake of different batches of laminin-targeting EVs into hepatocarcinoma cell line Huh7 were performed. 3 independent EV batches were tested in duplicates and normalized to MFI_{max}. 2.4x 10¹⁰ recombinant EVs harboring CD81-eGFP or respective targeting variants were incubated with 0.4x 10⁶ Huh7 cells for 3 hours. Cells were trypsinized, neutralized and resuspended in 120 µl of PBS. Flow cytometry was used to measure uptake levels of cells with incorporated

eGFP-positive EVs and the MFI values of the main population (gated through FSC/SSC) were determined. The values were normalized to MFI max. Means and standard deviations of duplicates are presented in Table 19. The EVs with overexpressed laminin-binding CD81 have shown a higher uptake into the Huh7-cells.

EV batch	construct	Mean % MFI max	St.dev. % MFI max
Batch I	CD81 wild-type-eGFP	29.0997817	0.38828317
	L2A_LU1-eGFP	95.0826781	4.91732192
	L3B_LU1-2-eGFP	75.5668285	1.66717452
Batch II	CD81 wild-type-eGFP	40.8614463	0.03708697
	L2A_LU1-eGFP	96.583803	3.41619699
	L3B_LU1-2-eGFP	71.1655595	1.82037455
Batch III	CD81 wild-type-eGFP	57.5887719	0.6223537
	L2A_LU1-eGFP	99.3962027	0.60379729
	L3B_LU1-2-eGFP	93.5651161	2.95038121

Table 19.

Example 19 CD81 LEL-based binders to human Her2/neu

19.1 Selections of CD81 LEL libraries L2A and L2B with human Her2/neu

Human Her2/neu-Fc was purchased from SinoBiological. For biotinylation, EZ-Link™ Sulfo-NHS-LC-LC-Biotin reagent (Thermo Scientific) was used in a 3:1 molar ratio of biotin to protein. Incubation with biotinylation reagent proceeded for 1 h at room temperature with shaking. Unbound biotin was removed with dialysis against 100x volume of PBS, using Snakeskin dialysis tubing (Thermo Scientific) with 10.000 Da MWCO, at 4°C overnight with stirring and aliquots of labelled antigen were stored at -80°C until further use.

For sorting, libraries 2A and 2B were first cultured in SD-CAA medium supplemented with penicillin-streptomycin at 30°C with shaking overnight and then the expression of recombinant protein was induced by resuspending the yeast cells in SG/R-CAA medium with penicillin-streptomycin and incubating them with shaking overnight at 37°C.

For MACS, 4×10^9 induced yeast cells were pelleted for 5 min at 1000g and blocked in 10% BSA-PBS, on a rotating wheel for 30 min at room temperature. Cells were pelleted and resuspended in 10% BSA-PBS with 0.5 μ M biotinylated antigen and incubated for 30 min at room temperature on a rotating wheel. Antigen binding was quenched with the addition of 10-fold volume of ice-cold PBS and the cells were pelleted

for 5 min at 1000g and 4°C. The cells were resuspended in 5 ml wash buffer plus 200 µl streptavidin microbeads (MACS Miltenyi) and incubated on ice for 15 min. Finally, 15 ml MACS wash buffer (0.5% BSA, 2 mM EDTA in PBS; pH 7.4) were added. Cell suspension was strained through a 40-µm cell strainer.

The LS column (MACS Miltenyi) was placed in the separator and 3 ml of wash buffer were applied to precondition the column. Then, the cell solution was loaded in 7 ml batches. Once the column stopped dripping it was briefly removed from the magnet and placed back in the magnet to reorient the beads in the column to allow trapped cells to flow through. Before loading the next 7 ml cell suspension, 1 ml of wash buffer was applied to the column. These steps were repeated until all cells were loaded. Then, the column was washed three times with 5 ml wash buffer. To elute the binding cells, the column was removed from the magnet and 5 ml wash buffer were added. Using the supplied plunger, the cells were pushed through the column into a new tube. The binding cell fraction was pelleted for 10 min at 2500g. The pellet was re-suspended in 10 ml SD-CAA medium and 10 µl of eluted cells were diluted in 990 µl SD-CAA and 100 µl were plated onto an MDL plate and incubated at 30°C for three days to estimate the output of MACS procedure, and the rest of the cells were incubated at 30°C and 180 rpm overnight.

In the FACS selection rounds that followed, the induced cell suspensions were diluted to 10^8 cells per 1 ml 10% BSA-PBS and centrifuged at 3000 rpm for 5 min at 20°C. To block the cells, the pellets were resuspended in 1 ml 10% BSA-PBS and incubated for 30 min at 20°C on a rotating platform. The cells were centrifuged at 3000 rpm for 5 min at 20°C and resuspended in 150 µl 10% BSA-PBS containing 0.5 µM biotinylated Her2/neu-Fc. After incubation for 1 h at room temperature on a rotating platform, antigen binding was quenched by adding 1 ml ice-cold PBS. The cells were centrifuged at 3000 rpm for 5 min at 4°C and resuspended in 800 µl 10% BSA-PBS with streptavidin-Alexa Fluor 647 (Thermo Fisher Scientific) (1:800) and anti-V5-FITC antibody (Thermo Scientific) (1:100) and incubated for 30 min on ice. The cells were centrifuged at 3000 rpm for 5 min at 4°C, resuspended in 250 µl ice-cold PBS and kept on ice until sorting with Sony SH8000 sorting apparatus. At least 20x output of the previous sorting round was processed and 0.1% top anti-V5-antibody positive yeast cells were collected. Enriched pools were subjected to five sorting rounds and single yeast clones were plated out for screening. 5 identified sequences were cloned to pDisplay expression vector, expressed in HEK293-6E cells and tested for binding to human

Her2/neu. Cells were harvested after 48 or 72h, blocked for 30 min in 4% BSA-PBS on ice and stained with an anti-c-myc antibody (A-14, sc789, Santa Cruz) at 10 µg/ml in 4% BSA- PBS, for 30 min on ice. Their binding was detected after incubation with secondary anti-rabbit (H+L) antibody conjugated with Alexa Fluor 488 (Thermo Scientific A11034), diluted 1:1000 in 4% BSA-PBS, for 30 min on ice. Antigen reactivity was determined after incubation with biotinylated human Her2/neu/Fc in 2-fold dilution series starting from 300 nM and detection with streptavidin-Alexa Fluor 647 at 1:1000 in 4% BSA-PBS. Cells were then resuspended in PBS and kept on ice until analysis with Guava EasyCyte flow cytometer (Merck Millipore).

Clone 81_H2_11 (SEQ ID NO:117) could specifically bind to human Her2/neu protein (Table 20). Residues different from the parental clone are in bold print.

SEQ ID NO:117:

FVNKDQIAKDVKQFYDQAL**VNCCYTRA**ANNACAVVKTFHETLDCCGSST**HYVDTH**CVL**NRN**LCP
SGSNIIISNLFKEDCHQKIDDLFSGK

EV batch	construct	Mean % MFI max	St.dev. % MFI max
Batch I	CD81 wild-type-eGFP	29.0997817	0.38828317
	L2A_LU1-eGFP	95.0826781	4.91732192
	L3B_LU1-2-eGFP	75.5668285	1.66717452
Batch II	CD81 wild-type-eGFP	40.8614463	0.03708697
	L2A_LU1-eGFP	96.583803	3.41619699
	L3B_LU1-2-eGFP	71.1655595	1.82037455
Batch III	CD81 wild-type-eGFP	57.5887719	0.6223537
	L2A_LU1-eGFP	99.3962027	0.60379729
	L3B_LU1-2-eGFP	93.5651161	2.95038121

Table 20.

19.2 Characterization of specificity of the CD81 LEL-based Her2/neu binder

HEK293-6E cells have been transfected with pDisplay construct encoding wild-type CD81 LEL and anti-Her2/neu directed CD81 LEL 81H2-11. 1×10^5 cells have been blocked in 2% BSA-PBS for 30 min on ice and then stained with each 500 nM biotinylated human EGFR-Fc, biotinylated human Her2/neu-Fc and biotinylated mouse EGFR-Fc for 30 min on ice in 2% BSA-PBS. After the centrifugation at 300g, 5 min at 4°C, binding of the antigens has been detected with streptavidin-Alexa Fluor 647, diluted 1:1000 for 30 min on ice. Cells were then centrifuged at 300g, 5 min at 4°C and re-suspended in 200 µl ice-cold PBS. The display was measured by staining the induced cultures with anti-c-myc antibody (A-14, sc789, Santa Cruz) at 10 µg/ml in 2% BSA-PBS and anti-rabbit (H+L) antibody conjugated with Alexa Fluor 488 (Thermo Scientific A11034), diluted 1:1000 in 2% BSA-PBS, for 30 min on ice. The fluorescence has been determined with Guava EasyCyte flow cytometer. The anti-Her2/neu clone has shown binding only to its cognate antigen (Table 21).

staining		construct		
1 st step	2 nd step		wild-type CD81 LEL	81_H2_11
		MFI		
anti-c-myc antibody	anti-rabbit(H+L)-Alexa Fluor 488		176.45	75.06
unstained	anti-rabbit (H+L)-Alexa Fluor 488		16.10	13.07
		% antigen binding cells		
Human EGFR-Fc-biotin	streptavidin-Alexa Fluor 647		10.36	5.16
Human Her2/neu-Fc-biotin	streptavidin-Alexa Fluor 647		10.49	44.22
Mouse EGFR-Fc-biotin	streptavidin-Alexa Fluor 647		12.80	5.68
unstained	streptavidin-Alexa Fluor 647		14.11	2.67

Table 21.

Example 20: Stabilization of human CD9 LEL

20.1 Design and construction of stabilized CD9 LEL mutants

The Genbank entries with the full-length human CD9 sequence were identified and BLAST comparison was performed to delineate the borders of the LEL region. Homology modelling of the CD9 LEL region, as defined in Seigneuret, 2006, was performed using Swissmodel (Waterhouse *et al.*, 2018) with CD81 LEL PDB:1iv5 as the closest model proposed. The resulting structure could be aligned with the CD81 LEL crystal structure 1g8q with an RMSD of 0.413 Å. The sequence of CD9 LEL (SEQ ID

NO:118) was cloned using oligonucleotides CD9hnhe1 (SEQ ID NO:119) and CD9p28_bste2 (SEQ ID NO:120) between the *NheI* and *BstEII* cloning sites of the pTT28 vector (CNRC).

SEQ ID NO:118:

HKDEVIKEVQEFYKDTYNKLTCKDEPQRETLKAIHYALNCCGLAGGVEQFISDICKPKD
VLETF TVKSCPDAIKEVFDNK

SEQ ID NO:119: ACGTGCTAGCCACAAGGATGAGGTGATTAAG

SEQ ID NO:120: ACGTGGTGACCGGTGCTGGAACCTTTATTGTCTGAAGACCTC

Mutagenesis reaction was performed to replace the amino acids at positions 20 and 28 for cysteine residues with QuickChange Lightning Mutagenesis Kit (Agilent) according to manufacturer's instructions using oligonucleotides CD9_L20C (SEQ ID NO:121) and CD9_20Ca (SEQ ID NO:122), and CD9_28C (SEQ ID NO:123) and CD9_28Ca (SEQ ID NO:124), to produce a stabilized variant CD9 LEL_20_28 with amino acid sequence as in (SEQ ID NO:125).

SEQ ID NO:121: GACACCTACAACAAGTGCAAAACCAAGGATGAG

SEQ ID NO:122: CTCATCCTTGGTTTTGCACTTGTTGTAGGTGTC

SEQ ID NO:123: AAGGATGAGCCCCAGTGTGAAACGCTGAAAGCC

SEQ ID NO:124: GGCTTTCAGCGTTTCACACTGGGGCTCATCCTT

SEQ ID NO:125:

HKDEVIKEVQEFYKDTYNKCKTKDEPQCETLKAIHYALNCCGLAGGVEQFISDICKPKDVLETF
TVKSCPDAIKEVFDNK

CD9 LEL and CD9 LEL_20_28 were expressed in ExpiCHO system according to MaxTiter protocol exactly according to manufacturer's instructions and purified with one-step Ni-NTA affinity chromatography at 35 and 70 mg/L supernatant, respectively. Purified proteins were analysed with SDS-PAGE and both were monomeric. Thermostability was determined using differential scanning calorimetry (DSC). The melting point of the wild-type protein was determined to be at 52.7 °C and of the stabilized variant 20-28 at 81.9 °C, indicating successful stabilization.

20.2 CD9 LEL-based yeast display library construction for binder selection

20.2.1 Yeast display of wild-type CD9 LEL

The sequence of CD9 LEL was cloned using oligonucleotides CD9YDbam1 (SEQ ID NO:126) and CD9YDnot2 (SEQ ID NO:127) between the *Bam*HI and *Not*I cloning sites of the pYD1 vector and the construct was transformed to *S. cerevisiae* using chemical transformation.

SEQ ID NO:126: ACGTGGATCCCACAAGGATGAGGTGATTAAG

SEQ ID NO:127: ACGTGCGGCCGCGCTTTATTGTCTGAAGACCTC

Transformants were selected on MDL-plates, cultured at 30°C in SD-CAA and induced for 48h at 20°C and for 24 h at 37°C. Subsequent measurement of yeast-surface displayed protein via the N-terminal tag was performed via detection with an 1:2000 dilution of anti-Xpress tag antibody (Thermo Scientific), followed by incubation with a goat anti-mouse (Fab')₂-Alexa Fluor 555 (Thermo Scientific) at 1:1000 and the determination of the correct reading frame of the displayed proteins was via detection of the C-terminal his-tag with an anti-his-Alexa Fluor 488 antibody (QIAGEN) at 1:200 and C-terminal V5-tag with an anti-V5-FITC antibody at 1:100 (Thermo Scientific), all in 2% BSA-PBS. Additionally, reactivity with an anti-CD9 specific antibody MEM-61 (Thermo Scientific) was determined after incubating the yeast cells first with the 10 µg/ml antibody in 2% BSA-PBS and detecting the binding with goat anti-mouse (Fab')₂-Alexa Fluor 555 (Thermo Scientific) at 1:1000 dilution in 2% BSA-PBS. After resuspending in 200 µl ice-cold PBS, the percentage of positive yeast cells was determined with Guava EasyFlow Cytometer (Table 22).

	Sample	wild-type CD9 LEL	wild-type CD9 LEL
	Induction temperature (°C)	20	37
staining			
1 st step	2 nd step		
unstained	unstained	0.38	4.90
Anti-Xpress	goat anti-mouse-Alexa Fluor 555	72.98	67.47
MEM-61	goat anti-mouse-Alexa Fluor 555	46.38	26.21
unstained	goat anti-mouse-Alexa Fluor 555	1.14	4.21
unstained	anti-his-tag-Alexa Fluor 488	83.87	84.65
unstained	anti-V5-tag-FITC	83.11	85.20

Table 22.

20.2.2 Mutagenesis of CD9 LEL for design of binding clones

Residues 18-19, 21-25 and 27 of the CD9 LEL_{20_28} sequence were mutagenized using oligonucleotide CD9PFOREC1 (SEQ ID NO:128) that was combined with oligonucleotide CD9NOTREC2 (SEQ ID NO:129) to produce a PCR recombination fragment.

SEQ ID NO:128:

CACAAGGATGAGGTGATTAAGGAAGTCCAGGAGTTTACAAGGACACCTACNNKNNKTGCNNKN
NKNKNNKNNKCCCNNKTGTGAAACGCTGAAAGCC

wherein N is any one of A, C, G, or T.

SEQ ID NO:129:

CTTCGAAGGGCCCTCTAGACTCGAGCGGCCGCGCTTTATTGTCTGAAGACCTC

This fragment is used together with vector pYD1CD9, linearized with enzymes *PfoI* and *NotI*, to transform *S. cerevisiae* EBY100. The resulting library of 1x10⁸ independent members in size is selected for binders with human EGFR-Fc with one MACS-based selection and several FACS-based selection rounds, until an enrichment of binding clones is achieved. The sequences of binders are recloned to mammalian pDisplay vector and resulting plasmids used to transfect HEK293-6E cells. After 48-72 h, cells are harvested and stained with antigen and a secondary reagent to detect specifically binding clones.

REFERENCES

1. Kalra, H.; Drummen, G.P.; Mathivanan, S. Focus on Extracellular Vesicles: Introducing the Next Small Big Thing. *Int J Mol Sci* **2016**, *17* (2), 170, 10.3390/ijms17020170.
2. Iraci, N.; Leonardi, T.; Gessler, F.; Vega, B.; Pluchino, S. Focus on Extracellular Vesicles: Physiological Role and Signalling Properties of Extracellular Membrane Vesicles. *Int J Mol Sci* **2016**, *17* (2), 171, 10.3390/ijms17020171.
3. Luan, X.; Sansanaphongpricha, K.; Myers, I.; Chen, H.; Yuan, H.; Sun, D. Engineering exosomes as refined biological nanoplateforms for drug delivery. *Acta Pharmacol Sin* **2017**, *38* (6), 754-763, 10.1038/aps.2017.12.
4. Tian, T.; Zhu, Y.L.; Zhou, Y.Y.; Liang, G.F.; Wang, Y.Y.; Hu, F.H.; Xiao, Z.D. Exosome uptake through clathrin-mediated endocytosis and macropinocytosis and mediating miR-21 delivery. *J Biol Chem* **2014**, *289* (32), 22258-67, 10.1074/jbc.M114.588046.

5. El Andaloussi, S.; Lakhali, S.; Mager, I.; Wood, M.J. Exosomes for targeted siRNA delivery across biological barriers. *Adv Drug Deliv Rev* **2013**, *65* (3), 391-7, 10.1016/j.addr.2012.08.008.
6. Rubinstein, E.; Le Naour, F.; Lagaudriere-Gesbert, C.; Billard, M.; Conjeaud, H.; Boucheix, C. CD9, CD63, CD81, and CD82 are components of a surface tetraspan network connected to HLA-DR and VLA integrins. *Eur J Immunol* **1996**, *26* (11), 2657-65, 10.1002/eji.1830261117.
7. Levy, S.; Shoham, T. Protein-protein interactions in the tetraspanin web. *Physiology (Bethesda)* **2005**, *20*, 218-24, 10.1152/physiol.00015.2005.
8. Morelli, A.E.; Larregina, A.T.; Shufesky, W.J.; Sullivan, M.L.; Stolz, D.B.; Papworth, G.D.; Zahorchak, A.F.; Logar, A.J.; Wang, Z.; Watkins, S.C.; Falo, L.D., Jr.; Thomson, A.W. Endocytosis, intracellular sorting, and processing of exosomes by dendritic cells. *Blood* **2004**, *104* (10), 3257-66, 10.1182/blood-2004-03-0824.
9. Svensson, K.J.; Christianson, H.C.; Wittrup, A.; Bourseau-Guilmain, E.; Lindqvist, E.; Svensson, L.M.; Morgelin, M.; Belting, M. Exosome uptake depends on ERK1/2-heat shock protein 27 signaling and lipid Raft-mediated endocytosis negatively regulated by caveolin-1. *J Biol Chem* **2013**, *288* (24), 17713-24, 10.1074/jbc.M112.445403.
10. Berditchevski, F.; Odintsova, E. Tetraspanins as regulators of protein trafficking. *Traffic* **2007**, *8* (2), 89-96, 10.1111/j.1600-0854.2006.00515.x.
11. Escola, J.M.; Kleijmeer, M.J.; Stoorvogel, W.; Griffith, J.M.; Yoshie, O.; Geuze, H.J. Selective enrichment of tetraspan proteins on the internal vesicles of multivesicular endosomes and on exosomes secreted by human B-lymphocytes. *J Biol Chem* **1998**, *273* (32), 20121-7,
12. Kitadokoro, K.; Galli, G.; Petracca, R.; Falugi, F.; Grandi, G.; Bolognesi, M. Crystallization and preliminary crystallographic studies on the large extracellular domain of human CD81, a tetraspanin receptor for hepatitis C virus. *Acta Crystallogr D Biol Crystallogr* **2001**, *57* (Pt 1), 156-8,
13. Kitadokoro, K.; Ponassi, M.; Galli, G.; Petracca, R.; Falugi, F.; Grandi, G.; Bolognesi, M. Subunit association and conformational flexibility in the head subdomain of human CD81 large extracellular loop. *Biol Chem* **2002**, *383* (9), 1447-52, 10.1515/BC.2002.164.
14. Kitadokoro, K.; Bordo, D.; Galli, G.; Petracca, R.; Falugi, F.; Abrignani, S.; Grandi, G.; Bolognesi, M. CD81 extracellular domain 3D structure: insight into the tetraspanin superfamily structural motifs. *EMBO J* **2001**, *20* (1-2), 12-8, 10.1093/emboj/20.1.12.
15. Higginbottom, A.; Quinn, E.R.; Kuo, C.C.; Flint, M.; Wilson, L.H.; Bianchi, E.; Nicosia, A.; Monk, P.N.; McKeating, J.A.; Levy, S. Identification of amino acid residues in CD81 critical for interaction with hepatitis C virus envelope glycoprotein E2. *J Virol* **2000**, *74* (8), 3642-9,
16. Imai, T.; Yoshie, O. C33 antigen and M38 antigen recognized by monoclonal antibodies inhibitory to syncytium formation by human T cell leukemia virus type 1 are both members of the transmembrane 4 superfamily and associate with each other and with CD4 or CD8 in T cells. *J Immunol* **1993**, *151* (11), 6470-81,

17. Seigneuret, M. Complete predicted three-dimensional structure of the facilitator transmembrane protein and hepatitis C virus receptor CD81: conserved and variable structural domains in the tetraspanin superfamily. *Biophys J* **2006**, 90 (1), 212-27, 10.1529/biophysj.105.069666.
18. Rajesh, S.; Sridhar, P.; Tews, B.A.; Feneant, L.; Cocquerel, L.; Ward, D.G.; Berditchevski, F.; Overduin, M. Structural basis of ligand interactions of the large extracellular domain of tetraspanin CD81. *J Virol* **2012**, 86 (18), 9606-16, 10.1128/JVI.00559-12.
19. Seigneuret, M.; Delaguillaumie, A.; Lagaudriere-Gesbert, C.; Conjeaud, H. Structure of the tetraspanin main extracellular domain. A partially conserved fold with a structurally variable domain insertion. *J Biol Chem* **2001**, 276 (43), 40055-64, 10.1074/jbc.M105557200.
20. Homsy, Y.; Schloetel, J.G.; Scheffer, K.D.; Schmidt, T.H.; Destainville, N.; Florin, L.; Lang, T. The extracellular delta-domain is essential for the formation of CD81 tetraspanin webs. *Biophys J* **2014**, 107 (1), 100-13, 10.1016/j.bpj.2014.05.028.
21. Schmidt, T.H.; Homsy, Y.; Lang, T. Oligomerization of the Tetraspanin CD81 via the Flexibility of Its delta-Loop. *Biophys J* **2016**, 110 (11), 2463-74, 10.1016/j.bpj.2016.05.003.
22. Homsy, Y.; Lang, T. The specificity of homomeric clustering of CD81 is mediated by its delta-loop. *FEBS Open Bio* **2017**, 7 (2), 274-283, 10.1002/2211-5463.12187.
23. Michel Seigneuret: Complete Predicted Three-Dimensional Structure of the Facilitator Transmembrane Protein and Hepatitis C Virus Receptor CD81: Conserved and Variable Structural Domains in the Tetraspanin Superfamily. *Biophys J*. 2006 Jan 1; 90(1): 212–227. doi: 10.1529/biophysj.105.069666
24. Andrew Waterhouse, Martino Bertoni, Stefan Bienert, Gabriel Studer, Gerardo Tauriello, Rafal Gumieny, Florian T Heer, Tjaart A P de Beer, Christine Rempfer, Lorenza Bordoli, Rosalba Lepore, Torsten Schwede: SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res*. 2018 Jul 2; 46(Web Server issue): W296–W303. Published online 2018 May 21. doi: 10.1093/nar/gky427

CLAIMS

1. A method of producing a protein comprising a target-specific extravesicular domain (TED) of an extracellular vesicle (EV) surface protein comprising modifying a polynucleotide comprising a nucleotide sequence encoding the extravesicular domain (ED) of an EV surface protein by a mutagenesis method within at least one modified region within the ED amino acid sequence with a length of 3-20 contiguous amino acids flanked by regions of the wild-type ED sequence at its N-terminus and C-terminus, to incorporate a target binding site within the ED, thereby producing a repertoire of polynucleotides encoding a variety of TEDs, each comprising a different target binding site, and selecting a TED specifically recognizing a predetermined target, and producing the protein comprising the selected TED.

2. The method of claim 1, wherein the repertoire of polynucleotides is comprised in genetic packages displaying the variety of TEDs on the outer surface, preferably employing a display system selected from the group consisting of a yeast, phage, bacterium, ribosome, mRNA or mammalian cell display.

3. The method of claim 1 or 2, wherein the modified region has a lower target binding affinity when isolated from the TED.

4. The method of any one of claims 1 to 3, wherein the target binding site comprises at least one further binding region which is within a further modified region distant at least 2 amino acids, or within the wild-type ED sequence.

5. The method of any one of claims 1 to 4, wherein the TED comprises at least 70% sequence identity to the wild-type ED.

6. The method of any one of claims 1 to 5, wherein the wild-type ED originates from an EV surface protein selected from the group consisting of tetraspan-like proteins, proteins of the integrin family, proteoglycans, five-transmembrane domains proteins, type I transmembrane proteins, Notch family proteins, enzymatic membrane proteins, immune regulatory surface proteins, surface markers of mesenchymal stem cells, glycoproteins, or channelling proteins.

7. The method of claim 6, wherein the tetraspan-like protein is a tetraspanin, preferably selected from the group consisting of:

- a) CD81 comprising the amino acid sequence identified as SEQ ID NO:87;
- b) CD9 comprising the amino acid sequence identified as SEQ ID NO:89;
- c) CD53 comprising the amino acid sequence identified as SEQ ID NO:90;
- d) TSPAN32 comprising the amino acid sequence identified as SEQ ID NO:91;
- e) CD82 comprising the amino acid sequence identified as SEQ ID NO:92;
- f) CD63 comprising the amino acid sequence identified as SEQ ID NO:93;
- g) CD151 comprising the amino acid sequence identified as SEQ ID NO:94; and
- h) CD37 comprising the amino acid sequence identified as SEQ ID NO:95

or wherein the tetraspan-like protein is a lysosome-associated membrane protein, preferably LAMP2 comprising the amino acid sequence identified as SEQ ID NO:96.

8. The method of any one of claims 1 to 7, wherein the wild-type ED is of any one of:

- a) CD81, wherein the ED comprises any one of the amino acid sequences identified as SEQ ID NO:130 or SEQ ID NO:131;
- b) CD9, wherein the ED comprises any one of the amino acid sequences identified as SEQ ID NO:132, SEQ ID NO:182 or SEQ ID NO:133;
- c) CD53, wherein the ED comprises any one of the amino acid sequences identified as SEQ ID NO:134 or SEQ ID NO:135;
- d) TSPAN32, wherein the ED comprises any one of the amino acid sequences identified as SEQ ID NO:136 or SEQ ID NO:137;
- e) CD82, wherein the ED comprises any one of the amino acid sequences identified as SEQ ID NO:138 or SEQ ID NO:139;
- f) CD63, wherein the ED comprises any one of the amino acid sequences identified as SEQ ID NO:140 or SEQ ID NO:141;
- g) CD151, wherein the ED comprises any one of the amino acid sequences identified as SEQ ID NO:142 or SEQ ID NO:143;
- h) CD37, wherein the ED comprises any one of the amino acid sequences identified as SEQ ID NO:144 or SEQ ID NO:145;

or

- i) LAMP2, wherein the ED comprises the amino acid sequence identified as SEQ ID NO:146:

9. The method of any one of claims 1 to 8, wherein the protein comprises a loop structure in the ED amino acid sequence which is stabilized by one or more cysteine(s) at position(s) to allow the formation of one or more disulfide bonds.

10. The method of any one of claims 1 to 9, wherein the modified region is positioned within a loop region of the ED.

11. The method of any one of claims 1 to 10, wherein the ED is

a) of CD81 and the amino acid sequence is modified to introduce cysteines to allow formation of one or more disulfide bonds not naturally-occurring in the wild-type ED sequence, preferably between positions 134 and 144 and/or 130 and 146 and/or 135 and 168, wherein numbering is of human CD81 identified as SEQ ID NO:87, or

b) of CD9 and the amino acid sequence is modified to introduce cysteines to allow formation of one or more disulfide bonds not naturally-occurring in the wild-type ED sequence, preferably between positions 20 and 28, wherein numbering of the positions is of the CD9 large extracellular loop (LEL, SEQ ID NO:118).

12. The method of any one of claims 1 to 11, wherein the ED is of CD81, and the modified region is positioned within positions 160 and 172, wherein numbering is of human CD81 identified as SEQ ID NO:87.

13. The method of claim 12, wherein the TED comprises at least one further binding region positioned between positions 132 and 141, or between positions 180 and 189, wherein numbering is of human CD81 identified as SEQ ID NO:87.

14. The method of any one of claims 1 to 11, wherein the ED is of CD9 and the modified region is positioned within any one of positions 155-166, positions 128-142, positions 130-140, or positions 169-180, wherein numbering is of human CD9 identified as SEQ ID NO:89.

15. The method of any one of claims 1 to 14, wherein the target is selected from the group consisting of cellular targets, preferably mitogenic receptors, cytokine receptors, asialoglycoprotein receptors, membrane transporters, lipoproteins,

liposaccharides, glycoproteins, proteoglycans, or acellular targets, preferably cytokines, artificial proteins or artificial surface structures.

16. The method of any one of claims 1 to 15, wherein the protein comprising the TED is a target-specific EV surface protein (TSP) comprising said TED and at least one transmembrane domain.

17. The method of claim 16, wherein the transmembrane domain comprises at least 70% sequence identity to a transmembrane domain originating from a mammalian EV surface protein.

18. The method of claim 16 or 17, wherein the transmembrane domain is of any one of:

a) CD81, wherein the transmembrane domain comprises any one of the amino acid sequences identified as SEQ ID NO:147, SEQ ID NO:148, SEQ ID NO:149 or SEQ ID NO:150;

b) CD9, wherein the transmembrane domain comprises any one of the amino acid sequences identified as SEQ ID NO:151, SEQ ID NO:152, SEQ ID NO:153 or SEQ ID NO:154;

c) CD53, wherein the transmembrane domain comprises any one of the amino acid sequences identified as SEQ ID NO:155, SEQ ID NO:156, SEQ ID NO:157 or SEQ ID NO:158;

d) TSPAN32, wherein the transmembrane domain comprises any one of the amino acid sequences identified as SEQ ID NO:159, SEQ ID NO:160, SEQ ID NO:161 or SEQ ID NO:162;

e) CD82, wherein the transmembrane domain comprises any one of the amino acid sequences identified as SEQ ID NO:163, SEQ ID NO:164, SEQ ID NO:165 or SEQ ID NO:166;

f) CD63, wherein the transmembrane domain comprises any one of the amino acid sequences identified as SEQ ID NO:167, SEQ ID NO:168, SEQ ID NO:169 or SEQ ID NO:170;

g) CD151, wherein the transmembrane domain comprises any one of the amino acid sequences identified as SEQ ID NO:171, SEQ ID NO:172, SEQ ID NO:173 or SEQ ID NO:174;

h) CD37, wherein the transmembrane domain comprises any one of the amino acid sequences identified as SEQ ID NO:175, SEQ ID NO:176, SEQ ID NO:177 or SEQ ID NO:178;

or

i) LAMP2, wherein the transmembrane domain comprises the amino acid sequence identified as SEQ ID NO:179.

19. The method of any one of claims 16 to 18, wherein both, the ED and the transmembrane domain, originate from the same EV surface protein, preferably wherein the EV surface protein is selected from the group consisting of:

- a) CD81 comprising the amino acid sequence identified as SEQ ID NO:87;
- b) CD9 comprising the amino acid sequence identified as SEQ ID NO:89;
- c) CD53 comprising the amino acid sequence identified as SEQ ID NO:90;
- d) TSPAN32 comprising the amino acid sequence identified as SEQ ID NO:91;
- e) CD82 comprising the amino acid sequence identified as SEQ ID NO:92;
- f) CD63 comprising the amino acid sequence identified as SEQ ID NO:93;
- g) CD151 comprising the amino acid sequence identified as SEQ ID NO:94;
- h) CD37 comprising the amino acid sequence identified as SEQ ID NO:95; and
- i) LAMP2 comprising the amino acid sequence identified as SEQ ID NO:96.

20. A method of producing a target-specific extracellular vesicle (TEV) preparation by

- a) providing a polynucleotide encoding a protein comprising the TED obtainable by a method of any one of claims 1 to 19;
- b) introducing said polynucleotide into a source cell or source cell mixture;
- b) culturing said cell(s) under conditions producing extracellular vesicles;
- c) isolating a fraction comprising a TEV comprising the target binding site of the TED; and
- d) producing a preparation of the TEV comprised in said fraction.

21. The method of claim 20, wherein the protein comprising the TED is a target-specific EV surface protein (TSP) comprising said TED and at least one transmembrane domain, and the TEV is displaying the target binding site on the outer surface of its membrane.

22. The method of claim 20 or 21, which further comprises loading the TEV with an intravesicular load, wherein the load comprises any one or more of peptides, polypeptides, protein domains, proteins, lipids, genes, nucleic acids such as mRNAs, miRNAs, RNAi mediating molecules in particular locked nucleic acids or phosphorothioates, DNA, DNA fragments, plasmids such as minicircle DNA, drugs such as small molecules, in particular chemotherapeutics or senolytics.

23. The method of any one of claims 20 to 22, wherein the source cell or source cell mixture is originating from a eukaryotic or prokaryotic source, preferably of body tissue, body fluid, or a cell culture.

24. The method of any one of claims 20 to 23, wherein the source cell or source cell mixture is obtained from a subject and the TEV preparation is formulated for autologous use.

25. A TEV preparation obtainable by the method of any one of claims 20 to 24.

26. An autologous TEV preparation produced by the method of claim 24, wherein the source cell or source cell mixture is obtained from a subject and the TEV preparation formulated for autologous use to the same subject.

27. A protein comprising a target-specific extravesicular domain (TED) of an extracellular vesicle (EV) surface protein obtainable by a method of any one of claims 1 to 19.

28. A target-specific extravesicular domain (TED) of an extracellular vesicle (EV) surface protein comprising at least 70% sequence identity to a wild-type extravesicular domain (ED) of a mammalian EV surface protein sequence and at least one modified region with a length of 3-20 contiguous amino acids flanked by regions of the wild-type ED sequence at its N-terminus and C-terminus, which modified region is at least part of a target binding site not naturally-occurring in the wild-type ED, wherein the modified region has a lower target binding affinity when isolated from the TED and/or wherein the target binding site comprises at least one further binding region within a further modified region distant at least 2 amino acids, or within the wild-type ED sequence.

29. A target-specific EV surface protein (TSP) comprising at least one transmembrane domain and an extravesicular domain (ED) comprising at least 70% sequence identity to a wild-type ED of a mammalian EV surface protein sequence and at least one modified region with a length of 3-20 contiguous amino acids flanked by regions of the wild-type ED sequence at its N-terminus and C-terminus, which modified region is at least part of a target binding site not naturally-occurring in the wild-type ED, wherein the modified region has a lower target binding affinity when isolated from the TSP and/or wherein the target binding site comprises at least one further binding region within a further modified region distant at least 2 amino acids, or within the wild-type ED sequence.

30. A polynucleotide encoding any one of the protein of claim 27, the TED of claim 28, or the TSP of claim 29.

31. A target-specific extracellular vesicle (TEV) comprising a membrane and displaying a target-specific molecule on the outer surface of the membrane, wherein the target-specific molecule is any one of the protein of claim 27, the TED of claim 28, or the TSP of claim 29.

32. A pharmaceutical preparation comprising any one of the protein of claim 27, the TED of claim 28, the TSP of claim 29, or the TEV of any one of claims 25, 26 or 31, and a pharmaceutically acceptable carrier, preferably in a formulation for intradermal, subcutaneous, intravenous, topical, or oral use.

33. A target-specific extracellular vesicle (TEV) library comprising a variety of at least 10^2 TEVs of claim 31, each with a different modified region flanked by the same regions of the same wild-type extravesicular domain (ED) at its N-terminus and C-terminus.

34. A method of producing a library of target-specific extracellular vesicles (TEVs), comprising:

a) providing a repertoire of polynucleotides encoding a variety of target-specific EV surface proteins (TSPs), each comprising a different target binding site;

b) introducing said repertoire into said source cell(s); and

b) isolating a fraction comprising a repertoire of target-specific extracellular vesicles (TEVs), with a different target binding specificity, to produce a library of TEVs,

wherein the repertoire of polynucleotides is produced by mutagenizing a polynucleotide encoding an EV surface protein by a mutagenesis method to obtain mutations of the extravesicular domain (ED) of said EV surface protein within at least one modified region with a length of 3-20 contiguous amino acids flanked by regions of the wild-type ED sequence at its N-terminus and C-terminus, to incorporate a target binding site within the ED.

35. A TEV library obtainable by the method of claim 34, preferably comprising at least 10^2 TEVs with different target specificity.