

AGENTS AND METHODS FOR TARGETED DELIVERY OF IMMUNE EFFECTOR CELLS**Field of the invention**

The invention relates to agents and methods for targeted delivery of immune effector cells to target cells. In some embodiments, the invention involves providing to a subject immune effector cells genetically modified to express a chimeric antigen receptor (CAR) and a compound (docking compound) comprising a binding moiety for an antigen on target cells and a binding moiety for the CAR. The CAR and the binding moiety for the CAR bind to each other by means of an ALFA-tag and a single-domain antibody binding to the ALFA-tag.

Background

Chimeric antigen receptors (CARs) - also known as chimeric immunoreceptors, chimeric T cell receptors or artificial T cell receptors - are receptor proteins that have been engineered to give T cells the new ability to target a specific antigen. The receptors are chimeric in that they combine both antigen-binding and T cell activating functions into a single receptor. CAR-T cell therapy uses T cells engineered with CARs to treat cancer, for example. CAR-T cell therapy aims to modify T cells to recognize target cells, e.g., cancer cells, in order to attack and destroy them more effectively. However, for this purpose, T cells engineered to express different CARs may need to be generated to recognize different targets on cells to be targeted, which can involve a significant amount of work and difficulty. This workload could be significantly reduced if the CAR-modified T cells are no longer targeted directly to target cells, but indirectly via a compound that binds to target cells. This allows the use of universal CAR-T cells, but requires the use of one or more compounds that bind to target cells and with which the CAR-T cells interact. Thus, a system of two interacting members, one placed on the CAR and the other placed on the compound that binds to target cells, useful for this purpose is required. In 2019, Götzke *et al.* published the article '*The ALFA-tag is a highly versatile tool for nanobody-based bioscience applications*' (Götzke *et al.*, Nat Commun. 2019 Sep 27;10(1):4403. doi: 10.1038/s41467-019-12301-7). The ALFA®-tag is a rationally designed peptide with a unique sequence that forms a stable alpha helix in solution, is electroneutral and highly soluble at physiological conditions. It can be fused to a binding partner in different

orientations (N- or C-terminal or in between two domains of a fusion protein) without loss of structure and function and can be detected with a high affinity (double-digit pM Kd) single-domain antibody (VHH) that the authors have identified from an immunized alpaca via a selection method based on the isolation of target specific B-cells (CELLINE). A detailed description of the ALFA[®]/NbALFA epitope tag system can be found in the PCT patent publication WO 2020/053239 A1. Thus, the ALFA[®]-tag and the NbALFA single-domain antibody specific for the tag have a great potential in bioscience applications.

NbALFA is derived from immunized alpaca and alpaca-derived single-domain antibodies (VHHs) in general share a low degree of homology to human sequences. However, for *in vivo* applications of the ALFA[®]/NbALFA tool in the above described CAR-T cell therapy in humans, higher sequence homology to human counterpart sequence would be desirable.

Summary

The invention relates to agents and methods for targeted delivery of immune effector cells to target cells, e.g., diseased cells. Targeted delivery of the immune effector cells to target cells may be achieved using a docking compound binding to the immune effector cells via their antigen receptor, said docking compound comprising a targeting moiety for targeting target cells. The docking compound binds to the antigen receptor via a tag/binder system comprising an ALFA-tag and a single-domain antibody described herein.

In some embodiments, the invention involves immune effector cells genetically modified to express a chimeric antigen receptor (CAR) and a compound (docking compound) comprising a binding moiety for an antigen on target cells and a further binding moiety for a CAR. Providing to a subject the immune effector cells and the docking compound may deliver the immune effector cells to target cells. The CAR and the binding moiety for a CAR bind to each other by means of an ALFA-tag and a single-domain antibody binding to the ALFA-tag. In some embodiments, the moiety of the CAR binding to the binding moiety for a CAR comprises an ALFA-tag and the binding moiety for a CAR comprises the single-domain antibody. In some embodiments, the moiety of the CAR binding to the binding moiety for a CAR comprises the single-domain antibody and the binding moiety for a CAR comprises an ALFA-tag.

The immune effector cells may be genetically modified *in vitro/ex vivo* or *in vivo* to express the antigen receptor. Genetic modification may be achieved using a nucleic acid encoding an antigen receptor for transfection of immune effector cells.

In some embodiments, the immune effector cells genetically modified to express a CAR are provided to a subject by administering the immune effector cells. In some embodiments, the immune effector cells genetically modified to express a CAR are provided to a subject by generating the immune effector cells in the subject, e.g., by administering nucleic acid, e.g., DNA or RNA, encoding the CAR for transfection of immune effector cells.

The docking compound may be provided by adding it to target cells, e.g., the docking compound may be administered to a subject, or by adding nucleic acid, e.g., RNA, encoding the docking compound, e.g., RNA encoding the docking compound may be administered to a subject for expression of the docking compound.

In some embodiments, the docking compound is provided to a subject by administering nucleic acid, e.g., RNA, encoding the docking compound. In some embodiments, the docking compound is provided to a subject by administering the docking compound. In some embodiments, a preformed complex wherein the docking compound is bound to the immune effector cells genetically modified to express a CAR is provided to a subject by administration. The binding moiety for an antigen on target cells may bind to a target antigen such as a cancer antigen on cancer cells and the immune effector cells may target the binding moiety for a CAR to thereby precisely deliver the immune effector cells to the target cells such as cancer cells. Accordingly, the present disclosure relates to an approach wherein a docking compound that labels target cells, e.g., by binding to a cell surface antigen, is used. The docking compound comprises a moiety which will be targeted by immune effector cells, equipped with an antigen receptor targeting the moiety.

The approach described herein may deliver immune effector cells to target cells *in vitro/ex vivo* as well as *in vivo*. Immune effector cells genetically modified to express an antigen receptor described herein are useful in the treatment of diseases wherein targeting cells such as diseased cells expressing an antigen such as a tumor antigen is beneficial. The target cells may express the antigen on the cell surface for recognition by a docking compound. The treatments described herein may provide for the selective eradication of such cells expressing an antigen, thereby minimizing adverse effects to normal cells not expressing the antigen. The methods and agents described herein are, in particular, useful for the treatment of diseases characterized by diseased cells expressing an antigen the docking compound is directed to.

The concept described herein allows to use a single type of immune effector cells for targeting a wide range of target cells, i.e., by using a single type of immune effector cells in combination with different docking compounds targeting different antigen targets and comprising an identical CAR binding moiety.

In one aspect the invention provides a system comprising:

- (i) an immune effector cell genetically modified to express a chimeric antigen receptor (CAR), or a nucleic acid encoding a CAR useful for genetically modifying an immune effector cell to express the CAR; and

(ii) a compound comprising a binding moiety for the CAR and a binding moiety for an antigen on target cells, or a nucleic acid encoding the compound, wherein the CAR and the binding moiety for the CAR comprise a binding system comprising an ALFA-tag and a single-domain antibody binding to the ALFA-tag, wherein the antibody comprises the CDR1 sequence as shown in SEQ ID NO: 94 or a sequence having 1 or 2 mutations relative to said sequence, the CDR2 sequence as shown in SEQ ID NO: 97 or a sequence having 1 or 2 mutations relative to said sequence, and the CDR3 sequence as shown in SEQ ID NO: 98 or a sequence having 1 or 2 mutations relative to said sequence; wherein the antibody further comprises at positions corresponding to sequence position 47, 48, and 50 of SEQ ID NO: 1 the following amino acid residues: Arg at position 47, Arg at position 48, Met at position 50; and wherein the antibody comprises a sequence having at least 90% sequence identity to any one of SEQ ID NOs: 118-125.

In some embodiments, the antibody comprises the CDR1 sequence as shown in SEQ ID NO: 94, the CDR2 sequence as shown in SEQ ID NO: 97, and the CDR3 sequence as shown in SEQ ID NO: 98.

In some embodiments, the immune effector cell comprises a T cell.

In some embodiments, the immune effector cell expresses the CAR on its cell surface.

In some embodiments, the CAR comprises a moiety binding to the binding moiety for the CAR.

In some embodiments, the moiety of the CAR binding to the binding moiety for the CAR is comprised by an extracellular moiety of the CAR.

In some embodiments, the moiety of the CAR binding to the binding moiety for the CAR comprises an ALFA-tag and the binding moiety for the CAR comprises the antibody.

In some embodiments, the moiety of the CAR binding to the binding moiety for the CAR comprises the antibody and the binding moiety for the CAR comprises an ALFA-tag.

In some embodiments, the antibody specifically binds to an epitope tag comprising the sequence SRLEEELRRRLTE (SEQ ID NO: 2).

In some embodiments, the antibody is humanized and/or wherein the antibody comprises a sequence that has at least about 70% V segment identity to the closest human germline sequence.

In some embodiments, the antibody has a higher T_m as compared to the antibody of SEQ ID NO: 1.

In some embodiments, the antibody comprises a His residue at the position corresponding to sequence position 99 of SEQ ID NO: 1, a Val residue at the position corresponding to sequence position 100 of SEQ ID NO: 1, and/or a Tyr residue at the position corresponding to sequence position 40 of SEQ ID NO: 1.

In some embodiments, the antibody comprises FR1, FR2, FR3, and FR4 sequences, that have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 102, 108, and 109, respectively.

In some embodiments, the antibody comprises a protein A binding site.

In some embodiments, the antibody comprises a sequence having at least 95% sequence identity to any one of SEQ ID NOs: 118-125, or wherein the antibody comprises a sequence set forth in any one of SEQ ID NOs: 118-125.

In some embodiments, the ALFA-tag comprises the amino acid sequence of X1-X2-L-E-X5-E-X7-R-R-R-L-X12-X13 (SEQ ID NO: 3), wherein X1 is G or S or P or T, X2 is R or G or P, X5 is E or Q, X7 is L or I, X12 is S or T or P or A or D or E, and wherein X13 is P or A or S or K or D or E or no amino acid.

In some embodiments, the binding moiety for the CAR and the binding moiety for an antigen on target cells are connected through a linking moiety.

In some embodiments, the linking moiety comprises a polymer.

In some embodiments, the linking moiety comprises a non-peptidic moiety.

In some embodiments, the linking moiety comprises a polymer selected from the group consisting of poly(ethylene glycol) (PEG), polysarcosine (pSar) (poly(N-methylglycine)), polyoxazoline (POX), polyoxazine (POZ), and poly-2-(2-(2-aminoethoxy)ethoxy)acetic acid (pAEEA), derivatives and combinations thereof.

In some embodiments, the linking moiety comprises a peptidic moiety.

In some embodiments, the binding moiety for an antigen on target cells comprises a moiety selected from the group consisting of an antibody binding to the antigen, an antibody derivative binding to the antigen and a ligand of the antigen.

In some embodiments, the antibody derivative binding to the antigen comprises an antibody fragment.

In some embodiments, the binding moiety for an antigen on target cells comprises a peptidic moiety.

In some embodiments, the compound comprises a fusion polypeptide comprising the binding moiety for the CAR and the binding moiety for an antigen on target cells.

In some embodiments, the ligand of the antigen comprises DUPA or FAPI.

In some embodiments, the target cells are diseased cells.

In some embodiments, the target cells are cancer cells.

In some embodiments, the antigen on target cells comprises a cell surface antigen.

In some embodiments, the antigen on target cells comprises a tumor antigen.

In some embodiments, the nucleic acid encoding a CAR comprises DNA.

In some embodiments, the nucleic acid encoding the compound comprises RNA.

In some embodiments, binding of a complex comprising:

(i) the immune effector cell genetically modified to express a chimeric antigen receptor (CAR); and

(ii) the compound comprising a binding moiety for the CAR and a binding moiety for an antigen on target cells

to cells expressing the antigen results in killing of the cells expressing the antigen.

In a further aspect the invention provides a method for treating a subject having a disease, disorder or condition characterized by cells expressing an antigen, comprising:

(i) providing to the subject immune effector cells genetically modified to express a chimeric antigen receptor (CAR); and

(ii) providing to the subject a compound comprising a binding moiety for the CAR and a binding moiety for the antigen,

wherein the CAR and the binding moiety for the CAR comprise a binding system comprising an ALFA-tag and a single-domain antibody binding to the ALFA-tag,

wherein the antibody comprises the CDR1 sequence as shown in SEQ ID NO: 94 or a sequence having 1 or 2 mutations relative to said sequence, the CDR2 sequence as shown in SEQ ID NO:

97 or a sequence having 1 or 2 mutations relative to said sequence, and the CDR3 sequence as shown in SEQ ID NO: 98 or a sequence having 1 or 2 mutations relative to said sequence; wherein the antibody further comprises at positions corresponding to sequence position 47, 48, and 50 of SEQ ID NO: 1 the following amino acid residues: Arg at position 47, Arg at position 48, Met at position 50;

and wherein the antibody comprises a sequence having at least 90% sequence identity to any one of SEQ ID NOs: 118-125.

In some embodiments, the antibody comprises the CDR1 sequence as shown in SEQ ID NO: 94, the CDR2 sequence as shown in SEQ ID NO: 97, and the CDR3 sequence as shown in SEQ ID NO: 98.

In some embodiments, the immune effector cells comprise T cells.

In some embodiments, the immune effector cells express the CAR on their cell surface.

In some embodiments, the CAR comprises a moiety binding to the binding moiety for the CAR.

In some embodiments, the moiety of the CAR binding to the binding moiety for the CAR is comprised by an extracellular moiety of the CAR.

In some embodiments, the moiety of the CAR binding to the binding moiety for the CAR comprises an ALFA-tag and the binding moiety for the CAR comprises the antibody.

In some embodiments, the moiety of the CAR binding to the binding moiety for the CAR comprises the antibody and the binding moiety for the CAR comprises an ALFA-tag.

In some embodiments, the antibody specifically binds to an epitope tag comprising the sequence SRLEEELRRRLTE (SEQ ID NO: 2).

In some embodiments, the antibody is humanized and/or wherein the antibody comprises a sequence that has at least about 70% V segment identity to the closest human germline sequence.

In some embodiments, the antibody has a higher T_m as compared to the antibody of SEQ ID NO: 1.

In some embodiments, the antibody comprises a His residue at the position corresponding to sequence position 99 of SEQ ID NO: 1, a Val residue at the position corresponding to sequence position 100 of SEQ ID NO: 1, and/or a Tyr residue at the position corresponding to sequence position 40 of SEQ ID NO: 1.

In some embodiments, the antibody comprises FR1, FR2, FR3, and FR4 sequences, that have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 102, 108, and 109, respectively.

In some embodiments, the antibody comprises a protein A binding site.

In some embodiments, the antibody comprises a sequence having at least 95% sequence identity to any one of SEQ ID NOs: 118-125, or wherein the antibody comprises a sequence set forth in any one of SEQ ID NOs: 118-125.

In some embodiments, the ALFA-tag comprises the amino acid sequence of X1-X2-L-E-X5-E-X7-R-R-R-L-X12-X13 (SEQ ID NO: 3), wherein X1 is G or S or P or T, X2 is R or G or P, X5 is E or Q, X7 is L or I, X12 is S or T or P or A or D or E, and wherein X13 is P or A or S or K or D or E or no amino acid.

In some embodiments, the binding moiety for the CAR and the binding moiety for an antigen on target cells are connected through a linking moiety.

In some embodiments, the linking moiety comprises a polymer.

In some embodiments, the linking moiety comprises a non-peptidic moiety.

In some embodiments, the linking moiety comprises a polymer selected from the group consisting of poly(ethylene glycol) (PEG), polysarcosine (pSar) (poly(N-methylglycine)), polyoxazoline (POX), polyoxazine (POZ), and poly-2-(2-(2-aminoethoxy)ethoxy)acetic acid (pAEEA), derivatives and combinations thereof.

In some embodiments, the linking moiety comprises a peptidic moiety.

In some embodiments, the binding moiety for the antigen comprises a moiety selected from the group consisting of an antibody binding to the antigen, an antibody derivative binding to the antigen and a ligand of the antigen.

In some embodiments, the antibody derivative binding to the antigen comprises an antibody fragment.

In some embodiments, the binding moiety for the antigen comprises a peptidic moiety.

In some embodiments, the compound comprises a fusion polypeptide comprising the binding moiety for the CAR and the binding moiety for the antigen.

In some embodiments, the ligand of the antigen comprises DUPA or FAPL.

In some embodiments, the cells expressing an antigen are diseased cells.

In some embodiments, the cells expressing an antigen are cancer cells.

In some embodiments, the antigen comprises a cell surface antigen.

In some embodiments, the antigen comprises a tumor antigen.

In some embodiments, binding of a complex comprising:

(i) the immune effector cells genetically modified to express a chimeric antigen receptor (CAR); and

(ii) the compound comprising a binding moiety for the CAR and a binding moiety for an antigen

to cells expressing the antigen results in killing of the cells expressing the antigen.

In some embodiments, the method comprises administering the immune effector cells genetically modified to express a CAR to the subject.

In some embodiments, the method comprises generating the immune effector cells genetically modified to express a CAR in the subject.

In some embodiments, the method comprises administering nucleic acid encoding the CAR to the subject.

In some embodiments, the nucleic acid encoding the CAR comprises DNA.

In some embodiments, the method comprises administering the compound to the subject.

In some embodiments, the method comprises generating the compound in the subject.

In some embodiments, the method comprises administering nucleic acid encoding the compound to the subject.

In some embodiments, the nucleic acid encoding the compound comprises RNA.

In some embodiments, the method comprises allowing expression of the compound by cells in the subject such that the compound becomes associated with cells expressing the antigen and the binding moiety for the CAR is available for binding by the immune effector cells.

In some embodiments, the cells expressing the compound are transfected with the nucleic acid encoding the compound.

In some embodiments, the cells expressing the compound secrete the compound.

In some embodiments, the cells expressing the compound express the compound such that it is released into the bloodstream.

In some embodiments, the nucleic acid encoding the CAR and/or the nucleic acid encoding the compound is administered as particulate formulation such as formulated as lipid nanoparticles.

In some embodiments, the disease, disorder or condition is cancer.

In a further aspect the invention provides a chimeric antigen receptor (CAR) comprising a single-domain antibody binding to an ALFA-tag,

wherein the antibody comprises the CDR1 sequence as shown in SEQ ID NO: 94 or a sequence having 1 or 2 mutations relative to said sequence, the CDR2 sequence as shown in SEQ ID NO: 97 or a sequence having 1 or 2 mutations relative to said sequence, and the CDR3 sequence as shown in SEQ ID NO: 98 or a sequence having 1 or 2 mutations relative to said sequence; wherein the antibody further comprises at positions corresponding to sequence position 47, 48, and 50 of SEQ ID NO: 1 the following amino acid residues: Arg at position 47, Arg at position 48, Met at position 50;

and wherein the antibody comprises a sequence having at least 90% sequence identity to any one of SEQ ID NOs: 118-125.

In some embodiments, the antibody comprises the CDR1 sequence as shown in SEQ ID NO: 94, the CDR2 sequence as shown in SEQ ID NO: 97, and the CDR3 sequence as shown in SEQ ID NO: 98.

In some embodiments, the antibody specifically binds to an epitope tag comprising the sequence SRLEEELRRRLTE (SEQ ID NO: 2).

In some embodiments, the antibody is humanized and/or wherein the antibody comprises a sequence that has at least about 70% V segment identity to the closest human germline sequence.

In some embodiments, the antibody has a higher T_m as compared to the antibody of SEQ ID NO: 1.

In some embodiments, the antibody comprises a His residue at the position corresponding to sequence position 99 of SEQ ID NO: 1, a Val residue at the position corresponding to sequence position 100 of SEQ ID NO: 1, and/or a Tyr residue at the position corresponding to sequence position 40 of SEQ ID NO: 1.

In some embodiments, the antibody comprises FR1, FR2, FR3, and FR4 sequences, that have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 102, 108, and 109, respectively.

In some embodiments, the antibody comprises a protein A binding site.

In some embodiments, the antibody comprises a sequence having at least 95% sequence identity to any one of SEQ ID NOs: 118-125, or wherein the antibody comprises a sequence set forth in any one of SEQ ID NOs: 118-125.

In some embodiments, the antibody is comprised by an extracellular domain of the CAR.

In a further aspect the invention provides a nucleic acid encoding the CAR described herein.

In some embodiments, the nucleic acid is DNA.

In a further aspect the invention provides a host cell comprising the nucleic acid encoding the CAR described herein and/or expressing the CAR described herein.

In some embodiments, the host cell comprises an immune effector cell.

In some embodiments, the immune effector cell comprises a T cell.

In some embodiments, the host cell is genetically modified to comprise the nucleic acid and/or expresses the CAR.

In some embodiments, the host cell expresses the CAR on its cell surface.

In a further aspect the invention provides a compound comprising a single-domain antibody binding to an ALFA-tag and a binding moiety for an antigen on target cells, wherein the antibody comprises the CDR1 sequence as shown in SEQ ID NO: 94 or a sequence having 1 or 2 mutations relative to said sequence, the CDR2 sequence as shown in SEQ ID NO: 97 or a sequence having 1 or 2 mutations relative to said sequence, and the CDR3 sequence as shown in SEQ ID NO: 98 or a sequence having 1 or 2 mutations relative to said sequence; wherein the antibody further comprises at positions corresponding to sequence position 47, 48, and 50 of SEQ ID NO: 1 the following amino acid residues: Arg at position 47, Arg at position 48, Met at position 50;

and wherein the antibody comprises a sequence having at least 90% sequence identity to any one of SEQ ID NOs: 118-125.

In some embodiments, the antibody comprises the CDR1 sequence as shown in SEQ ID NO: 94, the CDR2 sequence as shown in SEQ ID NO: 97, and the CDR3 sequence as shown in SEQ ID NO: 98.

In some embodiments, the antibody specifically binds to an epitope tag comprising the sequence SRLEELRRRLTE (SEQ ID NO: 2).

In some embodiments, the antibody is humanized and/or wherein the antibody comprises a sequence that has at least about 70% V segment identity to the closest human germline sequence.

In some embodiments, the antibody has a higher T_m as compared to the antibody of SEQ ID NO: 1.

In some embodiments, the antibody comprises a His residue at the position corresponding to sequence position 99 of SEQ ID NO: 1, a Val residue at the position corresponding to sequence position 100 of SEQ ID NO: 1, and/or a Tyr residue at the position corresponding to sequence position 40 of SEQ ID NO: 1.

In some embodiments, the antibody comprises FR1, FR2, FR3, and FR4 sequences, that have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 102, 108, and 109, respectively.

In some embodiments, the antibody comprises a protein A binding site.

In some embodiments, the antibody comprises a sequence having at least 95% sequence identity to any one of SEQ ID NOs: 118-125, or wherein the antibody comprises a sequence set forth in any one of SEQ ID NOs: 118-125.

In some embodiments, the antibody and the binding moiety for an antigen on target cells are connected through a linking moiety.

In some embodiments, the linking moiety comprises a polymer.

In some embodiments, the linking moiety comprises a non-peptidic moiety.

In some embodiments, the linking moiety comprises a polymer selected from the group consisting of poly(ethylene glycol) (PEG), polysarcosine (pSar) (poly(N-methylglycine)),

polyoxazoline (POX), polyoxazine (POZ), and poly-2-(2-(2-aminoethoxy)ethoxy)acetic acid (pAEEA), derivatives and combinations thereof.

In some embodiments, the linking moiety comprises a peptidic moiety.

In some embodiments, the binding moiety for an antigen on target cells comprises a moiety selected from the group consisting of an antibody binding to the antigen, an antibody derivative binding to the antigen and a ligand of the antigen.

In some embodiments, the antibody derivative binding to the antigen comprises an antibody fragment.

In some embodiments, the binding moiety for an antigen on target cells comprises a peptidic moiety.

In some embodiments, the compound comprises a fusion polypeptide comprising the antibody and the binding moiety for an antigen on target cells.

In some embodiments, the ligand of the antigen comprises DIPA or FAPI.

In some embodiments, the target cells are diseased cells.

In some embodiments, the target cells are cancer cells.

In some embodiments, the antigen on target cells comprises a cell surface antigen.

In some embodiments, the antigen on target cells comprises a tumor antigen.

In a further aspect the invention provides a nucleic acid encoding the compound comprising a single-domain antibody binding to an ALFA-tag and a binding moiety for an antigen on target cells described herein.

In some embodiments, the nucleic acid is RNA.

In a further aspect the invention provides a host cell comprising the nucleic acid encoding the compound comprising a single-domain antibody binding to an ALFA-tag and a binding moiety for an antigen on target cells described herein and/or expressing the compound comprising a single-domain antibody binding to an ALFA-tag and a binding moiety for an antigen on target cells described herein.

In a further aspect the invention provides the agents, compositions and systems described herein for use in the methods described herein.

Brief description of the drawings**Figure 1: Affinity of humanized NbALFA variants generated by CDR grafting**

A, affinity measurements of the original alpaca-derived NbALFA (wtNbALFA) for the peptide of SEQ ID NO: 2. **B**, affinity measurements of variant 3 for the peptide of SEQ ID NO: 2. **C**, affinity measurements of variant 4 for the peptide of SEQ ID NO: 2. **D**, affinity measurements of variant 5 for the peptide of SEQ ID NO: 2. **E**, affinity measurements of variant 6 for the peptide of SEQ ID NO: 2. **F**, affinity measurements of variant 7 for the peptide of SEQ ID NO: 2. Affinity was determined using biolayer interferometry as described in Method 1 of the Examples.

Figure 2: Affinity of further engineered humanized NbALFA variants

A, affinity measurements of variant 8 for the Peptide of SEQ ID NO: 2. **B**, affinity measurements of variant 9 for the peptide of SEQ ID NO: 2. **C**, affinity measurements of variant 10 for the peptide of SEQ ID NO: 2. **D**, affinity measurements of variant 11 for the peptide of SEQ ID NO: 2. **E**, affinity measurements of variant 12 for the peptide of SEQ ID NO: 2. **F**, affinity measurements of variant 13 for the peptide of SEQ ID NO: 2. **G**, affinity measurements of variant 14 for the peptide of SEQ ID NO: 2. **H**, affinity measurements of variant 15 for the peptide of SEQ ID NO: 2. Affinity was determined using biolayer interferometry as described in Method 1 of the Examples.

Figure 3: SpA-assisted purification of humanized NbALFA variants

SpA purification of SpA binding-optimized variants. Spiking experiments of the variants into cell culture supernatant were performed followed by SpA chromatography using spin columns.

Figure 4: Modular CAR-T cell approach

The figure shows a schematic representation of a universal CAR-T approach based on a modular interaction pair with ALFA tag/NbALFA. A CAR-T cell carries the tag-binding unit NbALFA VHH on its surface. A second binding unit, consisting of the ALFA-tag fused to a tumor antigen-specific ligand (here: CLDN6-specific scFv), can bind to a tumor cell due to its specificity for the tumor antigen and is bound by NbALFA-CAR-T cells, which can lead to specific lysis of the tumor cell.

Figure 5: T cells were genetically modified to stably express the wildtype ModCAR (NbALFA-CAR) or humanized ModCARs (variant 11/variant 13). On day 7, CAR positivity was adjusted to 50 % prior to co-culture with CLDN6⁺ tumor cells in a CAR⁺ effector to target cell ratio (E:T) of

5:1. Either no (0 %) or different dilutions (0.125 %, 0.03 %, 0.5 %, 2 %) of adaptor-containing supernatant were added to the co-culture at time point 0 hours. Surface expression (% of parent or mean fluorescence intensity (MFI)) of activation marker (CD25, CD69) and degranulation marker (CD107a) were analyzed after 24 hours of co-culture via flow cytometry. Mean and standard deviation of technical duplicates are shown.

Figure 6: T cells were genetically modified to stably express the wildtype ModCAR (NbALFA-CAR) or humanized ModCARs (variant 11/variant 13). On day 7, CAR positivity was adjusted to 50 % prior to co-culture with CLDN6⁺ tumor cells in a CAR⁺ effector to target cell ratio (E:T) of 5:1. Either no (0 %) or different dilutions (0.125 %, 0.03 %, 0.5 %, 2 %) of adaptor-containing supernatant were added to the co-culture at time point 0 hours. Cell culture supernatants were collected after 24 hours of co-culture and analyzed for GM-CSF, IFN- γ , IL-2 and TNF- α . Mean and standard deviation of technical duplicates are shown.

Figure 7: T cells were genetically modified to stably express the wildtype ModCAR (NbALFA-CAR) or humanized ModCARs (variant 11/variant 13). On day 7, CAR positivity was adjusted to 50 % prior to co-culture with CLDN6⁺ tumor cells in a CAR⁺ effector to target cell ratio (E:T) of 5:1. Either no (0 %) or different dilutions (0.125 %, 0.5 %, 2 %) of adaptor-containing supernatant were added to the co-culture at time point 0 hours. Tumor cell lysis was determined over 96 hours via xCELLigence. Data were normalized to the Mock (non-genetically modified) T cell control and to the maximum lysis by Triton-X-100. Mean and standard deviation of technical triplicates are shown.

Figure 8: Genetically modified wildtype ModCAR (NbALFA-CAR) T cells and humanized ModCAR (variant 11/variant 13) T cells were adjusted to 70 % CAR positivity and added to GFP⁺ CLDN6⁺ tumor spheroids, generated 24 hours prior to co-culture, in a CAR⁺ effector to target cell ratio of 7:1. Either no (0 %) or different dilutions (0.125 %, 0.5 %, 2 %) of adaptor-containing supernatant were added to the co-culture at time point 0 hours. A second and third CLDN6⁺ tumor spheroid was added at 166 hours and 334 hours, respectively (indicated with * Target). Tumor spheroid lysis was determined over 480 hours via Incucyte. Data were normalized to the Mock (non-genetically modified) T cell control. Mean and standard deviation of technical triplicates are shown.

Figure 9: Genetically modified wildtype ModCAR (NbALFA-CAR) T cells and humanized ModCAR (variant 11/variant 13) T cells were adjusted to 70 % CAR positivity and added to GFP⁺ CLDN6⁻ tumor spheroids, generated 24 hours prior to co-culture, in a CAR⁺ effector to target cell ratio of 7:1. No adaptor-containing supernatant was added to the co-culture. Tumor spheroid lysis and T cell expansion were determined over 168 hours via Incucyte. Data were normalized to the Mock (non-genetically modified) T cell control. Mean and standard deviation of technical triplicates are shown.

Figure 10: CD19⁺, CLDN6⁺ and CLDN18.2⁺ tumor spheroid lysis by adaptor-functionalized ModCAR T cells. Lysis of CD19⁺ (A), CLDN6⁺ (B) or CLDN18.2⁺ (C) spheroids in co-culture with Mock (non-genetically modified) or ModCAR T cells (NbALFA-CAR/variant 13) in an effector to target cell ratio of 5:1 with 0.5 % supernatant containing the matching adaptor. Adaptor was administered at 0 hours. ModCAR-T cells were repetitively challenged with antigen⁺ spheroids at the indicated time points (----Spheroid). Data were normalized to the Mock T cell control and represent mean $\pm$ standard deviation of technical triplicates from two donors.

Figure 11: Schematic in vivo study workflow for functionality comparison of NbALFA-CAR variants. CLDN6⁺ OV90-SC12 tumor in NSG mice was established by subcutaneous tumor cell injection. On day 38 of tumor engraftment 1×10^6 of NbALFA-CAR or variant 13 expressing T cells or corresponding numbers of T cells, derived from human PBMCs, were adoptively transferred intravenously into the mice. 5 μ g, 30 μ g or no Riboadaptor LNPs were administered weekly at the indicated time points and the anti-tumoral activity was assessed. n = 10 mice/group.

Figure 12: Anti-tumoral response of NbALFA-CAR variants upon in vivo adaptor functionalization. Tumor volumes in NbALFA-CAR, variant 13 and control T cell-treated animals after adoptive cell transfer (ACT) and additional treatment without (A), with 5 μ g (B) or 30 μ g (C) Riboadaptor are shown. Vertical lines indicate ACT (solid line) and Riboadaptor LNP (dotted line) treatment. Data are shown for n $\geq$ 4 and presented as mean $\pm$ SEM of all mice/group.

Figure 13: Peripheral CAR T cell expansion of NbALFA-CAR variants during anti-tumoral response. Kinetic of CAR T cell number in the blood of NbALFA-CAR or variant 13 treated mice during the tumor study is shown. CAR T cell expansion is displayed for each individual mouse

receiving no (A), 5 μg (B) or 30 μg (C) Riboadaptor. Mean expansion of the respective treatment group is presented (D, E, F). Vertical dotted lines indicate Riboadaptor LNP treatment.

Detailed description

Although the present disclosure is further described in more detail below, it is to be understood that this disclosure is not limited to the particular methodologies, protocols and reagents described herein as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present disclosure which will be limited only by the appended claims. Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art.

In the following, the elements of the present disclosure will be described in more detail. These elements are listed with specific embodiments, however, it should be understood that they may be combined in any manner and in any number to create additional embodiments. The variously described examples and preferred embodiments should not be construed to limit the present disclosure to only the explicitly described embodiments. This description should be understood to support and encompass embodiments which combine the explicitly described embodiments with any number of the disclosed and/or preferred elements. Furthermore, any permutations and combinations of all described elements in this application should be considered disclosed by the description of the present application unless the context indicates otherwise.

The practice of the present disclosure will employ, unless otherwise indicated, conventional chemistry, biochemistry, pharmaceutical, cell biology, immunology, and recombinant DNA techniques which are explained in the literature in the field.

Throughout this specification and the claims which follow, unless the context requires otherwise, the word "comprise", and variations such as "comprises" and "comprising", will be understood to imply the inclusion of a stated feature, element, member, integer or step or group of features, elements, members, integers or steps but not the exclusion of any other feature, element, member, integer or step or group of features, elements, members, integers or steps. The term "consisting essentially of" limits the scope of a claim or disclosure to the specified features, elements, members, integers, or steps and those that do not materially affect the basic and novel characteristic(s) of the claim or disclosure. The term "consisting of" limits the scope of a claim or disclosure to the specified features, elements, members,

integers, or steps. The term "comprising" encompasses the term "consisting essentially of" which, in turn, encompasses the term "consisting of". Thus, at each occurrence in the present application, the term "comprising" may be replaced with the term "consisting essentially of" or "consisting of". Likewise, at each occurrence in the present application, the term "consisting essentially of" may be replaced with the term "consisting of".

The terms "a", "an" and "the" and similar references used in the context of describing the present disclosure (especially in the context of the claims) are to be construed to cover both the singular and the plural, unless otherwise indicated herein or clearly contradicted by the context.

All methods described herein can be performed in any suitable order unless otherwise indicated herein or otherwise clearly contradicted by the context.

The use of any and all examples, or exemplary language (e.g., "such as"), provided herein is intended merely to better illustrate the present disclosure and does not pose a limitation on the scope of the present disclosure otherwise claimed. No language in the specification should be construed as indicating any non-claimed element essential to the practice of the present disclosure.

The term "optional" or "optionally" as used herein means that the subsequently described event, circumstance or condition may or may not occur, and that the description includes instances where said event, circumstance, or condition occurs and instances in which it does not occur.

Where used herein, "and/or" is to be taken as specific disclosure of each of the two specified features or components with or without the other. For example, "X and/or Y" is to be taken as specific disclosure of each of (i) X, (ii) Y, and (iii) X and Y, just as if each is set out individually herein.

In the context of the present disclosure, the term "about" denotes an interval of accuracy that the person of ordinary skill will understand to still ensure the technical effect of the feature in question. The term typically indicates deviation from the indicated numerical value by $\pm 10\%$, $\pm 5\%$, $\pm 4\%$, $\pm 3\%$, $\pm 2\%$, $\pm 1\%$, $\pm 0.9\%$, $\pm 0.8\%$, $\pm 0.7\%$, $\pm 0.6\%$, $\pm 0.5\%$, $\pm 0.4\%$, $\pm 0.3\%$, $\pm 0.2\%$, $\pm 0.1\%$, $\pm 0.05\%$, and for example $\pm 0.01\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 10\%$. In some embodiments, "about" indicates deviation from

the indicated numerical value by $\pm 5\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 4\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 3\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 2\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 1\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 0.9\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 0.8\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 0.7\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 0.6\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 0.5\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 0.4\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 0.3\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 0.2\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 0.1\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 0.05\%$. In some embodiments, "about" indicates deviation from the indicated numerical value by $\pm 0.01\%$. As will be appreciated by the person of ordinary skill, the specific such deviation for a numerical value for a given technical effect will depend on the nature of the technical effect. For example, a natural or biological technical effect may generally have a larger such deviation than one for a man-made or engineering technical effect.

Recitation of ranges of values herein is merely intended to serve as a shorthand method of referring individually to each separate value falling within the range. Unless otherwise indicated herein, each individual value is incorporated into the specification as if it were individually recited herein.

Several documents are cited throughout the text of this specification. Each of the documents cited herein (including all patents, patent applications, scientific publications, manufacturer's specifications, instructions, etc.), whether supra or infra, are hereby incorporated by reference in their entirety. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

In the following, definitions and embodiments will be provided which apply to all aspects of the present disclosure. Terms which are defined in the following have the meanings as defined unless otherwise indicated. Any undefined terms have their art recognized meanings.

Terms such as "reduce" or "inhibit" as used herein means the ability to cause an overall decrease, for example, of about 5% or greater, about 10% or greater, about 15% or greater, about 20% or greater, about 25% or greater, about 30% or greater, about 40% or greater, about 50% or greater, or about 75% or greater, in the level. The term "inhibit" or similar phrases includes a complete or essentially complete inhibition, i.e., a reduction to zero or essentially to zero.

Terms such as "enhance" as used herein means the ability to cause an overall increase, or enhancement, for example, by at least about 5% or greater, about 10% or greater, about 15% or greater, about 20% or greater, about 25% or greater, about 30% or greater, about 40% or greater, about 50% or greater, about 75% or greater, or about 100% or greater in the level.

"Physiological pH" as used herein refers to a pH of about 7.4. In some embodiments, physiological pH is from 7.3 to 7.5. In some embodiments, physiological pH is from 7.35 to 7.45. In some embodiments, physiological pH is 7.3, 7.35, 7.4, 7.45, or 7.5.

As used in the present disclosure, "% w/v" refers to weight by volume percent, which is a unit of concentration measuring the amount of solute in grams (g) expressed as a percent of the total volume of solution in milliliters (mL).

As used in the present disclosure, "% by weight" refers to weight percent, which is a unit of concentration measuring the amount of a substance in grams (g) expressed as a percent of the total weight of the total composition in grams (g).

As used in the present disclosure, "mol %" is defined as the ratio of the number of moles of one component to the total number of moles of all components, multiplied by 100.

As used in the present disclosure, "mol % of the total lipid" is defined as the ratio of the number of moles of one lipid component to the total number of moles of all lipids, multiplied by 100. In this context, in some embodiments, the term "total lipid" includes lipids and lipid-like material.

The term "ionic strength" refers to the mathematical relationship between the number of different kinds of ionic species in a particular solution and their respective charges. Thus, ionic strength I is represented mathematically by the formula:

$$I = \frac{1}{2} \cdot \sum_i z_i^2 \cdot c_i$$

in which c is the molar concentration of a particular ionic species and z the absolute value of its charge. The sum Σ is taken over all the different kinds of ions (i) in solution.

According to the disclosure, the term "ionic strength" in some embodiments relates to the presence of monovalent ions. Regarding the presence of divalent ions, in particular divalent cations, their concentration or effective concentration (presence of free ions) due to the presence of chelating agents is, in some embodiments, sufficiently low so as to prevent degradation of a nucleic acid. In some embodiments, the concentration or effective concentration of divalent ions is below the catalytic level for hydrolysis of the phosphodiester bonds between nucleotides such as RNA nucleotides. In some embodiments, the concentration of free divalent ions is 20 μM or less. In some embodiments, there are no or essentially no free divalent ions.

"Osmolality" refers to the concentration of a particular solute expressed as the number of osmoles of solute per kilogram of solvent.

The term "lyophilizing" or "lyophilization" refers to the freeze-drying of a substance by freezing it and then reducing the surrounding pressure (*e.g.*, below 15 Pa, such as below 10 Pa, below 5 Pa, or 1 Pa or less) to allow the frozen medium in the substance to sublime directly from the solid phase to the gas phase. Thus, the terms "lyophilizing" and "freeze-drying" are used herein interchangeably.

The term "spray-drying" refers to spray-drying a substance by mixing (heated) gas with a fluid that is atomized (sprayed) within a vessel (spray dryer), where the solvent from the formed droplets evaporates, leading to a dry powder.

The term "reconstitute" relates to adding a solvent such as water to a dried product to return it to a liquid state such as its original liquid state.

The term "recombinant" in the context of the present disclosure means "made through genetic engineering". In some embodiments, a "recombinant object" in the context of the present disclosure is not occurring naturally.

The term "naturally occurring" as used herein refers to the fact that an object can be found in nature. For example, a peptide or nucleic acid that is present in an organism (including viruses) and can be isolated from a source in nature and which has not been intentionally modified by man in the laboratory is naturally occurring. The term "found in nature" means "present in nature" and includes known objects as well as objects that have not yet been discovered and/or isolated from nature, but that may be discovered and/or isolated in the future from a natural source.

As used herein, the terms "room temperature" and "ambient temperature" are used interchangeably herein and refer to temperatures from at least about 15°C, e.g., from about 15°C to about 35°C, from about 15°C to about 30°C, from about 15°C to about 25°C, or from about 17°C to about 22°C. Such temperatures will include 15°C, 16°C, 17°C, 18°C, 19°C, 20°C, 21°C and 22°C.

The term "EDTA" refers to ethylenediaminetetraacetic acid disodium salt. All concentrations are given with respect to the EDTA disodium salt.

The term "cryoprotectant" relates to a substance that is added to a formulation in order to protect the active ingredients during the freezing stages.

The term "lyoprotectant" relates to a substance that is added to a formulation in order to protect the active ingredients during the drying stages.

According to the present disclosure, the term "peptide" refers to substances which comprise about two or more, about 3 or more, about 4 or more, about 6 or more, about 8 or more, about 10 or more, about 13 or more, about 16 or more, about 20 or more, and up to about 50, about 100 or about 150, consecutive amino acids linked to one another via peptide bonds.

The term "polypeptide" refers to large peptides, in particular peptides having at least about 151 amino acids. "Peptides" and "polypeptides" are both protein molecules. Thus, the terms "peptide", "protein" and "polypeptide" are used herein usually as synonyms.

Peptides and polypeptides disclosed herein may comprise a linear or a cyclized peptide sequence.

In some embodiments, the peptides disclosed herein comprises at least one cyclic portion, i.e., a polypeptide chain that contains a circular sequence of bonds that is referred to herein as a "cyclic peptide." The circular sequence can occur through a connection between the

amino and carboxyl ends of the peptide; a connection between the amino end and a side chain; a connection between the carboxyl end and a side chain; or a connection between two side chains including sulfur groups of two cysteine amino acids by forming a disulfide bond, or more complicated arrangements.

In some embodiments, the peptides and polypeptides disclosed herein are composed of naturally occurring amino acids, non-naturally occurring amino acids, amino acid derivatives and non-amino acid components, or a mixture thereof. In some embodiments, the peptides and polypeptides disclosed herein comprise amino acid mimetics and amino acid analogs. In some embodiments, the peptides and polypeptides disclosed herein comprise non-naturally occurring amino acid sequences that are resistant to enzymatic cleavage.

In some embodiments, one or more positions of a peptide or polypeptide disclosed herein are substituted with a non-naturally occurring amino acid. In some embodiments, the substituted amino acid is chemically related to the original residue (e.g., aliphatic, charged, basic, acidic, aromatic, hydrophilic) or an isostere of the original residue.

In its broadest sense, as used herein, the term "amino acid" or "amino acid residue" refers to a compound and/or substance that can be, is, or has been incorporated into a peptide, e.g., through formation of one or more peptide bonds. In some embodiments, the term "amino acid" refers to an α -amino carboxylic acid. In some embodiments, an amino acid has the general structure $\text{H}_2\text{N}-\text{C}(\text{H})(\text{R})-\text{COOH}$. In some embodiments, an amino acid is a naturally-occurring amino acid. In some embodiments, an amino acid is a non-natural amino acid. In some embodiments, an amino acid is a D-amino acid. In some embodiments, an amino acid is an L-amino acid.

"Standard amino acid" refers to any of the twenty standard L-amino acids commonly found in naturally occurring peptides and polypeptides. "Nonstandard amino acid" refers to any amino acid, other than the standard amino acids, regardless of whether it is prepared synthetically or obtained from a natural source. In some embodiments, an amino acid, including a carboxy- and/or amino-terminal amino acid in a peptide or polypeptide, can contain a structural modification as compared with the general structure above. For example, in some embodiments, an amino acid may be modified by methylation, amidation, acetylation, pegylation, glycosylation, phosphorylation, and/or substitution (e.g., of the amino group, the

carboxylic acid group, one or more protons, and/or the hydroxyl group) as compared with the general structure. In some embodiments, such modification may, for example, alter the circulating half-life of a peptide or polypeptide containing the modified amino acid as compared with one containing an otherwise identical unmodified amino acid. In some embodiments, such modification does not significantly alter a relevant activity of a peptide or polypeptide containing the modified amino acid, as compared with one containing an otherwise identical unmodified amino acid. As will be clear from context, in some embodiments, the term "amino acid" may be used to refer to a free amino acid. In some embodiments it may be used to refer to an amino acid residue of a peptide or polypeptide. In some embodiments, an amino acid is selected from the group consisting of: L-alanine (Ala or A); L-arginine (Arg or R); L-asparagine (Asn or N); L-aspartic acid (Asp or D); L-cysteine (Cys or C); L-glutamine (Gln or Q); L-glutamic acid (Glu or E); glycine (Gly or G); L-histidine (His or H); L-isoleucine (Ile or I); L-leucine (Leu or L); L-lysine (Lys or K); L-methionine (Met or M); L-phenylalanine (Phe or F); L-proline (Pro or P); L-serine (Ser or S); L-threonine (Thr or T); L-tryptophan (Trp or W); L-tyrosine (Tyr or Y); and L-valine (Val or V), although modified, synthetic, or rare amino acids such as e.g. taurine, ornithine, selenocysteine, homocystine, hydroxyproline, thioproline, iodo-tyrosine, 3-nitro-tyrosine, ornithine, citrulline, canavanine, 5-hydroxytryptophane, carnosine, cycloleucine, 3,4-dihydroxy phenylalanine, N-acetylcysteine, prolinol, allylglycine or acetidine-2-carboxylic acid may be used as desired.

Generally, amino acids are L-amino acids while D-amino acids are denoted by the prefix "D". The prefix "homo" or "h" designates an α -amino acid that is otherwise similar to one of the common ones, but that contains one more methylene group in the carbon chain.

As used herein, "Orn" means ornithine or 2,5-diaminopentanoic acid, "Dab" means 2,4-diaminobutanoic acid, "Dap" means 2,3-diaminopropanoic acid, "hLys" means 2,7-diaminoheptanoic acid, "hCys" means 2-amino-4-mercaptobutanoic acid, and "Pen" means penicillamine or 2-amino-3-methyl-3-sulfanylbutoic acid.

It may also be possible to include non-peptide linkages and other chemical modification. For example, part or all of the peptide or polypeptide may be synthesized as a peptidomimetic, e.g., a peptoid (see, e.g., Simon et al. (1992) Proc. Natl. Acad. Sci. USA 89:9367-71 and Horwell (1995) Trends Biotechnol. 13:132-4). A peptide or polypeptide may include one or more (e.g.,

all) non-hydrolyzable bonds. Many non-hydrolyzable peptide bonds are known in the art, along with procedures for synthesis of peptides containing such bonds. Exemplary non-hydrolyzable bonds include $-\text{CH}_2\text{NH}-$ reduced amide peptide bonds, $-\text{COCH}_2-$ ketomethylene peptide bonds, $-\text{CH}(\text{CN})\text{NH}-$ (cyanomethylene)amino peptide bonds, $-\text{CH}_2\text{CH}(\text{OH})-$ hydroxyethylene peptide bonds, $-\text{CH}_2\text{O}-$ oxymethylene peptide bonds, and $-\text{CH}_2\text{S}-$ thiomethylene peptide bonds (see e.g., U.S. Pat. No. 6,172,043).

The term "amide" as used herein, represents a group of formula $-\text{NHC(O)}-$.

The term "thioamide" represents a group of formula $-\text{NHC(S)}-$.

As used herein the term "disulfide bond", "disulfide bridge" or "disulfide" includes the covalent bond formed between two sulfur atoms. The amino acid cysteine comprises a thiol group that can form a disulfide bond or bridge with a second thiol group.

The term "ether" refers to a group or compound having an oxygen between two carbon atoms.

The term "thioether" refers to a group or compound having a sulfur between two carbon atoms.

The term "ester" refers a compound derived from an carboxylic acid and an alcohol by linking with formal loss of water the hydroxyl group of the $-\text{C(=O)OH}$ group in the former and a hydroxy group of the latter. Thus, the term refers to the group $-\text{C(O)O}-$.

The term "thioester" refers to the group $-\text{C(O)S}-$ or $-\text{C(S)O}-$.

The term "triazole" refers to chemical compounds that incorporate in their structure any heterocyclic structure having a five-membered ring of two carbon atoms and three nitrogen atoms (e.g., 1,2,3-triazole).

The term "portion" refers to a fraction. With respect to a particular structure such as an amino acid sequence or protein the term "portion" thereof may designate a continuous or a discontinuous fraction of said structure.

The terms "part" and "fragment" are used interchangeably herein and refer to a continuous element. For example, a part of a structure such as an amino acid sequence or protein refers to a continuous element of said structure. When used in context of a composition, the term "part" means a portion of the composition. For example, a part of a composition may be any

portion from 0.1% to 99.9% (such as 0.1%, 0.5%, 1%, 5%, 10%, 50%, 90%, or 99%) of said composition.

"Fragment", with reference to an amino acid sequence (peptide or polypeptide), relates to a part of an amino acid sequence, i.e., a sequence which represents the amino acid sequence shortened at the N-terminus and/or C-terminus. A fragment shortened at the C-terminus (N-terminal fragment) is obtainable, e.g., by translation of a truncated open reading frame that lacks the 3'-end of the open reading frame. A fragment shortened at the N-terminus (C-terminal fragment) is obtainable, e.g., by translation of a truncated open reading frame that lacks the 5'-end of the open reading frame, as long as the truncated open reading frame comprises a start codon that serves to initiate translation. A fragment of an amino acid sequence comprises, e.g., at least 50 %, at least 60 %, at least 70 %, at least 80%, at least 90% of the amino acid residues from an amino acid sequence. A fragment of an amino acid sequence comprises, e.g., at least 6, in particular at least 8, at least 10, at least 12, at least 15, at least 20, at least 30, at least 50, or at least 100 consecutive amino acids from an amino acid sequence. A fragment of an amino acid sequence comprises, e.g., a sequence of up to 8, in particular up to 10, up to 12, up to 15, up to 20, up to 30 or up to 55, consecutive amino acids of the amino acid sequence.

"Variant," as used herein and with reference to an amino acid sequence (peptide or polypeptide), is meant an amino acid sequence that differs from a parent amino acid sequence by virtue of at least one amino acid (e.g., a different amino acid, or a modification of the same amino acid). The parent amino acid sequence may be a naturally occurring or wild type (WT) amino acid sequence, or may be a modified version of a wild type amino acid sequence. In some embodiments, the variant amino acid sequence has at least one amino acid difference as compared to the parent amino acid sequence, e.g., from 1 to about 20 amino acid differences, such as from 1 to about 10 or from 1 to about 5 amino acid differences compared to the parent.

By "wild type" or "WT" or "native" as used herein and with reference to an amino acid sequence (peptide or polypeptide) is meant an amino acid sequence that is found in nature, including allelic variations. A wild type amino acid sequence, peptide or polypeptide has an amino acid sequence that has not been intentionally modified.

For the purposes of the present disclosure, "variants" of an amino acid sequence (peptide or polypeptide) may comprise amino acid insertion variants, amino acid addition variants, amino acid deletion variants and/or amino acid substitution variants. The term "variant" includes all mutants, splice variants, post-translationally modified variants, conformations, isoforms, allelic variants, species variants, and species homologs, in particular those which are naturally occurring. The term "variant" includes, in particular, fragments of an amino acid sequence.

Amino acid insertion variants comprise insertions of single or two or more amino acids in a particular amino acid sequence. In the case of amino acid sequence variants having an insertion, one or more amino acid residues are inserted into a particular site in an amino acid sequence, although random insertion with appropriate screening of the resulting product is also possible. Amino acid addition variants comprise amino- and/or carboxy-terminal fusions of one or more amino acids, such as 1, 2, 3, 5, 10, 20, 30, 50, or more amino acids. Amino acid deletion variants are characterized by the removal of one or more amino acids from the sequence, such as by removal of 1, 2, 3, 5, 10, 20, 30, 50, or more amino acids. The deletions may be in any position of the protein. Amino acid deletion variants that comprise the deletion at the N-terminal and/or C-terminal end of the protein are also called N-terminal and/or C-terminal truncation variants. Amino acid substitution variants are characterized by at least one residue in the sequence being removed and another residue being inserted in its place. Preference is given to the modifications being in positions in the amino acid sequence which are not conserved between homologous peptides or polypeptides and/or to replacing amino acids with other ones having similar properties. In some embodiments, amino acid changes in peptide and polypeptide variants are conservative amino acid changes, i.e., substitutions of similarly charged or uncharged amino acids. A conservative amino acid change involves substitution of one of a family of amino acids which are related in their side chains. Naturally occurring amino acids are generally divided into four families: acidic (aspartate, glutamate), basic (lysine, arginine, histidine), non-polar (alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), and uncharged polar (glycine, asparagine, glutamine, cysteine, serine, threonine, tyrosine) amino acids. Phenylalanine, tryptophan, and tyrosine are sometimes classified jointly as aromatic amino acids. In some embodiments, conservative amino acid substitutions include substitutions within the following groups:

glycine, alanine;
valine, isoleucine, leucine;
aspartic acid, glutamic acid;
asparagine, glutamine;
serine, threonine;
lysine, arginine; and
phenylalanine, tyrosine.

In some embodiments, the degree of similarity, such as identity between a given amino acid sequence and an amino acid sequence which is a variant of said given amino acid sequence will be at least about 60%, 70%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99%. In some embodiments, the degree of similarity or identity is given for an amino acid region which is at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90% or about 100% of the entire length of the reference amino acid sequence. For example, if the reference amino acid sequence consists of 200 amino acids, the degree of similarity or identity is given, e.g., for at least about 20, at least about 40, at least about 60, at least about 80, at least about 100, at least about 120, at least about 140, at least about 160, at least about 180, or about 200 amino acids, in some embodiments continuous amino acids. In some embodiments, the degree of similarity or identity is given for the entire length of the reference amino acid sequence. The alignment for determining sequence similarity, such as sequence identity can be done with art known tools, such as using the best sequence alignment, for example, using Align, using standard settings, preferably EMBOSS::needle, Matrix: Blosum62, Gap Open 10.0, Gap Extend 0.5.

"Sequence similarity" indicates the percentage of amino acids that either are identical or that represent conservative amino acid substitutions. "Sequence identity" between two amino acid sequences indicates the percentage of amino acids that are identical between the sequences. "Sequence identity" between two nucleic acid sequences indicates the percentage of nucleotides that are identical between the sequences.

The terms "% identical" and "% identity" or similar terms are intended to refer, in particular, to the percentage of nucleotides or amino acids which are identical in an optimal alignment

between the sequences to be compared. Said percentage is purely statistical, and the differences between the two sequences may be but are not necessarily randomly distributed over the entire length of the sequences to be compared. Comparisons of two sequences are usually carried out by comparing the sequences, after optimal alignment, with respect to a segment or "window of comparison", in order to identify local regions of corresponding sequences. The optimal alignment for a comparison may be carried out manually or with the aid of algorithms, e.g., the local homology algorithm by Smith and Waterman, 1981, *Adv. App. Math.* 2, 482, the local homology algorithm by Needleman and Wunsch, 1970, *J. Mol. Biol.* 48, 443, the similarity search algorithm by Pearson and Lipman, 1988, *Proc. Natl Acad. Sci. USA* 88, 2444, or with the aid of computer programs using said algorithms (GAP, BESTFIT, FASTA, BLAST P, BLAST N and TFASTA in Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Drive, Madison, Wis.). In some embodiments, percent identity of two sequences is determined using the BLASTN or BLASTP algorithm, as available on the United States National Center for Biotechnology Information (NCBI) website (e.g., at blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch&BLAST_SPEC=blast2seq&LINK_LOC=align2seq). In some embodiments, the algorithm parameters used for BLASTN algorithm on the NCBI website include: (i) Expect Threshold set to 10; (ii) Word Size set to 28; (iii) Max matches in a query range set to 0; (iv) Match/Mismatch Scores set to 1, -2; (v) Gap Costs set to Linear; and (vi) the filter for low complexity regions being used. In some embodiments, the algorithm parameters used for BLASTP algorithm on the NCBI website include: (i) Expect Threshold set to 10; (ii) Word Size set to 3; (iii) Max matches in a query range set to 0; (iv) Matrix set to BLOSUM62; (v) Gap Costs set to Existence: 11 Extension: 1; and (vi) conditional compositional score matrix adjustment.

Percentage identity is obtained by determining the number of identical positions at which the sequences to be compared correspond, dividing this number by the number of positions compared (e.g., the number of positions in the reference sequence) and multiplying this result by 100.

In some embodiments, the degree of similarity or identity is given for a region which is at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90% or about 100% of the entire length of the reference sequence. For example, if the reference

nucleic acid sequence consists of 200 nucleotides, the degree of identity is given for at least about 100, at least about 120, at least about 140, at least about 160, at least about 180, or about 200 nucleotides, in some embodiments continuous nucleotides. In some embodiments, the degree of similarity or identity is given for the entire length of the reference sequence.

Homologous amino acid sequences exhibit according to the disclosure at least 40%, in particular at least 50%, at least 60%, at least 70%, at least 80%, at least 90% and, e.g., at least 95%, at least 98 or at least 99% identity of the amino acid residues.

The amino acid sequence variants described herein may readily be prepared by the skilled person, for example, by recombinant DNA manipulation. The manipulation of DNA sequences for preparing peptides or polypeptides having substitutions, additions, insertions or deletions, is described in detail in *Molecular Cloning: A Laboratory Manual*, 4th Edition, M.R. Green and J. Sambrook et al. (1989), eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor 2012, for example. Furthermore, the peptides, polypeptides and amino acid variants described herein may be readily prepared with the aid of known peptide synthesis techniques such as, for example, by solid phase synthesis and similar methods.

In some embodiments, a fragment or variant of an amino acid sequence (peptide or polypeptide) is a "functional fragment" or "functional variant". The term "functional fragment" or "functional variant" of an amino acid sequence relates to any fragment or variant exhibiting one or more functional properties identical or similar to those of the amino acid sequence from which it is derived, i.e., it is functionally equivalent. With respect to sequences of binding agents such as antibodies, one particular function is one or more binding activities displayed by the amino acid sequence from which the fragment or variant is derived. The term "functional fragment" or "functional variant", as used herein, in particular refers to a variant molecule or sequence that comprises an amino acid sequence that is altered by one or more amino acids compared to the amino acid sequence of the parent molecule or sequence and that is still capable of fulfilling one or more of the functions of the parent molecule or sequence, e.g., binding to a target molecule. In some embodiments, the modifications in the amino acid sequence of the parent molecule or sequence do not significantly affect or alter the characteristics of the molecule or sequence. In different embodiments, the function of the functional fragment or functional variant may be reduced but still significantly present, e.g.,

function of the functional fragment or functional variant may be at least 50%, at least 60%, at least 70%, at least 80%, or at least 90% of the parent molecule or sequence. However, in other embodiments, function of the functional fragment or functional variant may be enhanced compared to the parent molecule or sequence.

An amino acid sequence (peptide or polypeptide) "derived from" a designated amino acid sequence (peptide or polypeptide) refers to the origin of the first amino acid sequence. In some embodiments, the amino acid sequence which is derived from a particular amino acid sequence has an amino acid sequence that is identical, essentially identical or homologous to that particular sequence or a fragment thereof. Amino acid sequences derived from a particular amino acid sequence may be variants of that particular sequence or a fragment thereof. For example, it will be understood by one of ordinary skill in the art that the sequences suitable for use herein may be altered such that they vary in sequence from the naturally occurring or native sequences from which they were derived, while retaining the desirable activity of the native sequences.

In some embodiments, "isolated" means removed (e.g., purified) from the natural state or from an artificial composition, such as a composition from a production process. For example, a nucleic acid, peptide or polypeptide naturally present in a living animal is not "isolated", but the same nucleic acid, peptide or polypeptide partially or completely separated from the coexisting materials of its natural state is "isolated". An isolated nucleic acid, peptide or polypeptide can exist in substantially purified form, or can exist in a non-native environment such as, for example, a host cell.

The term "bind" or "binding" relates to the non-covalent interaction with a target. In some embodiments, the term "bind" or "binding" relates to a specific binding.

The term " k_d " (sec^{-1}), as used herein, refers to the dissociation rate constant of a particular interaction, e.g., antibody-antigen interaction. Said value is also referred to as the k_{off} value.

The term " K_D " (M), as used herein, refers to the dissociation equilibrium constant of a particular interaction, e.g., antibody-antigen interaction.

Generally, the terms "bind" or "binding" and "target" or "targeting" are used interchangeably herein.

The term "genetic modification" or simply "modification" includes the transfection of cells

with nucleic acid. The term "transfection" relates to the introduction of nucleic acids, e.g., DNA and/or RNA, into a cell. For purposes of the present disclosure, the term "transfection" also includes the introduction of a nucleic acid into a cell or the uptake of a nucleic acid by such cell, wherein the cell may be present in a subject, e.g., a patient, or the cell may be *in vitro*, e.g., outside of a patient. Thus, according to the present disclosure, a cell for transfection of a nucleic acid described herein can be present *in vitro* or *in vivo*, e.g. the cell can form part of an organ, a tissue and/or the body of a patient. According to the disclosure, transfection can be transient or stable. For some applications of transfection, it is sufficient if the transfected genetic material is only transiently expressed. RNA can be transfected into cells to transiently express its coded protein. Since the nucleic acid introduced in the transfection process is usually not integrated into the nuclear genome, the foreign nucleic acid will be diluted through mitosis or degraded. Cells allowing episomal amplification of nucleic acids greatly reduce the rate of dilution. If it is desired that the transfected nucleic acid actually remains in the genome of the cell and its daughter cells, a stable transfection must occur. Such stable transfection can be achieved by using virus-based systems or transposon-based systems for transfection, for example. Generally, cells that are genetically modified to express an antigen receptor are stably transfected with nucleic acid encoding the antigen receptor. Generally, cells that are transfected with nucleic acid encoding a docking compound are transiently transfected with nucleic acid encoding the docking compound. RNA can be transfected into cells to transiently express its coded protein.

As used herein, the terms "linked", "fused", or "fusion" are used interchangeably. These terms refer to the joining together of two or more elements or components or domains.

The term "fusion polypeptide" or "fusion protein" as used herein refers to a polypeptide or protein comprising two or more subunits. Preferably, the fusion protein is a translational fusion between the two or more subunits. The translational fusion may be generated by genetically engineering the coding nucleotide sequence for one subunit in a reading frame with the coding nucleotide sequence of a further subunit. Subunits may be interspersed by a linker.

As used herein "endogenous" refers to any material from or produced inside an organism, cell, tissue or system.

As used herein, the term "exogenous" refers to any material introduced from or produced outside an organism, cell, tissue or system.

The term "autologous" is used to describe anything that is derived from the same subject. For example, "autologous transplant" refers to a transplant of tissue or organs derived from the same subject. Such procedures are advantageous because they overcome the immunological barrier which otherwise results in rejection.

The term "allogeneic" is used to describe anything that is derived from different individuals of the same species. Two or more individuals are said to be allogeneic to one another when the genes at one or more loci are not identical.

The term "syngeneic" is used to describe anything that is derived from individuals or tissues having identical genotypes, i.e., identical twins or animals of the same inbred strain, or their tissues.

The term "heterologous" is used to describe something consisting of multiple different elements. As an example, the transfer of one individual's bone marrow into a different individual constitutes a heterologous transplant. A heterologous gene is a gene derived from a source other than the subject.

According to various embodiments of the present disclosure, a nucleic acid encoding a peptide or polypeptide is taken up by or introduced, *i.e.* transfected or transduced, into a cell which cell may be present *in vitro* or in a subject, resulting in expression of said peptide or polypeptide. The cell may, *e.g.*, express the encoded peptide or polypeptide intracellularly (*e.g.* in the cytoplasm and/or in the nucleus), may secrete the encoded peptide or polypeptide, and/or may express it on the surface. In some embodiments, if the encoded peptide or polypeptide is an antigen receptor, the cell expresses the antigen receptor on the cell surface. In some embodiments, if the encoded peptide or polypeptide is a docking compound, the cell secretes the encoded peptide or polypeptide.

According to the present disclosure, terms such as "nucleic acid expressing" and "nucleic acid encoding" or similar terms are used interchangeably herein and with respect to a particular peptide or polypeptide mean that the nucleic acid, if present in the appropriate environment, *e.g.* within a cell, can be expressed to produce said peptide or polypeptide.

The term "expression" as used herein includes the transcription and/or translation of a

particular nucleotide sequence.

In the context of the present disclosure, the term "transcription" relates to a process, wherein the genetic code in a DNA sequence is transcribed into RNA (especially mRNA). Subsequently, the RNA may be translated into peptide or polypeptide.

With respect to RNA, the term "expression" or "translation" relates to the process in the ribosomes of a cell by which a strand of mRNA directs the assembly of a sequence of amino acids to make a peptide or polypeptide.

The term "host cell" includes prokaryotic or eukaryotic cells. Prokaryotic host cells include bacterial cells such as *Escherichia coli* (*E. coli*) or *Bacillus subtilis*. Eukaryotic host cells may be microbial cells, such as yeast cells, such as *Saccharomyces cerevisiae*, *Pichia pastoris*, plant cells, or animal cells, such as insect cells, such as SF9 or High5 insect cells, avian cells, or mammalian cells. Mammalian cells may be derived from any species, such as mouse, rat, rabbit, monkey, ape, or human. In some embodiments, the host cell is a non-human host cell. A host cell may be an immortalized mammalian cell (e.g., HeLa cells or CHO cells) or a primary mammalian cell. A host cell is preferably an isolated host cell.

In some embodiments, in a system comprising multiple components, these components are present in close spatial relationship, e.g. in a preparation comprising the components, in particular a kit. A medical preparation, in particular kit, described herein may comprise instructional material or instructions. As used herein, "instructional material" or "instructions" includes a publication, a recording, a diagram, or any other medium of expression which can be used to communicate the usefulness of the compositions and methods of the present disclosure. The instructional material of the kit of the present disclosure may, for example, be affixed to a container which contains the compositions/formulations of the present disclosure or be shipped together with a container which contains the compositions/formulations. Alternatively, the instructional material may be shipped separately from the container with the intention that the instructional material and the compositions be used cooperatively by the recipient.

The term "average diameter" refers to the mean hydrodynamic diameter of particles as measured by dynamic light scattering (DLS) with data analysis using the so-called cumulant algorithm, which provides as results the so-called Z_{average} with the dimension of a length, and

the polydispersity index (PDI), which is dimensionless (Koppel, D., J. Chem. Phys. 57, 1972, pp 4814-4820, ISO 13321). Here "average diameter", "diameter" or "size" for particles is used synonymously with this value of the Z_{average} .

In some embodiments, the "polydispersity index" is calculated based on dynamic light scattering measurements by the so-called cumulant analysis as mentioned in the definition of the "average diameter". Under certain prerequisites, it can be taken as a measure of the size distribution of an ensemble of nanoparticles.

The term "alkyl" refers to a monoradical of a saturated straight or branched hydrocarbon. Preferably, the alkyl group comprises from 1 to 12 (such as 1 to 10) carbon atoms, *i.e.*, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 carbon atoms, abbreviated as C_{1-12} alkyl, (such as 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 carbon atoms, abbreviated as C_{1-10} alkyl), more preferably 1 to 8 carbon atoms, such as 1 to 6 or 1 to 4 carbon atoms. Exemplary alkyl groups include methyl, ethyl, propyl, isopropyl (also called 2-propyl or 1-methylethyl), butyl, iso-butyl, tert-butyl, n-pentyl, iso-pentyl, sec-pentyl, neo-pentyl, 1,2-dimethyl-propyl, iso-amyl, n-hexyl, iso-hexyl, sec-hexyl, n-heptyl, iso-heptyl, n-octyl, 2-ethyl-hexyl, n-nonyl, n-decyl, n-undecyl, n-dodecyl, and the like. A "substituted alkyl" means that one or more (such as 1 to the maximum number of hydrogen atoms bound to an alkyl group, *e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, 9, or up to 10, such as between 1 to 5, 1 to 4, or 1 to 3, or 1 or 2) hydrogen atoms of the alkyl group are replaced with a substituent other than hydrogen (when more than one hydrogen atom is replaced the substituents may be the same or different). Preferably, the substituent other than hydrogen is a 1st level substituent as specified herein.

The term "alkylene" refers to a diradical of a saturated straight or branched hydrocarbon. Preferably, the alkylene comprises from 1 to 12 (such as 1 to 10) carbon atoms, *i.e.*, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 carbon atoms (such as 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 carbon atoms), more preferably 1 to 8 carbon atoms, such as 1 to 6 or 1 to 4 carbon atoms. Exemplary alkylene groups include methylene, ethylene (*i.e.*, 1,1-ethylene, 1,2-ethylene), propylene (*i.e.*, 1,1-propylene, 1,2-propylene ($-\text{CH}(\text{CH}_3)\text{CH}_2-$), 2,2-propylene ($-\text{C}(\text{CH}_3)_2-$), and 1,3-propylene), the butylene isomers (*e.g.*, 1,1-butylene, 1,2-butylene, 2,2-butylene, 1,3-butylene, 2,3-butylene (*cis* or *trans* or a mixture thereof), 1,4-butylene, 1,1-iso-butylene, 1,2-iso-butylene, and 1,3-iso-butylene), the pentylene isomers (*e.g.*, 1,1-pentylene, 1,2-pentylene, 1,3-pentylene, 1,4-

pentylene, 1,5-pentylene, 1,1-iso-pentylene, 1,1-sec-pentyl, 1,1-neo-pentyl), the hexylene isomers (*e.g.*, 1,1-hexylene, 1,2-hexylene, 1,3-hexylene, 1,4-hexylene, 1,5-hexylene, 1,6-hexylene, and 1,1-isohexylene), the heptylene isomers (*e.g.*, 1,1-heptylene, 1,2-heptylene, 1,3-heptylene, 1,4-heptylene, 1,5-heptylene, 1,6-heptylene, 1,7-heptylene, and 1,1-isohexylene), the octylene isomers (*e.g.*, 1,1-octylene, 1,2-octylene, 1,3-octylene, 1,4-octylene, 1,5-octylene, 1,6-octylene, 1,7-octylene, 1,8-octylene, and 1,1-isooctylene), and the like. The straight alkylene moieties having at least 3 carbon atoms and a free valence at each end can also be designated as a multiple of methylene (*e.g.*, 1,4-butylene can also be called tetramethylene). Generally, instead of using the ending "ylene" for alkylene moieties as specified above, one can also use the ending "diyl" (*e.g.*, 1,2-butylene can also be called butan-1,2-diyl). A "substituted alkylene" means that one or more (such as 1 to the maximum number of hydrogen atoms bound to an alkylene group, *e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, 9, or up to 10, such as between 1 to 5, 1 to 4, or 1 to 3, or 1 or 2) hydrogen atoms of the alkylene group are replaced with a substituent other than hydrogen (when more than one hydrogen atom is replaced the substituents may be the same or different). Preferably, the substituent other than hydrogen is a 1st level substituent as specified herein.

The term "cycloalkylene" represents cyclic non-aromatic versions of "alkylene" and may be saturated or unsaturated. A cycloalkylene is a geminal, vicinal or isolated diradical. In certain embodiments, the cycloalkylene (i) is monocyclic or polycyclic (such as bi- or tricyclic) and/or (ii) is 3- to 14-membered (*i.e.*, 3-, 4-, 5-, 6-, 7-, 8-, 9-, 10-, 11-, 12-, 13-, or 14-membered, such as 3- to 12-membered or 3- to 10-membered). In one embodiment the cycloalkylene is a mono-, bi- or tricyclic 3- to 14-membered (*i.e.*, 3-, 4-, 5-, 6-, 7-, 8-, 9-, 10-, 11-, 12-, 13-, or 14-membered, such as 3- to 12-membered or 3- to 10-membered) cycloalkylene. Generally, instead of using the ending "ylene" for cycloalkylene moieties as specified above, one can also use the ending "diyl" (*e.g.*, 1,2-cyclopropylene can also be called cyclopropan-1,2-diyl). Exemplary cycloalkylene groups include cyclohexylene, cycloheptylene, cyclopropylene, cyclobutylene, cyclobutenylene, cyclopentylene, cyclooctylene, bicyclo[3.2.1]octylene, bicyclo[3.2.2]nonylene, and adamantanylene (*e.g.*, tricyclo[3.3.1.1^{3,7}]decan-2,2-diyl). A "substituted cycloalkylene" means that one or more (such as 1 to the maximum number of hydrogen atoms bound to an cycloalkylene group, *e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, 9, or up to 10, such

as between 1 to 5, 1 to 4, or 1 to 3, or 1 or 2) hydrogen atoms of the cycloalkylene group are replaced with a substituent other than hydrogen (when more than one hydrogen atom is replaced the substituents may be the same or different). Preferably, the substituent other than hydrogen is a 1st level substituent as specified herein.

The term "aryl" refers to a monoradical of an aromatic cyclic hydrocarbon. Preferably, the aryl group contains 3 to 14 (e.g., 5, 6, 7, 8, 9, or 10, such as 5, 6, or 10) carbon atoms which can be arranged in one ring (e.g., phenyl) or two or more condensed rings (e.g., naphthyl). Exemplary aryl groups include cyclopropenyl, cyclopentadienyl, phenyl, indenyl, naphthyl, azulenyl, fluorenyl, anthryl, and phenanthryl. Preferably, "aryl" refers to a monocyclic ring containing 6 carbon atoms or an aromatic bicyclic ring system containing 10 carbon atoms. Preferred examples are phenyl and naphthyl. Aryl does not encompass fullerenes. A "substituted aryl" means that one or more (such as 1 to the maximum number of hydrogen atoms bound to an aryl group, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, or up to 10, such as between 1 to 5, 1 to 4, or 1 to 3, or 1 or 2) hydrogen atoms of the aryl group are replaced with a substituent other than hydrogen (when more than one hydrogen atom is replaced the substituents may be the same or different). Preferably, the substituent other than hydrogen is a 1st level substituent as specified herein.

The term "arylene" refers to a diradical of an aryl as specified above. Preferably, the arylene group contains 3 to 14 carbon atoms which can be arranged in one ring (e.g., phenylene), or two or more condensed rings (e.g., naphthylene). Exemplary arylene groups are derived from cyclopropenyl, cyclopentadienyl, benzene, indene, naphthalene, azulene, fluorene, anthracene, or phenanthracene by removing two hydrogen atoms. Preferably, "arylene" refers to a monocyclic ring containing 6 carbon atoms or an aromatic bicyclic ring system containing 10 carbon atoms. Preferred examples are phenylene and naphthylene. A "substituted arylene" means that one or more (such as 1 to the maximum number of hydrogen atoms bound to an arylene group, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, or up to 10, such as between 1 to 5, 1 to 4, or 1 to 3, or 1 or 2) hydrogen atoms of the arylene group are replaced with a substituent other than hydrogen (when more than one hydrogen atom is replaced the substituents may be the same or different). Preferably, the substituent other than hydrogen is a 1st level substituent as specified herein.

The term "heteroaryl" or "heteroaromatic ring" means an aryl group as defined above in which one or more carbon atoms in the aryl group are replaced by heteroatoms of O, S, or N. Preferably, heteroaryl refers to a five or six-membered aromatic monocyclic ring wherein 1, 2, or 3 carbon atoms are replaced by the same or different heteroatoms of O, N, or S. Alternatively, it means an aromatic bicyclic or tricyclic ring system wherein 1, 2, 3, 4, or 5 carbon atoms are replaced with the same or different heteroatoms of O, N, or S. Preferably, in each ring of the heteroaryl group the maximum number of O atoms is 1, the maximum number of S atoms is 1, and the maximum total number of O and S atoms is 2. Exemplary heteroaryl groups include furanyl, thienyl, oxazolyl, isoxazolyl, oxadiazolyl, pyrrolyl, imidazolyl, pyrazolyl, triazolyl, tetrazolyl, thiazolyl, isothiazolyl, thiadiazolyl, pyridyl, pyrimidinyl, pyrazinyl, triazinyl, benzofuranyl, indolyl, isoindolyl, benzothienyl, 1H-indazolyl, benzimidazolyl, benzoxazolyl, indoxazinyl, benzisoxazolyl, benzothiazolyl, benzisothiazolyl, benzotriazolyl, quinolinyl, isoquinolinyl, benzodiazinyl, quinoxalinyl, quinazolinyl, benzotriazinyl, pyridazinyl, phenoxazinyl, thiazolopyridinyl, pyrrolothiazolyl, phenothiazinyl, isobenzofuranyl, chromenyl, xanthenyl, pyrrolizinyl, indolizinyl, indazolyl, purinyl, quinolizinyl, phthalazinyl, naphthyridinyl, cinnolinyl, pteridinyl, carbazolyl, phenanthridinyl, acridinyl, perimidinyl, phenanthrolinyl, and phenazinyl. Exemplary 5- or 6-membered heteroaryl groups include furanyl, thienyl, oxazolyl, isoxazolyl, oxadiazolyl, pyrrolyl, imidazolyl (e.g., 2-imidazolyl), pyrazolyl, triazolyl, tetrazolyl, thiazolyl, isothiazolyl, thiadiazolyl, pyridyl (e.g., 4-pyridyl), pyrimidinyl, pyrazinyl, triazinyl, and pyridazinyl. A "substituted heteroaryl" means that one or more (such as 1 to the maximum number of hydrogen atoms bound to a heteroaryl group, e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, or up to 10, such as between 1 to 5, 1 to 4, or 1 to 3, or 1 or 2) hydrogen atoms of the heteroaryl group are replaced with a substituent other than hydrogen (when more than one hydrogen atom is replaced the substituents may be the same or different). Preferably, the substituent other than hydrogen is a 1st level substituent as specified herein.

The term "heterocyclyl" or "heterocyclic ring" means a cycloalkyl group as defined above in which from 1, 2, 3, or 4 carbon atoms in the cycloalkyl group are replaced by heteroatoms of oxygen, nitrogen, silicon, selenium, phosphorous, or sulfur, preferably O, S, or N. A heterocyclyl group has preferably 1 or 2 rings containing from 3 to 10, such as 3, 4, 5, 6, or 7,

ring atoms. Preferably, in each ring of the heterocyclyl group the maximum number of O atoms is 1, the maximum number of S atoms is 1, and the maximum total number of O and S atoms is 2. The term "heterocyclyl" is also meant to encompass partially or completely hydrogenated forms (such as dihydro, tetrahydro or perhydro forms) of the above-mentioned heteroaryl groups. Exemplary heterocyclyl groups include morpholinyl, pyrrolidinyl, imidazolidinyl, pyrazolidinyl, piperidinyl (also called piperidyl), piperazinyl, di- and tetrahydrofuranyl, di- and tetrahydrothienyl, di- and tetrahydropyranyl, urotropinyl, lactones, lactams, cyclic imides, and cyclic anhydrides. A "substituted heterocyclyl" means that one or more (such as 1 to the maximum number of hydrogen atoms bound to a heterocyclyl group, *e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, 9, or up to 10, such as between 1 to 5, 1 to 4, or 1 to 3, or 1 or 2) hydrogen atoms of the heterocyclyl group are replaced with a substituent other than hydrogen (when more than one hydrogen atom is replaced the substituents may be the same or different). Preferably, the substituent other than hydrogen is a 1st level substituent as specified herein.

The term "heterocycloalkylene" as used herein means a heterocyclyl group as defined above which contains at least one ring heteroatom (such as those selected from the group consisting of O, S, N, B, Si, and P) and in which one hydrogen atom has been removed resulting in a geminal, vicinal or isolated diradical. In some embodiments, the heteroatoms of the heterocycloalkylene group are selected from the group consisting of O, S, and N. For example, the heterocycloalkylene may be O/S-heterocycloalkylene, such as O-heterocycloalkylene. In some embodiments, in each ring of the heterocycloalkylene group the maximum number of O atoms is 1, the maximum number of S atoms is 1, and the maximum total number of O and S atoms is 2. The heterocycloalkylene may be monocyclic or polycyclic (such as bi- or tricyclic). In some embodiments, the heterocycloalkylene is a mono-, bi- or tricyclic 4- to 14-membered (*i.e.*, 4-, 5-, 6-, 7-, 8-, 9-, 10-, 11-, 12-, 13-, or 14-membered, such as 4- to 12-membered or 4- to 10-membered) heterocycloalkylene. The term "heterocycloalkylene" is also meant to encompass partially or completely hydrogenated forms (such as dihydro, tetrahydro or perhydro forms) of the above-mentioned heteroaryl groups (preferably partially or completely hydrogenated forms of the above-mentioned mono-, bi-, or tricyclic heteroaryl groups) in which one hydrogen atom has been removed from the same carbon atom resulting in a geminal diradical. Thus, in some embodiments, a heterocycloalkylene is saturated or

unsaturated (*i.e.*, it contains one or more double bonds within the ring) but cannot be aromatic. A "substituted heterocycloalkylene" means that one or more (such as 1 to the maximum number of hydrogen atoms bound to a heterocycloalkylene group, *e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, 9, or up to 10, such as between 1 to 5, 1 to 4, or 1 to 3, or 1 or 2) hydrogen atoms of the heterocycloalkylene group are replaced with a substituent other than hydrogen (when more than one hydrogen atom is replaced the substituents may be the same or different). Preferably, the substituent other than hydrogen is a 1st level substituent as specified herein. The expression "partially hydrogenated form" of an unsaturated compound or group as used herein means that part of the unsaturation has been removed by formally adding hydrogen to the initially unsaturated compound or group without removing all unsaturated moieties. The phrase "completely hydrogenated form" of an unsaturated compound or group is used herein interchangeably with the term "perhydro" and means that all unsaturation has been removed by formally adding hydrogen to the initially unsaturated compound or group. For example, partially hydrogenated forms of a 5-membered heteroaryl group (containing 2 double bonds in the ring, such as furan) include dihydro forms of said 5-membered heteroaryl group (such as 2,3-dihydrofuran or 2,5-dihydrofuran), whereas the tetrahydro form of said 5-membered heteroaryl group (*e.g.*, tetrahydrofuran, *i.e.*, THF) is a completely hydrogenated (or perhydro) form of said 5-membered heteroaryl group. Likewise, for a 6-membered heteroaryl group having 3 double bonds in the ring (such as pyridyl), partially hydrogenated forms include di- and tetrahydro forms (such as di- and tetrahydropyridyl), whereas the hexahydro form (such as piperidinyl in case of the heteroaryl pyridyl) is the completely hydrogenated (or perhydro) derivative of said 6-membered heteroaryl group. Consequently, a hexahydro form of an aryl or heteroaryl can only be considered a partially hydrogenated form according to the present disclosure if the aryl or heteroaryl contains at least 4 unsaturated moieties consisting of double and triple bonds between ring atoms.

The term "aromatic" as used in the context of hydrocarbons means that the whole molecule has to be aromatic. For example, if a monocyclic aryl is hydrogenated (either partially or completely) the resulting hydrogenated cyclic structure is classified as cycloalkyl for the purposes of the present disclosure. Likewise, if a bi- or polycyclic aryl (such as naphthyl) is hydrogenated the resulting hydrogenated bi- or polycyclic structure (such as 1,2-

dihydronaphthyl) is classified as cycloalkyl for the purposes of the present disclosure (even if one ring, such as in 1,2-dihydronaphthyl, is still aromatic). A similar distinction is made within the present application between heteroaryl and heterocyclyl. For example, indolinyl, *i.e.*, a dihydro variant of indolyl, is classified as heterocyclyl for the purposes of the present disclosure, since only one ring of the bicyclic structure is aromatic and one of the ring atoms is a heteroatom.

Typical 1st level substituents are preferably selected from the group consisting of C₁₋₃ alkyl, phenyl, halogen, -CF₃, -OH, -OCH₃, -SCH₃, -NH_{2-z}(CH₃)_z, -C(=O)OH, oxo, and -C(=O)OCH₃, wherein z is 0, 1, or 2 and C₁₋₃ alkyl is methyl, ethyl, propyl or isopropyl. Particularly preferred 1st level substituents are selected from the group consisting of methyl, ethyl, propyl, isopropyl, halogen (such as F, Cl, or Br), and -CF₃, such as halogen (*e.g.*, F, Cl, or Br), and -CF₃.

Docking compound

According to the disclosure, an immune effector cell is delivered specifically to a target cell by providing a docking compound with a moiety that binds to a target on target cells, *e.g.*, an antigen on target cells, and a moiety that binds to an antigen receptor which is an integral part of the immune effector cell.

A "docking compound" is used to form a connection, such as a non-covalent connection, between a target cell or an antigen on target cells, and the docking compound. The docking compound may form a connection, such as a non-covalent connection, to an immune effector cell to be delivered to a target cell through an antigen receptor on the immune effector cell. The antigen receptor comprises a binding moiety for binding to the docking compound and forms part of said immune effector cell.

In some embodiments, a docking compound comprises a binding moiety for an antigen on target cells, *e.g.*, a moiety targeting a cell surface antigen on target cells, that is capable of binding to the target of interest, *e.g.*, a cell surface antigen on target cells. A "binding moiety for an antigen on target cells" as used herein relates to the part of the docking compound which binds to an antigen on target cells. Such targeting moieties are typically moieties that have affinity for cell surface targets. These moieties can be any peptide or protein (*e.g.* antibodies or antibody fragments) binding to the antigen on target cells. Particular

embodiments of suitable binding moieties for an antigen on target cells for use herein include cell surface antigen binding moieties, such as antibodies, antibody fragments and DARPs. Other examples of binding moieties for an antigen on target cells are peptides or proteins which bind to a receptor.

A binding moiety for an antigen on target cells preferably binds with high specificity and/or high affinity and the bond with the antigen on target cells is preferably stable within the body. In order to allow specific targeting of antigens on target cells, the binding moiety for an antigen on target cells of the docking compound can comprise compounds including but not limited to antibodies, antibody fragments, e.g. Fab2, Fab, scFV, VHH domains, and other proteins or peptides.

According to some embodiments, the antigen on target cells is a cell surface antigen such as a cancer antigen, and suitable binding moieties for an antigen on target cells include but are not limited to, peptides and polypeptides targeting the cell surface antigen, e.g., antibodies, antibody fragments and DARPs.

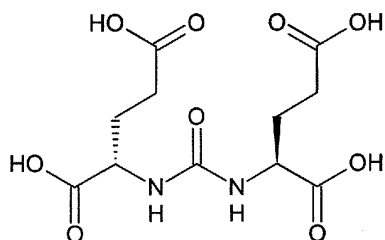
According to some embodiments, the antigen on target cells is a receptor and suitable binding moieties for an antigen on target cells include but are not limited to, the ligand of such a receptor or a part thereof which still binds to the receptor, e.g., a receptor binding peptide in the case of receptor binding protein ligands.

In some embodiments, the binding moiety for an antigen on target cells comprises a compound which is not a protein or peptide. In some embodiments, the binding moiety for an antigen on target cells comprises a small compound.

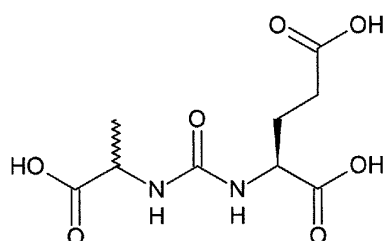
In some embodiments, the binding moiety for an antigen on target cells is a tumor receptor ligand. In some embodiments, the binding moiety for an antigen on target cells is folate, 2-[3-(1,3-dicarboxypropyl)ureido] pentanedioic acid (DUPA), or cholecystokinin 2 receptor (CCK2R) ligand.

For example, DUPA, which belongs to a class of glutamate ureas, may be used as the targeting moiety to selectively deliver immune effector cells to prostate-specific membrane antigen (PSMA) expressing prostate cancer.

The term "DUPA" as used herein means 2-[3-(1,3-dicarboxypropyl)ureido]pentanedioic acid and has the following structure:

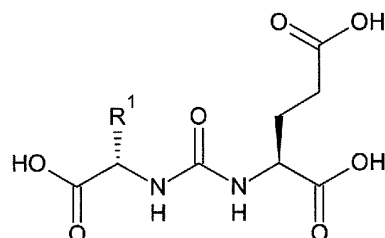


The term "DOPA derivative" as used herein means a compound which comprises a substructure of DOPA and which still binds to prostate-specific membrane antigen (PSMA). Preferably, the DOPA substructure comprises the following formula:



wherein  represents the bond by which the DOPA substructure is attached to the remainder of the DOPA derivative. Preferably, the C atom to which  is attached has S configuration.

In some embodiments, the binding moiety for an antigen on target cells comprises a DOPA derivative. In some embodiments, the DOPA derivative comprises the following formula:



wherein R^1 is $-(L^1)-(L^2)_s-(L^3)_t-H$; wherein each of L^1 , L^2 , and L^3 is a linking moiety; s is 0 or 1; and t is 0 or 1, wherein when s is 0, then t is 0.

In some embodiments, s is 0 and t is 0, i.e., R^1 is $-(L^1)-H$.

In some alternative embodiments, s is 1 and t is 1, i.e., R^1 is $-(L^1)-(L^2)-(L^3)-H$.

In some embodiments, L^1 is selected from the group consisting of -alkylene- $N(R^{10})$ -, -alkylene-O-, -alkylene-C(O)-, -alkylene-C(O)O-, -alkylene-OC(O)-, -alkylene-C(O) $N(R^{10})$ -, and -alkylene- $N(R^{10})C(O)$ -, wherein each R^{10} is independently H or alkyl; and each of the alkylene moieties is optionally substituted with 1, 2, or 3 R^{11} , wherein each R^{11} is independently selected from the

group consisting of OH, halogen, C₁₋₆ alkyl, and =O. In some embodiments, L¹ is selected from the group consisting of -alkylene-N(R¹⁰)-, -alkylene-O-, and -alkylene-C(O)-, wherein each R¹⁰ is independently H or alkyl; and each of the alkylene moieties is optionally substituted with 1, 2, or 3 R¹¹, wherein each R¹¹ is independently selected from the group consisting of OH, halogen, C₁₋₆ alkyl, and =O.

In some embodiments, L¹ is selected from the group consisting of *-alkylene-N(R¹⁰)-, *-alkylene-O-, *-alkylene-C(O)-, *-alkylene-C(O)O-, *-alkylene-OC(O)-, *-alkylene-C(O)N(R¹⁰)-, and *-alkylene-N(R¹⁰)C(O)-, wherein * represents the attachment point to the DUPA substructure; each R¹⁰ is independently H or alkyl; and each of the alkylene moieties is optionally substituted with 1, 2, or 3 R¹¹, wherein each R¹¹ is independently selected from the group consisting of OH, halogen, C₁₋₆ alkyl, and =O. In some embodiments, L¹ is selected from the group consisting of *-alkylene-N(R¹⁰)-, *-alkylene-O-, and *-alkylene-C(O)-, wherein * represents the attachment point to the DUPA substructure; each R¹⁰ is independently H or alkyl; and each of the alkylene moieties is optionally substituted with 1, 2, or 3 R¹¹, wherein each R¹¹ is independently selected from the group consisting of OH, halogen, C₁₋₆ alkyl, and =O.

In some embodiments, each of the alkylene moieties contained in L¹ is independently a C₁₋₁₂ alkylene moiety, preferably a C₂₋₆ alkylene moiety, such as a C₃, C₄ or C₅ alkylene moiety, wherein each of the alkylene moieties is optionally substituted with 1, 2, or 3 R¹¹, wherein each R¹¹ is independently selected from the group consisting of OH, halogen, C₁₋₆ alkyl, and =O (such as from the group consisting of OH, halogen, C₁₋₃ alkyl, and =O). In some embodiments, each of the alkylene moieties contained in L¹ is independently a C₂₋₆ alkylene moiety, such as a C₃, C₄ or C₅ alkylene moiety.

In some embodiments, each of the alkyl moieties of R¹⁰ is independently a C₁₋₁₂ alkyl moiety, such as a C₁₋₆ alkyl moiety or a C₁₋₃ alkyl moiety. In some embodiments, R¹⁰ is H or methyl. In some embodiments, R¹⁰ is H.

In some embodiments, L¹ is derived from any one of the side chains of lysine, glutamic acid, and aspartic acid.

In some embodiments, L¹ is *-tetramethylene-NH-, wherein * represents the attachment point to the DUPA substructure.

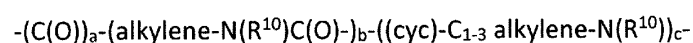
In some embodiments, L^2 is a linking moiety which improves the binding of the DUPA substructure/derivative to PSMA and/or improves the pharmacokinetics of the DUPA derivative. In some embodiments, L^2 comprises at least one aromatic and/or cycloalkylene moiety, such as 1 to 4 aromatic or cycloalkylene moieties. In some embodiments, L^2 comprises 1 to 4 aromatic or cycloalkylene moieties and optionally 1 alkylene moiety. The aromatic and cycloalkylene moieties may be attached to each other by any means, e.g., by a functional group (such as an amino, amide, or ester group) and/or a C_{1-6} alkylene moiety.

In some embodiments, the at least one aromatic moiety is selected from the group consisting of C_{6-12} aromatic moieties (such as phenyl, biphenyl, and naphthyl). In some embodiments, the aromatic group is divalent (such as phenylene). In some embodiments, the aromatic group is monovalent (e.g., biphenyl or naphthyl).

In some embodiments, the at least one cycloalkylene moiety is selected from the group consisting of C_{3-10} cycloalkylene moieties (such as C_5 , C_6 , or C_7 cycloalkylene moieties). In some embodiments, the cycloalkylene moiety is divalent (such as cyclohexylene).

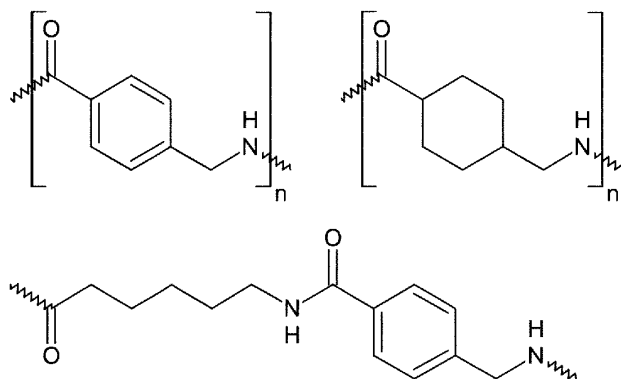
In some embodiments, L^2 comprises the divalent structure $-\text{CH}(\text{CH}_2\text{-arom})-$, wherein arom is a monovalent C_{6-12} aryl group, such as naphthyl or biphenyl. Optionally, L^2 may further comprise 1 or 2 additional divalent moieties independently selected from the group consisting of alkylene, cycloalkylene, and arylene moieties, wherein when one additional divalent moiety is present it can be positioned before or after the $-\text{CH}(\text{CH}_2\text{-arom})-$ structure and wherein when 2 additional divalent moieties are present one is positioned before the $-\text{CH}(\text{CH}_2\text{-arom})-$ structure and the other is positioned after the $-\text{CH}(\text{CH}_2\text{-arom})-$ structure.

In some embodiments, L^2 comprises the formula



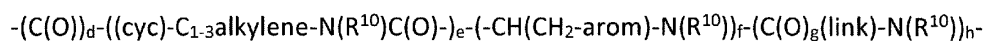
wherein each R^{10} is independently H or alkyl; cyc is cycloalkylene or arylene; a is 0 or 1; b is 0 or 1; and c is 1, 2, 3, or 4. In some embodiments, in particular those, where L^1 ends with an amino (such as $-\text{NH}-$) group, a is 1. In some embodiments, each R^{10} is H or C_{1-6} alkyl, such as H or C_{1-3} alkyl, e.g., H or methyl; preferably each R^{10} is H. In some embodiments, the alkylene moiety in the $\text{alkylene}-\text{N}(\text{R}^{10})\text{C}(\text{O})-$ group is a C_{1-12} alkylene moiety, preferably a C_{2-6} alkylene moiety, such as a C_4 or C_5 alkylene moiety. In some embodiments, cyc is C_{3-10} cycloalkylene (such as C_5 , C_6 , or C_7 cycloalkylene) or C_{6-12} arylene (such as phenylene). In some embodiments,

cyc is cyclohexylene or phenylene. In some embodiments, C₁₋₃ alkylene is methylene, ethylene, or trimethylene, such as methylene. In specific examples of these embodiments, L² comprises one of the following formulas:



wherein each n is independently 1, 2, 3, or 4; and  at the carbonyl end preferably represents the bond by which L² is attached to L¹.

In some embodiments, L² comprises the formula



wherein d is 0 or 1; cyc is cycloalkylene or arylene; each R¹⁰ is independently H or alkyl; e is 0 or 1; arom is a monovalent C₆₋₁₂ aryl group; f is 0 or 1; g is 0 or 1; link is a divalent moiety selected from the group consisting of alkylene, cycloalkylene, arylene, and combinations thereof (such as any two- or three-way combination thereof, e.g., a combination of one arylene moiety and one alkylene moiety or a combination of one cycloalkylene moiety and 2 alkylene moieties); and h is 0 or 1; wherein at least one of e and f is 1; and when e is 1 and f is 0, then g is 0. In some embodiments, in particular those, where L¹ ends with an amino (such as -NH-) group, d is 1. In some embodiments, cyc is C₃₋₁₀ cycloalkylene (such as C₅, C₆, or C₇ cycloalkylene) or C₆₋₁₂ arylene (such as phenylene). In some embodiments, cyc is cyclohexylene or phenylene. In some embodiments, C₁₋₃ alkylene is methylene, ethylene, or trimethylene, such as methylene. In some embodiments, each R¹⁰ is H or C₁₋₆ alkyl, such as H or C₁₋₃ alkyl, e.g., H or methyl; preferably each R¹⁰ is H. In some embodiments, arom is naphthyl or biphenyl. In some embodiments, link is selected from the group consisting of C₁₋₆ alkylene (such as C₂₋₆ alkylene, e.g., C₃, C₄, C₅, or C₆ alkylene), C₃₋₁₀ cycloalkylene (such as C₅, C₆, or C₇ cycloalkylene), C₆₋₁₂ arylene (such as phenylene), and combinations thereof (such as any two-

or three-way combination thereof, e.g., -methylene-cyclohexylene-methylene- or -phenylene-methylene-).

In some embodiments, e is 0, f is 1, and h is 0, and optionally d is 1. In these embodiments, it is preferred that arom is naphthyl (e.g., 2-naphthyl) or biphenyl; and/or R¹⁰ is H or methyl (preferably H).

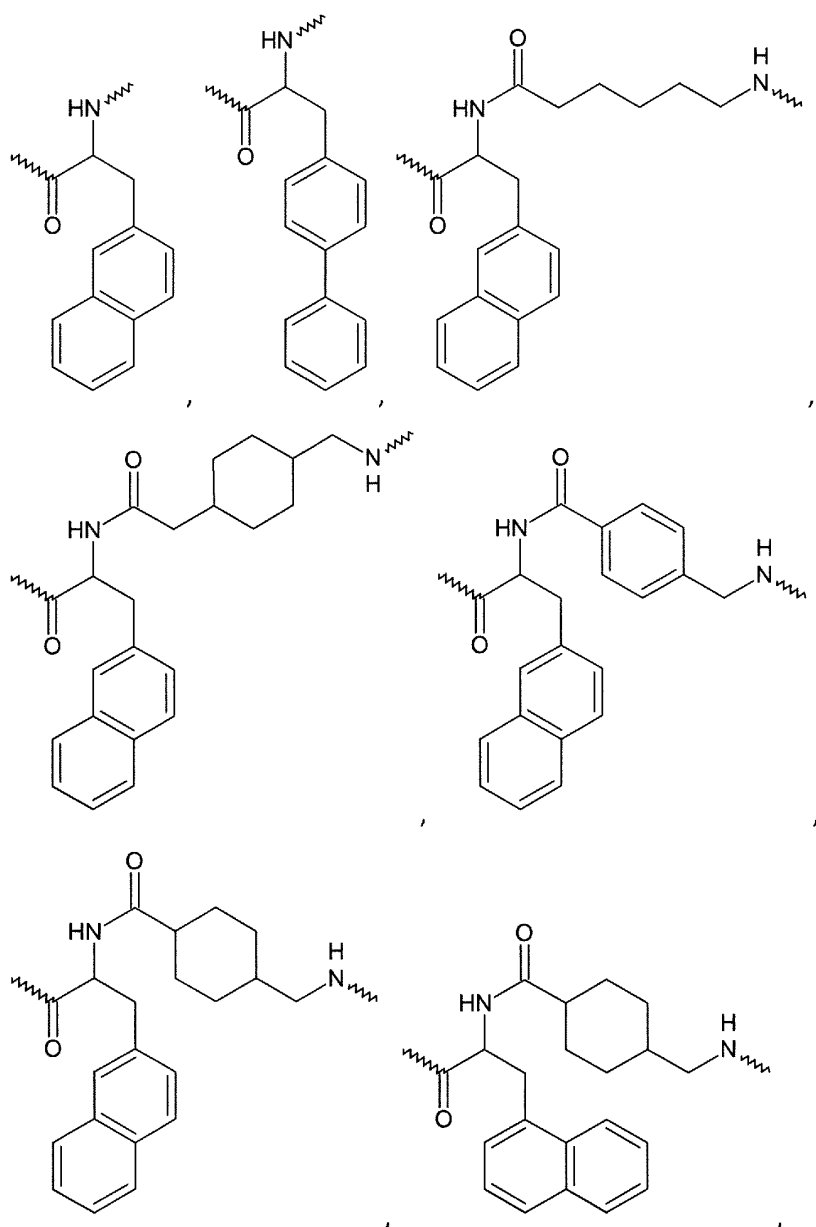
In some embodiments, e is 0, f is 1, g is 1, and h is 1, and optionally d is 1. In these embodiments, it is preferred that (i) arom is naphthyl (e.g., 2-naphthyl) or biphenyl; (ii) each R¹⁰ is H or methyl (preferably H); or (iii) link is selected from the group consisting of C₂₋₆ alkylene (such as C₃, C₄, C₅, or C₆ alkylene), C₃₋₁₀ cycloalkylene (such as C₅, C₆, or C₇ cycloalkylene), C₆₋₁₂ arylene (such as phenylene), and combinations thereof (such as any two- or three-way combination thereof, e.g., -methylene-cyclohexylene-methylene- or -phenylene-methylene-); or combinations of any one of (i) to (iii). Particularly preferred embodiments are the following: (i) and (ii); (i) and (iii); (ii) and (iii) and (i), (ii) and (iii).

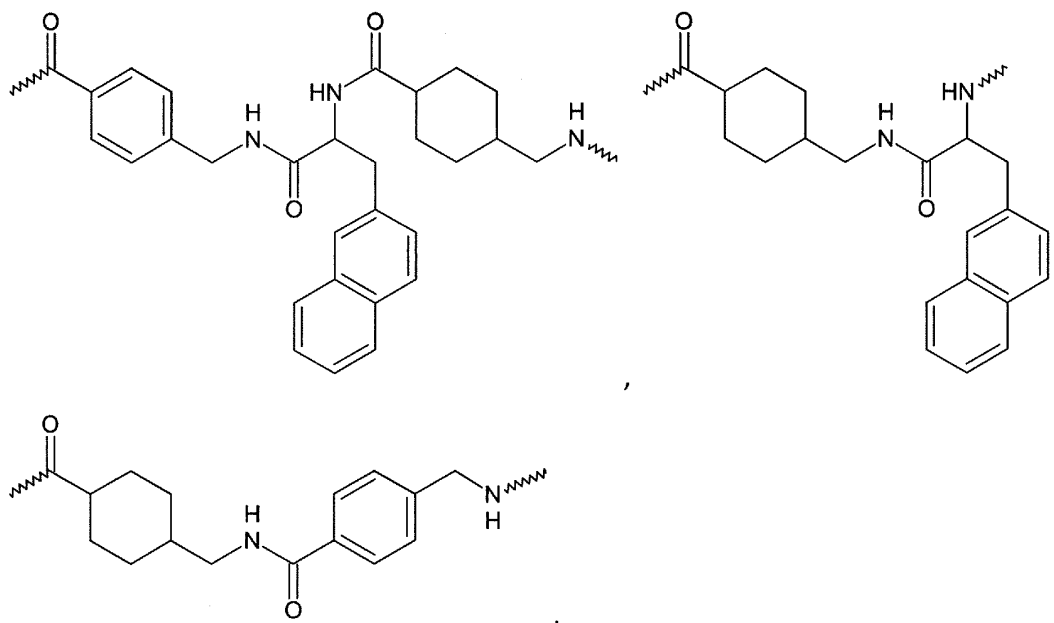
In some embodiments, e is 1, f is 1, g is 1, and h is 1, and optionally d is 1. In these embodiments, it is preferred that (i) cyc is C₃₋₁₀ cycloalkylene or C₆₋₁₂ arylene (preferably cyc is cyclohexylene or phenylene); (ii) C₁₋₃ alkylene is methylene, ethylene, or trimethylene (such as methylene); (iii) each R¹⁰ is H or methyl (preferably H); (iv) arom is naphthyl (e.g., 2-naphthyl) or biphenyl; or (v) link is selected from the group consisting of C₂₋₆ alkylene (such as C₃, C₄, C₅, or C₆ alkylene), C₃₋₁₀ cycloalkylene (such as C₅, C₆, or C₇ cycloalkylene), C₆₋₁₂ arylene (such as phenylene), and combinations thereof (such as any two- or three-way combination thereof, e.g., -methylene-cyclohexylene-methylene- or -phenylene-methylene-); or combinations of any one of (i) to (v). Particularly preferred embodiments are the following: (i) and (ii); (i) and (iii); (i) and (iv); (i) and (v); (ii) and (iii); (ii) and (iv); (ii) and (v); (iii) and (iv); (iii) and (v); (iv) and (v); (i), (ii), and (iii); (i), (ii), and (iv); (i), (ii), and (v); (i), (iii), and (iv); (i), (iii), and (v); (i), (iv), and (v); (ii), (iii), and (iv); (ii), (iii), and (v); (ii), (iv), and (v); and (iii), (iv), and (v).

In some embodiments, e is 1, f is 0, g is 0, and h is 1, and optionally d is 1. In these embodiments, it is preferred that (i) cyc is C₃₋₁₀ cycloalkylene or C₆₋₁₂ arylene (preferably cyc is cyclohexylene or phenylene); (ii) C₁₋₃ alkylene is methylene, ethylene, or trimethylene (such as methylene); (iii) each R¹⁰ is H or methyl (preferably H); or (iv) link is selected from the group consisting of C₂₋₆ alkylene (such as C₃, C₄, C₅, or C₆ alkylene), C₃₋₁₀ cycloalkylene (such as C₅, C₆,

or C₇ cycloalkylene), C₆₋₁₂ arylene (such as phenylene), and combinations thereof (such as any two- or three-way combination thereof, e.g., -methylene-cyclohexylene-methylene- or -phenylene-methylene-); or combinations of any one of (i) to (iv). Particularly preferred embodiments are the following: (i) and (ii); (i) and (iii); (i) and (iv); (ii) and (iii); (ii) and (iv); (iii) and (iv); (i), (ii), and (iii); (i), (ii), and (iv); (i), (iii), and (iv); and (ii), (iii), and (iv).

In specific examples of these embodiments, L² comprises one of the following formulas:





wherein ~~~~ at the carbonyl end preferably represents the bond by which L² is attached to L¹. In some embodiments, the C atom to which the methylene-naphthyl or methylene-biphenyl group is attached has S configuration.

In some embodiments, L³ is present and is -C(O)-alkylene. In some embodiments, the alkylene moiety is C₁₋₁₂ alkylene, such as C₁₋₆ alkylene, e.g., C₂, C₃, C₄, or C₅ alkylene. In some alternative embodiments, L³ is absent.

In some embodiments,

(I) L¹ is selected from the group consisting of *-alkylene-N(R¹⁰)-, *-alkylene-O-, and *-alkylene-C(O)-, wherein * represents the attachment point to the DUPA substructure; each R¹⁰ is independently H or alkyl (preferably H or methyl, more preferably H); and each of the alkylene moieties is optionally substituted with 1, 2, or 3 R¹¹, wherein each R¹¹ is independently selected from the group consisting of OH, halogen, C₁₋₆ alkyl, and =O (preferably each of the alkylene moieties contained in L¹ is independently a C₁₋₁₂ alkylene moiety, more preferably a C₂₋₆ alkylene moiety, such as a C₃, C₄ or C₅ alkylene moiety, wherein the alkylene moiety is optionally substituted with 1, 2, or 3 R¹¹, wherein each R¹¹ is independently selected from the group consisting of OH, halogen, C₁₋₆ alkyl, and =O (such as from the group consisting of OH, halogen, C₁₋₃ alkyl, and =O), more preferably, each of the alkylene moieties contained in L¹ is independently a C₂₋₆ alkylene moiety, such as a C₃, C₄ or C₅ alkylene moiety); and

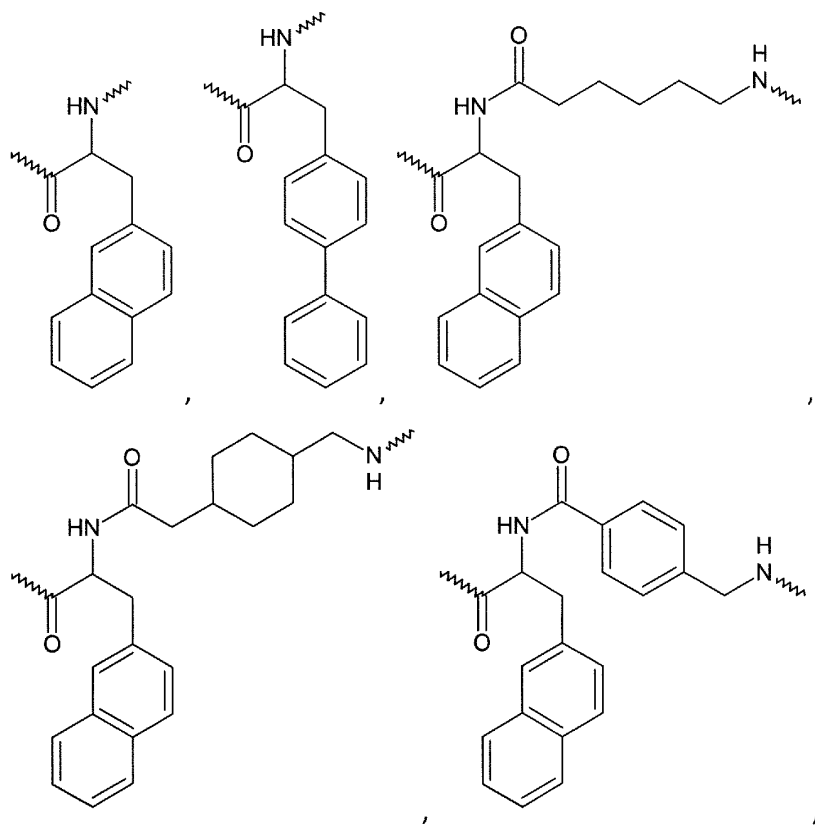
(II) L^2 comprises at least one aromatic and/or cycloalkylene moiety (preferably at least one moiety selected from the group consisting of phenyl, biphenyl, naphthyl, and cyclohexyl), such as the divalent structure $-\text{CH}(\text{CH}_2\text{-arom})-$, wherein arom is a monovalent C_{6-12} aryl group (such as naphthyl or biphenyl); and

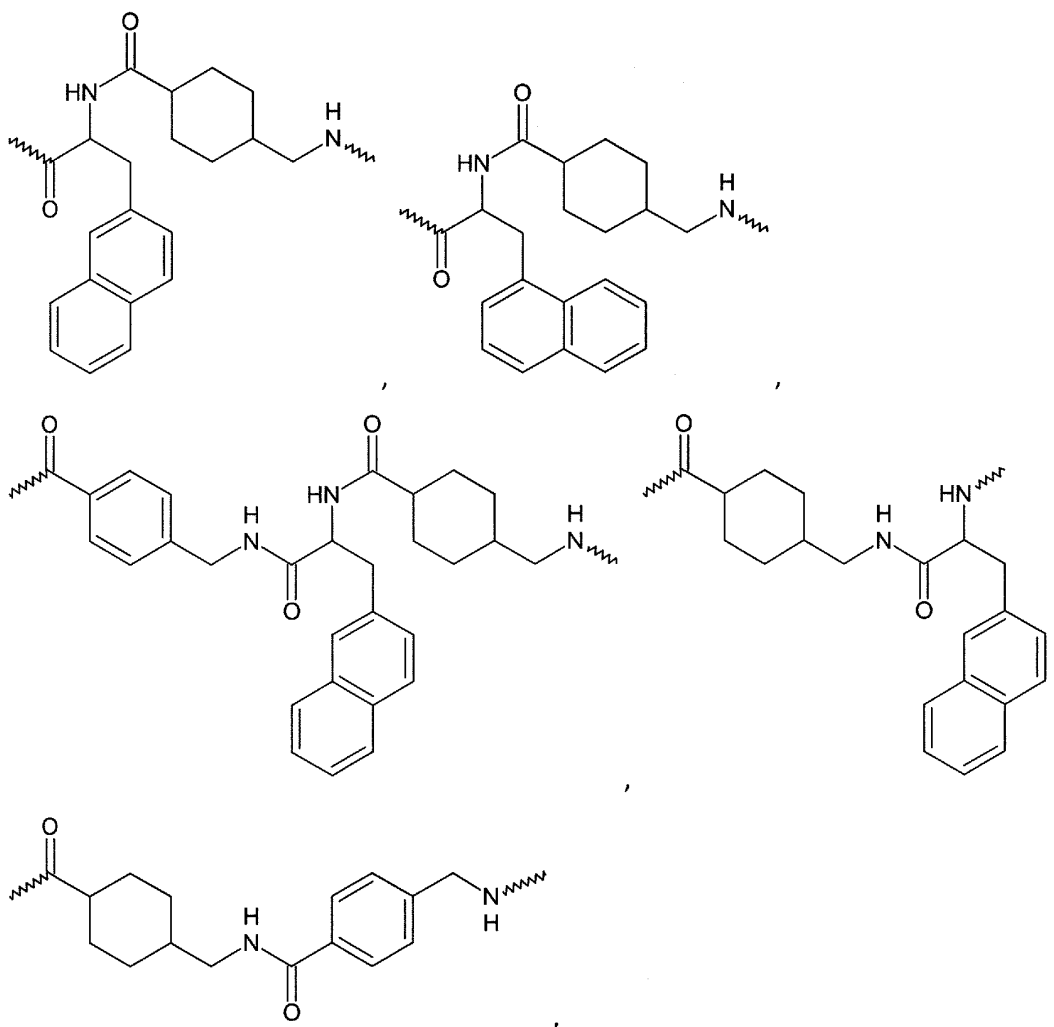
(III) L^3 is present and is $-\text{C}(\text{O})-\text{C}_{1-6}$ alkylene (wherein preferably the alkylene moiety is a C_2 , C_3 , C_4 , or C_5 alkylene moiety) or L^3 is not present.

In some embodiments,

(I) L^1 is derived from any one of the side chains of lysine, glutamic acid, and aspartic acid, preferably L^1 is $^*\text{-tetramethylene-NH-}$, wherein * represents the attachment point to the DUPA substructure;

(II) L^2 comprises one of the following formulas:

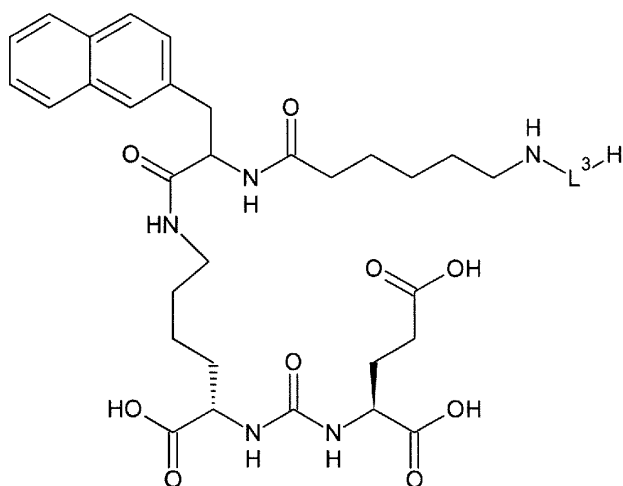
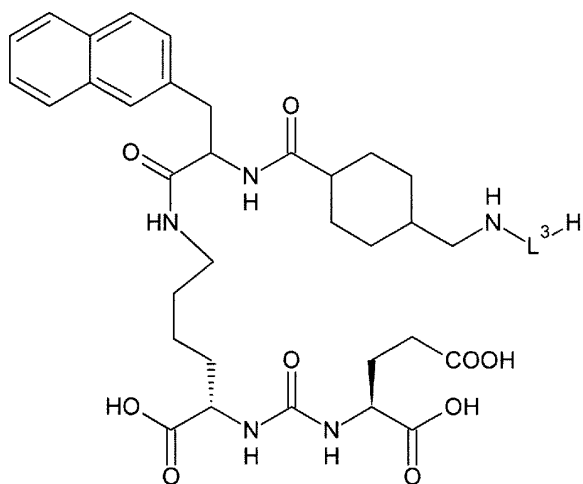
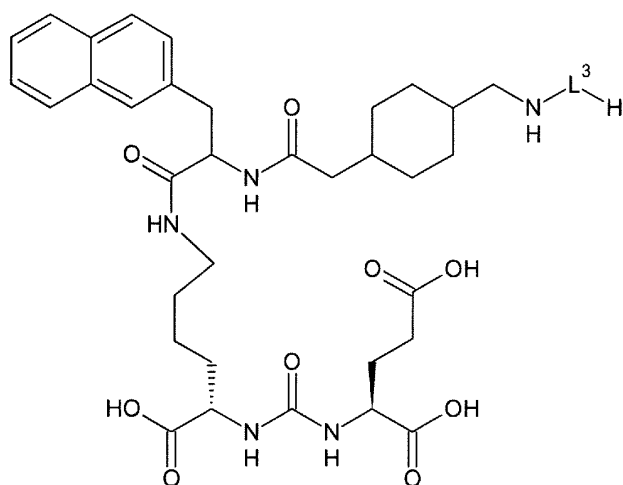


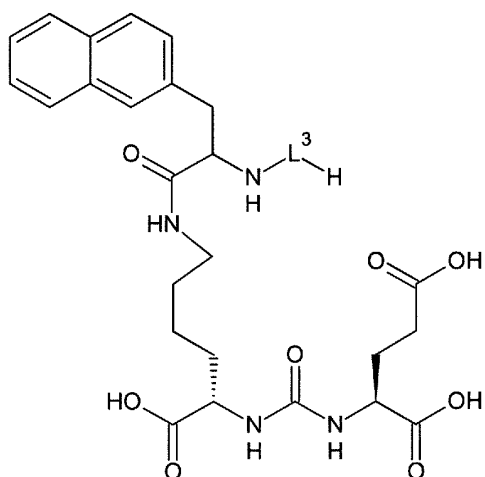


wherein $\sim$ at the carbonyl end represents the bond by which L² is attached to L¹ (preferably, the C atom to which the methylene-naphthyl or methylene-biphenyl group is attached has S configuration); and

(III) L³ is present and is -C(O)-C₁₋₆ alkylene (wherein preferably the alkylene moiety is a C₂, C₃, C₄, or C₅ alkylene moiety) or L³ is not present.

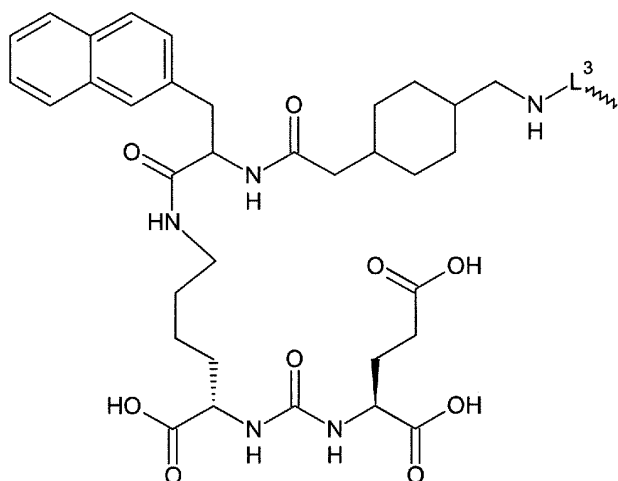
In some preferred embodiments, the DUPA derivative comprises one of the following formulas:





wherein either L^3 is present and is $-C(O)-C_{1-6}$ alkylene (wherein preferably the alkylene moiety is a C_2 , C_3 , C_4 , or C_5 alkylene moiety) or L^3 is absent; and optionally the C atom to which the methylene-naphthyl group is attached has S configuration.

In case the DOPA derivative is (in particular covalently) attached to a different molecule, the DOPA derivative is preferably a DOPA moiety, i.e., a monovalent radical of the DOPA derivative, preferably that in which the terminal hydrogen atom of the R^1 group has been removed (in these embodiments, the DOPA moiety may comprise any of the above formulas, wherein R^1 is $-(L^1)-(L^2)_s-(L^3)_t-$). Thus, in some preferred embodiments, the DOPA moiety comprises one of the following formulas:





wherein either L^3 is present and is $-C(O)-C_{1-6}$ alkylene (wherein preferably the alkylene moiety is a C_2 , C_3 , C_4 , or C_5 alkylene moiety) or L^3 is absent; and optionally the C atom to which the methylene-naphthyl group is attached has S configuration.

It is to be understood that any of the above formulas/structures also includes the salts (in particular pharmaceutically acceptable salts), tautomers, solvates (*e.g.*, hydrates), and isotopically labelled forms thereof.

Fibroblast activation protein (FAP), a type II transmembrane serine protease, is highly expressed in more than 90% of epithelial tumors and is closely associated with various tumor invasion, metastasis, and prognosis. Using FAP as a target, various FAP inhibitors (FAPIs) have been developed, most of which have nanomolar levels of FAP affinity and high selectivity. Such ligands may be used for targeting different tumors.

In some embodiments, the binding moiety for an antigen on target cells comprises a compound that binds to fibroblast activation protein (FAP), *e.g.*, comprises a FAP inhibitor (FAPi).

According to some embodiments, the antigen on target cells and binding moiety for an antigen on target cells are selected so as to result in the specific or increased targeting of certain cells, *e.g.*, diseased cells, such as cells involved in and characteristic for a disease such as cancer, an inflammation, an infection, a cardiovascular disease, *e.g.* thrombus, atherosclerotic lesion, hypoxic site, *e.g.* stroke, tumor, cardiovascular disorder, brain disorder, apoptosis, and angiogenesis. This can be achieved by selecting antigens on target cells with cell-specific expression. For example, cancer antigens, *e.g.*, those described herein, may be expressed in cancer cells while they are not expressed or expressed in a lower amount in normal non-cancerous cells.

The docking compound further comprises a group which serves as a binding partner for an antigen receptor on immune effector cells. The moiety of the docking compound binding to the antigen receptor and the binding moiety for an antigen on target cells are linked to each other, preferably by a covalent linkage.

The antigen receptor and the moiety binding to the antigen receptor of the docking compound bind to each other by means of an ALFA-tag and a single-domain antibody binding to the ALFA-tag.

In some embodiments, the antigen receptor comprises an ALFA-tag and the docking compound comprises a single-domain antibody binding to an ALFA-tag. In some embodiments, the moiety of the antigen receptor binding to the moiety of the docking compound binding to the antigen receptor comprises an ALFA-tag and the moiety of the docking compound binding to the antigen receptor comprises a single-domain antibody binding to an ALFA-tag.

In some embodiments, the antigen receptor comprises a single-domain antibody binding to an ALFA-tag and the docking compound comprises an ALFA-tag. In some embodiments, the moiety of the antigen receptor binding to the moiety of the docking compound binding to the antigen receptor comprises a single-domain antibody binding to an ALFA-tag and the moiety of the docking compound binding to the antigen receptor comprises an ALFA-tag.

In some embodiments, the docking compound comprises a single-domain antibody binding to an ALFA-tag. In some embodiments, the moiety of the docking compound binding to the antigen receptor comprises a single-domain antibody binding to an ALFA-tag. In some embodiments, the docking compound comprises a bispecific molecule such as a bispecific polypeptide, e.g., a bispecific antibody. In some embodiments, the docking compound comprises a single-domain antibody binding to an ALFA-tag and an antibody or antibody fragment binding to an antigen on target cells. In some embodiments, the binding moiety for an antigen on target cells comprises a heavy chain variable region (VH) and a light chain variable region (VL) of an antibody. In some embodiments, the binding moiety for an antigen on target cells comprises a single-domain antibody such as a VHH.

In some embodiments, the docking compound comprises an ALFA-tag. In some embodiments, the moiety of the docking compound binding to the antigen receptor comprises an ALFA-tag. In some embodiments, the docking compound comprises an ALFA-tag and an antibody or antibody fragment binding to an antigen on target cells. In some embodiments, the binding moiety for an antigen on target cells comprises a heavy chain variable region (VH) and a light chain variable region (VL) of an antibody. In some embodiments, the binding moiety for an antigen on target cells comprises a single-domain antibody such as a VHH.

In some embodiments, the docking compound comprises a single peptide chain. In some embodiments, the single peptide chain comprises a portion, e.g., antibody, antibody fragment

or DARPin, binding to an antigen on target cells, and a portion, e.g., single-domain antibody or tag binding to the antigen receptor. In some embodiments, the single peptide chain comprises a portion, e.g., antibody, antibody fragment or DARPin, binding to an antigen on target cells, and a single-domain antibody binding to the antigen receptor. In some embodiments, the single peptide chain comprises a portion, e.g., antibody, antibody fragment or DARPin, binding to an antigen on target cells, and a tag binding to the antigen receptor. In some embodiments, the antibody fragments are VHH, scFv, or a mixture thereof. In different embodiments, the docking compound comprises one of the following structures (from N- to C-terminus):

VHH (α antigen receptor)-optional linker-VHH (α antigen)

VHH (α antigen)-optional linker-VHH (α antigen receptor)

VHH (α antigen receptor)-optional linker-scFv (α antigen)

scFv (α antigen)-optional linker-VHH (α antigen receptor)

α tag-optional linker-VHH (α antigen)

VHH (α antigen)-optional linker- α tag

α tag-optional linker-scFv (α antigen)

scFv (α antigen)-optional linker- α tag

A binding moiety for a CAR and a binding moiety for an antigen on target cells of a docking compound described herein are linked, e.g., through a covalent linkage. In some embodiments, such linkage may comprise a linking moiety. In some embodiments, a binding moiety for a CAR and a binding moiety for an antigen on target cells of a docking compound are linked through a linking moiety comprising a polymer.

In some embodiments, the polymer is not a polymer of proteinogenic amino acids or their D-isomers. In some embodiments, the polymer is a hydrophilic polymer. In some embodiments, the polymer portion of the docking compound contributes to conferring stealth properties on the docking compound. In some embodiments, the plasmatic half-life of the docking compound described herein is greater than 2 hours, e.g., between 3 and 10 hours. This characteristic advantageously allows the docking compound to accumulate at the target cells and to remain at the target cells for reasonable amounts of time. The effectiveness of the targeted delivery described herein therefore increases as a result.

The term "stealth" is used herein to describe the ability of compounds and complexes described herein not to be detected and then sequestered and/or degraded, or to be hardly detected and then sequestered and/or degraded, and/or to be detected and then sequestered and/or degraded late, by the immune system of the host to which they are administered.

Macrophages constitute one of the most important components of the immune system and play a predominant role in eliminating foreign particles, including liposomes and other colloidal particles, from the blood circulation. At the molecular level, the clearance of particles takes place in two steps: opsonization by the depositing of serum proteins (or "opsonins") at the surface of the particles followed by recognition and capture of the opsonized particles by macrophages.

Modification of the surface of particles with chains of hydrophilic and flexible polymers confers them a steric protection by preventing the opsonins reaching the surface of the particles.

In some embodiments, the polymer for use herein is selected from the group consisting of poly(ethylene glycol) (PEG), polysarcosine (pSar) (poly(N-methylglycine), polyoxazoline (POX), polyoxazine (POZ), and poly-2-(2-(2-aminoethoxy)ethoxy)acetic acid (pAEEA) (including derivatives thereof).

In some embodiments, a polymer is designed to sterically stabilize a docking compound by forming a protective hydrophilic layer. In some embodiments, a polymer can reduce association of a docking compound with serum proteins and/or the resulting uptake by the reticuloendothelial system when such docking compounds are administered *in vivo*.

In some embodiments, the PEG is an optionally substituted linear or branched polymer of ethylene glycol or ethylene oxide. In some embodiments, the PEG is unsubstituted. In some embodiments, the PEG is substituted, e.g., by one or more alkyl, alkoxy, acyl, hydroxy or aryl groups. In some embodiments, the PEG has a molecular weight of from about 130 to about 50,000, in another embodiment about 150 to about 30,000, in another embodiment about 150 to about 20,000, in another embodiment about 150 to about 15,000, in another embodiment about 150 to about 10,000, in another embodiment about 150 to about 6000, in another embodiment about 150 to about 5000, in another embodiment about 150 to about 4000, in another embodiment about 150 to about 3000, in another embodiment about 300 to

about 3000, in another embodiment about 1000 to about 3000, and in still another embodiment about 1500 to about 2500.

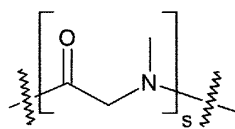
In some embodiments, the PEG moiety has a molecular weight of 1000 or more. In some embodiments, the PEG moiety comprises 10 units or more of formula $(\text{O}-\text{CH}_2-\text{CH}_2)_n$. In some embodiments, the PEG comprises from 20 to 200 ethylene oxide units, such as about 45 ethylene oxide units.

In some embodiments, the PEG comprises "PEG2k", also termed "PEG 2000", which has an average molecular weight of about 2000 Daltons.

In some embodiments, PEG₂₀₀₀, PEG₃₀₀₀ and PEG₅₀₀₀ are used as the polymer.

In some embodiments, a pSar comprises between 2 and 200 sarcosine units, such as between 5 and 100 sarcosine units, between 10 and 50 sarcosine units, between 15 and 40 sarcosine units, e.g., about 23 sarcosine units.

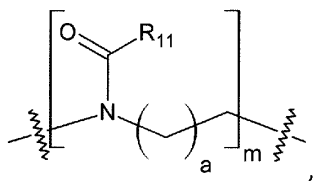
In some embodiments, a pSar comprises the structure of the following general formula:



wherein s is the number of sarcosine units.

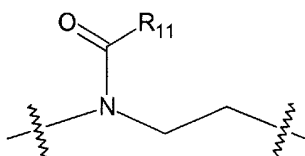
In some embodiments, the POX and/or POZ polymer comprises between 2 and 200, between 2 and 190, between 2 and 180, between 2 and 170, between 2 and 160, between 2 and 150, between 2 and 140, between 2 and 130, between 2 and 120, between 2 and 110, between 2 and 100, between 2 and 90, between 2 and 80, between 2 and 70, between 5 and 200, between 5 and 190, between 5 and 180, between 5 and 170, between 5 and 160, between 5 and 150, between 5 and 140, between 5 and 130, between 5 and 120, between 5 and 110, between 5 and 100, between 5 and 90, between 5 and 80, between 5 and 70, between 10 and 200, between 10 and 190, between 10 and 180, between 10 and 170, between 10 and 160, between 10 and 150, between 10 and 140, between 10 and 130, between 10 and 120, between 10 and 110, between 10 and 100, between 10 and 90, between 10 and 80, or between 10 and 70 POX and/or POZ repeating units.

In some embodiments, the POX and/or POZ polymer comprises the following general formula:

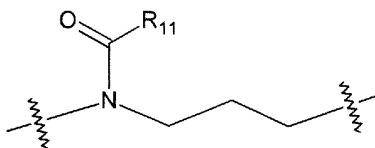


wherein a is an integer between 1 and 2; R_{11} is alkyl, in particular C_{1-3} alkyl, such as methyl, ethyl, iso-propyl, or n -propyl, and is independently selected for each repeating unit; and m refers to the number of POX and/or POZ repeating units.

In some embodiments, the POX and/or POZ polymer is a polymer of POX and comprises repeating units of the following general formula:



In some embodiments, the POX and/or POZ polymer is a polymer of POZ and comprises repeating units of the following general formula:



In any of the above embodiments of formulas, m (*i.e.*, the number of repeating units in the polymer) preferably is between 2 and 190, such as between 2 and 180, between 2 and 170, between 2 and 160, between 2 and 150, between 2 and 140, between 2 and 130, between 2 and 120, between 2 and 110, between 2 and 100, between 2 and 90, between 2 and 80, between 2 and 70, between 5 and 200, between 5 and 190, between 5 and 180, between 5 and 170, between 5 and 160, between 5 and 150, between 5 and 140, between 5 and 130, between 5 and 120, between 5 and 110, between 5 and 100, between 5 and 90, between 5 and 80, between 5 and 70, between 10 and 200, between 10 and 190, between 10 and 180, between 10 and 170, between 10 and 160, between 10 and 150, between 10 and 140, between 10 and 130, between 10 and 120, between 10 and 110, between 10 and 100, between 10 and 90, between 10 and 80, or between 10 and 70. In certain embodiments, m is 2 to 180, such as 4 to 160, 6 to 140, 8 to 120 or 10 to 100, *e.g.*, 20 to 80, 30 to 70, or 40 to 50.

In some embodiments, the POX and/or POZ polymer is a copolymer comprising repeating units of the following general formulas:



wherein the number of repeating units shown on the left in the copolymer is 1 to 199; the number of repeating units of formula on the right in the copolymer is 1 to 199; and the sum of the number of repeating units of formula on the left and the number of repeating units of formula on the right in the copolymer is 2 to 200.

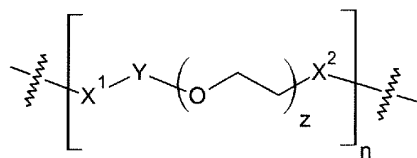
In some embodiments, the number of repeating units of formula on the left in the copolymer is 1 to 179, such as 1 to 159, 1 to 139, 1 to 119 or 1 to 99; the number of repeating units of formula on the right in the copolymer is 1 to 179, such as 1 to 159, 1 to 139, 1 to 119 or 1 to 99; and the sum of the number of repeating units of formula on the left and the number of repeating units of formula on the right in the copolymer is 2 to 180, such as 4 to 160, 6 to 140, 8 to 120 or 10 to 100, *e.g.*, 20 to 80, 30 to 70, or 40 to 50.

In some of the above embodiments, R_{11} at each occurrence (*i.e.*, in each repeating unit) may be the same alkyl group (*e.g.*, R_{11} may be methyl in each repeating unit). In some alternative embodiments, R_{11} in at least one repeating unit differs from R_{11} in another repeating unit (*e.g.*, for at least one repeating unit R_{11} is one specific alkyl (such as ethyl), and for at least one different repeating unit R_{11} is a different specific alkyl (such as methyl)). For example, each R_{11} may be selected from two different alkyl groups (such as methyl and ethyl) and not all R_{11} are the same alkyl.

In any of the above embodiments, R_{11} preferably is methyl or ethyl, more preferably methyl. Thus, in some embodiments, each R_{11} is methyl or each R_{11} is ethyl. In some alternative embodiments, R_{11} is independently selected from methyl and ethyl for each repeating unit, wherein in at least one repeating unit R_{11} is methyl, and in at least one repeating unit R_{11} is ethyl.

In some embodiments, the polymer comprises poly-2-(2-(2-aminoethoxy)ethoxy)acetic acid (pAEEA) or poly-2-(2-(2-methylaminoethoxy)ethoxy)acetic acid (pMAEEA), or a derivative thereof.

In some embodiments, the pAEEA moiety or a derivative thereof comprises the following general formula:



wherein

X² and X¹ taken together are optionally substituted amide, optionally substituted thioamide, ester or thioester;

Y is -CH₂-, -(CH₂)₂-, or -(CH₂)₃-;

z is 2 to 24; and

n is the number of repeating units of 2-(2-(2-aminoethoxy)ethoxy)acetic acid (AEEA) or a derivative thereof, e.g., 1 to 100.

In some embodiments,

(i) when X¹ is -C(O)- then X² is -NR¹-;

(ii) when X¹ is -NR¹- then X² is -C(O)-;

(iii) when X¹ is -C(S)- then X² is -NR¹-;

(iv) when X¹ is -NR¹- then X² is -C(S)-;

(v) when X¹ is -C(O)- then X² is -O-; or

(vi) when X¹ is -O- then X² is -C(O)-;

(vii) when X¹ is -C(S)- then X² is -O-;

(viii) when X¹ is -O- then X² is -C(S)-;

(ix) when X¹ is -C(O)- then X² is -S-; or

(x) when X¹ is -S- then X² is -C(O)-;

wherein R¹ is hydrogen or C₁₋₈ alkyl; preferably

(i) when X¹ is -C(O)- then X² is -NR¹-;

(ii) when X¹ is -NR¹- then X² is -C(O)-;

(iii) when X¹ is -C(S)- then X² is -NR¹-;

(iv) when X¹ is -NR¹- then X² is -C(S)-;

(v) when X¹ is -C(O)- then X² is -O-; or

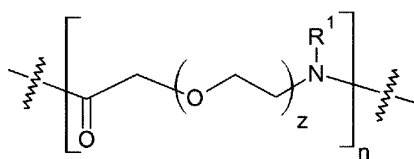
(vi) when X¹ is -O- then X² is -C(O)-;

wherein R^1 is hydrogen or C_{1-8} alkyl.

In some embodiments, X^1 is $-C(O)-$ and X^2 is $-NR^1-$, wherein R^1 is hydrogen or C_{1-8} alkyl. In some embodiments, X^1 is $-C(O)-$ and X^2 is $-NR^1-$, wherein R^1 is hydrogen or methyl. In some embodiments, X^1 is $-C(O)-$ and X^2 is $-NR^1-$, wherein R^1 is hydrogen.

In some embodiments, Y is $-CH_2-$ or $-(CH_2)_2-$. In some embodiments, Y is $-CH_2-$.

In some embodiments, the polymer comprises the following general formula:



wherein

R^1 is hydrogen or C_{1-8} alkyl;

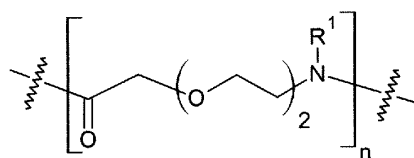
z is 2 to 24; and

n is the number of repeating units of 2-(2-(2-aminoethoxy)ethoxy)acetic acid (AEEA) or a derivative thereof, e.g., 1 to 100.

In some embodiments of the above formulas, z is 2 to 10. In some embodiments, z is 2 to 7.

In some embodiments, z is 2 to 5. In some embodiments, z is 2 or 3. In some embodiments, z is 2.

In some embodiments, the polymer comprises the following general formula:



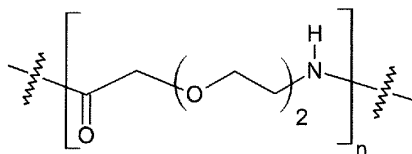
wherein

R^1 is hydrogen or C_{1-8} alkyl; and

n is the number of repeating units of 2-(2-(2-aminoethoxy)ethoxy)acetic acid (AEEA) or a derivative thereof, e.g., 1 to 100.

In some embodiments of the above formulas, R^1 is hydrogen or methyl. In some embodiments, R^1 is hydrogen.

In some embodiments, the polymer comprises the following general formula:



wherein

n is the number of repeating units of 2-(2-(2-aminoethoxy)ethoxy)acetic acid (AEEA), e.g., 1 to 100.

In some embodiments of the above formulas, n is 5 to 50. In some embodiments, n is 5 to 25. In some embodiments, n is 7 to 14. In some embodiments, n is 10 to 25. In some embodiments, n is 14 to 17. In some embodiments, n is 8 or 14.

In some embodiments, the number of repeating units of 2-(2-(2-aminoethoxy)ethoxy)acetic acid (AEEA) or a derivative thereof of all poly-2-(2-(2-aminoethoxy)ethoxy)acetic acid (pAEEA) moieties or derivatives thereof in the chain connecting the binding moiety for a CAR and the binding moiety for an antigen on target cells (main chain) is no more than 100, no more than 90, no more than 80, no more than 70, no more than 60, no more than 50, no more than 40, or no more than 30.

In some embodiments, the number of repeating units of 2-(2-(2-aminoethoxy)ethoxy)acetic acid (AEEA) or a derivative thereof of all poly-2-(2-(2-aminoethoxy)ethoxy)acetic acid (pAEEA) moieties or derivatives thereof in the chain connecting the binding moiety for a CAR and the binding moiety for an antigen on target cells (main chain) is at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, or at least 15.

In some embodiments, the number of repeating units of 2-(2-(2-aminoethoxy)ethoxy)acetic acid (AEEA) or a derivative thereof of all poly-2-(2-(2-aminoethoxy)ethoxy)acetic acid (pAEEA) moieties or derivatives thereof in the chain connecting the binding moiety for a CAR and the binding moiety for an antigen on target cells (main chain) is between 2 and 100, 4 and 80, 4 and 70, 4 and 60, 4 and 50, 4 and 40, or 4 and 30.

In some embodiments, the number of repeating units of 2-(2-(2-aminoethoxy)ethoxy)acetic acid (AEEA) or a derivative thereof of all poly-2-(2-(2-aminoethoxy)ethoxy)acetic acid (pAEEA) moieties or derivatives thereof in all side chains is no more than 150, no more than 140, no more than 130, no more than 120, no more than 110, no more than 100, no more than 90, no

more than 80, no more than 70, no more than 60, no more than 50, no more than 40, or no more than 30.

In some embodiments, the number of repeating units of 2-(2-(2-aminoethoxy)ethoxy)acetic acid (AEEA) or a derivative thereof of all poly-2-(2-(2-aminoethoxy)ethoxy)acetic acid (pAEEA) moieties or derivatives thereof in all side chains is at least 10, at least 12, at least 14, at least 16, at least 18, at least 20, at least 22, at least 24, at least 26, at least 28, or at least 30.

In some embodiments, the number of repeating units of 2-(2-(2-aminoethoxy)ethoxy)acetic acid (AEEA) or a derivative thereof of all poly-2-(2-(2-aminoethoxy)ethoxy)acetic acid (pAEEA) moieties or derivatives thereof in all side chains is between 10 and 150, 20 and 100, 20 and 80, 20 and 70, 30 and 60, or 30 and 50.

In some embodiments, the number of repeating units of 2-(2-(2-aminoethoxy)ethoxy)acetic acid (AEEA) or a derivative thereof of all poly-2-(2-(2-aminoethoxy)ethoxy)acetic acid (pAEEA) moieties or derivatives thereof in the chain connecting the binding moiety for a CAR and the binding moiety for an antigen on target cells (main chain) is between 4 and 60 and the number of repeating units of 2-(2-(2-aminoethoxy)ethoxy)acetic acid (AEEA) or a derivative thereof of all poly-2-(2-(2-aminoethoxy)ethoxy)acetic acid (pAEEA) moieties or derivatives thereof in all side chains is between 20 and 100.

In some embodiments, the number of repeating units of 2-(2-(2-aminoethoxy)ethoxy)acetic acid (AEEA) or a derivative thereof of all poly-2-(2-(2-aminoethoxy)ethoxy)acetic acid (pAEEA) moieties or derivatives thereof in the chain connecting the binding moiety for a CAR and the binding moiety for an antigen on target cells (main chain) is between 4 and 40 and the number of repeating units of 2-(2-(2-aminoethoxy)ethoxy)acetic acid (AEEA) or a derivative thereof of all poly-2-(2-(2-aminoethoxy)ethoxy)acetic acid (pAEEA) moieties or derivatives thereof in all side chains is between 20 and 80.

In some embodiments, the number of repeating units of 2-(2-(2-aminoethoxy)ethoxy)acetic acid (AEEA) or a derivative thereof of all poly-2-(2-(2-aminoethoxy)ethoxy)acetic acid (pAEEA) moieties or derivatives thereof in the chain connecting the binding moiety for a CAR and the binding moiety for an antigen on target cells (main chain) is between 4 and 30 and the number of repeating units of 2-(2-(2-aminoethoxy)ethoxy)acetic acid (AEEA) or a derivative thereof of

all poly-2-(2-(2-aminoethoxy)ethoxy)acetic acid (pAEEA) moieties or derivatives thereof in all side chains is between 20 and 60.

In some embodiments, a docking compound may be provided by administering to a subject nucleic acid encoding the docking compound and allowing expression of the docking compound by cells of the subject. Delivery of nucleic acid encoding a docking compound to cells for expression may be effected by using particles comprising the nucleic acid. The particles may comprise a targeting molecule that binds to a target, e.g., an antigen on target cells for expression. In some embodiments, the docking compound is secreted from the cells expressing the nucleic acid. In some embodiments, the docking compound comprises a signal peptide, e.g., an N-terminal signal peptide, which allows secretion of the docking compound from the cell expressing the nucleic acid. In some embodiments, the cells expressing the nucleic acid encoding a docking compound are the same cells as those to which an immune effector cell is to be delivered herein. In some embodiments, the cells expressing the nucleic acid encoding a docking compound are different to the cells to which an immune effector cell is to be delivered herein. In some embodiments, the cells expressing the nucleic acid encoding a docking compound are liver cells. In some embodiments, the cells expressing the nucleic acid secrete the docking compound into the bloodstream. In some preferred embodiments, the nucleic acid encoding the docking compound is RNA. The RNA-encoded docking compound is also called "RiboDocker" herein.

Interacting moieties on the antigen receptor and on the docking compound

The moieties on the antigen receptor and on the docking compound interacting with each other comprise an ALFA system comprising an ALFA-tag and an ALFA-specific single-domain antibody (sdAb) (NbALFA-nanobody). The ALFA system is an epitope tag/binder system comprising the ALFA-tag and an ALFA-specific single-domain antibody (sdAb), NbALFA-nanobody (Götzke, H. et al, Nat Commun 10, 4403 (2019)).

In some embodiments, the moiety of the antigen receptor on the immune effector cell and the moiety of the docking compound (binding moiety for a CAR) interacting with each other non-covalently bind to each other.

In some embodiments, the moiety of the antigen receptor on the immune effector cell and the moiety of the docking compound (binding moiety for a CAR) interacting with each other bind to each other under physiological conditions.

Humanized antibodies offer the advantage of reduced immunogenicity, as they are engineered to closely resemble human antibodies, minimizing the risk of triggering an immune response when administered to patients. Therefore, humanized antibodies have advantages for *in vivo* applications, such as for imaging techniques. Furthermore, antibodies binding to epitope tags are becoming increasingly interesting for various therapeutic applications (Mitwasi *et al.*, Sci. Rep. 2020; doi:10.1038/s41598-020-59082-4).

Humanization of antibodies can be conducted for classical VH/VL-based molecules as well as their single domain counterparts. One approach that is commonly used is CDR grafting onto humanized acceptor frameworks, in order to retain the spatial orientation of the original variable region from which the donor CDRs are derived. For humanization of VHH regions, the CDRs can be grafted into a compatible human light chain variable region framework. However, sequence grafting can have unexpected and undesired effects on antibody conformation (Fernández-Quintero *et al.*, Prot. Eng. Des. and Sel., 2019; doi: 10.1093/protein/gzaa004). As a result, the provision of a humanized antibody that has at least comparable binding affinity as the original antibody remains challenging.

When the inventors of the present invention attempted to humanize the wt NbALFA single domain antibody, they initially generated seven humanized variants (variants 1-7) by CDR grafting into various humanized acceptor frameworks. However, the resulting single-domain antibodies either could not be expressed or their affinity for the target antigen was completely lost or significantly impaired (see Example 1).

Having realized that CDR grafting does not straightforwardly result in the generation of high-affinity humanized versions of NbALFA, the inventors of the present application have analysed the structure of the antibody in detail and tested the effect of retaining residues that may be relevant for high affinity binding. Thereby, they have unexpectedly found that, in addition to CDR grafting, several key residues of wt NbALFA should be retained in order to arrive at humanized versions of NbALFA, which retain their high affinity for peptide of SEQ ID NO: 2 or which surprisingly have an even improved affinity (see Example 2). Furthermore, the inventors

of the present application have surprisingly found, that so-generated humanized versions of NbALFA can even have an improved thermal stability (see Example 4).

Without wishing to be bound by theory, the inventors of the present application consider that the arginine residue at position 47 of wt NbALFA is important for the interaction of the antibody with the C-terminus of the peptide of SEQ ID NO: 2, and, therefore, retaining Arg 47 of the original camelid VHH framework was considered to improve binding behavior. Furthermore, without wishing to be bound by theory, the inventors of the present application consider that the methionine at position 50 is important for the formation the hydrophobic pocket which would be shielded by the peptide of SEQ ID NO: 2 upon binding, adding free energy to the interaction of antibody and antigen. Thus, retaining Met 50 of the original camelid VHH framework was also considered to improve binding behavior. Further, without wishing to be bound by theory, the inventors of the present application consider that the arginine residue at position 48 of wt NbALFA is also important for binding of the peptide of SEQ ID NO: 2. Hence, retaining Arg 48 of the original camelid VHH framework was also considered to improve binding behavior. These three positions, Arg 47, Arg 48, and Met 50 where thus considered as key residues, which should be retained in a humanized version of NbALFA. In fact, tested variants 8-15 in Example 2 all comprise these three key residues and all variants 8-15 have a binding affinity that is at least comparable to wt NbALFA.

Without wishing to be bound by theory, the inventors of the present application further considered that His 99 and/or Val 100 may be involved in NbALFA binding. These two amino acid residues are believed to influence the conformation of CDR3.

A further residue that is considered to improve the binding of the peptide of SEQ ID NO: 2 is Tyr 40 of wt NbALFA. Without wishing to be bound by theory, the inventors of the present application consider that Tyr 40 forms a hydrogen bond with the R11' side chain on the peptide, which is found in wt NbALFA. By retaining Tyr 40 in a humanized version of NbALFA, antigen binding can be further improved. However, as shown in variant 8, which comprises a Phe at position 40, a binding affinity that is comparable to wt NbALFA can also be achieved without a Tyr residue at position 40. However, by replacing the Phe at position 40 with Tyr, as shown in variant 9, the binding affinity to the peptide of SEQ ID NO: 2 can be further improved.

On the other hand, the inventors of the present application have surprisingly found that Met 61 and Glu 64 which would form part of CDR2 as determined according to Kabat are not essential for binding to the peptide of SEQ ID NO: 2. These residues can be mutated to their human germline counterparts, e.g., to further humanize the antibody, while the high binding affinity to the peptide of SEQ ID NO: 2 can be retained.

Following this rational, the inventors of the present application have created humanized versions of the NbALFA single domain antibody with reduced immunogenic potential and which can be preferably used in *in vivo* applications in human patients, such as diagnostic or therapeutic applications.

Furthermore, the inventors of the present invention have surprisingly found that in addition to their reduced immunogenicity, several humanized versions of NbALFA can be bound by Staphylococcal protein A (SpA; see Example 3). Antibodies that can be bound by SpA have the advantage, that GMP production of the antibody is facilitated. Therefore, the binding affinity towards SpA is a further advantage of some humanized antibodies described herein. Surprisingly, the inventors of the present application found that introducing mutations into CDR2 as determined according to IMGT, in particular the mutation Ala 60 → Thr, and/or into CDR2 as determined according to Kabat, in particular the mutation Gln 67 → Lys, retains the high binding affinity to the peptide of SEQ ID NO: 2, while promoting SpA binding.

Accordingly, the present describes a single-domain antibody (for use in the compositions and methods described herein), that specifically binds to an epitope comprising the sequence SRLEELRRRLTE (SEQ ID NO: 2), wherein the antibody comprises the CDR1 sequence as shown in SEQ ID NO: 94 or a sequence having 1 or 2 mutations relative to said sequence, the CDR2 sequence as shown in SEQ ID NO: 97 or a sequence having 1 or 2 mutations relative to said sequence, and the CDR3 sequence as shown in SEQ ID NO: 98 or a sequence having 1 or 2 mutations relative to said sequence. In some embodiments, the antibody further comprises at positions corresponding to sequence position 47, 48, and/or 50 of SEQ ID NO: 1 at least one of the following amino acid residues: Arg at position 47, Arg at position 48, and/or Met at position 50.

Single-domain antibodies are antibodies whose complementary determining regions are part of a single domain polypeptide. Examples include, but are not limited to, heavy chain

antibodies, antibodies naturally devoid of light chains, single-domain antibodies derived from conventional 4-chain antibodies, engineered antibodies and single domain scaffolds other than those derived from antibodies. Single-domain antibodies may be any of the art, or any future single-domain antibodies. Single-domain antibodies may be derived from any species including, but not limited to mouse, human, camel, llama, goat, rabbit, or cattle. A single-domain antibody as used herein can be derived from a naturally occurring antibody known as heavy chain antibody devoid of light chains. Such single-domain antibodies are disclosed in WO 94/04678 for example. For clarity reasons, this variable domain derived from a heavy chain antibody naturally devoid of light chain is known herein as a VHH to distinguish it from the conventional VH of four chain immunoglobulins. Such a VHH molecule can be derived from antibodies raised in Camelidae species, for example in camel, dromedary, llama, vicuña, alpaca and guanaco. Other species besides Camelidae may produce heavy chain antibodies naturally devoid of light chain. As an illustrative example, it is known that sharks produce heavy chain antibodies naturally devoid of light chains (commonly named IgNAR), which also comprise a VHH domain. Furthermore, some species have been genetically modified to artificially produce heavy chain antibodies naturally devoid of light chain, such as certain mouse strains. In addition, VHHs may be obtained from synthetic libraries. All such VHHs are within the scope of the disclosure. Single-domain antibodies as used herein also includes humanized versions of single-domain antibodies, such as humanized versions of a VHH region.

VHHs, according to the present disclosure, and as known to the skilled addressee are preferably heavy chain variable domains derived from immunoglobulins naturally devoid of light chains such as those derived from Camelidae as described in WO 94/04678 (and referred to hereinafter as VHH domains or nanobodies). VHH molecules are about 10x smaller than IgG molecules. They are single polypeptides and very stable, resisting extreme pH and temperature conditions. Moreover, they are highly resistant to the action of proteases, which is not the case for conventional antibodies. Furthermore, in vitro expression of VHHs or expression in prokaryotic or eukaryotic organisms suited for recombinant protein expression produces high yield, properly folded functional VHHs.

The term "variable" refers to the portions of the immunoglobulin domains that exhibit variability in their sequence and that are involved in determining the specificity and binding

affinity of a particular antibody (i.e., the "variable domain(s)"). Variability is not evenly distributed throughout the variable domains of antibodies; it is concentrated in sub-domains of each of the heavy and light chain variable regions. These sub-domains are called "hypervariable regions", "HVR," or "HV," or "complementarity determining regions" (CDRs). The more conserved (i.e., non-hypervariable) portions of the variable domains are called the "framework" regions (FR). The variable domains of naturally occurring heavy and light chains each include four FR regions, largely adopting a β -sheet configuration, connected by three hypervariable regions, which form loops connecting, and in some cases forming part of, the β -sheet structure. The hypervariable regions in each chain are held together in close proximity by the FR and, with the hypervariable regions from the other chain, contribute to the formation of the antigen-binding site (see Kabat et al., see below). Generally, naturally occurring immunoglobulins include six CDRs (see below); three in the VH (H1, H2, H3), and three in the VL (L1, L2, L3). In naturally occurring immunoglobulins, H3 and L3 display the most diversity of the six CDRs, and H3 in particular is believed to play a unique role in conferring fine specificity to immunoglobulins. Immunoglobulins naturally devoid of light chains, however, include three CDRs that are in the VHH region. The constant domains are not directly involved in antigen binding, but exhibit various effector functions, such as, for example, antibody-dependent, cell-mediated cytotoxicity and complement activation.

Unless otherwise indicated, explicitly or implicitly, CDRs sequences of the present disclosure follow a hybrid Kabat/IMGT definition. Heavy chain CDR1 follows a hybrid Kabat/IMGT definition, the begin of heavy chain CDR1 follows the definition of IMGT, whereas the end of heavy chain CDR1 follows the definition of Kabat. As an illustrative example, the CDR1 of wtNbALFA according to IMGT is GVTISALNAMA (SEQ ID NO: 127), while the CDR1 of wtNbALFA according to Kabat is ALNAMAMG (SEQ ID NO: 128). According to the hybrid Kabat/IMGT definition that is preferably used in the context of the present disclosure, the CDR1 sequence of wtNbALFA is GVTISALNAMAMG (SEQ ID NO: 94). Heavy chain CDR2 and CDR3 are defined according to IMGT definition. Light chain CDR sequences are likewise defined according to IMGT definition. The reason for the hybrid definition of heavy chain CDR1 is to contain a longer sequence to assure that the CDR1 graft includes all important residues. For CDR definition see, generally, e.g., Dondelinger et al., 'Understanding the Significance and Implications of

Antibody Numbering and Antigen-Binding Surface/Residue Definition', 2018; doi: [10.3389/fimmu.2018.02278](https://doi.org/10.3389/fimmu.2018.02278). Other standards for defining CDRs exist as well, such as the definition according to Maass 2007 (Journal of Immunological Methods 324 (2007) 13-25). Another standard for characterizing the antigen binding site is to refer to the hypervariable loops as described by Chothia (see, e.g., Chothia, et al. (1992); J. Mol. Biol. 227:799-817; and Tomlinson et al. (1995) EMBO J. 14:4628-4638).

In line with the CDR definition as described herein, unless otherwise indicated, explicitly or implicitly, the FR regions as used herein follow a hybrid Kabat/IMGT definition. For heavy chain variable regions, FR1, FR3, and FR4 are defined according to IMGT definition. For FR2 of the heavy chain variable region, the beginning is defined according to Kabat definition, while the end is defined according to IMGT definition. Accordingly, as an illustrative example, FR1 of wtNbALFA is preferably defined by amino acid positions 1-25 of SEQ ID NO: 1, FR2 of wtNbALFA is preferably defined by amino acid positions 39-53 of SEQ ID NO: 1, FR3 of wtNbALFA is preferably defined by amino acid positions 61-98 of SEQ ID NO: 1, and FR4 of wtNbALFA is preferably defined by amino acid positions 112-122 of SEQ ID NO: 1.

Each VHH, VH and VL has three CDRs and four FRs, arranged from amino-terminus (N-terminus) to carboxy-terminus (C-terminus) in the following order: FR1, CDR1, FR2, CDR2, FR3, CDR3, and FR4. The variable regions of the heavy and/or light chains contain a binding domain that interacts with an epitope of an antigen.

The term "epitope", also known as the "antigenic determinant", refers to the portion of an antigen to which an antibody specifically binds, thereby forming a complex. Thus, the term "epitope" includes any molecule or protein determinant capable of specific binding to an immunoglobulin or T-cell receptor. The binding site(s) (paratope) of an antibody molecule described herein may specifically bind to/interact with conformational or continuous epitopes, which are unique for the target structure. Epitopic determinants usually consist of chemically active surface groupings of molecules such as amino acids or sugar side chains and usually have specific three-dimensional structural characteristics, as well as specific charge characteristics. Epitope determinants may include chemically active surface groupings of molecules such as amino acids, sugar side chains, phosphoryl, or sulfonyl, and, in certain embodiments, may have specific three-dimensional structural characteristics, and/or specific

charge characteristics. With regard to polypeptide antigens a conformational or discontinuous epitope is characterized by the presence of two or more discrete amino acid residues, separated in the primary sequence, but assembling to a consistent structure on the surface of the molecule when the polypeptide folds into the native protein/antigen (Sela, M., *Science* (1969) 166, 1365-1374; Laver, W.G., et al. *Cell* (1990) 61, 553-556). The two or more discrete amino acid residues contributing to the epitope may be present on separate sections of one or more polypeptide chain(s). These residues come together on the surface of the molecule when the polypeptide chain(s) fold(s) into a three-dimensional structure to constitute the epitope. In contrast, a continuous or linear epitope consists of two or more discrete amino acid residues, which are present in a single linear segment of a polypeptide chain. As an illustrative example, a "context-dependent" CD3 epitope refers to the conformation of said epitope. Such a context-dependent epitope, localized on the epsilon chain of CD3, can only develop its correct conformation if it is embedded within the rest of the epsilon chain and held in the right position by heterodimerization of the epsilon chain with either CD3 gamma or delta chain. In contrast thereto, a context-independent CD3 epitope may be an N-terminal 1-27 amino acid residue polypeptide or a functional fragment thereof of CD3 epsilon. Generally, epitopes can be linear in nature or can be a discontinuous epitope. Thus, as used herein, the term "conformational epitope" refers to a discontinuous epitope formed by a spatial relationship between amino acids of an antigen other than an unbroken series of amino acids. The term "epitope" also includes an antigenic determinant of a hapten, which is known as a small molecule that can serve as an antigen by displaying one or more immunologically recognized epitopes upon binding to larger matter such as a larger molecule e.g. a protein. As used herein, an "epitope tag" refers to a stretch of amino acids to which a specific antibody or proteinaceous molecule with antibody-like function can be raised. Such an epitope tag may allow for specifically identifying and/or tracking of the tagged polypeptide or protein that may be present in a living organism or cultured cells. Detection of the tagged molecule can be achieved using a number of different techniques. Examples of such techniques include: immunohistochemistry, immunoprecipitation, flow cytometry, immunofluorescence microscopy, electron microscopy, ELISA, immunoblotting ("Western blot"), and affinity chromatography. The epitope tag adds a known epitope (antibody binding site) on the subject

polypeptide, to provide binding of a known and often high-affinity antibody. An epitope tag may also be used for isolation and/or purification of the tagged molecule, e.g. by pull-down applications.

The term "mutation" as used herein refers to changes in sequences of biomolecules, such as amino acids sequence of a protein or nucleic acid sequence of DNA or RNA. Mutations include deletions, insertions, and substitutions, and these can apply to single nucleotides/amino acids as well as large sections of nucleotides/amino acid sequences.

As mentioned herein, an antibody described herein comprises at positions corresponding to sequence position 47, 48, and/or 50 of SEQ ID NO: 1 at least one of the following amino acid residues: Arg at position 47, Arg at position 48, and/or Met at position 50.

Preferably, the antibody described herein comprises at the position corresponding to sequence position 47 of SEQ ID NO: 1 an Arg residue. Preferably, the antibody described herein comprises at the position corresponding to sequence position 48 of SEQ ID NO: 1 an Arg residue. Preferably, the antibody described herein comprises at the position corresponding to sequence position 50 of SEQ ID NO: 1 a Met residue. Any one of these residues are believed to contribute to the binding of NbALFA to the target of SEQ ID NO: 2.

Preferably, the antibody described herein comprises at positions corresponding to sequence positions 47, 48, and/or 50 of SEQ ID NO: 1 at least two of the following amino acid residues: Arg at position 47, Arg at position 48, and/or Met at position 50. More preferably, the antibody described herein preferably comprises at positions corresponding to sequence positions 47, 48, and 50 of SEQ ID NO: 1 the following amino acid residues: Arg at position 47, Arg at position 48, and Met at position 50.

Apart from positions 47, 48, and 50, also retaining a Tyr residue at position 40 is considered to improve the affinity of an antibody described herein to SEQ ID NO: 2. While also single-domain antibodies not having a Tyr residue at position 40 may have comparably binding affinity to SEQ ID NO: 2 as compared to wtNbALFA, such as variant 8, other variants having a Tyr at said position may show further improved affinity. Accordingly, an antibody described herein may preferably comprise a Tyr residue at the position corresponding to sequence position 40 of SEQ ID NO: 1. Thus, the antibody described herein preferably comprises at positions corresponding to sequence positions 40, 47, 48, and 50 of SEQ ID NO: 1 the following

amino acid residues: Tyr at position 40, Arg at position 47, Arg at position 48, and Met at position 50.

Further, it is also postulated that the His residue at position 99 and/or Val residue at position 100 may positively influence the binding affinity to SEQ ID NO: 2 of an antibody described herein. Accordingly, an antibody described herein may preferably comprise a His residue at the position corresponding to sequence position 99 of SEQ ID NO: 1. Accordingly, an antibody described herein may preferably comprise a Val residue at the position corresponding to sequence position 100 of SEQ ID NO: 1. Thus, an antibody described herein may preferably comprise a His residue at the position corresponding to sequence position 99 of SEQ ID NO: 1 and a Val residue at the position corresponding to sequence position 100 of SEQ ID NO: 1. Thus, an antibody described herein preferably comprises at positions corresponding to sequence positions 47, 48, 50, 99, and 100 of SEQ ID NO: 1 the following amino acid residues: Arg at position 47, Arg at position 48, Met at position 50, His at position 99, and Val at position 100. Thus, an antibody described herein preferably comprises at positions corresponding to sequence positions 40, 47, 48, 50, 99, and 100 of SEQ ID NO: 1 the following amino acid residues: Tyr at position 40, Arg at position 47, Arg at position 48, Met at position 50, His at position 99, and Val at position 100. In the antibodies described herein, further native amino acids of wtNbALFA may be retained.

An antibody described herein preferably comprises a Val residue at the position corresponding to sequence position 5 of SEQ ID NO: 1. However, without wishing to be bound by theory, position 5 is believed to be of minor importance for binding the peptide of SEQ ID NO: 2. Further, an antibody described herein preferably comprises a Gln residue at the position corresponding to sequence position 117 of SEQ ID NO: 1. However, without wishing to be bound by theory, Gln at position 117 is believed to be of minor importance for binding the peptide of SEQ ID NO: 2, however, Gln 117 is believed to increase solubility. Thus, an antibody described herein preferably comprises a Val residue at the position corresponding to sequence position 5 of SEQ ID NO: 1 and a Gln residue at the position corresponding to sequence position 117 of SEQ ID NO: 1. Preferably, the antibody described herein comprises at positions corresponding to sequence positions 5, 47, 48, 50, and 117 of SEQ ID NO: 1 the following amino acid residues: Val at position 5, Arg at position 47, Arg at position 48, Met at

position 50, and Gln at position 117. Preferably, the antibody described herein comprises at positions corresponding to sequence positions 5, 47, 48, 50, 99, 100, and 117 of SEQ ID NO: 1 the following amino acid residues: Val at position 5, Arg at position 47, Arg at position 48, Met at position 50, His at position 99, Val at position 100, and Gln at position 117. Preferably, the antibody described herein comprises at positions corresponding to sequence positions 5, 40, 47, 48, 50, 99, 100, and 117 of SEQ ID NO: 1 the following amino acid residues: Val at position 5, Tyr at position 40, Arg at position 47, Arg at position 48, Met at position 50, His at position 99, Val at position 100, and Gln at position 117. In wt NbALFA, Met 61 and Glu 64 form part of CDR2 determined according to Kabat. Nevertheless, the inventors of the present application have surprisingly found that these two residues are not crucial for binding of the peptide of SEQ ID NO: 2. By mutating the two residues to the human IGHV3-23 derived sequence counterparts, Tyr 61 and/or Asp 64, high binding affinity to the peptide of SEQ ID NO: 2 can be retained, as shown in variants 13 and 15. An antibody described herein may also comprise a Tyr residue at the position corresponding to sequence position 61 of SEQ ID NO: 1. An antibody described herein may also comprise an Asp residue at the positions corresponding to sequence position 64 of SEQ ID NO: 1. An antibody described herein may also comprises a Tyr residue at the position corresponding to sequence position 61 of SEQ ID NO: 1 and an Asp residue at the position corresponding to sequence position 64 of SEQ ID NO: 1.

A single-domain antibody described herein preferably comprises a sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% sequence identity or being identical to any one of SEQ ID NOs: 118-125. Preferably, an antibody described herein comprises a sequence having at least 90% sequence identity to any one of SEQ ID NOs: 118-125. Preferably, an antibody described herein comprises a sequence having at least 95% sequence identity to any one of SEQ ID NOs: 118-125. Preferably, an antibody described herein comprises a sequence having at least 98% sequence identity to any one of SEQ ID NOs: 118-125. Preferably, an antibody described herein comprises a sequence having at least 99% sequence identity to any one of SEQ ID NOs: 118-125. Preferably, an antibody described herein comprises or consist of an amino acid sequence selected from the group consisting of:

EVQLVESGGGLVQPGGSLRLSCAASGVTISALNAMAMGWFRQAPGKRREMVAAVSERGNAMYRESV
QGRFTISRDNARMVYLQMNSLRAEDTAVYYCHVLEDRVDSFHDYWGQGTQVTVSS (SEQ ID NO:
118);

EVQLVESGGGLVQPGGSLRLSCAASGVTISALNAMAMGWYRQAPGKRREMVAAVSERGNAMYRESV
QGRFTISRDNARMVYLQMNSLRAEDTAVYYCHVLEDRVDSFHDYWGQGTQVTVSS (SEQ ID NO:
119);

EVQLVESGGGLVQPGGSLRLSCAASGVTISALNAMAMGWYRQAPGKRREMVAAVSERGNTMYRESV
QGRFTISRDNARMVYLQMNSLRAEDTAVYYCHVLEDRVDSFHDYWGQGTQVTVSS (SEQ ID NO:
120);

EVQLVESGGGLVQPGGSLRLSCAASGVTISALNAMAMGWYRQAPGKRREMVAAVSERGNAMYRESV
KGRFTISRDNARMVYLQMNSLRAEDTAVYYCHVLEDRVDSFHDYWGQGTQVTVSS (SEQ ID NO:
121);

EVQLVESGGGLVQPGGSLRLSCAASGVTISALNAMAMGWYRQAPGKRREMVAAVSERGNTMYRESVK
GRFTISRDNARMVYLQMNSLRAEDTAVYYCHVLEDRVDSFHDYWGQGTQVTVSS (SEQ ID NO: 122);

EVQLVESGGGLVQPGGSLRLSCAASGVTISALNAMAMGWYRQAPGKRREMVAAVSERGNTYYRDSVK
GRFTISRDNARMVYLQMNSLRAEDTAVYYCHVLEDRVDSFHDYWGQGTQVTVSS (SEQ ID NO: 123);

EVQLVESGGGLVQPGGSLRLSCAASGVTISALNAMAMGWYRQAPGKRREMVAAVSERGNAMYRESV
KGRFTISRDNKNTLYQMNSLRAEDTAVYYCHVLEDRVDSFHDYWGQGTQVTVSS (SEQ ID NO: 124);

and

EVQLVESGGGLVQPGGSLRLSCAASGVTISALNAMAMGWYRQAPGKRREMVAAVSERGNTYYRDSVK
GRFTISRDNKNTLYQMNSLRAEDTAVYYCHVLEDRVDSFHDYWGQGTQVTVSS (SEQ ID NO: 125).

A single-domain antibody described herein preferably comprises a sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% sequence identity or being identical to SEQ ID NO: 118. Preferably, an antibody described herein comprises a sequence having at least 90% sequence identity to SEQ ID NO: 118. Preferably, an antibody described herein comprises a sequence having at least 95% sequence identity to SEQ ID NO: 118. Preferably, an antibody described herein comprises a sequence having at least 98% sequence identity to SEQ ID NO: 118. Preferably, an antibody described herein comprises a sequence having at least 99% sequence identity to SEQ ID NO: 118. Preferably, an antibody described herein comprises or consist of the amino acid sequence set forth in SEQ ID NO: 118.

A single-domain antibody described herein preferably comprises a sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% sequence identity or being identical to SEQ ID NO: 119. Preferably, an antibody described herein comprises a sequence having at least 90% sequence identity to SEQ ID NO: 119. Preferably, an antibody described herein comprises a sequence having at least 95% sequence identity to SEQ ID NO: 119. Preferably, an antibody described herein comprises a sequence having at least 98% sequence identity to SEQ ID NO: 119. Preferably, an antibody described herein comprises a sequence having at least 99% sequence identity to SEQ ID NO: 119. Preferably, an antibody described herein comprises or consist of the amino acid sequence set forth in SEQ ID NO: 119.

A single-domain antibody described herein preferably comprises a sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% sequence identity or being identical to SEQ ID NO: 120. Preferably, an antibody described herein comprises a sequence having at least 90% sequence identity to SEQ ID NO: 120. Preferably, an antibody described herein comprises a sequence having at least 95% sequence identity to SEQ ID NO: 120. Preferably, an antibody described herein comprises a sequence having at least 98% sequence identity to SEQ ID NO: 120. Preferably, an antibody described herein comprises a sequence having at least 99% sequence identity to SEQ ID NO: 120. Preferably, an antibody described herein comprises or consist of the amino acid sequence set forth in SEQ ID NO: 120.

A single-domain antibody described herein preferably comprises a sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% sequence identity or being identical to SEQ ID NO: 121. Preferably, an antibody described herein comprises a sequence having at least 90% sequence identity to SEQ ID NO: 121. Preferably, an antibody described herein comprises a sequence having at least 95% sequence identity to SEQ ID NO: 121. Preferably, an antibody described herein comprises a sequence having at least 98% sequence identity to SEQ ID NO: 121. Preferably, an antibody described herein comprises a sequence having at least 99% sequence identity to SEQ ID NO: 121. Preferably, an antibody described herein comprises or consist of the amino acid sequence set forth in SEQ ID NO: 121.

A single-domain antibody described herein preferably comprises a sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% sequence identity or being identical to SEQ ID NO: 122. Preferably, an antibody described herein comprises a

sequence having at least 90% sequence identity to SEQ ID NO: 122. Preferably, an antibody described herein comprises a sequence having at least 95% sequence identity to SEQ ID NO: 122. Preferably, an antibody described herein comprises a sequence having at least 98% sequence identity to SEQ ID NO: 122. Preferably, an antibody described herein comprises a sequence having at least 99% sequence identity to SEQ ID NO: 122. Preferably, an antibody described herein comprises or consist of the amino acid sequence set forth in SEQ ID NO: 122. A single-domain antibody described herein preferably comprises a sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% sequence identity or being identical to SEQ ID NO: 123. Preferably, an antibody described herein comprises a sequence having at least 90% sequence identity to SEQ ID NO: 123. Preferably, an antibody described herein comprises a sequence having at least 95% sequence identity to SEQ ID NO: 123. Preferably, an antibody described herein comprises a sequence having at least 98% sequence identity to SEQ ID NO: 123. Preferably, an antibody described herein comprises a sequence having at least 99% sequence identity to SEQ ID NO: 123. Preferably, an antibody described herein comprises or consist of the amino acid sequence set forth in SEQ ID NO: 123. A single-domain antibody described herein preferably comprises a sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% sequence identity or being identical to SEQ ID NO: 124. Preferably, an antibody described herein comprises a sequence having at least 90% sequence identity to SEQ ID NO: 124. Preferably, an antibody described herein comprises a sequence having at least 95% sequence identity to SEQ ID NO: 124. Preferably, an antibody described herein comprises a sequence having at least 98% sequence identity to SEQ ID NO: 124. Preferably, an antibody described herein comprises a sequence having at least 99% sequence identity to SEQ ID NO: 124. Preferably, an antibody described herein comprises or consist of the amino acid sequence set forth in SEQ ID NO: 124. A single-domain antibody described herein preferably comprises a sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 98%, or at least 99% sequence identity or being identical to SEQ ID NO: 125. Preferably, an antibody described herein comprises a sequence having at least 90% sequence identity to SEQ ID NO: 125. Preferably, an antibody described herein comprises a sequence having at least 95% sequence identity to SEQ ID NO: 125. Preferably, an antibody described herein comprises a sequence having at least 98%

sequence identity to SEQ ID NO: 125. Preferably, an antibody described herein comprises a sequence having at least 99% sequence identity to SEQ ID NO: 125. Preferably, an antibody described herein comprises or consist of the amino acid sequence set forth in SEQ ID NO: 125. "Percent (%) sequence identity" with respect to amino acid sequences disclosed herein is defined as the percentage of amino acid residues in a candidate sequence that are pair-wise identical with the amino acid residues in a reference sequence, i.e. an antibody molecule of the present disclosure, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, ALIGN, or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximum alignment over the full length of the sequences being compared. The same is true for nucleotide sequences disclosed herein.

As disclosed herein, an antibody described herein may comprise the CDR1 sequence GVTISALNAMAMG (SEQ ID NO: 94) or a sequence having 1 or 2 mutations relative to said sequence, the CDR2 sequence VSERGNX1, wherein X1 is A or T (SEQ ID NO: 97) or a sequence having 1 or 2 mutations relative to said sequence, and the CDR3 sequence HVLEDRVDSFHDY (SEQ ID NO: 98) or a sequence having 1 or 2 mutations relative to said sequence. Preferably, an antibody described herein may comprise the CDR1 sequence GVTISALNAMAMG (SEQ ID NO: 94) or a sequence having 1 mutation relative to said sequence, the CDR2 sequence VSERGNX1, wherein X1 is A or T (SEQ ID NO: 97) or a sequence having 1 mutation relative to said sequence, and the CDR3 sequence HVLEDRVDSFHDY (SEQ ID NO: 98) or a sequence having 1 mutation relative to said sequence. More preferably, an antibody described herein may comprise the CDR1 sequence GVTISALNAMAMG (SEQ ID NO: 94), the CDR2 sequence VSERGNX1, wherein X1 is A or T (SEQ ID NO: 97), and the CDR3 sequence HVLEDRVDSFHDY (SEQ ID NO: 98).

An antibody described herein preferably comprises a CDR2 sequence having at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or being identical to any one of SEQ

ID NO: 95 or 96. Such an antibody described herein may comprise the CDR2 sequence VSEGRNA (SEQ ID NO: 95) or a sequence having 1 or 2 mutations relative to said sequence, or the CDR2 sequence VSEGRNT (SEQ ID NO: 96) or a sequence having 1 or 2 mutations relative to said sequence. Preferably, an antibody described herein comprises the CDR2 sequence of SEQ ID NO: 95 or a sequence having 1 mutation relative to said sequence, or the CDR2 sequence of SEQ ID NO: 96 or a sequence having 1 mutation relative to said sequence. Preferably, an antibody described herein comprises the CDR2 sequence of SEQ ID NO: 95 or SEQ ID NO: 96.

An antibody described herein may comprise the CDR1 sequence set forth in SEQ ID NO: 94 or a sequence having 1 or 2 mutations relative to said sequence, the CDR2 sequence set forth in SEQ ID NO: 95 or a sequence having 1 or 2 mutations relative to said sequence, and the CDR3 sequence set forth in SEQ ID NO: 98 or a sequence having 1 or 2 mutations relative to said sequence. Preferably, an antibody described herein may comprise the CDR1 sequence set forth in SEQ ID NO: 94 or a sequence having 1 mutation relative to said sequence, the CDR2 sequence set forth in SEQ ID NO: 95 or a sequence having 1 mutation relative to said sequence, and the CDR3 sequence set forth in SEQ ID NO: 98 or a sequence having 1 mutation relative to said sequence. More preferably, an antibody described herein may comprise the CDR1 sequence set forth in SEQ ID NO: 94, the CDR2 sequence set forth in SEQ ID NO: 95, and the CDR3 sequence set forth in SEQ ID NO: 98.

An antibody described herein may comprise the CDR1 sequence set forth in SEQ ID NO: 94 or a sequence having 1 or 2 mutations relative to said sequence, the CDR2 sequence set forth in SEQ ID NO: 96 or a sequence having 1 or 2 mutations relative to said sequence, and the CDR3 sequence set forth in SEQ ID NO: 98 or a sequence having 1 or 2 mutations relative to said sequence. Preferably, an antibody described herein may comprise the CDR1 sequence set forth in SEQ ID NO: 94 or a sequence having 1 mutation relative to said sequence, the CDR2 sequence set forth in SEQ ID NO: 96 or a sequence having 1 mutation relative to said sequence, and the CDR3 sequence set forth in SEQ ID NO: 98 or a sequence having 1 mutation relative to said sequence. More preferably, an antibody described herein may comprise the CDR1 sequence set forth in SEQ ID NO: 94, the CDR2 sequence set forth in SEQ ID NO: 96, and the CDR3 sequence set forth in SEQ ID NO: 98.

It is understood that such an antibody is a single-domain antibody. It is further understood that an antibody described herein in which 1 or 2 mutations have been introduced to one, two, or all three of the CDR sequences is preferably still capable of specifically binding the peptide of SEQ ID NO: 2 with at least comparable binding affinity as wtNbALFA.

An antibody described herein preferably comprises an FR1 sequence having at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or being identical to the sequence of EVQLVESGGGLVQPGGSLRLSCAAS (SEQ ID NO: 99).

An antibody described herein preferably comprises a FR2 sequence having at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or being identical to the sequence of WX1RQAPGKRREMVA, wherein X1 is F or Y (SEQ ID NO: 102).

An antibody described herein preferably comprises a FR3 sequence having at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or being identical to the sequence of X1YRX2SVX3GRFTISRDNX4KX5X6X7YLMNSLRAEDTAVYYC, wherein X1 is M or Y, X2 is D or E, X3 is Q or K, X4 is A or S, X5 is N or R, X6 is M or T, and X7 is V or L (SEQ ID NO: 108).

An antibody described herein preferably comprises a FR4 sequence having at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or being identical to the sequence of WGQGTQVTVSS (SEQ ID NO: 109).

Accordingly, an antibody described herein preferably comprises FR1, FR2, FR3, and FR4 sequences, which each have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences of SEQ ID NOs: 99, 102, 108, and 109, respectively.

An antibody described herein preferably comprises an FR2 sequence having at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or being identical to any one of SEQ ID NO: 100 or 101. Such an antibody described herein may comprise the FR2 sequence having at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or being identical to WFRQAPGKRREMVA (SEQ ID NO: 100). Such an antibody described herein may also comprise the FR2 sequence having at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or being identical to WYRQAPGKRREMVA (SEQ ID NO: 101).

An antibody described herein preferably comprises a FR3 sequence having at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or being identical to any one of SEQ ID NOs: 103-107. Such an antibody described herein may comprise a FR3 sequence having at

least 80%, at least 85%, at least 90%, or at least 95% sequence identity or being identical to the sequence MYRESVQGRFTISRDN AKRMVYLQMNSLRAEDTAVYYC (SEQ ID NO: 103). Such an antibody described herein may comprise a FR3 sequence having at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or being identical to the sequence MYRESVKGRFTISRDN AKRMVYLQMNSLRAEDTAVYYC (SEQ ID NO: 104). Such an antibody described herein may comprise a FR3 sequence having at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or being identical to the sequence YYRDSVKGRFTISRDN AKRMVYLQMNSLRAEDTAVYYC (SEQ ID NO: 105). Such an antibody described herein may comprise a FR3 sequence having at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or being identical to the sequence MYRESVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYC (SEQ ID NO: 106). Such an antibody described herein may comprise a FR3 sequence having at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or being identical to the sequence YYRDSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYYC (SEQ ID NO: 107).

An antibody described herein preferably comprises FR1, FR2, FR3, and FR4 sequences, which each have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 100, 103, and 109, respectively.

An antibody described herein preferably comprises FR1, FR2, FR3, and FR4 sequences, which each have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 101, 103, and 109, respectively.

An antibody described herein preferably comprises FR1, FR2, FR3, and FR4 sequences, which each have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 101, 104, and 109, respectively.

An antibody described herein preferably comprises FR1, FR2, FR3, and FR4 sequences, which each have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 101, 105, and 109, respectively.

An antibody described herein preferably comprises FR1, FR2, FR3, and FR4 sequences, which each have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 101, 106, and 109, respectively.

An antibody described herein preferably comprises FR1, FR2, FR3, and FR4 sequences, which each have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 101, 107, and 109, respectively.

An antibody described herein preferably comprises the CDR1, CDR2, and CDR3 sequences set forth in SEQ ID NO: 94, 97, and 98, and further comprises FR1, FR2, FR3, and FR4 sequences which each have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 102, 108, and 109, respectively.

An antibody described herein preferably comprises the CDR1, CDR2, and CDR3 sequences set forth in SEQ ID NO: 94, 95, and 98, and further comprises FR1, FR2, FR3, and FR4 sequences which each have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 100, 103, and 109, respectively.

An antibody described herein preferably comprises the CDR1, CDR2, and CDR3 sequences set forth in SEQ ID NO: 94, 95, and 98, and further comprises FR1, FR2, FR3, and FR4 sequences which each have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 101, 103, and 109, respectively.

An antibody described herein preferably comprises the CDR1, CDR2, and CDR3 sequences set forth in SEQ ID NO: 94, 96, and 98, and further comprises FR1, FR2, FR3, and FR4 sequences which each have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 101, 103, and 109, respectively.

An antibody described herein preferably comprises the CDR1, CDR2, and CDR3 sequences set forth in SEQ ID NO: 94, 95, and 98, and further comprises FR1, FR2, FR3, and FR4 sequences which each have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 101, 104, and 109, respectively.

An antibody described herein preferably comprises the CDR1, CDR2, and CDR3 sequences set forth in SEQ ID NO: 94, 96, and 98, and further comprises FR1, FR2, FR3, and FR4 sequences which each have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 101, 104, and 109, respectively.

An antibody described herein preferably comprises the CDR1, CDR2, and CDR3 sequences set forth in SEQ ID NO: 94, 95, and 98, and further comprises FR1, FR2, FR3, and FR4 sequences

which each have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 101, 105, and 109, respectively.

An antibody described herein preferably comprises the CDR1, CDR2, and CDR3 sequences set forth in SEQ ID NO: 94, 96, and 98, and further comprises FR1, FR2, FR3, and FR4 sequences which each have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 101, 105, and 109, respectively.

An antibody described herein preferably comprises the CDR1, CDR2, and CDR3 sequences set forth in SEQ ID NO: 94, 95, and 98, and further comprises FR1, FR2, FR3, and FR4 sequences which each have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 101, 106, and 109, respectively.

An antibody described herein preferably comprises the CDR1, CDR2, and CDR3 sequences set forth in SEQ ID NO: 94, 96, and 98, and further comprises FR1, FR2, FR3, and FR4 sequences which each have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 101, 106, and 109, respectively.

An antibody described herein preferably comprises the CDR1, CDR2, and CDR3 sequences set forth in SEQ ID NO: 94, 95, and 98, and further comprises FR1, FR2, FR3, and FR4 sequences which each have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 101, 107, and 109, respectively.

An antibody described herein preferably comprises the CDR1, CDR2, and CDR3 sequences set forth in SEQ ID NO: 94, 96, and 98, and further comprises FR1, FR2, FR3, and FR4 sequences which each have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 101, 107, and 109, respectively.

Preferably, the framework sequence of the antibody is defined by the amino acid positions corresponding to positions 1-25, 39-53, 61-98, and 112-122 of SEQ ID NO: 1.

Preferably, the FR1, FR2, FR3, and/or FR4 regions of an antibody described herein correspond to the amino acid positions 1-25, 39-53, 61-98, and 112-122 of SEQ ID NO: 1.

The antibody described herein is preferably humanized.

The term "humanization" as used herein refers to a process of modifying non-human antibodies, typically derived from animals, to maximize their amino acid sequence identity to the closest human equivalent. This alteration involves genetic engineering that retains the

antibody's ability to recognize and bind to specific antigens while minimizing potential adverse reactions when administered to humans. Typically, degree of humanization of an antibody is defined as percent identity of V segment or framework regions with the human counterpart. It is known that the variable regions of both heavy and light chains contain three complementarity-determining regions (CDRs) which vary in response to the epitopes in question and determine binding capability, flanked by four framework regions (FRs) which are relatively conserved in a given species and which putatively provide a scaffolding for the CDRs. When nonhuman antibodies are prepared with respect to a particular epitope, the variable regions can be "reshaped" or "humanized" by grafting CDRs derived from nonhuman antibody on the FRs present in the human antibody to be modified. Application of this approach to various antibodies has been reported by Sato, K., et al., (1993) *Cancer Res* 53:851-856. Riechmann, L., et al., (1988) *Nature* 332:323-327; Verhoeyen, M., et al., (1988) *Science* 239:1534-1536; Kettleborough, C. A., et al., (1991) *Protein Engineering* 4:773-3783; Maeda, H., et al., (1991) *Human Antibodies Hybridoma* 2:124-134; Gorman, S. D., et al., (1991) *Proc Natl Acad Sci USA* 88:4181-4185; Tempest, P. R., et al., (1991) *Bio/Technology* 9:266-271; Co, M. S., et al., (1991) *Proc Natl Acad Sci USA* 88:2869-2873; Carter, P., et al., (1992) *Proc Natl Acad Sci USA* 89:4285-4289; and Co, M. S. et al., (1992) *J Immunol* 148:1149-1154.

In some embodiments, humanized antibodies preserve all CDR sequences (for example, a humanized mouse antibody which contains all six CDRs from the mouse antibodies). In other embodiments, humanized antibodies have one or more CDRs (one, two, three, four, five, six) which are altered with respect to the original antibody, which are also termed one or more CDRs "derived from" one or more CDRs from the original antibody.

In the context of the present invention, in particular in the context of humanized versions of wtNbALFA, the term "humanized" preferably refers to an antibody that has higher degree of framework sequence identity to a given human germline sequence than wtNbALFA. The given human germline sequence is preferably the human germline sequence(s) that has the highest sequence identity to the framework regions of wtNbALFA. More preferably, the given human germline sequence is IGHV3-23*4 set forth in SEQ ID NO: 110 for the V segment and IGHJ6*01 for the J segment set forth in SEQ ID NO: 126.

The term “germline sequence” refers to an amino-acid sequence that represents the unaltered, naturally occurring protein derived from the genetic sequence present in the germ cells (sperm or egg cells) of an organism. This sequence serves as the baseline genetic information and the term “germline sequence” is used to distinguish the protein derived from the original genetic makeup and the protein comprising any modifications to the protein sequence (i.e., mutations), that might be introduced for specific purposes.

As used herein “framework identity” refers to the percentage identity of the germline-encoded framework of the closest human germline (e.g., IGHV3-23 or one of its allotypes) to the framework of the humanized or wtNbALFA to which it was compared. The frameworks are defined as described herein (SEQ ID NOs: 99-109). The alignment was calculated using the global alignment algorithm of Needleman and Wunsch, accessible via the Pairwise Sequence Alignment tool (https://www.ebi.ac.uk/Tools/psa/emboss_needle/). As MATRIX, the Blosum62 matrix was utilized with a GAP OPEN value of 10, a GAP EXTENDED value of 0.5, the END GAP PENALTY was set to false, an END GAP OPEN value of 0 and an END GAP EXTEND value of 0.5. These values represent the standard running parameters of the tool. After running the algorithm, the identity value was given. As used herein, an “overall framework identity” refers to the weighted average framework identity that is calculated as average identity of FR1 to FR4 normalized by the sequence length of the respective region (i.e., 25, 15, 38, and 11 amino acids for FR1, FR2, FR3, and FR4, respectively; see Table 3).

An antibody described herein preferably comprises a sequence that has at least about 71%, at least about 75%, or preferably at least about 80%, overall framework identity to the closest human germline sequence(s). An antibody described herein may even comprise a sequence that has at least about 85%, or at least about 90% overall framework identity to the closest human germline sequence(s). An antibody described herein preferably comprises a sequence that has at least about 71%, at least about 75%, or preferably at least about 80% overall framework identity to the framework regions of IGHV3-23*4 as shown in SEQ ID NO: 110 (for FR1, FR2, and FR3) and IGHJ6*01 as shown in SEQ ID NO:126 (for FR4). An antibody described herein may even comprise a sequence that has at least about 85%, or at least about 90% overall framework identity to the framework regions of IGHV3-23*4 as shown in SEQ ID NO: 110 (for FR1, FR2, and FR3) and IGHJ6*01 as shown in SEQ ID NO:126 (for FR4).

As used herein “V segment identity” refers to the percentage identity of the germline-encoded FR1, CDR1, FR2, CDR2, FR3 of the closest human germline (e.g., IGHV3-23 or one of its allotypes) to the FR1, CDR1, FR2, CDR2, FR3 of the humanized or wtNbALFA to which it was compared. For this, the DomainGapAlign tool of IMGT was used (<https://www.imgt.org/3Dstructure-DB/cgi/DomainGapAlign.cgi>). The Domain type was set to “V”, as species “any” or “Homo sapiens (human)” was chosen and the Smith-Waterman score was set to be above 0. The E values was chosen as 200, the Gap penalty for query was set to -5 and the Gap penalty for the reference was set to -20. These values represent the standard running parameters of the tool. The antibody sequence of interest was inserted in the FASTA format. The tool calculated the closest germline sequences and provided a “V-REGION identity percentage” value for each one of them. The “V-REGION identity percentage” of the closest related human sequence was reported as V segment identity herein.

An antibody described herein preferably comprises a sequence that has at least about 60%, at least about 65% or preferably at least about 70% V segment identity to the closest human germline sequence. An antibody described herein preferably comprises a sequence that has at least about 60%, at least about 65% or preferably at least about 70% V segment identity to IGHV3-23*4 as shown in SEQ ID NO: 110.

An antibody described herein preferably has an affinity for the epitope tag shown in SEQ ID NO: 2 that is at least comparable with the affinity of the antibody of SEQ ID NO: 1. As used herein, “at least comparable” preferably means that the K_D of a given antibody is not more than by about factor 2.5 higher than the K_D of the reference antibody it is compared to, such as the antibody of SEQ ID NO: 1. The K_D of a given antibody and the reference antibody is preferably measured using the same method. Any suitable method for determining the K_D can be used in the art. Known methods for determining K_D include, but are not limited to, e.g., ELISA, surface plasmon resonance (SPR), or biolayer interferometry (BLI). Preferably, the affinity is measured by biolayer interferometry, preferably by a method as essentially described in Method 1.

Preferably, an antibody described herein has a K_D for the epitope tag shown in SEQ ID NO: 2 that is not more than by about factor 2.5 higher than the K_D of wtNbALFA for said epitope tag.

More preferably, an antibody described herein has a K_D for the epitope tag shown in SEQ ID NO: 2 that is not more than by about factor 2 higher than the K_D of wtNbALFA for said epitope tag. More preferably, an antibody described herein has a K_D for the epitope tag shown in SEQ ID NO: 2 that is not more than by about factor 1.5 higher than the K_D of wtNbALFA for said epitope tag. More preferably, an antibody described herein has a K_D for the epitope tag shown in SEQ ID NO: 2 that is not more than by about factor 1.2 higher than the K_D of wtNbALFA for said epitope tag.

An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2, that is not higher than the K_D of wtNbALFA for said epitope tag. Preferably, an antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2, that is lower than the K_D of wtNbALFA for said epitope tag. Preferably, an antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2, that is at least by about factor 1.5 lower than the K_D of wtNbALFA for said epitope tag. Preferably, an antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2, that is at least by about factor 2.0 lower than the K_D of wtNbALFA for said epitope tag. Preferably, an antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2, that is at least by about factor 2.5 lower than the K_D of wtNbALFA for said epitope tag.

An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 100 lower than the K_D of wtNbALFA for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 50 lower than the K_D of wtNbALFA for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 40 lower than the K_D of wtNbALFA for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 30 lower than the K_D of wtNbALFA for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 20 lower than the K_D of wtNbALFA for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 15 lower than the K_D of wtNbALFA for said epitope tag. An antibody described herein

may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 10 lower than the K_D of wtNbALFA for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 5 lower than the K_D of wtNbALFA for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 4 lower than the K_D of wtNbALFA for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 3 lower than the K_D of wtNbALFA for said epitope tag.

Preferably, an antibody described herein has a K_D for the epitope tag shown in SEQ ID NO: 2 that is not more than by about factor 2.5 higher than the K_D of the antibody of SEQ ID NO: 121 for said epitope tag. More preferably, an antibody described herein has a K_D for the epitope tag shown in SEQ ID NO: 2 that is not more than by about factor 2 higher than the K_D of the antibody of SEQ ID NO: 121 for said epitope tag. More preferably, an antibody described herein has a K_D for the epitope tag shown in SEQ ID NO: 2 that is not more than by about factor 1.5 higher than the K_D of the antibody of SEQ ID NO: 121 for said epitope tag. More preferably, an antibody described herein has a K_D for the epitope tag shown in SEQ ID NO: 2 that is not more than by about factor 1.2 higher than the K_D of the antibody of SEQ ID NO: 121 for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2, that is not higher than the K_D of the antibody of SEQ ID NO: 121 for said epitope tag.

An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 100 lower than the K_D of the antibody set forth in SEQ ID NO: 121 for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 50 lower than the K_D of the antibody set forth in SEQ ID NO: 121 for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 40 lower than the K_D of the antibody set forth in SEQ ID NO: 121 for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 30 lower than the K_D of the antibody set forth in SEQ ID NO: 121 for said epitope tag. An antibody described

herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 20 lower than the K_D of the antibody set forth in SEQ ID NO: 121 for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 15 lower than the K_D of the antibody set forth in SEQ ID NO: 121 for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 10 lower than the K_D of the antibody set forth in SEQ ID NO: 121 for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 5 lower than the K_D of the antibody set forth in SEQ ID NO: 121 for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 4 lower than the K_D of the antibody set forth in SEQ ID NO: 121 for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 3 lower than the K_D of the antibody set forth in SEQ ID NO: 121 for said epitope tag.

Preferably, an antibody described herein has a K_D for the epitope tag shown in SEQ ID NO: 2 that is not more than by about factor 2.5 higher than the K_D of the antibody of SEQ ID NO: 123 for said epitope tag. More preferably, an antibody described herein has a K_D for the epitope tag shown in SEQ ID NO: 2 that is not more than by about factor 2 higher than the K_D of the antibody of SEQ ID NO: 123 for said epitope tag. More preferably, an antibody described herein has a K_D for the epitope tag shown in SEQ ID NO: 2 that is not more than by about factor 1.5 higher than the K_D of the antibody of SEQ ID NO: 123 for said epitope tag. More preferably, an antibody described herein has a K_D for the epitope tag shown in SEQ ID NO: 2 that is not more than by about factor 1.2 higher than the K_D of the antibody of SEQ ID NO: 123 for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2, that is not higher than the K_D of the antibody of SEQ ID NO: 123 for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 100 lower than the K_D of the antibody set forth in SEQ ID NO: 123 for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 50

lower than the K_D of the antibody set forth in SEQ ID NO: 123 for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 40 lower than the K_D of the antibody set forth in SEQ ID NO: 123 for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 30 lower than the K_D of the antibody set forth in SEQ ID NO: 123 for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 20 lower than the K_D of the antibody set forth in SEQ ID NO: 123 for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 15 lower than the K_D of the antibody set forth in SEQ ID NO: 123 for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 10 lower than the K_D of the antibody set forth in SEQ ID NO: 123 for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 5 lower than the K_D of the antibody set forth in SEQ ID NO: 123 for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 4 lower than the K_D of the antibody set forth in SEQ ID NO: 123 for said epitope tag. An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 that is from about factor 2.5 higher to about factor 3 lower than the K_D of the antibody set forth in SEQ ID NO: 123 for said epitope tag.

An antibody described herein may have a K_D for the epitope tag shown in SEQ ID NO: 2 of about 5×10^{-10} M or lower, preferably 3×10^{-10} M or lower, as preferably measured by biolayer interferometry, preferably by a method as essentially described in Method 1. An antibody described herein may also have a K_D for the epitope tag shown in SEQ ID NO: 2 of about 2×10^{-10} M or lower or 2×10^{-10} M or lower, as preferably measured by biolayer interferometry, preferably by a method as essentially described in Method 1.

An antibody described herein preferably has a higher T_m as compared to the antibody of SEQ ID NO: 1.

The term "T_m" as used herein refers to the temperature at which half of all molecules are denatured and half of all molecules retain their conformation. For example, a T_m of 60°C means that at this temperature, 50% of antibodies present in the sample are not denatured while the other 50% of antibodies is denatured. T_m can be measured as essentially described in Method 2.

An antibody described herein preferably has a T_m of at least about 62°C, preferably at least about 63°C. The antibody described herein may also have a T_m that is at least about 64°C, at least about 65°C or even at least about 66°C. The T_m is preferably measured as essentially described in Method 2.

An antibody described herein preferably comprises a protein A binding site.

As used herein, staphylococcal Protein A (SpA) or protein A refers to a molecule derived from the cell wall of the bacteria *Staphylococcus aureus* with a strong affinity to immunoglobulins, particularly human IgG. SpA's unique ability to specifically interact with the Fc region of antibodies offers significant potential for various biotechnological applications, including antibody purification, immobilization, and detection in diagnostic and therapeutic settings. SpA affinity purification is currently the most applied affinity system for antibody purification. In the context of the present application, it is believed that up to 13 amino acid positions may be involved in the recognition of SpA: residues at positions corresponding to sequence positions 15, 17, 19, 60, 62, 67, 68, 69, 71, 73, 84, 86, and 87 of SEQ ID NO: 1. In particular, the residues Gly 15, Ser 17, Arg 19, Thr 60, Tyr 62, Lys 67, Gly 68, Arg 69, Thr 71, Ser 73, Gln 84, Asn 86, Ser 87 are important for SpA binding. The residues Gly 15, Ser 17, Arg 19, Tyr 62, Gly 68, Arg 69, Thr 71, and Gln 84 are already present in wt NbALFA. The residues Ser 73, Asn 86, and Ser 87, are not comprised in wtALFA, but in all acceptor frameworks that are used for generating the CDR variants 1-7. The two positions 60 and 67, however, are located within the CDR2 region (position 67 is located within the CDR2 region determined according to Kabat). Since wt NbALFA comprises the residues Ala 60 and Gln 67, these two residues would have to be mutated to improve SpA binding. Although these two residues are within the CDR2 region or within the CDR2 region determined according to Kabat, the inventors of the present application have surprisingly found that introducing the mutations Ala 60 → Thr and/or Gln 67 → Lys retains the high binding affinity to the epitope tag of SEQ ID NO: 2, as shown for

variants 10-13 and 15. However, introducing the mutations Ala 60 → Thr and/or Gln 67 → Lys can promote SpA binding, as shown in Example 3 and Table 4.

An antibody described herein may comprise a Thr residue at the position corresponding to sequence position 60 of SEQ ID NO: 1. An antibody described herein may comprise a Lys residue at the position corresponding to sequence position 67 of SEQ ID NO: 1.

An antibody described herein may comprise at positions corresponding to sequence positions 60 and 67 of SEQ ID NO: 1 the following amino acid residues: a Thr at position 60 and a Lys at position 67.

An antibody described herein may comprise a Gly residue at the position corresponding to sequence position 15. An antibody described herein may comprise a Ser residue at the position corresponding to sequence position 17. An antibody described herein may comprise an Arg residue at the position corresponding to sequence position 19. An antibody described herein may comprise a Tyr residue at the position corresponding to sequence position 62. An antibody described herein may comprise a Gly residue at the position corresponding to sequence position 68. An antibody described herein may comprise a Arg residue at the position corresponding to sequence position 69. An antibody described herein may comprise a Thr residue at the position corresponding to sequence position 71. An antibody described herein may comprise a Ser residue at the position corresponding to sequence position 73. An antibody described herein may comprise a Gln residue at the position corresponding to sequence position 84. An antibody described herein may comprise an Asn residue at the position corresponding to sequence position 86. An antibody described herein may comprise a Ser residue at the position corresponding to sequence position 87.

An antibody described herein may comprise at positions corresponding to sequence positions 15, 17, 19, 62, 68, 69, 71, 73, 84, 86, and 87 of SEQ ID NO: 1, at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, or preferably all of the following amino acid residues: a Gly at position 15, Ser at position 17, Arg at position 19, Tyr at position 62, Gly at position 68, Arg at position 69, Thr at position 71, Ser at position 73, Gln at position 84, Asn at position 86, and Ser at position 87.

An antibody described herein may comprise at positions corresponding to sequence positions 15, 17, 19, 60, 62, 67, 68, 69, 71, 73, 84, 86, and 87 of SEQ ID NO: 1, at least 1, at least 2, at

least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, or preferably all of the following amino acid residues: a Gly at position 15, Ser at position 17, Arg at position 19, a Thr at position 60, Tyr at position 62, Lys at position 67, Gly at position 68, Arg at position 69, Thr at position 71, Ser at position 73, Gln at position 84, Asn at position 86, and Ser at position 87.

As described herein, an ALFA-tag may be part of a CAR or a docking compound (depending on whether the single-domain antibody is located on the docking compound or on the CAR, respectively). Preferably, an ALFA-tag is part of a docking compound (in this case the single-domain antibody is located on the CAR). In some embodiments, an ALFA-tag described herein (for use in the compositions and methods described herein) comprises the amino acid sequence of X1-X2-L-E-X5-E-X7-R-R-L-X12-X13 (SEQ ID NO: 3), wherein X1 is G or S or P or T, X2 is R or G or P, X5 is E or Q, X7 is L or I, X12 is S or T or P or A or D or E, and wherein X13 is P or A or S or K or D or E or no amino acid. In some embodiments, an ALFA-tag described herein comprises a peptide set forth in any one of SEQ ID NOs: 2-93. Preferably, an ALFA-tag described herein comprises the amino acid sequence of SEQ ID NO: 2.

SEQ ID NO: 3 defines the core structure of a peptide that can be bound by the antibody described herein, which may further comprise up to two additional amino acids at the N terminus and up to two additional amino acids at the C terminus. Such additional amino acids at the ends of the core structure of the peptide usually do not necessarily influence the secondary structure of the of the peptide or specific binding of the peptide to an antibody specific for the peptide, but may serve as linker structures in a fusion protein. Accordingly, type and number of the additional amino acids may depend on the location of the peptide in a fusion protein and may vary depending on whether the peptide is located on N-terminus or C-terminus or somewhere in between of the polypeptide. Accordingly, the peptide sequence may comprise N-terminal of X1 the amino acids Xa-Xb, wherein Xa is D or S or G or M or P or no amino acid and Xb is S or D or P or M or R or G no amino acid. The peptide may also comprise C-terminal of X13 the amino acids Xy-Xz, wherein Xy is G or S or P or D or A or E or K or no amino acid, and Xz is S or P or no amino acid.

In some embodiments, an ALFA-tag described herein (for use in the compositions and methods described herein) comprises the sequence of X1-X2-L-E-X5-E-L-R-R-L-X12-X13 (SEQ

ID NO: 4), wherein X1 is S or T, X2 is R or G, X5 is E or Q, X12 is T or D or E, and wherein X13 is A or D or E or no amino acid. The ALFA-tag may comprise the sequence of S-R-L-E-E-L-R-R-R-L-T-E (SEQ ID NO: 2) or a variant thereof, wherein the variant has as compared to SEQ ID NO: 2 1 to 5 mutations selected from the group consisting of: S1→T, R2→G, E5→Q, T12→D, and T12→E, E13→A, E13→D, and deletion of E13. The variant may have as compared to SEQ ID NO: 2 the following mutations: (a) S1→T and E13→A; (b) R2→G; (c) R2→G and E5→Q; (d) R2→G, E5→Q and E13→A; (e) R2→G, E5→Q, T12→D, and E13→A; (f) R2→G, E5→Q, T12→E, and E13→A; (g) T12→D and E13→A; (h) T12→E and E13→A; (i) E13→A; (j) E13→D; or (k) deletion of E13. An ALFA-tag comprising as a core structure the sequence SEQ ID NO: 4 may comprise N-terminal of X1 the amino acids Xa-Xb, wherein Xa is S or G or M or P or no amino acid, and Xb is R or G or S or P or M or no amino acid. Xa-Xb may be selected from the group consisting of P, M-P, G-R, P-G, P-S, S-P, G-P, S-P, M, and M-S, preferably P or M-P. Such a peptide may also comprise C-terminal of X13 the amino acids Xy-Xz, wherein Xy-Xz is P or D or A or no amino acid, and Xz is P or S or no amino acid. Xy-Xz may be selected from the group consisting of no amino acid, P, D-P, A, and A-S, preferably no amino acid or P. The ALFA-tag may comprise a combination of Xa-Xb and Xy-Xz selected from the group consisting of: (a) M-P and P; (b) P and P; and (c) P and no amino acid.

An ALFA-tag described herein (for use in the compositions and methods described herein) preferably comprises the sequence of SRLEEELRRRLTE (SEQ ID NO: 2).

An ALFA-tag described herein (for use in the compositions and methods described herein) may comprise a sequence selected from the group consisting of:

- (a) MPSRLEEELRRRLTEP (SEQ ID NO: 5);
- (b) PSRLEEELRRRLTEP (SEQ ID NO: 6);
- (c) PSRLEEELRRRLTE (SEQ ID NO: 7);
- (d) GRSRLEEELRRRLTA (SEQ ID NO: 8);
- (e) PGSRLEEELRRRLTAP (SEQ ID NO: 9);
- (f) PSTRLEEELRRRLTAP (SEQ ID NO: 10);
- (g) SPSRLEEELRRRLTAP (SEQ ID NO: 11);
- (h) SPSRLEEELRRRLDAP (SEQ ID NO: 12);
- (i) SPSRLEEELRRRLEAP (SEQ ID NO: 13);

- (j) SPSRLEEEELRRRLTDP (SEQ ID NO: 14);
- (k) SPSRLEEEELRRRLTEP (SEQ ID NO: 15);
- (l) SPSRLEEEELRRRLTADP (SEQ ID NO: 16);
- (m) SPSGLEEEELRRRLTEP (SEQ ID NO: 17);
- (n) GPSRLEEEELRRRLT (SEQ ID NO: 18);
- (o) GPSRLEEEELRRRLTA (SEQ ID NO: 19);
- (p) GPSRLEEEELRRRLTAA (SEQ ID NO: 20);
- (q) GPSRLEEEELRRRLTAAS (SEQ ID NO: 21);
- (r) SPSGLEQELRRRLTAP (SEQ ID NO: 22);
- (s) SPSGLEQELRRRLDAP (SEQ ID NO: 23);
- (t) SPSGLEQELRRRLLEAP (SEQ ID NO: 24);
- (u) SPSGLEQELRRRLTEP (SEQ ID NO: 25);
- (v) GPSRLEEEELRRRLTAP (SEQ ID NO: 26);
- (w) GPSRLEEEELRRRLTEP (SEQ ID NO: 27);
- (x) GPSRLEEEELRRRLTE (SEQ ID NO: 28);
- (y) MSRLEEEELRRRLTEP (SEQ ID NO: 29); and
- (z) MSSRLEEEELRRRLTEP (SEQ ID NO: 30).

An ALFA-tag described herein (for use in the compositions and methods described herein) may comprise the sequence of X1-X2-L-E-X5-E-X7-R-R-R-L-X12-X13 (SEQ ID NO: 31), wherein X1 is G or S or P, X2 is R or G, X5 is E or Q, X7 is L or I, X12 is S or T or P or A, and X13 is P or A or S or no amino acid.

An ALFA-tag described herein (for use in the compositions and methods described herein) may comprise the sequence of G-R-L-E-E-E-L-R-R-L-S (SEQ ID NO: 32) or a variant thereof, wherein the variant has as compared to SEQ ID NO: 32 1 to 6 mutations selected from the group consisting of: G1→S, G1→P, R2→G, E5→Q, L7→I, S12→T, S12→P, S12→A, addition of P13, addition of A13, and addition of S13. The variant may have as compared to SEQ ID NO: 32 the following mutations: (a) G1→S, R2→G, E5→Q, and addition of P13; (b) R2→G, E5→Q, S12→T, and addition of A13; (c) G1→P, R2→G, E5→Q, S12→T, and addition of A13; (d) G1→S, R2→G, E5→Q, S12→T, and addition of P13; (e) G1→S, R2→G, S12→T, and addition of A13; (f) G1→S, R2→G, E5→Q, and S12→T; (g) G1→S, R2→G, E5→Q, S12→T, and addition of A13; (h)

G1→S and S12→P, and addition of P13; (i) G1→S, R2→G, E5→Q, S12→T, and addition of P13; (j) E5→Q L7→I, and addition of P13; (k) addition of P13; or (l) S12→A. An ALFA-tag comprising as a core structure the sequence SEQ ID NO: 31 may comprise N-terminal of X1 the amino acids Xa-Xb, wherein Xa is M or S or P or D or G or no amino acid, and Xb is S or D or P or no amino acid. Xa-Xb may be selected from the group consisting of M-S, S-D, P-D, P-S, D-S, S-P, and G-P, preferably M-S. Such an ALFA-tag may also comprise C-terminal of X13 the amino acids Xy-Xz, wherein Xy is G or P or A or E or K or S or no amino acid, and Xz is P or S or no amino acid. Xy-Xz may be selected from the group consisting of no amino acid, G, P, A, E-P, A-S, K, and S, preferably no amino acid. The ALFA-tag may comprise a combination of Xa-Xb and Xy-Xz that is M-S and no amino acid.

An ALFA-tag described herein (for use in the compositions and methods described herein) may comprise a sequence selected from the group consisting of:

- (a) GRLEEEELRRRLS (SEQ ID NO: 32);
- (b) MSGRLEEEELRRRLSP (SEQ ID NO: 33);
- (c) SDSGLEQELRRRLSPG (SEQ ID NO: 34);
- (d) PDGGLQELRRRLTAP (SEQ ID NO: 35);
- (e) PSGGLEQELRRRLTAP (SEQ ID NO: 36);
- (f) DSPGLEQELRRRLTAP (SEQ ID NO: 37);
- (g) PDSGLEQELRRRLTPA (SEQ ID NO: 38);
- (h) SPSGLEEEELRRRLTAEP (SEQ ID NO: 39);
- (i) GPSGLEQELRRRLT (SEQ ID NO: 40);
- (j) GPSGLEQELRRRLTAAS (SEQ ID NO: 41);
- (k) SPSRLEEEELRRRLPSK (SEQ ID NO: 42);
- (l) SPSGLEQELRRRLTPS (SEQ ID NO: 43);
- (m) SPGRLEQEIRRRRLSPS (SEQ ID NO: 44);
- (n) PSGRLEEEELRRRLSPS (SEQ ID NO: 45);
- (o) PSGRLEEEELRRRLS (SEQ ID NO: 46);
- (p) PSGRLEEEELRRRLA (SEQ ID NO: 47); and
- (q) PSGRLEEEELRRRLSP (SEQ ID NO: 48).

An ALFA-tag described herein (for use in the compositions and methods described herein) may comprise a peptide comprising as a core structure the sequence of X1-X2-L-E-X5-E-L-R-R-L-X12-X13 (SEQ ID NO: 49), wherein X1 is S or G or P, X2 is R or G or P, X5 is E or Q, X12 is S or T or D or E, and X13 is P or A or D or no amino acid. A peptide comprising as a core structure the sequence SEQ ID NO: 49 may comprise N-terminal of X1 the amino acids Xa-Xb, wherein Xa is P or D or S or G or no amino acid, and Xb is D or S or P or no amino acid. Such a peptide may also comprise C-terminal of X13 the amino acids Xy-Xz, wherein Xy is G or P or E or D or S or no amino acid, and Xz is P or no amino acid.

An ALFA-tag described herein (for use in the compositions and methods described herein) may comprise a sequence selected from the group consisting of:

- (a) PDSGLEQELRRRLSPG (SEQ ID NO: 50);
- (b) PDSGLEQELRRRLTAP (SEQ ID NO: 51);
- (c) PSSGLEQELRRRLTAP (SEQ ID NO: 52);
- (d) DPSGLEQELRRRLTAP (SEQ ID NO: 53);
- (e) DSGPLEQELRRRLTAP (SEQ ID NO: 54);
- (f) SPSRLEEEELRRRLTAEP (SEQ ID NO: 55);
- (g) SPSGLEEEELRRRLTAP (SEQ ID NO: 56);
- (h) SPSGLEEEELRRRLDAP (SEQ ID NO: 57);
- (i) SPSGLEEEELRRRLLEAP (SEQ ID NO: 58);
- (j) SPSGLEEEELRRRLTDP (SEQ ID NO: 59);
- (k) SPSGLEEEELRRRLTADP (SEQ ID NO: 60);
- (l) GPSGLEQELRRRLTA (SEQ ID NO: 61);
- (m) SPSGLEQELRRRLTDP (SEQ ID NO: 62);
- (n) SPSGLEQELRRRLTADP (SEQ ID NO: 63);
- (o) SPSGLEQELRRRLTAEP (SEQ ID NO: 64);
- (p) DPSGLEQELRRRLTAP (SEQ ID NO: 37); and
- (q) SPSGLEQELRRRLSPS (SEQ ID NO: 65).

An ALFA-tag described herein (for use in the compositions and methods described herein) may comprise a peptide comprising as a core structure the sequence of X1-X2-X3-X4-X5-X6-X7-X8-X9-R-L-X12-X13 (SEQ ID NO: 66), wherein X1 is G or S, X2 is R or G or A or E, X3 is L or V, X4 is

E or Q, X5 is E or Q, X6 is E or Q, X7 is L or I or V, X8 is R or A or Q or E, X9 is R or A or Q or E, X12 is S or T or L or no amino acid, and X13 is K or P or S or no amino acid. A peptide comprising as a core structure the sequence SEQ ID NO: 66 may comprise N-terminal of X1 the amino acids Xa-Xb, wherein Xa is D or S or G or M or no amino acid and Xb is S or D or P or M or no amino acid. Such a peptide may also comprise C-terminal of X13 the amino acids Xy-Xz, wherein Xy is G or S or P or no amino acid, Xz is S or no amino acid.

An ALFA-tag described herein (for use in the compositions and methods described herein) may comprise a sequence selected from the group consisting of:

- (a) DSGRLEEEELRRRLSKG (SEQ ID NO: 67);
- (b) DSGRLEEEELRRRLSPG (SEQ ID NO: 68);
- (c) SDSGLEEEELRRRLSPG (SEQ ID NO: 69);
- (d) SDSGVEEEELRRRLSPG (SEQ ID NO: 70);
- (f) SDSAVEEEELRRRLSPG (SEQ ID NO: 71);
- (g) SDSGLQEEELRRRLSPG (SEQ ID NO: 72);
- (h) SDSGLEEQLRRRLSPG (SEQ ID NO: 73);
- (i) SDSGLEEEIIRRLSPG (SEQ ID NO: 74);
- (j) SDSGLEEEVRRRLSPG (SEQ ID NO: 75);
- (k) DSGELEEEELRRRLSPG (SEQ ID NO: 76);
- (l) DSGRLEQELRRRLSPG (SEQ ID NO: 77);
- (m) DSGRLEEEIIRRLSPG (SEQ ID NO: 78);
- (n) DSGRLEQEIRRLSPG (SEQ ID NO: 79);
- (o) DSGRLEQEIARRLSPG (SEQ ID NO: 80);
- (p) DSGRLEQEIQRRRLSPG (SEQ ID NO: 81);
- (q) DSGRLEQEIERRLSPG (SEQ ID NO: 82);
- (r) DSGRLEQEIRARLSPG (SEQ ID NO: 83);
- (s) DSGRLEQEIRQRLSPG (SEQ ID NO: 84);
- (t) DSGRLEQEIRERLSPG (SEQ ID NO: 85);
- (u) GPSRLEEEELRRRL (SEQ ID NO: 86);
- (v) MSGLEQELRRRLTPS (SEQ ID NO: 87);
- (w) MSGRLEEEELRRRLSPS (SEQ ID NO: 88);

- (x) SPSAVEEEELRRRLSPS (SEQ ID NO: 89);
- (y) GPSAVEEEELRRRLS (SEQ ID NO: 90);
- (z) MPSGLEQELRRRLTPS (SEQ ID NO: 91);
- (aa) MSSGLEQELRRRLTPS (SEQ ID NO: 92); and
- (bb) MPSGRLEEEELRRRLSPS (SEQ ID NO: 93).

In some embodiments, an ALFA-tag described herein (for use in the compositions and methods described herein) comprises the amino acid sequence

-AA0-AA1-AA2-AA3-AA4-AA5-AA6-AA7-AA8-AA9-AA10-AA11-AA12-AA13-AA14-,

wherein the amino acids of AA0, AA1, AA2, AA3, AA4, AA5, AA6, AA7, AA8, AA9, AA10, AA11, AA12, AA13 and AA14 are:

AA0 is Pro or deleted;

AA1 is Ser, Gly, Thr, or Pro;

AA2 is Arg, Gly, Ala, Glu, or Pro;

AA3 is Leu, Ile, or Val;

AA4 is Glu or Gln;

AA5 is Glu or Gln;

AA6 is Glu or Gln;

AA7 is Leu, Ile, or Val;

AA8 is Arg, Ala, Gln, or Glu;

AA9 is Arg, Ala, Gln, or Glu;

AA10 is Arg;

AA11 is Leu;

AA12 is Thr, Ser, Asp, Glu, Pro, Ala, or deleted;

AA13 is Glu, Lys, Pro, Ser, Ala, Asp, or deleted; and

AA14 is Pro or deleted.

In some embodiments, an ALFA-tag comprises a sequence selected from the group consisting of SRLEEEELRRRLTE, PSRLEEEELRRRLTE, SRLEEEELRRRLTEP, and PSRLEEEELRRRLTEP.

In some embodiments, an ALFA-tag comprises the cyclized amino acid sequence

-AA0-AA1-AA2-AA3-AA4-AA5-AA6-AA7-AA8-AA9-AA10-AA11-AA12-AA13-AA14-,

wherein the side-chains of any two of the amino acids of AA0, AA1, AA2, AA3, AA4, AA5, AA6, AA7, AA8, AA9, AA10, AA11, AA12, AA13 and AA14 (X1, X2) are connected covalently; and wherein the amino acids of AA0, AA1, AA2, AA3, AA4, AA5, AA6, AA7, AA8, AA9, AA10, AA11, AA12, AA13 and AA14 which are not X1 and X2 are:

AA0 is Pro or deleted;

AA1 is Ser, Gly, Thr, or Pro;

AA2 is Arg, Gly, Ala, Glu, or Pro;

AA3 is Leu, Ile, or Val;

AA4 is Glu or Gln;

AA5 is Glu or Gln;

AA6 is Glu or Gln;

AA7 is Leu, Ile, or Val;

AA8 is Arg, Ala, Gln, or Glu;

AA9 is Arg, Ala, Gln, or Glu;

AA10 is Arg;

AA11 is Leu;

AA12 is Thr, Ser, Asp, Glu, Pro, Ala, or deleted;

AA13 is Glu, Lys, Pro, Ser, Ala, Asp, or deleted; and

AA14 is Pro or deleted.

In some embodiments, X1 and X2 are separated by 2 or 3 amino acids.

In some embodiments, AA5 is X1 and AA9 is X2, AA5 is X1 and AA8 is X2, AA9 is X1 and AA13 is X2, AA6 is X1 and AA9 is X2, AA9 is X1 and AA12 is X2, AA10 is X1 and AA13 is X2, AA6 is X1 and AA10 is X2 or AA4 is X1 and AA8 is X2.

In some embodiments, an ALFA-tag comprises a cyclized amino acid sequence selected from the group consisting of

- a. -AA0-AA1-AA2-AA3-AA4-cyclo(X1-AA6-AA7-AA8-X2)-Arg-Leu-AA12-AA13-AA14-,
- b. -AA0-AA1-AA2-AA3-AA4-cyclo(X1-AA6-AA7-X2)-AA9-Arg-Leu-AA12-AA13-AA14-,
- c. -AA0-AA1-AA2-AA3-AA4-AA5-AA6-AA7-AA8-cyclo(X1-Arg-Leu-AA12-X2)-AA14-,
- d. -AA0-AA1-AA2-AA3-AA4-AA5-cyclo(X1-AA7-AA8-X2)-Arg-Leu-AA12-AA13-AA14-,
- e. -AA0-AA1-AA2-AA3-AA4-AA5-AA6-AA7-AA8-cyclo(X1-Arg-Leu-X2)-AA13-AA14-,

- f. -AA0-AA1-AA2-AA3-AA4-AA5-AA6-AA7-AA8-AA9-cyclo(X₁-Leu-AA12-X₂)-AA14-,
- g. -AA0-AA1-AA2-AA3-AA4-AA5-cyclo(X₁-AA7-AA8-AA9-X₂)-Leu-AA12-AA13-AA14-, and
- h. -AA0-AA1-AA2-AA3-cyclo(X₁-AA5-AA6-AA7-X₂)-AA9-Arg-Leu-AA12-AA13-AA14-,

wherein the side-chains of X₁ and X₂ amino acid residues are connected covalently;

AA0 is Pro or deleted;

AA1 is Ser, Gly, Thr, or Pro;

AA2 is Arg, Gly, Ala, Glu, or Pro;

AA3 is Leu, Ile, or Val;

AA4 is Glu or Gln;

AA5 is Glu or Gln;

AA6 is Glu or Gln;

AA7 is Leu, Ile, or Val;

AA8 is Arg, Ala, Gln, or Glu;

AA9 is Arg, Ala, Gln, or Glu;

AA12 is Thr, Ser, Asp, Glu, Pro, Ala, or deleted;

AA13 is Glu, Lys, Pro, Ser, Ala, Asp, or deleted; and

AA14 is Pro or deleted.

In some embodiments, X₁ and X₂ in the peptides disclosed herein are connected covalently via an amide, disulfide, thioether, ether, ester, thioester, thioamide, alkylene, alkenylene, alkynylene, and/or 1,2,3-triazole.

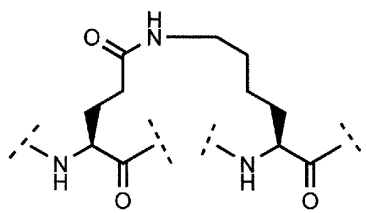
In some embodiments, a cyclized amino acid sequence described herein is generated by linking an amino group of a side-chain of one of X₁ and X₂ to the carboxyl group of a side-chain of the other of X₁ and X₂ via an amide bond. The amino group of the side chain of an amino acid that possesses a pendant amine group, e.g., lysine or a lysine derivative, and the carboxyl group of the side chain of an acidic amino acid, e.g., aspartic acid, glutamic acid or a derivative thereof, can be used to generate a cyclized amino acid sequence via an amide bond.

In some embodiments, a cyclized amino acid sequence described herein is generated by linking a sulfhydryl group of a side-chain of one of X₁ and X₂ to the sulfhydryl group of a side-chain of the other of X₁ and X₂ via a disulfide bond. Sulfhydryl group-containing amino acids

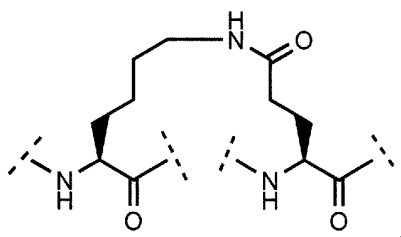
include cysteine and other sulfhydryl-containing amino acids as Pen.

In some embodiments, X_1 and X_2 are, independently, selected from the group consisting of Glu, DGLu, Asp, DAsp, Lys, DLys, hLys, DhLys, Orn, DOrn, Dab, DDab, Dap, DDap, Cys, DCys, hCys, DhCys, Pen, and DPen, with the proviso that when X_1 is Glu, DGLu, Asp, or DAsp, X_2 is Lys, DLys, hLys, DhLys, Orn, DOrn, Dab, DDab, Dap, or DDap; when X_1 is Lys, DLys, hLys, DhLys, Orn, DOrn, Dab, DDab, Dap, or DDap, X_2 is Glu, DGLu, Asp, or DAsp; and when X_1 is Cys, DCys, hCys, DhCys, Pen, or DPen, X_2 is Cys, DCys, hCys, DhCys, Pen, or DPen.

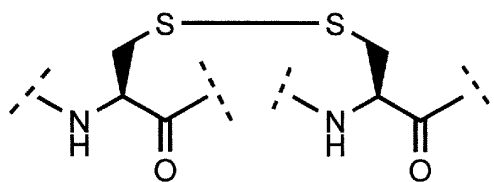
In some embodiments, X_1 is Glu and X_2 is Lys. In some embodiments, -cyclo(Glu-----Lys)-, -c(Glu-----Lys)-, -cyclo(E-----K)-, -c(E-----K)-, -E-----K- cyclo, or -cycloE-----cycloK- comprises the following structure:



In some embodiments, X_1 is Lys and X_2 is Glu. In some embodiments, -cyclo(Lys-----Glu)-, -c(Lys-----Glu)-, -cyclo(K-----E)-, -c(K-----E)-, -K-----E- cyclo, or cycloK-----cycloE- comprises the following structure:



In some embodiments, X_1 is Cys and X_2 is Cys. In some embodiments, -cyclo(Cys-----Cys)-, c(Cys-----Cys)-, -cyclo(C-----C)-, -c(C-----C)-, -C-----C- cyclo, or -cycloC-----cycloC- comprises the following structure:



Particular cyclized amino acid sequences of the above-identified generic formulas include, for example,

-Ser-Arg-Leu-Glu-cyclo(Glu-Glu-Leu-Arg-Lys)-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(Asp-Glu-Leu-Arg-Lys)-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(Glu-Glu-Leu-Lys)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Lys-Arg-Leu-Thr-Glu)-,
-Ser-Arg-Leu-Glu-cyclo(Cys-Glu-Leu-Arg-Cys)-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(Asp-Glu-Leu-Lys)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Lys-Arg-Leu-Thr-Asp)-,
-Ser-Arg-Leu-Glu-cyclo(Glu-Glu-Leu-Arg-DLys)-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Glu-Glu-Leu-Arg-Lys)-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DGlu-Glu-Leu-Arg-Lys)-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Glu-Glu-Leu-Arg-DLys)-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-Arg-Glu)-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-cyclo(Glu-Glu-Glu-Leu-Lys)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Cys-Glu-Leu-DCys)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Cys-Arg-Leu-Thr-Cys)-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Cys-Glu-Leu-Arg-Cys)-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-Cys)-,
-Pro-Ser-Arg-Leu-Glu-Glu-cyclo(Cys-Leu-Arg-Arg-Cys)-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Cys-Arg-Leu-Thr-Cys)-,
-Ser-Arg-Leu-Glu-cyclo(Cys-Glu-Leu-Arg-Cys)-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(Cys-Glu-Leu-Cys)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-Cys)-,
-Ser-Arg-Leu-Glu-Glu-cyclo(Cys-Leu-Arg-Arg-Cys)-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(DGlu-Glu-Leu-Arg-DLys)-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-Arg-Glu)-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-Arg-DGlu)-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-Arg-DGlu)-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DGlu-Glu-Leu-Arg-DLys)-Arg-Leu-Thr-Glu-,

-Pro-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-Arg-Glu)-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-Arg-DGlu)-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-Arg-DGlu)-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-Arg-Asp)-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(DAsp-Glu-Leu-Arg-DLys)-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(DAsp-Glu-Leu-Arg-Lys)-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(Asp-Glu-Leu-Arg-DLys)-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-Arg-DAsp)-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-Arg-DAsp)-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-Arg-Asp)-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DAsp-Glu-Leu-Arg-DLys)-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DAsp-Glu-Leu-Arg-Lys)-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Asp-Glu-Leu-Arg-DLys)-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-Arg-DAsp)-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-Arg-DAsp)-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(DGlu-Glu-Leu-Lys)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(Glu-Glu-Leu-DLys)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(DGlu-Glu-Leu-DLys)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-Glu)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-DGlu)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-DGlu)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(Asp-Glu-Leu-Lys)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(Asp-Glu-Leu-DLys)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-Asp)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-DAsp)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(DAsp-Glu-Leu-Lys)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(DAsp-Glu-Leu-DLys)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-DAsp)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-Asp)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Glu-Glu-Leu-DLys)-Arg-Arg-Leu-Thr-Glu-,

-Pro-Ser-Arg-Leu-Glu-cyclo(DGlu-Glu-Leu-DLys)-Arg-Arg-Leu-Thr-Glu-
-Pro-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-Glu)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-DGlu)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-DGlu)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Asp-Glu-Leu-Lys)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Asp-Glu-Leu-DLys)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-Asp)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-DAsp)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DAsp-Glu-Leu-Lys)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DAsp-Glu-Leu-DLys)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-DAsp)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-Asp)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-Glu-cyclo(Asp-Leu-Arg-Lys)-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-Glu-cyclo(Lys-Leu-Arg-Asp)-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-Glu-cyclo(DAsp-Leu-Arg-Lys)-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-Glu-cyclo(Lys-Leu-Arg-DAsp)-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-Glu-cyclo(Asp-Leu-Arg-DLys)-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-Glu-cyclo(DLys-Leu-Arg-Asp)-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-Glu-cyclo(DLys-Leu-Arg-DAsp)-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-Glu-cyclo(DAsp-Leu-Arg-DLys)-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-cyclo(Asp-Leu-Arg-Lys)-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-cyclo(Lys-Leu-Arg-Asp)-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-cyclo(DAsp-Leu-Arg-Lys)-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-cyclo(Lys-Leu-Arg-DAsp)-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-cyclo(Asp-Leu-Arg-DLys)-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-cyclo(DLys-Leu-Arg-Asp)-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-cyclo(DLys-Leu-Arg-DAsp)-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-cyclo(DAsp-Leu-Arg-DLys)-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(Cys-Glu-Leu-Cys)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-Cys)-,

-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Cys-Arg-Leu-Cys)-Glu-,
-Ser-Arg-Leu-Glu-Glu-cyclo(Cys-Leu-Arg-Cys)-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Cys-Glu-Leu-Cys)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-Cys)-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Cys-Arg-Leu-Cys)-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-cyclo(Cys-Leu-Arg-Cys)-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(DCys-Glu-Leu-DCys)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(DCys-Glu-Leu-Cys)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(hCys-Glu-Leu-Cys)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(hCys-Glu-Leu-hCys)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(Cys-Glu-Leu-hCys)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(hCys-Glu-Leu-DCys)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(DCys-Glu-Leu-hCys)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(hCys-Glu-Leu-DhCys)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(DhCys-Glu-Leu-DhCys)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(Cys-Glu-Leu-DhCys)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(DCys-Glu-Leu-DhCys)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(DhCys-Glu-Leu-Cys)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(DhCys-Glu-Leu-DCys)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DCys-Glu-Leu-DCys)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DCys-Glu-Leu-Cys)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(hCys-Glu-Leu-Cys)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(hCys-Glu-Leu-hCys)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Cys-Glu-Leu-hCys)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(hCys-Glu-Leu-DCys)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DCys-Glu-Leu-hCys)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(hCys-Glu-Leu-DhCys)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DhCys-Glu-Leu-DhCys)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Cys-Glu-Leu-DhCys)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DCys-Glu-Leu-DhCys)-Arg-Arg-Leu-Thr-Glu-,

-Pro-Ser-Arg-Leu-Glu-cyclo(DhCys-Glu-Leu-Cys)-Arg-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DhCys-Glu-Leu-DCys)-Arg-Arg-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-DCys)-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DCys-Leu-Thr-Cys)-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DCys-Leu-Thr-DCys)-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(hCys-Leu-Thr-Cys)-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(hCys-Leu-Thr-DCys)-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(hCys-Leu-Thr-hCys)-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-hCys)-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DCys-Leu-Thr-hCys)-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-hCys)-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DhCys-Leu-Thr-Cys)-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DhCys-Leu-Thr-DCys)-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DhCys-Leu-Thr-hCys)-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-DhCys)-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DCys-Leu-Thr-DhCys)-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-DhCys)-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Pen-Leu-Thr-Pen)-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Pen-Leu-Thr-DPen)-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DPen-Leu-Thr-Pen)-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DPen-Leu-Thr-DPen)-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(DCys-Arg-Leu-Cys)-Glu-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(DCys-Arg-Leu-DCys)-Glu-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Cys-Arg-Leu-DCys)-Glu-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(hCys-Arg-Leu-Cys)-Glu-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(hCys-Arg-Leu-DCys)-Glu-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Cys-Arg-Leu-hCys)-Glu-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(hCys-Arg-Leu-hCys)-Glu-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(hCys-Arg-Leu-DhCys)-Glu-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(DhCys-Arg-Leu-hCys)-Glu-,

-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-DCys)-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DCys-Leu-Thr-Cys)-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DCys-Leu-Thr-DCys)-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(hCys-Leu-Thr-Cys)-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(hCys-Leu-Thr-DCys)-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(hCys-Leu-Thr-hCys)-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-hCys)-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DCys-Leu-Thr-hCys)-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-hCys)-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DhCys-Leu-Thr-Cys)-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DhCys-Leu-Thr-DCys)-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DhCys-Leu-Thr-hCys)-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-DhCys)-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DCys-Leu-Thr-DhCys)-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-DhCys)-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Pen-Leu-Thr-Pen)-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Pen-Leu-Thr-DPen)-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DPen-Leu-Thr-Pen)-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DPen-Leu-Thr-DPen)-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(DCys-Arg-Leu-Cys)-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(DCys-Arg-Leu-DCys)-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Cys-Arg-Leu-DCys)-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(hCys-Arg-Leu-Cys)-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(hCys-Arg-Leu-DCys)-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Cys-Arg-Leu-hCys)-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(hCys-Arg-Leu-hCys)-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(hCys-Arg-Leu-DhCys)-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(DhCys-Arg-Leu-hCys)-Glu-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Lys-Arg-Leu-Asp)-Glu-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Asp-Arg-Leu-Lys)-Glu-,

-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Lys-Arg-Leu-Glu)-Glu-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Glu-Arg-Leu-Lys)-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Lys-Arg-Leu-Asp)-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Asp-Arg-Leu-Lys)-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Lys-Arg-Leu-Glu)-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Glu-Arg-Leu-Lys)-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Cys-Arg-Leu-Thr-Cys)-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Cys-Glu-Leu-Arg-Cys)-Arg-Leu-Thr-Glu-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-Cys)-,
-Pro-Ser-Arg-Leu-Glu-Glu-cyclo(Cys-Leu-Arg-Arg-Cys)-Leu-Thr-Glu-,
-Ser-Arg-Leu-Glu-cyclo(DGlu-Glu-Leu-Arg-DLys)-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-Arg-Glu)-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-Arg-DGlu)-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-Arg-DGlu)-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DGlu-Glu-Leu-Arg-DLys)-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-Arg-Glu)-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-Arg-DGlu)-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-Arg-DGlu)-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-Arg-Asp)-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(DAsp-Glu-Leu-Arg-DLys)-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(DAsp-Glu-Leu-Arg-Lys)-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(Asp-Glu-Leu-Arg-DLys)-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-Arg-DAsp)-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-Arg-DAsp)-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-Arg-Asp)-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DAsp-Glu-Leu-Arg-DLys)-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DAsp-Glu-Leu-Arg-Lys)-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Asp-Glu-Leu-Arg-DLys)-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-Arg-DAsp)-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-Arg-DAsp)-Arg-Leu-Thr-Glu-Pro-,

-Ser-Arg-Leu-Glu-cyclo(DGlu-Glu-Leu-Lys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(Glu-Glu-Leu-DLys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(DGlu-Glu-Leu-DLys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-Glu)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-DGlu)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-DGlu)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(Asp-Glu-Leu-Lys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(Asp-Glu-Leu-DLys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-Asp)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-DAsp)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(DAsp-Glu-Leu-Lys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(DAsp-Glu-Leu-DLys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-DAsp)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-Asp)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Glu-Glu-Leu-DLys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DGlu-Glu-Leu-DLys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-Glu)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-DGlu)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-DGlu)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Asp-Glu-Leu-Lys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Asp-Glu-Leu-DLys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-Asp)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-DAsp)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DAsp-Glu-Leu-Lys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DAsp-Glu-Leu-DLys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-DAsp)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DLys-Glu-Leu-Asp)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-Glu-cyclo(Asp-Leu-Arg-Lys)-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-Glu-cyclo(Lys-Leu-Arg-Asp)-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-Glu-cyclo(DAsp-Leu-Arg-Lys)-Arg-Leu-Thr-Glu-Pro-,

-Ser-Arg-Leu-Glu-Glu-cyclo(Lys-Leu-Arg-DAsp)-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-Glu-cyclo(Asp-Leu-Arg-DLys)-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-Glu-cyclo(DLys-Leu-Arg-Asp)-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-Glu-cyclo(DLys-Leu-Arg-DAsp)-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-Glu-cyclo(DAsp-Leu-Arg-DLys)-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-cyclo(Asp-Leu-Arg-Lys)-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-cyclo(Lys-Leu-Arg-Asp)-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-cyclo(DAsp-Leu-Arg-Lys)-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-cyclo(Lys-Leu-Arg-DAsp)-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-cyclo(Asp-Leu-Arg-DLys)-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-cyclo(DLys-Leu-Arg-Asp)-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-cyclo(DLys-Leu-Arg-DAsp)-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-cyclo(DAsp-Leu-Arg-DLys)-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(Cys-Glu-Leu-Cys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-Cys)-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Cys-Arg-Leu-Cys)-Glu-Pro-,
-Ser-Arg-Leu-Glu-Glu-cyclo(Cys-Leu-Arg-Cys)-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Cys-Glu-Leu-Cys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-Cys)-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Cys-Arg-Leu-Cys)-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-cyclo(Cys-Leu-Arg-Cys)-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(DCys-Glu-Leu-DCys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(DCys-Glu-Leu-Cys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(hCys-Glu-Leu-Cys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(hCys-Glu-Leu-hCys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(Cys-Glu-Leu-hCys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(hCys-Glu-Leu-DCys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(DCys-Glu-Leu-hCys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(hCys-Glu-Leu-DhCys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(DhCys-Glu-Leu-DhCys)-Arg-Arg-Leu-Thr-Glu-Pro-,

-Ser-Arg-Leu-Glu-cyclo(Cys-Glu-Leu-DhCys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(DCys-Glu-Leu-DhCys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(DhCys-Glu-Leu-Cys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-cyclo(DhCys-Glu-Leu-DCys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DCys-Glu-Leu-DCys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DCys-Glu-Leu-Cys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(hCys-Glu-Leu-Cys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(hCys-Glu-Leu-hCys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Cys-Glu-Leu-hCys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(hCys-Glu-Leu-DCys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DCys-Glu-Leu-hCys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(hCys-Glu-Leu-DhCys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DhCys-Glu-Leu-DhCys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(Cys-Glu-Leu-DhCys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DCys-Glu-Leu-DhCys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DhCys-Glu-Leu-Cys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-cyclo(DhCys-Glu-Leu-DCys)-Arg-Arg-Leu-Thr-Glu-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-DCys)-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DCys-Leu-Thr-Cys)-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DCys-Leu-Thr-DCys)-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(hCys-Leu-Thr-Cys)-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(hCys-Leu-Thr-DCys)-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(hCys-Leu-Thr-hCys)-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-hCys)-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DCys-Leu-Thr-hCys)-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-hCys)-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DhCys-Leu-Thr-Cys)-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DhCys-Leu-Thr-DCys)-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DhCys-Leu-Thr-hCys)-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-DhCys)-Pro-,

-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DCys-Leu-Thr-DhCys)-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-DhCys)-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Pen-Leu-Thr-Pen)-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Pen-Leu-Thr-DPen)-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DPen-Leu-Thr-Pen)-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DPen-Leu-Thr-DPen)-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(DCys-Arg-Leu-Cys)-Glu-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(DCys-Arg-Leu-DCys)-Glu-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Cys-Arg-Leu-DCys)-Glu-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(hCys-Arg-Leu-Cys)-Glu-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(hCys-Arg-Leu-DCys)-Glu-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Cys-Arg-Leu-hCys)-Glu-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(hCys-Arg-Leu-hCys)-Glu-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(hCys-Arg-Leu-DhCys)-Glu-Pro-,
-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(DhCys-Arg-Leu-hCys)-Glu-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-DCys)-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DCys-Leu-Thr-Cys)-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DCys-Leu-Thr-DCys)-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(hCys-Leu-Thr-Cys)-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(hCys-Leu-Thr-DCys)-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(hCys-Leu-Thr-hCys)-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-hCys)-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DCys-Leu-Thr-hCys)-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-hCys)-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DhCys-Leu-Thr-Cys)-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DhCys-Leu-Thr-DCys)-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DhCys-Leu-Thr-hCys)-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-DhCys)-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DCys-Leu-Thr-DhCys)-Pro-,
-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-DhCys)-Pro-,

-Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Pen-Leu-Thr-Pen)-Pro-,
 -Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Pen-Leu-Thr-DPen)-Pro-,
 -Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DPen-Leu-Thr-Pen)-Pro-,
 -Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(DPen-Leu-Thr-DPen)-Pro-,
 -Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(DCys-Arg-Leu-Cys)-Glu-Pro-,
 -Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(DCys-Arg-Leu-DCys)-Glu-Pro-,
 -Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Cys-Arg-Leu-DCys)-Glu-Pro-,
 -Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(hCys-Arg-Leu-Cys)-Glu-Pro-,
 -Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(hCys-Arg-Leu-DCys)-Glu-Pro-,
 -Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Cys-Arg-Leu-hCys)-Glu-Pro-,
 -Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(hCys-Arg-Leu-hCys)-Glu-Pro-,
 -Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(hCys-Arg-Leu-DhCys)-Glu-Pro-,
 -Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(DhCys-Arg-Leu-hCys)-Glu-Pro-,
 -Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Lys-Arg-Leu-Asp)-Glu-Pro-,
 -Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Asp-Arg-Leu-Lys)-Glu-Pro-,
 -Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Lys-Arg-Leu-Glu)-Glu-Pro-,
 -Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Glu-Arg-Leu-Lys)-Glu-Pro-,
 -Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Lys-Arg-Leu-Asp)-Glu-Pro-,
 -Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Asp-Arg-Leu-Lys)-Glu-Pro-,
 -Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Lys-Arg-Leu-Glu)-Glu-Pro-,
 -Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Glu-Arg-Leu-Lys)-Glu-Pro-,
 -Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Cys-Arg-Leu-Thr-Cys)-Pro-,
 -Pro-Ser-Arg-Leu-Glu-cyclo(Cys-Glu-Leu-Arg-Cys)-Arg-Leu-Thr-Glu-Pro-,
 -Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-Cys)-Pro-,
 -Pro-Ser-Arg-Leu-Glu-Glu-cyclo(Cys-Leu-Arg-Arg-Cys)-Leu-Thr-Glu-Pro-,
 -Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-Arg-Glu)-Arg-Leu-Thr-Glu-, and
 -Ser-Arg-Leu-cyclo(Glu-Glu-Glu-Leu-Lys)-Arg-Arg-Leu-Thr-Glu-.

In some embodiments, the cyclic peptide is attached to a 3-mercaptopropionyl moiety through an α -amine moiety of the leftmost amino acid in the cyclic peptide. In some embodiments, the rightmost amino acid in the cyclic peptide comprises an amide.

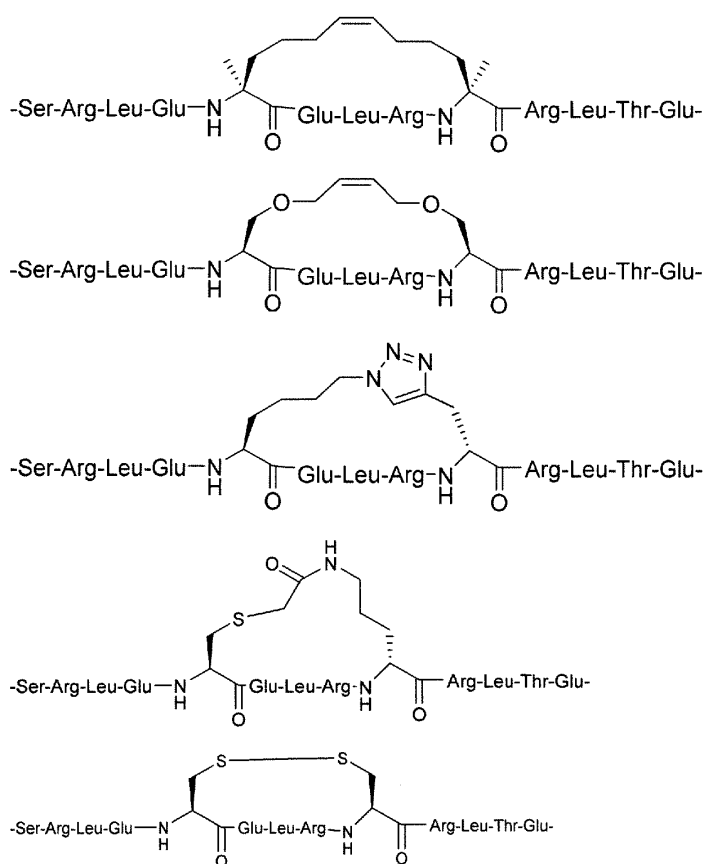
In some embodiments, the cyclized amino acid sequence is one selected from the group consisting of

-Ser-Arg-Leu-Glu-cyclo(Glu-Glu-Leu-Arg-Lys)-Arg-Leu-Thr-Glu-,
 -Ser-Arg-Leu-Glu-cyclo(Asp-Glu-Leu-Arg-Lys)-Arg-Leu-Thr-Glu-,
 -Ser-Arg-Leu-Glu-cyclo(Glu-Glu-Leu-Lys)-Arg-Arg-Leu-Thr-Glu-,
 -Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Lys-Arg-Leu-Thr-Glu)-,
 -Ser-Arg-Leu-Glu-cyclo(Cys-Glu-Leu-Arg-Cys)-Arg-Leu-Thr-Glu-,
 -Ser-Arg-Leu-Glu-cyclo(Asp-Glu-Leu-Lys)-Arg-Arg-Leu-Thr-Glu-,
 -Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Lys-Arg-Leu-Thr-Asp)-,
 -Ser-Arg-Leu-Glu-cyclo(Glu-Glu-Leu-Arg-DLys)-Arg-Leu-Thr-Glu-,
 -Pro-Ser-Arg-Leu-Glu-cyclo(Glu-Glu-Leu-Arg-Lys)-Arg-Leu-Thr-Glu-,
 -Pro-Ser-Arg-Leu-Glu-cyclo(DGlu-Glu-Leu-Arg-Lys)-Arg-Leu-Thr-Glu-,
 -Pro-Ser-Arg-Leu-Glu-cyclo(Glu-Glu-Leu-Arg-DLys)-Arg-Leu-Thr-Glu-,
 -Pro-Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-Arg-Glu)-Arg-Leu-Thr-Glu-,
 -Pro-Ser-Arg-Leu-cyclo(Glu-Glu-Glu-Leu-Lys)-Arg-Arg-Leu-Thr-Glu-,
 -Pro-Ser-Arg-Leu-Glu-cyclo(Cys-Glu-Leu-DCys)-Arg-Arg-Leu-Thr-Glu-,
 -Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Cys-Arg-Leu-Thr-Cys)-,
 -Pro-Ser-Arg-Leu-Glu-cyclo(Cys-Glu-Leu-Arg-Cys)-Arg-Leu-Thr-Glu-,
 -Pro-Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-Cys)-,
 -Pro-Ser-Arg-Leu-Glu-Glu-cyclo(Cys-Leu-Arg-Arg-Cys)-Leu-Thr-Glu-,
 -Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Cys-Arg-Leu-Thr-Cys)-,
 -Ser-Arg-Leu-Glu-cyclo(Cys-Glu-Leu-Arg-Cys)-Arg-Leu-Thr-Glu-,
 -Ser-Arg-Leu-Glu-cyclo(Cys-Glu-Leu-Cys)-Arg-Arg-Leu-Thr-Glu-,
 -Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-Arg-cyclo(Cys-Leu-Thr-Cys)-,
 -Ser-Arg-Leu-Glu-Glu-cyclo(Cys-Leu-Arg-Arg-Cys)-Leu-Thr-Glu-,
 -Ser-Arg-Leu-Glu-cyclo(Lys-Glu-Leu-Arg-Glu)-Arg-Leu-Thr-Glu-, and
 -Ser-Arg-Leu-cyclo(Glu-Glu-Glu-Leu-Lys)-Arg-Arg-Leu-Thr-Glu-.

In some embodiments, the cyclized amino acid sequence is -Ser-Arg-Leu-Glu-cyclo(Glu-Glu-Leu-Arg-Lys)-Arg-Leu-Thr-Glu-. In some other embodiments, the cyclized amino acid sequence is -Ser-Arg-Leu-Glu-cyclo(Asp-Glu-Leu-Arg-Lys)-Arg-Leu-Thr-Glu-. In yet some other

embodiments, the cyclized amino acid sequence is -Ser-Arg-Leu-Glu-cyclo(Glu-Glu-Leu-Lys)-Arg-Arg-Leu-Thr-Glu-. In still some other embodiments, the cyclized amino acid sequence is -Ser-Arg-Leu-Glu-Glu-Glu-Leu-Arg-cyclo(Lys-Arg-Leu-Thr-Glu)-.

The cyclic peptides may have different cyclic bridging moieties forming the ring structure. Preferably, chemically stable bridging moieties are included in the ring structure such as, for example, an amide group, a lactone group, an ether group, a thioether group, a disulfide group, an alkylene group, an alkenyl group, or a 1,2,3-triazole. The following are examples illustrating the variability of bridging moieties in a peptide:



The present disclosure provides in one aspect, a complex wherein an immune effector cell genetically modified to express a chimeric antigen receptor (CAR) is bound to a docking compound. Thus, the immune effector cell genetically modified to express a chimeric antigen receptor (CAR) and the docking compound comprise moieties interacting with each other.

Different embodiments of the immune effector cell genetically modified to express a chimeric antigen receptor (CAR) and the docking compound which are complexed are described herein. In some embodiments, the docking compound comprises an ALFA-tag. In these embodiments, the moiety binding to a docking compound of the antigen receptor on immune effector cells comprises a NbALFA-nanobody (NbALFA). In some embodiments, the docking compound may have a structure selected from the group consisting of ALFA-tag x anti-antigen on target cells DARPIn, ALFA-tag x anti-antigen on target cells VHH and ALFA-tag x anti-antigen on target cells scFv.

In some embodiments, a single-domain antibody (for use in the compositions and methods described herein) is characterized by the following items.

Item 1: A single-domain antibody that specifically binds to an epitope tag comprising the sequence SRLEEELRRRLTE (SEQ ID NO: 2), wherein the antibody comprises the CDR1 sequence as shown in SEQ ID NO: 94, the CDR2 sequence as shown in SEQ ID NO: 97, and the CDR3 sequence as shown in SEQ ID NO: 98; wherein the antibody further comprises at positions corresponding to sequence position 47, 48, and 50 of SEQ ID NO: 1 the following amino acid residues: Arg at position 47, Arg at position 48, Met at position 50; and wherein the antibody comprises a sequence having at least 90% sequence identity to any one of SEQ ID NOs: 118-125.

Item 2: The antibody of item 1, wherein the antibody is humanized.

Item 3: The antibody of any one of the preceding items, wherein the antibody comprises a sequence that has at least about 70% V segment identity to the closest human germline sequence.

Item 4: The antibody of item 3, wherein the human germline sequence is IGHV3-23*4 set forth in SEQ ID NO: 110 .

Item 5: The antibody of any one of the preceding items, wherein the antibody has an affinity for the epitope tag shown in SEQ ID NO: 2 that is at least comparable with the affinity of the antibody of SEQ ID NO: 1, wherein at least comparable preferably means that the K_D of the antibody is not more than by factor 2.5 higher than the K_D of SEQ ID NO: 1, as preferably measured by biolayer interferometry, preferably by a method as essentially described in Method 1.

Item 6: The antibody of any one of the preceding items, wherein the antibody has a K_D for the epitope tag shown in SEQ ID NO: 2 of about 5×10^{-10} M or lower, preferably 3×10^{-10} M or lower, as preferably measured by biolayer interferometry, preferably by a method as essentially described in Method 1.

Item 7: The antibody of any one of the preceding items, wherein the antibody has a higher T_m as compared to the antibody of SEQ ID NO: 1.

Item 8: The antibody of any one of the preceding items, wherein the antibody has a T_m of at least about 63°C, wherein the T_m is preferably measured as essentially described in Method 2.

Item 9: The antibody of any one the preceding items, wherein the antibody comprises a His residue at the position corresponding to sequence position 99 of SEQ ID NO: 1.

Item 10: The antibody of any one the preceding items, wherein the antibody comprises a Val residue at the position corresponding to sequence position 100 of SEQ ID NO: 1.

Item 11: The antibody of any one the preceding items, wherein the antibody comprises a Val residue at the position corresponding to sequence position 5 of SEQ ID NO: 1.

Item 12: The antibody of any one the preceding items, wherein the antibody comprises a Tyr residue at the position corresponding to sequence position 40 of SEQ ID NO: 1.

Item 13: The antibody of any one the preceding items, wherein the antibody comprises a Gln residue at the position corresponding to sequence position 117 of SEQ ID NO: 1.

Item 14: The antibody of any one of the preceding items, wherein the antibody comprises an FR1 sequence having at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or being identical to the sequence of SEQ ID NO: 99.

Item 15: The antibody of any one of the preceding items, wherein the antibody comprises an FR2 sequence having at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or being identical to the sequence of SEQ ID NO: 102.

Item 16: The antibody of any one of the preceding items, wherein the antibody comprises an FR3 sequence having at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or being identical to the sequence of SEQ ID NO: 108.

Item 17: The antibody of any one of the preceding items, wherein the antibody comprises an FR4 sequence having at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or being identical to the sequence of SEQ ID NO: 109.

Item 18: The antibody of any one of the preceding items, wherein the antibody comprises FR1, FR2, FR3, and FR4 sequences, that have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 102, 108, and 109, respectively.

Item 19: The antibody of any one of the preceding items, wherein the framework sequence of the antibody is defined by the amino acid positions corresponding to positions 1-25, 39-53, 61-98, and 112-122 of SEQ ID NO: 1.

Item 20: The antibody of any one of the preceding items, wherein the antibody comprises a protein A binding site.

Item 21: The antibody of any one of the preceding items, wherein the antibody comprises a Thr residue at the position corresponding to sequence position 60 of SEQ ID NO: 1.

Item 22: The antibody of any one of the preceding items, wherein the antibody comprises a Lys residue at the position corresponding to sequence position 67 of SEQ ID NO: 1.

Item 23: The antibody of any one of the preceding items, wherein the antibody comprises at positions corresponding to sequence positions 60 and 67 of SEQ ID NO: 1 the following amino acid residues: a Thr at position 60 and a Lys at position 67.

Item 24: The antibody of any one of the preceding items, wherein the antibody comprises a Tyr residue at the position corresponding to sequence position 61 of SEQ ID NO: 1.

Item 25: The antibody of any one of the preceding items, wherein the antibody comprises an Asp residue at the positions corresponding to sequence position 64 of SEQ ID NO: 1.

Item 26: The antibody of any one of the preceding items, wherein the antibody comprises a CDR2 sequence set forth in SEQ ID NOs: 95 or 96.

Item 27: The antibody of any one of the preceding items, wherein the antibody comprises a FR2 sequence having at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or being identical to any one of SEQ ID NOs: 100, 101.

Item 28: The antibody of any one of the preceding items, wherein the antibody comprises a FR3 sequence having at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or being identical to any one of SEQ ID NOs: 103-107.

Item 29: The antibody of any one of the preceding items, wherein the antibody comprises a sequence having at least 95% sequence identity to any one of SEQ ID NOs: 118-125.

Item 30: The antibody of any one of the preceding items, wherein the antibody comprises a sequence set forth in any one of SEQ ID NOs: 118-125.

In some embodiments, the antibody of any one of items 1-30 is used in combination with an ALFA-tag comprising the amino acid sequence of X1-X2-L-E-X5-E-X7-R-R-R-L-X12-X13 (SEQ ID NO: 3), wherein X1 is G or S or P or T, X2 is R or G or P, X5 is E or Q, X7 is L or I, X12 is S or T or P or A or D or E, and wherein X13 is P or A or S or K or D or E or no amino acid.

In some embodiments, the antibody of any one of items 1-30 is used in combination with an ALFA-tag comprising the sequence of X1-X2-L-E-X5-E-L-R-R-R-L-X12-X13 (SEQ ID NO: 4), wherein X1 is S or T, X2 is R or G, X5 is E or Q, X12 is T or D or E, and wherein X13 is A or D or E or no amino acid.

In some embodiments, the antibody of any one of items 1-30 is used in combination with an ALFA-tag comprising the sequence of SRLEEELRRRLTE (SEQ ID NO: 2).

In some embodiments, the antibody of any one of items 1-30 is used in combination with an ALFA-tag comprising a sequence selected from the group consisting of:

MPSRLEEELRRRLTEP (SEQ ID NO: 5);

PSRLEEELRRRLTEP (SEQ ID NO: 6);

PSRLEEELRRRLTE (SEQ ID NO: 7);

GRSRLEEELRRRLTA (SEQ ID NO: 8);

PGSRLEEELRRRLTAP (SEQ ID NO: 9);

PSTRLEEELRRRLTAP (SEQ ID NO: 10);

SPSRLEEELRRRLTAP (SEQ ID NO: 11);

SPSRLEEELRRRLDAP (SEQ ID NO: 12);

SPSRLEEELRRRLEAP (SEQ ID NO: 13);

SPSRLEEELRRRLTDP (SEQ ID NO: 14);

SPSRLEEELRRRLTEP (SEQ ID NO: 15);

SPSRLEEEELRRRLTADP (SEQ ID NO: 16);
SPSGLEEEELRRRLTEP (SEQ ID NO: 17);
GPSRLEEEELRRRLT (SEQ ID NO: 18);
GPSRLEEEELRRRLTA (SEQ ID NO: 19);
GPSRLEEEELRRRLTAA (SEQ ID NO: 20);
GPSRLEEEELRRRLTAAS (SEQ ID NO: 21);
SPSGLEQELRRRLTAP (SEQ ID NO: 22);
SPSGLEQELRRRLDAP (SEQ ID NO: 23);
SPSGLEQELRRRLEAP (SEQ ID NO: 24);
SPSGLEQELRRRLTEP (SEQ ID NO: 25);
GPSRLEEEELRRRLTAP (SEQ ID NO: 26);
GPSRLEEEELRRRLTEP (SEQ ID NO: 27);
GPSRLEEEELRRRLTE (SEQ ID NO: 28);
MSRLEEEELRRRLTEP (SEQ ID NO: 29); and
MSSRLEEEELRRRLTEP (SEQ ID NO: 30).

In some embodiments, the antibody of any one of items 1-30 is used in combination with an ALFA-tag comprising the sequence of X1-X2-L-E-X5-E-X7-R-R-R-L-X12-X13 (SEQ ID NO: 31), wherein X1 is G or S or P, X2 is R or G, X5 is E or Q, X7 is L or I, X12 is S or T or P or A, and X13 is P or A or S or no amino acid.

In some embodiments, the antibody of any one of items 1-30 is used in combination with an ALFA-tag comprising the sequence of G-R-L-E-E-L-R-R-R-L-S (SEQ ID NO: 32) or a variant thereof, wherein the variant has as compared to (SEQ ID NO: 32) 1 to 6 mutations selected from the group consisting of: G1→S, G1→P, R2→G, E5→Q, L7→I, S12→T, S12→P, S12→A, addition of P13, addition of A13, and addition of S13.

In some embodiments, the antibody of any one of items 1-30 is used in combination with an ALFA-tag comprising a sequence selected from the group consisting of:

GRLEEEELRRRLS (SEQ ID NO: 32);
MSGRLEEEELRRRLSP (SEQ ID NO: 33);
SDSGLEQELRRRLSPG (SEQ ID NO: 34);
PDGGLEQELRRRLTAP (SEQ ID NO: 35);

PSGGLEQELRRRLTAP (SEQ ID NO: 36);
DSPGGLEQELRRRLTAP (SEQ ID NO: 37);
PDSGGLEQELRRRLTPA (SEQ ID NO: 38);
SPSGLEELRRRLTAEP (SEQ ID NO: 39);
GPSGGLEQELRRRLT (SEQ ID NO: 40);
GPSGGLEQELRRRLTAAS (SEQ ID NO: 41);
SPSRLEELRRRLPSK (SEQ ID NO: 42);
SPSGLEQELRRRLTPS (SEQ ID NO: 43);
SPGRLEQEIRRLSPS (SEQ ID NO: 44);
PSGRLEELRRRLSPS (SEQ ID NO: 45);
PSGRLEELRRRLS (SEQ ID NO: 46);
PSGRLEELRRRLA (SEQ ID NO: 47); and
PSGRLEELRRRLSP (SEQ ID NO: 48).

Immune effector cells

The immune effector cells used in connection with the methods and agents described herein include, in particular, immune effector cells such as cells with lytic potential, in particular lymphoid cells, and are preferably T cells, in particular cytotoxic lymphocytes, preferably selected from cytotoxic T cells, natural killer (NK) cells, and lymphokine-activated killer (LAK) cells. Upon activation, each of these cytotoxic lymphocytes triggers the destruction of target cells. For example, cytotoxic T cells trigger the destruction of target cells by either or both of the following means. First, upon activation T cells release cytotoxins such as perforin, granzymes, and granulysin. Perforin and granulysin create pores in the target cell, and granzymes enter the cell and trigger a caspase cascade in the cytoplasm that induces apoptosis (programmed cell death) of the cell. Second, apoptosis can be induced via Fas-Fas ligand interaction between the T cells and target cells. The cells used in connection with the present disclosure will preferably be autologous cells, although heterologous cells or allogenic cells can be used.

The term "effector functions" in the context of the present disclosure includes any functions mediated by components of the immune system that result, for example, in the killing of

diseased cells such as tumor cells, or in the inhibition of tumor growth and/or inhibition of tumor development, including inhibition of tumor dissemination and metastasis. Preferably, the effector functions in the context of the present disclosure are T cell mediated effector functions. Such functions comprise in the case of a helper T cell (CD4⁺ T cell) the release of cytokines and/or the activation of CD8⁺ lymphocytes (CTLs) and/or B cells, and in the case of CTL the elimination of cells, i.e., cells characterized by expression of an antigen, for example, via apoptosis or perforin-mediated cell lysis, production of cytokines such as IFN- γ and TNF- α , and specific cytolytic killing of antigen expressing target cells.

The term "immune effector cell" or "immunoreactive cell" in the context of the present disclosure relates to a cell which exerts effector functions during an immune reaction. An "immune effector cell" in some embodiments is capable of binding an antigen such as an antigen presented by in the context of MHC on a cell or expressed on the surface of a cell and mediating an immune response. For example, immune effector cells comprise T cells (cytotoxic T cells, helper T cells, tumor infiltrating T cells), B cells, natural killer cells, neutrophils, macrophages, and dendritic cells. Preferably, in the context of the present disclosure, "immune effector cells" are T cells, preferably CD4⁺ and/or CD8⁺ T cells, most preferably CD8⁺ T cells. The term "immune effector cell" also includes a cell which can mature into an immune cell (such as T cell, in particular T helper cell, or cytolytic T cell) with suitable stimulation. Immune effector cells comprise CD34⁺ hematopoietic stem cells, immature and mature T cells and immature and mature B cells. The differentiation of T cell precursors into a cytolytic T cell, when exposed to an antigen, is similar to clonal selection of the immune system.

In some embodiments, the genetically modified immune effector cells are CAR-expressing immune effector cells.

The immune effector cells to be used herein may express an endogenous antigen receptor such as T cell receptor or B cell receptor or may lack expression of an endogenous antigen receptor.

A "lymphoid cell" is a cell which, optionally after suitable modification, e.g. after transfer of an antigen receptor such as a TCR or a CAR, is capable of producing an immune response such as a cellular immune response, or a precursor cell of such cell, and includes lymphocytes,

preferably T lymphocytes, lymphoblasts, and plasma cells. A lymphoid cell may be an immune effector cell as described herein. A preferred lymphoid cell is a T cell which can be modified to express an antigen receptor on the cell surface. In some embodiments, the lymphoid cell lacks endogenous expression of a T cell receptor.

The terms "T cell" and "T lymphocyte" are used interchangeably herein and include T helper cells (CD4⁺ T cells) and cytotoxic T cells (CTLs, CD8⁺ T cells) which comprise cytolytic T cells.

T cells belong to a group of white blood cells known as lymphocytes, and play a central role in cell-mediated immunity. They can be distinguished from other lymphocyte types, such as B cells and natural killer cells by the presence of a special receptor on their cell surface called T cell receptors (TCR). The thymus is the principal organ responsible for the maturation of T cells. Several different subsets of T cells have been discovered, each with a distinct function.

T helper cells assist other white blood cells in immunologic processes, including maturation of B cells into plasma cells and activation of cytotoxic T cells and macrophages, among other functions. These cells are also known as CD4⁺ T cells because they express the CD4 glycoprotein on their surface. Helper T cells become activated when they are presented with peptide antigens by MHC class II molecules that are expressed on the surface of antigen presenting cells (APCs). Once activated, they divide rapidly and secrete small proteins called cytokines that regulate or assist in the active immune response.

Cytotoxic T cells destroy virally infected cells and tumor cells, and are also implicated in transplant rejection. These cells are also known as CD8⁺ T cells since they express the CD8 glycoprotein on their surface. These cells recognize their targets by binding to antigen associated with MHC class I, which is present on the surface of nearly every cell of the body.

"Regulatory T cells" or "Tregs" are a subpopulation of T cells that modulate the immune system, maintain tolerance to self-antigens, and prevent autoimmune disease. Tregs are immunosuppressive and generally suppress or downregulate induction and proliferation of effector T cells. Tregs express the biomarkers CD4, FoxP3, and CD25.

As used herein, the term "naïve T cell" refers to mature T cells that, unlike activated or memory T cells, have not encountered their cognate antigen within the periphery. Naïve T cells are commonly characterized by the surface expression of L-selectin (CD62L), the absence of the activation markers CD25, CD44 or CD69 and the absence of the memory CD45RO isoform.

As used herein, the term "memory T cells" refers to a subgroup or subpopulation of T cells that have previously encountered and responded to their cognate antigen. At a second encounter with the antigen, memory T cells can reproduce to mount a faster and stronger immune response than the first time the immune system responded to the antigen. Memory T cells may be either CD4⁺ or CD8⁺ and usually express CD45RO.

As used herein, the term "T cell" also includes a cell which can mature into a T cell with suitable stimulation.

A majority of T cells have a T cell receptor (TCR) existing as a complex of several proteins. The actual T cell receptor is composed of two separate peptide chains, which are produced from the independent T cell receptor alpha and beta (TCR α and TCR β) genes and are called α - and β -TCR chains. $\gamma\delta$ T cells (gamma delta T cells) represent a small subset of T cells that possess a distinct T cell receptor (TCR) on their surface. However, in $\gamma\delta$ T cells, the TCR is made up of one γ -chain and one δ -chain. This group of T cells is much less common (2% of total T cells) than the $\alpha\beta$ T cells.

All T cells originate from hematopoietic stem cells in the bone marrow. Hematopoietic progenitors derived from hematopoietic stem cells populate the thymus and expand by cell division to generate a large population of immature thymocytes. The earliest thymocytes express neither CD4 nor CD8, and are therefore classed as double-negative (CD4⁻CD8⁻) cells. As they progress through their development they become double-positive thymocytes (CD4⁺CD8⁺), and finally mature to single-positive (CD4⁺CD8⁻ or CD4⁻CD8⁺) thymocytes that are then released from the thymus to peripheral tissues.

T cells may generally be prepared *in vitro* or *ex vivo*, using standard procedures. For example, T cells may be isolated from bone marrow, peripheral blood or a fraction of bone marrow or peripheral blood of a mammal, such as a patient, using a commercially available cell separation system. Alternatively, T cells may be derived from related or unrelated humans, non-human animals, cell lines or cultures. A sample comprising T cells may, for example, be peripheral blood mononuclear cells (PBMC).

As used herein, the term "NK cell" or "Natural Killer cell" refers to a subset of peripheral blood lymphocytes defined by the expression of CD56 or CD16 and the absence of the T cell receptor. As provided herein, the NK cell can also be differentiated from a stem cell or progenitor cell.

Chimeric antigen receptors (CAR)

T cells usually have a T cell receptor (TCR) for recognition of target cells through binding to antigen presented in the context of MHC. The term "T cell receptor" or "TCR" as used herein refers to a protein receptor on T cells that is composed of a heterodimer of an alpha (α) and beta (β) chain, although in some cells the TCR consists of gamma and delta ($\gamma\delta$) chains. Each α , β , γ , and δ chain is composed of two Ig-like domains: a variable domain (V) that confers antigen recognition through the complementarity determining regions (CDR), followed by a constant domain (C) that is anchored to cell membrane by a connecting peptide and a transmembrane (TM) region. The TM region associates with the invariant subunits of the CD3 signaling apparatus. Each of the V domains has three CDRs. These CDRs interact with a complex between an antigenic peptide bound to a protein encoded by the major histocompatibility complex (MHC).

T cells can be engineered to express chimeric antigen receptors (CAR) targeting virtually any antigen, preferably in an MHC-independent manner. For example, patient's T cells may be genetically engineered (genetically modified) to express CARs specifically directed towards antigens on the patient's tumor cells. Adoptive cell transfer therapy with CAR-engineered T cells expressing chimeric antigen receptors is a promising anti-cancer therapeutic.

As used herein, the term "CAR" (or "chimeric antigen receptor") is synonymous with the terms "chimeric T cell receptor" and "artificial T cell receptor" and relates to an artificial receptor comprising a single molecule or a complex of molecules which recognizes, i.e., binds to, a target structure (e.g. an antigen) (e.g. by binding of an antigen binding domain to an antigen) and may confer specificity onto an immune effector cell such as a T cell expressing said CAR on the cell surface. Such cells do not necessarily require processing and presentation of an antigen for recognition of the target cell but rather may recognize preferably with specificity any antigen. Preferably, recognition of the target structure by a CAR results in activation of an immune effector cell expressing said CAR. A CAR may comprise one or more protein units said protein units comprising one or more domains as described herein. The term "CAR" does not include T cell receptors.

A CAR generally comprises a target-specific binding element otherwise referred to as an antigen binding moiety or antigen binding domain that is generally part of the extracellular domain of the CAR.

In some embodiments, an antigen binding domain comprises a variable region of a heavy chain of an immunoglobulin (VH) with a specificity for the antigen and a variable region of a light chain of an immunoglobulin (VL) with a specificity for the antigen. In some embodiments, an immunoglobulin is an antibody. In some embodiments, said heavy chain variable region (VH) and the corresponding light chain variable region (VL) are connected via a peptide linker. The antigen binding moiety portion in the CAR may be a scFv. In some embodiments, an antigen binding domain comprises a VHH domain. The target-specific binding element or extracellular domain of a CAR described herein may comprise an ALFA-tag-binding single domain antibody described herein or an ALFA-tag. In some embodiments, the target-specific binding element or extracellular domain of a CAR described herein comprises an ALFA-tag-binding single domain antibody described herein

The CAR is preferably designed to comprise a transmembrane domain that is fused to the extracellular domain of the CAR. In some embodiments, the transmembrane domain is not naturally associated with one of the domains in the CAR. In some embodiments, the transmembrane domain is naturally associated with one of the domains in the CAR. In some embodiments, the transmembrane domain is modified by amino acid substitution to avoid binding of such domains to the transmembrane domains of the same or different surface membrane proteins to minimize interactions with other members of the receptor complex. The transmembrane domain may be derived either from a natural or from a synthetic source. Where the source is natural, the domain may be derived from any membrane-bound or transmembrane protein. Transmembrane regions of particular use herein may be derived from (i.e. comprise at least the transmembrane region(s) of) the alpha, beta or zeta chain of the T cell receptor, CD28, CD3 epsilon, CD45, CD4, CD5, CD8, CD9, CD16, CD22, CD33, CD37, CD64, CD80, CD86, CD134, CD137, CD154. Alternatively, the transmembrane domain may be synthetic, in which case it will comprise predominantly hydrophobic residues such as leucine and valine. Preferably a triplet of phenylalanine, tryptophan and valine will be found at each end of a synthetic transmembrane domain.

In some instances, the CAR comprises a hinge domain which forms the linkage between the transmembrane domain and the extracellular domain.

The cytoplasmic domain or otherwise the intracellular signaling domain of the CAR is responsible for activation of at least one of the normal effector functions of the immune cell in which the CAR has been placed in. The term "effector function" refers to a specialized function of a cell. Effector function of a T cell, for example, may be cytolytic activity or helper activity including the secretion of cytokines. Thus, the term "intracellular signaling domain" refers to the portion of a protein which transduces the effector function signal and directs the cell to perform a specialized function. While usually the entire intracellular signaling domain can be employed, in many cases it is not necessary to use the entire chain. To the extent that a truncated portion of the intracellular signaling domain is used, such truncated portion may be used in place of the intact chain as long as it transduces the effector function signal. The term intracellular signaling domain is thus meant to include any truncated portion of the intracellular signaling domain sufficient to transduce the effector function signal.

It is known that signals generated through the TCR alone are insufficient for full activation of the T cell and that a secondary or co-stimulatory signal is also required. Thus, T cell activation can be said to be mediated by two distinct classes of cytoplasmic signaling sequence: those that initiate antigen-dependent primary activation through the TCR (primary cytoplasmic signaling sequences) and those that act in an antigen-independent manner to provide a secondary or co-stimulatory signal (secondary cytoplasmic signaling sequences).

In some embodiments, the CAR comprises a primary cytoplasmic signaling sequence derived from CD3-zeta. Further, the cytoplasmic domain of the CAR may comprise the CD3-zeta signaling domain combined with a costimulatory signaling region.

The identity of the co-stimulation domain is limited only in that it has the ability to enhance cellular proliferation and survival upon binding of the targeted moiety by the CAR. Suitable co-stimulation domains include CD28, CD137 (4-1BB), a member of the tumor necrosis factor receptor (TNFR) superfamily, CD134 (OX40), a member of the TNFR-superfamily of receptors, and CD278 (ICOS), a CD28-superfamily co-stimulatory molecule expressed on activated T cells. The skilled person will understand that sequence variants of these noted co-stimulation domains can be used without adversely impacting the disclosure, where the variants have the

same or similar activity as the domain on which they are modeled. Such variants will have at least about 80% sequence identity to the amino acid sequence of the domain from which they are derived. In some embodiments, the CAR constructs comprise two co-stimulation domains. While the particular combinations include all possible variations of the four noted domains, specific examples include CD28+CD137 (4-1BB) and CD28+CD134 (OX40).

The cytoplasmic signaling sequences within the cytoplasmic signaling portion of the CAR may be linked to each other in a random or specified order. Optionally, a short oligo- or polypeptide linker, preferably between 2 and 10 amino acids in length may form the linkage. A glycine-serine doublet provides a particularly suitable linker.

In some embodiments, the CAR comprises a signal peptide which directs the nascent protein into the endoplasmic reticulum. In some embodiments, the signal peptide precedes the antigen binding domain. In some embodiments, the signal peptide is derived from an immunoglobulin such as IgG.

A CAR may comprise the above domains, together in the form of a fusion protein. Such fusion proteins will generally comprise a target-specific binding domain, one or more co-stimulation domains, and a signaling sequence, linked in a N-terminal to C-terminal direction. However, the CARs are not limited to this arrangement and other arrangements are acceptable and include a binding domain, a signaling domain, and one or more co-stimulation domains. It will be understood that because the binding domain must be free to bind the target, the placement of the binding domain in the fusion protein will generally be such that display of the region on the exterior of the cell is achieved. In the same manner, because the co-stimulation and signaling domains serve to induce activity and proliferation of the cytotoxic lymphocytes, the fusion protein will generally display these two domains in the interior of the cell.

In some embodiments, a CAR molecule comprises:

- i) a target binding domain;
- ii) a transmembrane domain; and
- iii) an intracellular domain that comprises a signaling domain, e.g., a CD3-zeta signaling domain, optionally in combination with one or more costimulatory domains, e.g., an intracellular domain that comprises a 4-1BB costimulatory domain.

In some embodiments, the target binding domain comprises an ALFA-tag-binding single-domain antibody described herein. In some embodiments, the target binding domain comprises an ALFA-tag. In some embodiments, the transmembrane domain comprises a transmembrane domain of a protein selected from the group consisting of the alpha, beta or zeta chain of the T cell receptor, CD28, CD3 epsilon, CD45, CD4, CD5, CD8, CD9, CD16, CD22, CD33, CD37, CD64, CD80, CD86, CD134, CD154, KIRDS2, OX40, CD2, CD27, LFA-1 (CD11a, CD18), ICOS (CD278), 4-1BB (CD137), GITR, CD40, BAFRR, HVEM (LIGHTR), SLAMF7, NKp80 (KLRF1), CD160, CD19, IL2R beta, IL2R gamma, IL7Ra, ITGA1, VLA1, CD49a, ITGA4, IA4, CD49D, ITGA6, VLA-6, CD49f, ITGAD, CD11d, ITGAE, CD103, ITGAL, CD11a, LFA-1, ITGAM, CD11b, ITGAX, CD11c, ITGB1, CD29, ITGB2, CD18, LFA-1, ITGB7, TNFR2, DNAMI (CD226), SLAMF4 (CD244, 2B4), CD84, CD96 (Tactile), CEACAM1, CRT AM, Ly9 (CD229), CD160 (BY55), PSGLI, CD100 (SEMA4D), SLAMF6 (NTB-A, Ly108), SLAM (SLAMF1, CD150, IPO-3), BLAME (SLAMF8), SELPLG (CD162), LTBR, PAG/Cbp, NKp44, NKp30, NKp46, NKG2D, and NKG2C, or a functional variant thereof. In some embodiments, the transmembrane domain comprises a CD8 α transmembrane domain. In some embodiments, the target binding domain is connected to the transmembrane domain by a hinge domain. In some embodiments, the hinge domain is a CD8 α hinge domain.

In some embodiments, the CAR molecule comprises:

- i) a target binding domain;
- ii) a CD8 α hinge domain;
- iii) a CD8 α transmembrane domain; and
- iv) an intracellular domain that comprises a 4-1BB costimulatory domain, and a CD3-zeta signaling domain.

Genetic modification of immune effector cells

Immune effector cells described herein are cells expressing an antigen receptor, e.g., a CAR, targeting cells through a docking compound. The immune effector cells are generally cells which are genetically modified *ex vivo* or *in vivo* to express an antigen receptor. The immune effector cells may be provided to a subject such as by administration of immune effector cells

such as genetically modified immune effector cells to the subject or generation of genetically modified immune effector cells in the subject.

Nucleic acid encoding an antigen receptor may be delivered *ex vivo/in vitro* or *in vivo* to immune effector cells such as T cells, in particular CD8⁺ T cells, thus producing immune effector cells genetically modified to express an antigen receptor.

In some embodiments, immune effector cells genetically modified to express a chimeric antigen receptor (CAR) are provided to a subject by administering nucleic acid, e.g., RNA, encoding the antigen receptor. Delivery of nucleic acid encoding the antigen receptor to immune effector cells for expression may be effected by using particles comprising the nucleic acid. The particles may comprise a targeting molecule that binds to a target, e.g., an antigen, on immune effector cells. In some embodiments, the target comprises CD3, such as CD3e, CD4 or CD8. In some embodiments, immune effector cells genetically modified to express a chimeric antigen receptor (CAR) are provided to a subject by administering the immune effector cells genetically modified to express a chimeric antigen receptor (CAR). In some embodiments, a preformed complex wherein the docking compound is bound to the immune effector cells genetically modified to express a chimeric antigen receptor (CAR) is provided to a subject by administration.

Genetic modification of immune effector cells to express an antigen receptor described herein includes non-viral-based DNA transfection, non-viral-based RNA transfection, e.g., mRNA transfection, transposon-based systems, and viral-based systems. Non-viral-based DNA transfection has low risk of insertional mutagenesis. Transposon-based systems can integrate transgenes more efficiently than plasmids that do not contain an integrating element. Viral-based systems include the use of γ -retroviruses and lentiviral vectors. γ -Retroviruses are relatively easy to produce, efficiently and permanently transduce cells such as T cells, and have preliminarily proven safe from an integration standpoint in primary human T cells. Lentiviral vectors also efficiently and permanently transduce cells such as T cells but are more expensive to manufacture. They are also potentially safer than retrovirus based systems.

In some embodiments, T cells or T cell progenitors are transfected either *ex vivo* or *in vivo* with nucleic acid encoding an antigen receptor. In some embodiments, a combination of *ex vivo* and *in vivo* transfection may be used. In some embodiments, the T cells or T cell progenitors

are from the subject to be treated. In some embodiments, the T cells or T cell progenitors are from a subject which is different to the subject to be treated.

In some embodiments, CAR T cells may be produced *in vivo*, and therefore nearly instantaneously, using particles such as nanoparticles targeted to T cells. Upon binding to T cells, these particles may be endocytosed. Their contents, for example nucleic acid encoding antigen receptor, e.g., plasmid DNA encoding a CAR, may be directed to the T cell nucleus due to, for example, the inclusion of peptides containing microtubule-associated sequences (MTAS) and nuclear localization signals (NLSs). The inclusion of transposons flanking the nucleic acid encoding antigen receptor, e.g., the CAR gene expression cassette, and a separate nucleic acid, e.g., plasmid, encoding a hyperactive transposase, may allow for the efficient integration of the nucleic acid encoding antigen receptor, e.g., the CAR vector, into chromosomes.

Another possibility is to use the CRISPR/Cas9 method to deliberately place a peptide/polypeptide coding sequence, e.g., an antigen receptor coding sequence such as a CAR coding sequence, at a specific locus. For example, existing T cell receptors (TCR) may be knocked out, while knocking in the CAR and placing it under the dynamic regulatory control of the endogenous promoter that would otherwise moderate TCR expression.

Accordingly, besides nucleic acid encoding an antigen receptor particles may also deliver as cargo gene editing tools like CRISPR/Cas9 (or related) or transposon systems like sleeping beauty or piggy bag. Such tools (e.g. transposase, gene editing tools like CRISPR/Cas9) for genomic integration/editing may be delivered as protein or coding nucleic acid (DNA or RNA). Nevertheless, also delivery of mRNA is an option to induce transient expression of antigen receptors like CAR.

In some embodiments, the cells genetically modified to express an antigen receptor are stably or transiently transfected with nucleic acid encoding the antigen receptor. Thus, the nucleic acid encoding the antigen receptor is integrated or not integrated into the genome of the cells. In some embodiments, the cells genetically modified to express an antigen receptor are inactivated for expression of an endogenous T cell receptor and/or endogenous HLA.

In some embodiments, the cells described herein may be autologous, allogeneic or syngeneic to the subject to be treated. In some embodiments, the present disclosure envisions the

removal of cells from a patient and the subsequent re-delivery of the cells to the patient. In some embodiments, the present disclosure does not envision the removal of cells from a patient. In the latter case all steps of genetic modification of cells may be performed *in vivo*.

Cells for targeted delivery

According to the disclosure, an immune effector cell is delivered specifically to a target cell by targeting a target on target cells, e.g., an antigen on target cells.

In some embodiments, the antigen on target cells is a structure such as a protein present on the surface of a target cell such as a cell surface antigen including a cell surface receptor.

Terms such as "expressed on the cell surface", "associated with the cell surface" or "cell surface molecule" mean that a molecule such as a receptor or antigen is associated with and located at the plasma membrane of a cell, wherein at least a part of the molecule faces the extracellular space of said cell and is accessible from the outside of said cell, e.g., by a binding molecule such as an antibody located outside the cell. In this context, a part is preferably at least 4, preferably at least 8, preferably at least 12, more preferably at least 20 amino acids. The association may be direct or indirect. For example, the association may be by one or more transmembrane domains, one or more lipid anchors, or by the interaction with any other protein, lipid, saccharide, or other structure that can be found on the outer leaflet of the plasma membrane of a cell. For example, a molecule associated with the surface of a cell may be a transmembrane protein having an extracellular portion or may be a protein associated with the surface of a cell by interacting with another protein that is a transmembrane protein. "Cell surface" or "surface of a cell" is used in accordance with its normal meaning in the art, and thus includes the outside of the cell which is accessible to binding by proteins and other molecules. An antigen is expressed on the surface of cells if it is located at the surface of said cells and is accessible to binding by e.g. antigen-specific antibodies added to the cells. In one embodiment, an antigen expressed on the surface of cells is an integral membrane protein having an extracellular portion recognized by a binding molecule such as an antibody.

The term "extracellular portion" or "exodomain" in the context of the present invention refers to a part of a molecule such as a protein that is facing the extracellular space of a cell and

preferably is accessible from the outside of said cell, e.g., by binding molecules such as antibodies located outside the cell.

In some embodiments, an antigen on target cells may be present on a diseased cell.

The antigen on target cells may be upregulated during a disease, e.g. infection or cancer. In diseased tissues, markers can differ from healthy tissue and offer unique possibilities for therapy, especially targeted therapy.

In some embodiments, the antigen on target cells is a disease-associated antigen, such as a tumor antigen, a viral antigen, or a bacterial antigen. This allows diseased cells to be targeted by the methods and agents described herein.

The term "disease-associated antigen" is used in its broadest sense to refer to any antigen associated with a disease. Disease-associated antigens may be associated with infection by microbes, typically microbial antigens, or associated with cancer, typically tumors.

In some embodiments, the antigen on target cells is a tumor antigen. In the context of the present disclosure, the term "tumor antigen" or "tumor-associated antigen" relates to proteins that are under normal conditions specifically expressed in a limited number of tissues and/or organs or in specific developmental stages, for example, the tumor antigen may be under normal conditions specifically expressed in stomach tissue, preferably in the gastric mucosa, in reproductive organs, e.g., in testis, in trophoblastic tissue, e.g., in placenta, or in germ line cells, and are expressed or aberrantly expressed in one or more tumor or cancer tissues. In this context, "a limited number" preferably means not more than 3, more preferably not more than 2. The tumor antigens in the context of the present disclosure include, for example, differentiation antigens, preferably cell type specific differentiation antigens, i.e., proteins that are under normal conditions specifically expressed in a certain cell type at a certain differentiation stage, cancer/testis antigens, i.e., proteins that are under normal conditions specifically expressed in testis and sometimes in placenta, and germ line specific antigens. In the context of the present disclosure, the tumor antigen is preferably associated with the cell surface of a cancer cell and is preferably not or only rarely expressed in normal tissues. Preferably, the tumor antigen or the aberrant expression of the tumor antigen identifies cancer cells. In the context of the present disclosure, the tumor antigen that is expressed by a cancer cell in a subject, e.g., a patient suffering from a cancer disease, is

preferably a self-protein in said subject. In preferred embodiments, the tumor antigen in the context of the present disclosure is expressed under normal conditions specifically in a tissue or organ that is non-essential, i.e., tissues or organs which when damaged by the immune system do not lead to death of the subject, or in organs or structures of the body which are not or only hardly accessible by the immune system. Preferably, the amino acid sequence of the tumor antigen is identical between the tumor antigen which is expressed in normal tissues and the tumor antigen which is expressed in cancer tissues.

Examples for tumor antigens include p53, ART-4, BAGE, beta-catenin/m, Bcr-abl CAMEL, CAP-1, CASP-8, CDC27/m, CDK4/m, CEA, the cell surface proteins of the claudin family, such as CLAUDIN-6, CLAUDIN-18.2 and CLAUDIN-12, c-MYC, CT, Cyp-B, DAM, ELF2M, ETV6-AML1, G250, GAGE, GnT-V, Gap100, HAGE, HER-2/neu, HPV-E7, HPV-E6, HAST-2, hTERT (or hTRT), LAGE, LDLR/FUT, MAGE-A, preferably MAGE-A1, MAGE-A2, MAGE-A3, MAGE-A4, MAGE-A5, MAGE-A6, MAGE-A7, MAGE-A8, MAGE-A9, MAGE-A10, MAGE-A11, or MAGE-A12, MAGE-B, MAGE-C, MART-1/Melan-A, MC1R, Myosin/m, MUC1, MUM-1, -2, -3, NA88-A, NF1, NY-ESO-1, NY-BR-1, p190 minor BCR-abL, Pm1/RARa, PRAME, proteinase 3, PSA, PSM, RAGE, RU1 or RU2, SAGE, SART-1 or SART-3, SCGB3A2, SCP1, SCP2, SCP3, SSX, SURVIVIN, TEL/AML1, TPI/m, TRP-1, TRP-2, TRP-2/INT2, TPTE and WT. Particularly preferred tumor antigens include CLAUDIN-18.2 (CLDN18.2) and CLAUDIN-6 (CLDN6).

In some embodiments, the antigen on target cells is a structure such as a protein present on the surface of a target cell such as a cell surface antigen or cell surface receptor the presence or amount of which is characteristic for certain cell types compared to others. This allows certain cell types characterized by the presence or increased amounts to be targeted by the methods and agents described herein.

Binding moieties and agents

The present disclosure describes binding moieties or agents such as antibodies or antibody derivatives. Moreover, the disclosure describes bispecific or multispecific binding agents such as bispecific molecules comprising a first and a second binding domain, wherein the first binding domain is capable of binding to a antigen on target cells and the second binding domain is capable of binding to an antigen receptor on immune effector cells.

The term "binding agent" as used herein refers to any agent capable of binding to desired antigens. In certain embodiments, the binding agent is or comprises an antibody, antibody fragment, or any other binding protein, or any combination thereof.

The term "binding moiety" as used herein refers to any moiety, group or domain capable of binding to desired antigens. In certain embodiments, the binding moiety is or comprises an antibody, antibody fragment, or any other binding protein, or any combination thereof.

As used herein, the term "antigen" is a molecule capable of being bound by a binding moiety or agent, such as an antibody. An antigen may additionally be capable of inducing a humoral immune response and/or cellular immune response leading to the production of B- and/or T-lymphocytes. An antigen may have one or more epitopes (B-cell and T-cell epitopes).

The term "immunoglobulin" refers to a class of structurally related glycoproteins consisting of two pairs of polypeptide chains, one pair of light (L) low molecular weight chains and one pair of heavy (H) chains, all four inter-connected by disulfide bonds. The structure of immunoglobulins has been well characterized. See for instance Fundamental Immunology Ch. 7 (Paul, W., ed., 2nd ed. Raven Press, N.Y. (1989)). Briefly, each heavy chain typically is comprised of a heavy chain variable region (abbreviated herein as V_H or VH) and a heavy chain constant region (abbreviated herein as C_H or CH). The heavy chain constant region typically is comprised of three domains, $CH1$, $CH2$, and $CH3$. The hinge region is the region between the $CH1$ and $CH2$ domains of the heavy chain and is highly flexible. Disulphide bonds in the hinge region are part of the interactions between two heavy chains in an IgG molecule. Each light chain typically is comprised of a light chain variable region (abbreviated herein as V_L or VL) and a light chain constant region (abbreviated herein as C_L or CL). The light chain constant region typically is comprised of one domain, CL . The VH and VL regions may be further subdivided into regions of hypervariability (or hypervariable regions which may be hypervariable in sequence and/or form of structurally defined loops), also termed complementarity determining regions (CDRs), interspersed with regions that are more conserved, termed framework regions (FRs). Each VH and VL is typically composed of three CDRs and four FRs, arranged from amino-terminus to carboxy-terminus in the following order: FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4 (see also Chothia and Lesk J. Mol. Biol. 196, 901-917 (1987)).

The term "antibody" (Ab) as used herein refers to an immunoglobulin molecule, a fragment of an immunoglobulin molecule, or a derivative of either thereof, which has the ability to bind, preferably specifically bind to an antigen. In some embodiments, binding takes place under typical physiological conditions with a half-life of significant periods of time, such as at least about 30 minutes, at least about 45 minutes, at least about one hour, at least about two hours, at least about four hours, at least about 8 hours, at least about 12 hours, about 24 hours or more, about 48 hours or more, about 3, 4, 5, 6, 7 or more days, etc., or any other relevant functionally-defined period (such as a time sufficient to induce, promote, enhance, and/or modulate a physiological response associated with antibody binding to the antigen). The variable regions of the heavy and light chains of the immunoglobulin molecule contain a binding domain that interacts with an antigen. The term "antigen-binding region", "binding region" or "binding domain", as used herein, refers to the region or domain which interacts with the antigen and typically comprises both a VH region and a VL region. The term antibody when used herein comprises not only monospecific antibodies, but also multispecific antibodies which comprise multiple, such as two or more, e.g. three or more, different antigen-binding regions. The constant regions of the antibodies (Abs) may mediate the binding of the immunoglobulin to host tissues or factors, including various cells of the immune system (such as effector cells) and components of the complement system such as C1q, the first component in the classical pathway of complement activation. As indicated above, the term antibody as used herein, unless otherwise stated or clearly contradicted by context, includes fragments of an antibody that are antigen-binding fragments, *i.e.*, retain the ability to specifically bind to the antigen, and antibody derivatives, *i.e.*, constructs that are derived from an antibody. It has been shown that the antigen-binding function of an antibody may be performed by fragments of a full-length antibody. Examples of antigen-binding fragments encompassed within the term "antibody" include (i) a Fab' or Fab fragment, a monovalent fragment consisting of the VL, VH, CL and CH1 domains, or a monovalent antibody as described in WO2007059782 (Genmab); (ii) F(ab')₂ fragments, bivalent fragments comprising two Fab fragments linked by a disulfide bridge at the hinge region; (iii) a Fd fragment consisting essentially of the VH and CH1 domains; (iv) a Fv fragment consisting essentially of the VL and VH domains of a single arm of an antibody, (v) a dAb fragment (Ward et al., Nature 341,

544-546 (1989)), which consists essentially of a VH domain and also called domain antibodies (Holt et al; Trends Biotechnol. 2003 Nov;21(11):484-90); (vi) camelid or Nanobody molecules (Revets et al; Expert Opin Biol Ther. 2005 Jan;5(1):111-24) and (vii) an isolated complementarity determining region (CDR). Furthermore, although the two domains of the Fv fragment, VL and VH, are coded for by separate genes, they may be joined, using recombinant methods, by a synthetic linker that enables them to be made as a single protein chain in which the VL and VH regions pair to form monovalent molecules (known as single chain antibodies or single chain Fv (scFv), see for instance Bird et al., Science 242, 423-426 (1988) and Huston et al., PNAS USA 85, 5879-5883 (1988)). Such single chain antibodies are encompassed within the term antibody unless otherwise noted or clearly indicated by context. Although such fragments are generally included within the meaning of antibody, they collectively and each independently are unique features of the present disclosure, exhibiting different biological properties and utility. These and other useful antibody fragments in the context of the present disclosure, as well as bispecific formats of such fragments, are discussed further herein. It also should be understood that the term antibody, unless specified otherwise, also includes polyclonal antibodies, monoclonal antibodies (mAbs), antibody-like polypeptides, such as chimeric antibodies and humanized antibodies, and antibody fragments retaining the ability to specifically bind to the antigen (antigen-binding fragments) provided by any known technique, such as enzymatic cleavage, peptide synthesis, and recombinant techniques.

The terms "Fab", "Fab region", "Fab portion" or "Fab fragment" are understood to define a polypeptide that includes a V_H, a C_{H1}, a V_L, and a C_L immunoglobulin domain. Fab may refer to this region in isolation, or this region in the context of an antibody molecule according to the invention, as well as a full-length immunoglobulin or immunoglobulin fragment. Typically a Fab region contains an entire light chain of an antibody. A Fab region can be taken to define "an arm" of an immunoglobulin molecule. It contains the epitope-binding portion of that Ig. The Fab region of a naturally occurring immunoglobulin can be obtained as a proteolytic fragment by a partial papain-digestion. A "F(ab')₂ portion" is the proteolytic fragment of a partially pepsin-digested immunoglobulin. A "Fab' portion" is the product resulting from reducing the disulfide bonds of an F(ab')₂ portion. As used herein the terms "Fab", "Fab

region", "Fab portion" or "Fab fragment" may further include a hinge region that defines the C-terminal end of the antibody arm. This hinge region corresponds to the hinge region found C-terminally of the C_H1 domain within a full-length immunoglobulin at which the arms of the antibody molecule can be taken to define a Y. The term hinge region is used in the art because immunoglobulin has some flexibility at this region.

An "Fv" or "Fv fragment" consists of only the VL and VH domains of a "single arm" of an immunoglobulin. Thus an "Fv" is the minimum antibody fragment which contains a complete antigen-recognition and binding site. A "two-chain" Fv fragment consists of a dimer of one heavy- and one light-chain variable domain in a tight, non-covalent association. A single-chain Fv species (scFv) includes a VH and a VL domain of an immunoglobulin, with these domains being present in a single polypeptide chain in which they are covalently linked to each other by a flexible peptide linker. Typically, in a scFv fragment the variable domains of the light and heavy chain associate in a dimeric structure analogous to that in a two-chain Fv species. In single chain Fv fragments, it is possible to either have the variable domain of the light chain arranged at the N-terminus of the single polypeptide chain, followed by the linker and the variable domain of the heavy chain arranged at the C-terminus of the polypeptide chain or vice versa, having the variable domain of the heavy chain arranged on the N-terminus and the variable domain of the light chain at the C-terminus with the peptide linker arranged in between. The peptide linker can be any flexible linker known in the art, for example, made from glycine and serine residues. It is also possible to additionally stabilize the domain association between the VH and the VL domain by introducing disulfide bonds into conserved framework regions (see Reiter et al. Stabilization of the Fv fragments in recombinant immunotoxins by disulfide bonds engineered into conserved framework regions, *Biochemistry* 1994, 33, 6551-5459). Such scFv fragments are also known as disulfide-stabilized scFv fragments (ds-scFv).

The term "Fc region" or "Fc fragment" is used herein to define a C-terminal region of an immunoglobulin heavy chain, including native-sequence Fc regions and variant Fc regions. The Fc part mediates the effector function of antibodies, e.g. the activation of the complement system and of Fc-receptor bearing immune effector cells, such as NK cells. In human IgG molecules, the Fc region is generated by papain cleavage N-terminal to Cys226. Although the

boundaries of the Fc region of an immunoglobulin heavy chain might vary, the human IgG heavy-chain Fc region is usually defined to stretch from an amino acid residue at position Cys226, or from Pro230, to the carboxyl-terminus thereof. The C-terminal lysine (residue 447 according to the EU numbering system) of the Fc region may be removed, for example, during production or purification of the antibody molecule, or by recombinantly engineering the nucleic acid encoding a heavy chain of the antibody molecule. Accordingly, a composition of intact antibodies may include antibody populations with all K447 residues removed, antibody populations with no K447 residues removed, and antibody populations having a mixture of antibodies with and without the K447 residue. Suitable native-sequence Fc regions for use in the antibodies described herein include mammalian, e.g. human or murine, IgG1, IgG2 (IgG2A, IgG2B), IgG3 and IgG4. The Fc region contains two or three constant domains, depending on the class of the antibody. In embodiments where the immunoglobulin is an IgG the Fc region has a CH2 and a CH3 domain.

An antibody can possess any isotype. As used herein, the term "isotype" refers to the immunoglobulin class (for instance IgG1, IgG2, IgG3, IgG4, IgD, IgA, IgE, or IgM) that is encoded by heavy chain constant region genes. When a particular isotype, e.g. IgG1, is mentioned herein, the term is not limited to a specific isotype sequence, e.g. a particular IgG1 sequence, but is used to indicate that the antibody is closer in sequence to that isotype, e.g. IgG1, than to other isotypes. Thus, e.g. an IgG1 antibody may be a sequence variant of a naturally-occurring IgG1 antibody, including variations in the constant regions.

In various embodiments, an antibody is an IgG1 antibody, more particularly an IgG1, kappa or IgG1, lambda isotype (i.e. IgG1, κ , λ), an IgG2a antibody (e.g. IgG2a, κ , λ), an IgG2b antibody (e.g. IgG2b, κ , λ), an IgG3 antibody (e.g. IgG3, κ , λ) or an IgG4 antibody (e.g. IgG4, κ , λ).

The term "full-length" when used in the context of an antibody indicates that the antibody is not a fragment, but contains all of the domains of the particular isotype normally found for that isotype in nature, e.g. the VH, CH1, CH2, CH3, hinge, VL and CL domains for an IgG1 antibody.

An antibody or antibody molecule/fragment is said to "specifically" bind to an antigen when it recognizes its target antigen within a complex mixture of proteins and/or macromolecules. Typically, the antibody is capable of specifically interacting with and/or binding to its target

but does not essentially bind to another (preferably non-related) epitope or antigen. Antibodies are said to "bind to the same epitope" if the antibodies cross-compete so that only one antibody can bind to the epitope at a given point of time, i.e. one antibody prevents the binding or modulating effect of the other. An antibody or antibody molecule/fragment that specifically binds to a certain target, can however be cross-reactive with structures that are similar, such as with closely related variants of the target it specifically binds to.

Typically, binding that is considered specific may also have a high affinity, e.g. when the binding affinity is higher than 10^{-6} M (in terms of K_D). In particular, the binding affinity may be about 10^{-8} to 10^{-11} M (K_D), or of about 10^{-9} to 10^{-11} M or even higher. Thus, antibody molecules with an affinity in the picomolar range (with a K_D of 9.9×10^{-10} M to 10^{-12} M) are also encompassed in the present invention. If necessary, nonspecific binding of a binding site can be reduced without substantially affecting specific binding by varying the binding conditions. The term "DARPin" refers to designed ankyrin repeat proteins. DARPins are based on naturally occurring ankyrin repeat proteins, yet contain one or more amino acid mutations that can affect, for example, their binding affinity to a target molecule, their cell surface expression, and the like. DARPins preferably include 2 to 3 ankyrin repeat modules flanked by N- and C-capping repeats. Each ankyrin repeat module includes about 33 amino acid residues.

Ankyrin repeat proteins have been identified in 1987 through sequence comparisons between four such proteins in *Saccharomyces cerevisiae*, *Drosophila melanogaster* and *Caenorhabditis elegans*. Breeden and Nasmyth reported multiple copies of a repeat unit of approximately 33 residues in the sequences of swi6p, cddOp, notch and lin-12 (Breeden et al., Nature 329, 651–654 (1987)). The subsequent discovery of 24 copies of this repeat unit in the ankyrin protein led to the naming of this repeat unit as the ankyrin repeat (Lux et al., Nature 344, 36–42 (1990)). Later, this repeat unit has been identified in several hundreds of proteins of different organisms and viruses (Bork, Proteins 17(4), 363–74 (1993)). These proteins are located in the nucleus, the cytoplasm or the extracellular space. This is consistent with the fact that the ankyrin repeat domain of these proteins is independent of disulfide bridges and thus independent of the oxidation state of the environment. The number of repeat units per protein varies from two to more than twenty. Tertiary structures of ankyrin repeat units share a characteristic fold (Sedgwick and Smerdon, Trends Biochem Sci. 24(8), 311–6 (1999))

composed of a β -hairpin followed by two antiparallel α -helices and ending with a loop connecting the repeat unit with the next one. Domains built of ankyrin repeat units are formed by stacking the repeat units to an extended and curved structure. Proteins containing ankyrin repeat domains often contain additional domains. While the latter domains have variable functions, the function of the ankyrin repeat domain is most often the binding of other proteins. When analysing the repeat units of these proteins, the target interaction residues are mainly found in the β -hairpin and the exposed part of the first α -helix. These target interaction residues are hence forming a large contact surface on the ankyrin repeat domain. This contact surface is exposed on a framework built of stacked units of α -helix 1, α -helix 2 and the loop.

DARPin that bind to specific targets can be identified by screening combinatorial libraries of DARPins and selecting those with desired binding properties for the target. Such screening methods are described in, e.g., Muench et al., *Molecular Therapy*, 16(4), 686-693, 2011. For example, ribosomal display or phage display methods can be used to select target-specific DARPins from diverse libraries.

The term "repeat protein" refers to a (poly)peptide/protein comprising one or more repeat domains. In one embodiment, a repeat protein comprises up to four repeat domains. In one embodiment, a repeat protein comprises up to three repeat domains. In one embodiment, a repeat protein comprises up to two repeat domains. In the most preferred embodiment, a repeat protein comprises one repeat domain.

The individual domains of a repeat protein may be connected to each other directly or via (poly)peptide linkers. The term "(poly)peptide linker" refers to an amino acid sequence which is able to link two protein domains. Such linkers include, for example, glycine-serine-linkers of variable lengths and are known to the person skilled in the relevant art.

The term "repeat domain" refers to a protein domain comprising two or more consecutive repeat units (modules). In one embodiment, said repeat units are structural units having the same or a similar folding structure, and preferably stack tightly to preferably create a superhelical structure having a joint hydrophobic core.

The term "structural unit" refers to a locally ordered part of a (poly)peptide, formed by three-dimensional interactions between two or more segments of secondary structure that are near one another along the (poly)peptide chain. Such a structural unit comprises a structural motif. The term "structural motif" refers to a three-dimensional arrangement of secondary structure elements present in at least one structural unit. Structural motifs are well known to the person skilled in the relevant art. Said structural units may alone not be able to acquire a defined three-dimensional arrangement; however, their consecutive arrangement as repeat modules in a repeat domain leads to a mutual stabilization of neighbouring units which may result in a superhelical structure.

The term "repeat modules" refers to the repeated amino acid sequences of the repeat proteins, which are derived from the repeat units of naturally occurring proteins. Each repeat module comprised in a repeat domain is derived from one or more repeat units of a family of naturally occurring repeat proteins, e.g., ankyrin repeat proteins.

The term "set of repeat modules" refers to the total number of repeat modules present in a repeat domain. Such "set of repeat modules" present in a repeat domain comprises two or more consecutive repeat modules, and may comprise just one type of repeat module in two or more copies, or two or more different types of modules, each present in one or more copies. Such set of repeat modules comprising, for example, 3 repeat modules may comprise consecutively, from N- to C-terminus, repeat module 1, repeat module 2, and repeat module 3.

Different repeat domains may have an identical number of repeat modules per repeat domain or may differ in the number of repeat modules per repeat domain.

Preferably, the repeat modules comprised in a set are homologous repeat modules. In the context of the present disclosure, the term "homologous repeat modules" refers to repeat modules, wherein more than 70% of the framework residues of said repeat modules are homologous. Preferably, more than 80% of the framework residues of said repeat modules are homologous. Most preferably, more than 90% of the framework residues of said repeat modules are homologous. Computer programs to determine the percentage of homology between polypeptides, such as Fasta, Blast or Gap, are known to the person skilled in the relevant art.

The term "repeat unit" refers to amino acid sequences comprising sequence motifs of one or more naturally occurring proteins, wherein said "repeat units" are found in multiple copies, and which exhibit a defined folding topology common to all said motifs determining the fold of the protein. Such repeat units comprise framework residues and interaction residues.

One example of such repeat units is an ankyrin repeat unit. Naturally occurring proteins containing two or more such repeat units are referred to as "naturally occurring repeat proteins". The amino acid sequences of the individual repeat units of a repeat protein may have a significant number of mutations, substitutions, additions and/or deletions when compared to each other, while still substantially retaining the general pattern, or motif, of the repeat units.

The term "repeat sequence motif" or "repeat consensus sequence" refers to an amino acid sequence, which is deduced from one or more repeat units. Such repeat sequence motifs comprise framework residue positions and target interaction residue positions. Said framework residue positions correspond to the positions of framework residues of said repeat units. Said target interaction residue positions correspond to the positions of target interaction residues of said repeat units. Such repeat sequence motifs comprise fixed positions and randomized positions. The term "fixed position" refers to an amino acid position in a repeat sequence motif, wherein said position is set to a particular amino acid. Frequently, such fixed positions correspond to the positions of framework residues.

The term "randomized position" refers to an amino acid position in a repeat sequence motif, wherein two or more amino acids are allowed at said amino acid position. Frequently, such randomized positions correspond to the positions of target target interaction residues. However, some positions of framework residues may also be randomized.

The term "folding topology" refers to the tertiary structure of said repeat units. The folding topology will be determined by stretches of amino acids forming at least parts of α -helices or β -sheets, or amino acid stretches forming linear polypeptides or loops, or any combination of α -helices, β -sheets and/or linear polypeptides/loops.

The term "consecutive" refers to an arrangement, wherein said modules are arranged in tandem.

In repeat proteins, there are at least 2, frequently 6 or more, 10 or more, or 20 or more repeat units, usually about 2 to 6 repeat units. For the most part, the repeat proteins are structural proteins and/or adhesive proteins, being present in prokaryotes and eukaryotes, including vertebrates and non-vertebrates.

In most cases, said repeat units will exhibit a high degree of sequence identity (same amino acid residues at corresponding positions) or sequence similarity (amino acid residues being different, but having similar physicochemical properties), and some of the amino acid residues might be key residues being strongly conserved in the different repeat units found in naturally occurring proteins.

However, a high degree of sequence variability by amino acid insertions and/or deletions, and/or substitutions between the different repeat units found in naturally occurring proteins will be possible as long as the common folding topology is maintained.

The term "framework residues" relates to amino acid residues of the repeat units, or the corresponding amino acid residues of the repeat modules, which contribute to the folding topology, i.e. which contribute to the fold of said repeat unit (or module) or which contribute to the interaction with a neighboring unit (or module). Such contribution might be the interaction with other residues in the repeat unit (module), or the influence on the polypeptide backbone conformation as found in α -helices or β -sheets, or amino acid stretches forming linear polypeptides or loops.

The term "target interaction residues" refers to amino acid residues of the repeat units, or the corresponding amino acid residues of the repeat modules, which contribute to the interaction with target substances. Such contribution might be the direct interaction with the target substances, or the influence on other directly interacting residues, e.g. by stabilising the conformation of the (poly)peptide of said repeat unit (module) to allow or enhance the interaction of said directly interacting residues with said target.

A "target" may be an individual molecule such as a nucleic acid molecule, a (poly)peptide protein, a carbohydrate, or any other naturally occurring molecule, including any part of such individual molecule, or complexes of two or more of such molecules. The target may be, in particular, a molecule on immune effector cells, in particular CD8.

In one embodiment, the repeat modules are directly connected. In the context of the present invention, the term "directly connected" refers to repeat modules, which are arranged as direct repeats in a repeat protein without an intervening amino acid sequence.

In another embodiment, the repeat modules are connected by a (poly)peptide linker. Thus, the repeat modules may be linked indirectly via a (poly)peptide linker as intervening sequence separating the individual modules. An "intervening sequence" may be any amino acid sequence, which allows to connect the individual modules without interfering with the folding topology or the stacking of the modules. Preferentially, said intervening sequences are short (poly)peptide linkers of less than 10, and even more preferably, of less than 5 amino acid residues.

In one embodiment, a repeat protein further comprises an N- and/or a C-terminal capping module having an amino acid sequence different from any one of said repeat modules. The term "capping module" refers to a polypeptide fused to the N- or C- terminal repeat module of a repeat domain, wherein said capping module forms tight tertiary interactions with said repeat module thereby providing a cap that shields the hydrophobic core of said repeat module at the side not in contact with the consecutive repeat module from the solvent.

Said N- and/or C-terminal capping module may be, or may be derived from, a capping unit or other domain found in a naturally occurring repeat protein adjacent to a repeat unit.

The term "capping unit" refers to a naturally occurring folded (poly)peptide, wherein said (poly)peptide defines a particular structural unit which is N- or C-terminally fused to a repeat unit, wherein said (poly)peptide forms tight tertiary interactions with said repeat unit thereby providing a cap that shields the hydrophobic core of said repeat unit at one side from the solvent. Such capping units may have sequence similarities to said repeat sequence motif.

Nucleic acids

The term "nucleic acid" comprises deoxyribonucleic acid (DNA), ribonucleic acid (RNA), combinations thereof, and modified forms thereof. The term comprises genomic DNA, cDNA, mRNA, recombinantly produced and chemically synthesized molecules. In some embodiments, a nucleic acid is DNA. In some embodiments, a nucleic acid is RNA. In some embodiments, a nucleic acid is a mixture of DNA and RNA. A nucleic acid may be present as a

single-stranded or double-stranded and linear or covalently circularly closed molecule. A nucleic acid can be isolated. The term "isolated nucleic acid" means, according to the present disclosure, that the nucleic acid (i) was amplified *in vitro*, for example via polymerase chain reaction (PCR) for DNA or *in vitro* transcription (using, *e.g.*, an RNA polymerase) for RNA, (ii) was produced recombinantly by cloning, (iii) was purified, for example, by cleavage and separation by gel electrophoresis, or (iv) was synthesized, for example, by chemical synthesis. The term "nucleoside" (abbreviated herein as "N") relates to compounds which can be thought of as nucleotides without a phosphate group. While a nucleoside is a nucleobase linked to a sugar (*e.g.*, ribose or deoxyribose), a nucleotide is composed of a nucleoside and one or more phosphate groups. Examples of nucleosides include cytidine, uridine, pseudouridine, adenosine, and guanosine.

The five standard nucleosides which usually make up naturally occurring nucleic acids are uridine, adenosine, thymidine, cytidine and guanosine. The five nucleosides are commonly abbreviated to their one letter codes U, A, T, C and G, respectively. However, thymidine is more commonly written as "dT" ("d" represents "deoxy") as it contains a 2'-deoxyribofuranose moiety rather than the ribofuranose ring found in uridine. This is because thymidine is found in deoxyribonucleic acid (DNA) and not ribonucleic acid (RNA). Conversely, uridine is found in RNA and not DNA. The remaining three nucleosides may be found in both RNA and DNA. In RNA, they would be represented as A, C and G, whereas in DNA they would be represented as dA, dC and dG.

A modified purine (A or G) or pyrimidine (C, T, or U) base moiety is, in some embodiments, modified by one or more alkyl groups, *e.g.*, one or more C₁₋₄ alkyl groups, *e.g.*, one or more methyl groups. Particular examples of modified purine or pyrimidine base moieties include N⁷-alkyl-guanine, N⁶-alkyl-adenine, 5-alkyl-cytosine, 5-alkyl-uracil, and N(1)-alkyl-uracil, such as N⁷-C₁₋₄ alkyl-guanine, N⁶-C₁₋₄ alkyl-adenine, 5-C₁₋₄ alkyl-cytosine, 5-C₁₋₄ alkyl-uracil, and N(1)-C₁₋₄ alkyl-uracil, preferably N⁷-methyl-guanine, N⁶-methyl-adenine, 5-methyl-cytosine, 5-methyl-uracil, and N(1)-methyl-uracil.

DNA

Herein, the term "DNA" relates to a nucleic acid molecule which includes deoxyribonucleotide residues. In preferred embodiments, the DNA contains all or a majority of deoxyribonucleotide residues. As used herein, "deoxyribonucleotide" refers to a nucleotide which lacks a hydroxyl group at the 2'-position of a β -D-ribofuranosyl group. DNA encompasses without limitation, double stranded DNA, single stranded DNA, isolated DNA such as partially purified DNA, essentially pure DNA, synthetic DNA, recombinantly produced DNA, as well as modified DNA that differs from naturally occurring DNA by the addition, deletion, substitution and/or alteration of one or more nucleotides. Such alterations may refer to addition of non-nucleotide material to internal DNA nucleotides or to the end(s) of DNA. It is also contemplated herein that nucleotides in DNA may be non-standard nucleotides, such as chemically synthesized nucleotides or ribonucleotides. For the present disclosure, these altered DNAs are considered analogs of naturally-occurring DNA. A molecule contains "a majority of deoxyribonucleotide residues" if the content of deoxyribonucleotide residues in the molecule is more than 50% (such as at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%), based on the total number of nucleotide residues in the molecule. The total number of nucleotide residues in a molecule is the sum of all nucleotide residues (irrespective of whether the nucleotide residues are standard (*i.e.*, naturally occurring) nucleotide residues or analogs thereof).

DNA may be recombinant DNA and may be obtained by cloning of a nucleic acid, in particular cDNA. The cDNA may be obtained by reverse transcription of RNA.

Nucleic acids may be comprised in a vector. The term "vector" as used herein includes any vectors known to the skilled person including plasmid vectors, cosmid vectors, phage vectors such as lambda phage, viral vectors such as retroviral, adenoviral or baculoviral vectors, or artificial chromosome vectors such as bacterial artificial chromosomes (BAC), yeast artificial chromosomes (YAC), or P1 artificial chromosomes (PAC). Said vectors include expression as well as cloning vectors. Expression vectors comprise plasmids as well as viral vectors and generally contain a desired coding sequence and appropriate DNA sequences necessary for the expression of the operably linked coding sequence in a particular host organism (*e.g.*,

bacteria, yeast, plant, insect, or mammal) or in *in vitro* expression systems. Cloning vectors are generally used to engineer and amplify a certain desired DNA fragment and may lack functional sequences needed for expression of the desired DNA fragments.

RNA

The term "RNA" relates to a nucleic acid molecule which includes ribonucleotide residues. In preferred embodiments, the RNA contains all or a majority of ribonucleotide residues. As used herein, "ribonucleotide" refers to a nucleotide with a hydroxyl group at the 2'-position of a β -D-ribofuranosyl group. RNA encompasses without limitation, double stranded RNA, single stranded RNA, isolated RNA such as partially purified RNA, essentially pure RNA, synthetic RNA, recombinantly produced RNA, as well as modified RNA that differs from naturally occurring RNA by the addition, deletion, substitution and/or alteration of one or more nucleotides. Such alterations may refer to addition of non-nucleotide material to internal RNA nucleotides or to the end(s) of RNA. It is also contemplated herein that nucleotides in RNA may be non-standard nucleotides, such as chemically synthesized nucleotides or deoxynucleotides. For the present disclosure, these altered/modified nucleotides can be referred to as analogs of naturally occurring nucleotides, and the corresponding RNAs containing such altered/modified nucleotides (*i.e.*, altered/modified RNAs) can be referred to as analogs of naturally occurring RNAs. A molecule contains "a majority of ribonucleotide residues" if the content of ribonucleotide residues in the molecule is more than 50% (such as at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%), based on the total number of nucleotide residues in the molecule. The total number of nucleotide residues in a molecule is the sum of all nucleotide residues (irrespective of whether the nucleotide residues are standard (*i.e.*, naturally occurring) nucleotide residues or analogs thereof).

"RNA" includes mRNA, tRNA, ribosomal RNA (rRNA), small nuclear RNA (snRNA), self-amplifying RNA (saRNA), trans-amplifying RNA (taRNA), single-stranded RNA (ssRNA), dsRNA, inhibitory RNA (such as antisense ssRNA, small interfering RNA (siRNA), or microRNA (miRNA)), activating RNA (such as small activating RNA) and immunostimulatory RNA (isRNA). In some embodiments, "RNA" refers to mRNA.

The term "*in vitro* transcription" or "IVT" as used herein means that the transcription (*i.e.*, the generation of RNA) is conducted in a cell-free manner. *I.e.*, IVT does not use living/cultured cells but rather the transcription machinery extracted from cells (*e.g.*, cell lysates or the isolated components thereof, including an RNA polymerase (preferably T7, T3 or SP6 polymerase)).

According to the present disclosure, the term "RNA" includes "mRNA". According to the present disclosure, the term "mRNA" means "messenger-RNA" and includes a "transcript" which may be generated by using a DNA template. Generally, mRNA encodes a peptide or polypeptide.

mRNA is single-stranded but may contain self-complementary sequences that allow parts of the mRNA to fold and pair with itself to form double helices.

According to the present disclosure, "dsRNA" means double-stranded RNA and is RNA with two partially or completely complementary strands.

In some embodiments, the mRNA which preferably encodes a peptide or polypeptide has a length of at least 45 nucleotides (such as at least 60, at least 90, at least 100, at least 200, at least 300, at least 400, at least 500, at least 600, at least 700, at least 800, at least 900, at least 1,000, at least 1,500, at least 2,000, at least 2,500, at least 3,000, at least 3,500, at least 4,000, at least 4,500, at least 5,000, at least 6,000, at least 7,000, at least 8,000, at least 9,000 nucleotides), preferably up to 15,000, such as up to 14,000, up to 13,000, up to 12,000 nucleotides, up to 11,000 nucleotides or up to 10,000 nucleotides.

As established in the art, mRNA generally contains a 5' untranslated region (5'-UTR), a peptide/polypeptide coding region and a 3' untranslated region (3'-UTR). In some embodiments, the mRNA is produced by *in vitro* transcription or chemical synthesis. In some embodiments, the mRNA is produced by *in vitro* transcription using a DNA template. The *in vitro* transcription methodology is known to the skilled person; cf., *e.g.*, Molecular Cloning: A Laboratory Manual, 4th Edition, M.R. Green and J. Sambrook eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor 2012. Furthermore, a variety of *in vitro* transcription kits is commercially available, *e.g.*, from Thermo Fisher Scientific (such as TranscriptAid™ T7 kit, MEGAscript® T7 kit, MAXIsript®), New England BioLabs Inc. (such as HiScribe™ T7 kit, HiScribe™ T7 ARCA mRNA kit), Promega (such as RiboMAX™, HeLaScribe®, Riboprobe®

systems), Jena Bioscience (such as SP6 or T7 transcription kits), and Epicentre (such as AmpliScribe™). For providing modified mRNA, correspondingly modified nucleotides, such as modified naturally occurring nucleotides, non-naturally occurring nucleotides and/or modified non-naturally occurring nucleotides, can be incorporated during synthesis (preferably *in vitro* transcription), or modifications can be effected in and/or added to the mRNA after transcription.

In some embodiments, RNA is *in vitro* transcribed RNA (IVT-RNA) and may be obtained by *in vitro* transcription of an appropriate DNA template. The promoter for controlling transcription can be any promoter for any RNA polymerase. Particular examples of RNA polymerases are the T7, T3, and SP6 RNA polymerases. Preferably, the *in vitro* transcription is controlled by a T7 or SP6 promoter. A DNA template for *in vitro* transcription may be obtained by cloning of a nucleic acid, in particular cDNA, and introducing it into an appropriate vector for *in vitro* transcription. The cDNA may be obtained by reverse transcription of RNA.

In some embodiments of the present disclosure, the RNA is "replicon RNA" or simply a "replicon", in particular "self-replicating RNA" or "self-amplifying RNA". In certain embodiments, the replicon or self-replicating RNA is derived from or comprises elements derived from an ssRNA virus, in particular a positive-stranded ssRNA virus such as an alphavirus. Alphaviruses are typical representatives of positive-stranded RNA viruses. Alphaviruses replicate in the cytoplasm of infected cells (for review of the alphaviral life cycle see José *et al.*, Future Microbiol., 2009, vol. 4, pp. 837–856). The total genome length of many alphaviruses typically ranges between 11,000 and 12,000 nucleotides, and the genomic RNA typically has a 5'-cap, and a 3' poly(A) tail. The genome of alphaviruses encodes non-structural proteins (involved in transcription, modification and replication of viral RNA and in protein modification) and structural proteins (forming the virus particle). There are typically two open reading frames (ORFs) in the genome. The four non-structural proteins (nsP1–nsP4) are typically encoded together by a first ORF beginning near the 5' terminus of the genome, while alphavirus structural proteins are encoded together by a second ORF which is found downstream of the first ORF and extends near the 3' terminus of the genome. Typically, the first ORF is larger than the second ORF, the ratio being roughly 2:1. In cells infected by an alphavirus, only the nucleic acid sequence encoding non-structural proteins is translated from

the genomic RNA, while the genetic information encoding structural proteins is translatable from a subgenomic transcript, which is an RNA molecule that resembles eukaryotic messenger RNA (mRNA; Gould *et al.*, 2010, Antiviral Res., vol. 87 pp. 111–124). Following infection, *i.e.* at early stages of the viral life cycle, the (+) stranded genomic RNA directly acts like a messenger RNA for the translation of the open reading frame encoding the non-structural poly-protein (nsP1234).

Alphavirus-derived vectors have been proposed for delivery of foreign genetic information into target cells or target organisms. In simple approaches, the open reading frame encoding alphaviral structural proteins is replaced by an open reading frame encoding a protein of interest. Alphavirus-based trans-replication (trans-amplification) systems rely on alphavirus nucleotide sequence elements on two separate nucleic acid molecules: one nucleic acid molecule encodes a viral replicase, and the other nucleic acid molecule is capable of being replicated by said replicase in trans (hence the designation trans-replication system). Trans-replication requires the presence of both these nucleic acid molecules in a given host cell. The nucleic acid molecule capable of being replicated by the replicase in trans must comprise certain alphaviral sequence elements to allow recognition and RNA synthesis by the alphaviral replicase.

In some embodiments of the present disclosure, the RNA (in particular, mRNA) described herein contains one or more modifications, *e.g.*, in order to increase its stability and/or increase translation efficiency and/or decrease immunogenicity and/or decrease cytotoxicity. For example, in order to increase expression of the RNA (in particular, mRNA), it may be modified within the coding region, *i.e.*, the sequence encoding the expressed peptide or polypeptide, preferably without altering the sequence of the expressed peptide or polypeptide. Such modifications are described, for example, in WO 2007/036366 and PCT/EP2019/056502, and include the following: a 5'-cap structure; an extension or truncation of the naturally occurring poly(A) tail; an alteration of the 5'- and/or 3'-untranslated regions (UTR) such as introduction of a UTR which is not related to the coding region of said RNA; the replacement of one or more naturally occurring nucleotides with synthetic nucleotides; and codon optimization (*e.g.*, to alter, preferably increase, the GC content of the RNA).

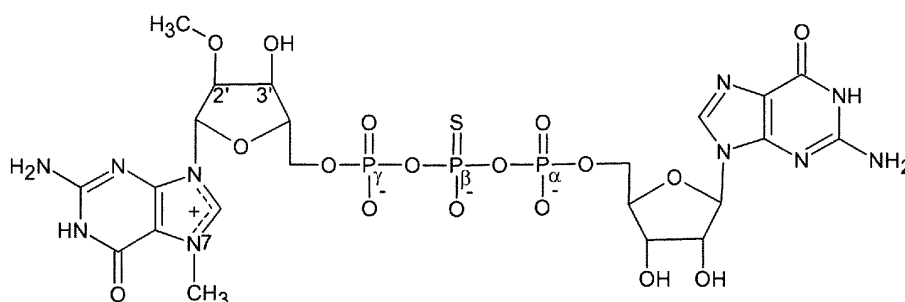
In some embodiments, the RNA (in particular, mRNA) described herein comprises a 5'-cap structure. In some embodiments, the RNA does not have uncapped 5'-triphosphates. In some embodiments, the RNA (in particular, mRNA) may comprise a conventional 5'-cap and/or a 5'-cap analog. The term "conventional 5'-cap" refers to a cap structure found on the 5'-end of an RNA molecule and generally comprises a guanosine 5'-triphosphate (Gppp) which is connected via its triphosphate moiety to the 5'-end of the next nucleotide of the RNA (*i.e.*, the guanosine is connected via a 5' to 5' triphosphate linkage to the rest of the RNA). The guanosine may be methylated at position N⁷ (resulting in the cap structure m⁷Gppp). The term "5'-cap analog" includes a 5'-cap which is based on a conventional 5'-cap but which has been modified at either the 2'- or 3'-position of the m⁷guanosine structure in order to avoid an integration of the 5'-cap analog in the reverse orientation (such 5'-cap analogs are also called anti-reverse cap analogs (ARCA)). Particularly preferred 5'-cap analogs are those having one or more substitutions at the bridging and non-bridging oxygen in the phosphate bridge, such as phosphorothioate modified 5'-cap analogs at the β -phosphate (such as m₂^{7,2'-O}G(5')ppSp(5')G (referred to as beta-S-ARCA or β -S-ARCA)), as described in PCT/EP2019/056502. Providing an RNA (in particular, mRNA) with a 5'-cap structure as described herein may be achieved by *in vitro* transcription of a DNA template in presence of a corresponding 5'-cap compound, wherein said 5'-cap structure is co-transcriptionally incorporated into the generated RNA (in particular, mRNA) strand, or the RNA (in particular, mRNA) may be generated, for example, by *in vitro* transcription, and the 5'-cap structure may be attached to the RNA post-transcriptionally using capping enzymes, for example, capping enzymes of vaccinia virus.

In some embodiments, the RNA (in particular, mRNA) comprises a 5'-cap structure selected from the group consisting of m₂^{7,2'-O}G(5')ppSp(5')G (in particular its D1 diastereomer), m₂^{7,3'-O}G(5')ppp(5')G, and m₂^{7,3'-O}Gppp(m₁^{2'-O})ApG.

In some embodiments, the RNA (in particular, mRNA) comprises a cap0, cap1, or cap2, preferably cap1 or cap2. According to the present disclosure, the term "cap0" means the structure "m⁷GpppN", wherein N is any nucleoside bearing an OH moiety at position 2'. According to the present disclosure, the term "cap1" means the structure "m⁷GpppNm", wherein Nm is any nucleoside bearing an OCH₃ moiety at position 2'. According to the present

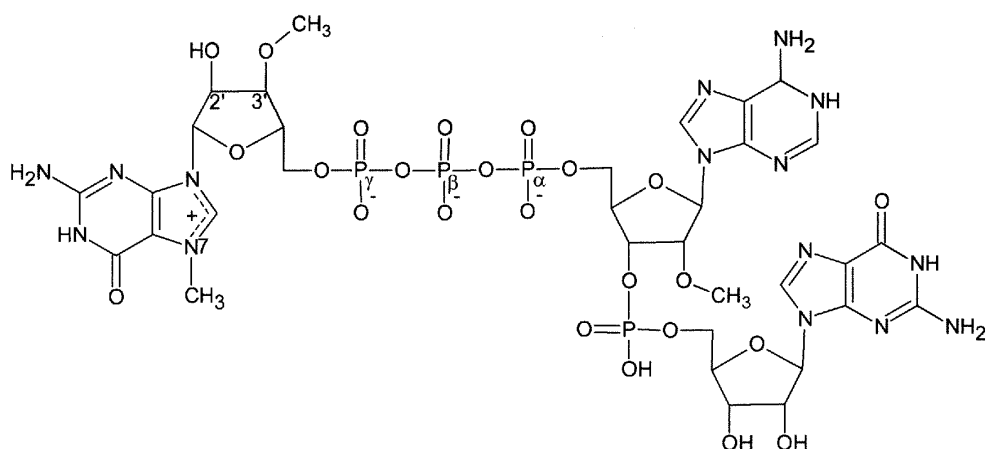
disclosure, the term "cap2" means the structure "m⁷GpppNmNm", wherein each Nm is independently any nucleoside bearing an OCH₃ moiety at position 2'.

The 5'-cap analog beta-S-ARCA (β -S-ARCA) has the following structure:

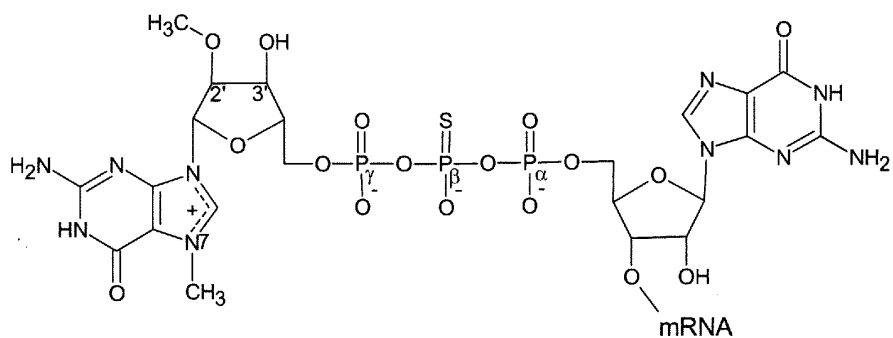


The "D1 diastereomer of beta-S-ARCA" or "beta-S-ARCA(D1)" is the diastereomer of beta-S-ARCA which elutes first on an HPLC column compared to the D2 diastereomer of beta-S-ARCA (beta-S-ARCA(D2)) and thus exhibits a shorter retention time. The HPLC preferably is an analytical HPLC. In some embodiments, a Supelcosil LC-18-T RP column, preferably of the format: 5 μ m, 4.6 x 250 mm is used for separation, whereby a flow rate of 1.3 ml/min can be applied. In some embodiments, a gradient of methanol in ammonium acetate, for example, a 0-25% linear gradient of methanol in 0.05 M ammonium acetate, pH = 5.9, within 15 min is used. UV-detection (VWD) can be performed at 260 nm and fluorescence detection (FLD) can be performed with excitation at 280 nm and detection at 337 nm.

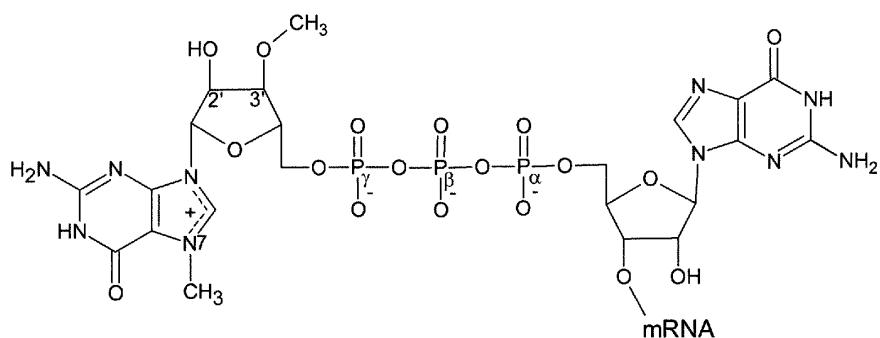
The 5'-cap analog m₂^{7,3'-O}Gppp(m₁^{2'-O})ApG (also referred to as m₂^{7,3'-O}G(5')ppp(5')m₁^{2'-O}ApG) which is a building block of a cap1 has the following structure:



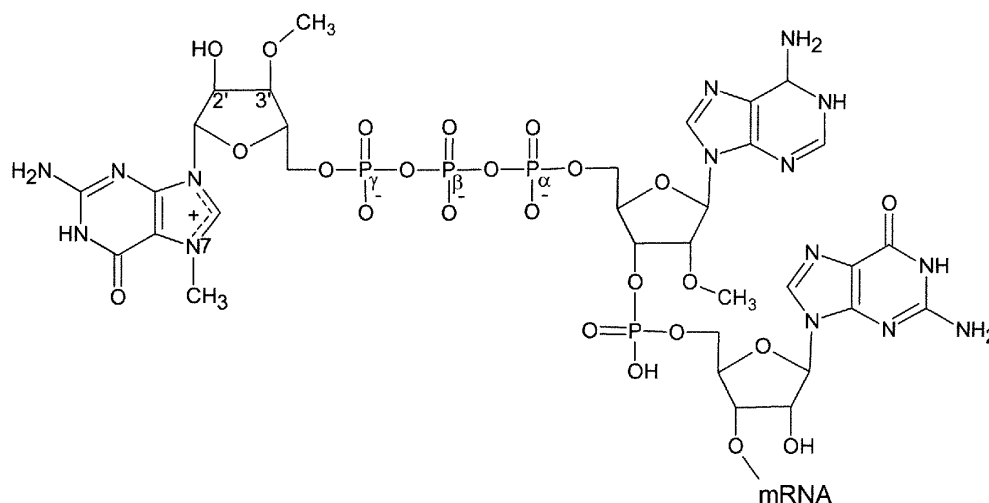
An exemplary cap0 mRNA comprising β -S-ARCA and mRNA has the following structure:



An exemplary cap0 mRNA comprising $m_2^{7,3^{10}}G(5')ppp(5')G$ and mRNA has the following structure:



An exemplary cap1 mRNA comprising $m_2^{7,3'-O}Gppp(m_1^{2'-O})ApG$ and mRNA has the following structure:



As used herein, the term "poly-A tail" or "poly-A sequence" refers to an uninterrupted or interrupted sequence of adenylate residues which is typically located at the 3'-end of an RNA (in particular, mRNA) molecule. Poly-A tails or poly-A sequences are known to those of skill in the art and may follow the 3'-UTR in the RNAs (in particular, mRNAs) described herein. An uninterrupted poly-A tail is characterized by consecutive adenylate residues. In nature, an uninterrupted poly-A tail is typical. RNAs (in particular, mRNAs) disclosed herein can have a poly-A tail attached to the free 3'-end of the RNA by a template-independent RNA polymerase after transcription or a poly-A tail encoded by DNA and transcribed by a template-dependent RNA polymerase.

It has been demonstrated that a poly-A tail of about 120 A nucleotides has a beneficial influence on the levels of RNA in transfected eukaryotic cells, as well as on the levels of protein that is translated from an open reading frame that is present upstream (5') of the poly-A tail (Holtkamp *et al.*, 2006, Blood, vol. 108, pp. 4009-4017).

The poly-A tail may be of any length. In some embodiments, a poly-A tail comprises, essentially consists of, or consists of at least 20, at least 30, at least 40, at least 80, or at least 100 and up to 500, up to 400, up to 300, up to 200, or up to 150 A nucleotides, and, in particular, about 120 A nucleotides. In this context, "essentially consists of" means that most nucleotides in the poly-A tail, typically at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% by number of nucleotides in the poly-A tail are A nucleotides, but permits that remaining nucleotides are nucleotides other than A

nucleotides, such as U nucleotides (uridylate), G nucleotides (guanylate), or C nucleotides (cytidylate). In this context, "consists of" means that all nucleotides in the poly-A tail, *i.e.*, 100% by number of nucleotides in the poly-A tail, are A nucleotides. The term "A nucleotide" or "A" refers to adenylate.

In some embodiments, a poly-A tail is attached during RNA transcription, *e.g.*, during preparation of *in vitro* transcribed RNA, based on a DNA template comprising repeated dT nucleotides (deoxythymidylate) in the strand complementary to the coding strand. The DNA sequence encoding a poly-A tail (coding strand) is referred to as poly(A) cassette.

In some embodiments, the poly(A) cassette present in the coding strand of DNA essentially consists of dA nucleotides, but is interrupted by a random sequence of the four nucleotides (dA, dC, dG, and dT). Such random sequence may be 5 to 50, 10 to 30, or 10 to 20 nucleotides in length. Such a cassette is disclosed in WO 2016/005324 A1, hereby incorporated by reference. Any poly(A) cassette disclosed in WO 2016/005324 A1 may be used in the present disclosure. A poly(A) cassette that essentially consists of dA nucleotides, but is interrupted by a random sequence having an equal distribution of the four nucleotides (dA, dC, dG, dT) and having a length of *e.g.*, 5 to 50 nucleotides shows, on DNA level, constant propagation of plasmid DNA in *E. coli* and is still associated, on RNA level, with the beneficial properties with respect to supporting RNA stability and translational efficiency is encompassed. Consequently, in some embodiments, the poly-A tail contained in an RNA (in particular, mRNA) molecule described herein essentially consists of A nucleotides, but is interrupted by a random sequence of the four nucleotides (A, C, G, U). Such random sequence may be 5 to 50, 10 to 30, or 10 to 20 nucleotides in length.

In some embodiments, the poly(A) tail comprises 30 adenine nucleotides followed by 70 adenine nucleotides, wherein the 30 adenine nucleotides and 70 adenine nucleotides are separated by a linker sequence of 10 nucleotides.

In some embodiments, no nucleotides other than A nucleotides flank a poly-A tail at its 3'-end, *i.e.*, the poly-A tail is not masked or followed at its 3'-end by a nucleotide other than A.

In some embodiments, a poly-A tail may comprise at least 20, at least 30, at least 40, at least 80, or at least 100 and up to 500, up to 400, up to 300, up to 200, or up to 150 nucleotides. In some embodiments, the poly-A tail may essentially consist of at least 20, at least 30, at least

40, at least 80, or at least 100 and up to 500, up to 400, up to 300, up to 200, or up to 150 nucleotides. In some embodiments, the poly-A tail may consist of at least 20, at least 30, at least 40, at least 80, or at least 100 and up to 500, up to 400, up to 300, up to 200, or up to 150 nucleotides. In some embodiments, the poly-A tail comprises at least 100 nucleotides. In some embodiments, the poly-A tail comprises about 150 nucleotides. In some embodiments, the poly-A tail comprises about 120 nucleotides.

In some embodiments, RNA (in particular, mRNA) described in present disclosure comprises a 5'-UTR and/or a 3'-UTR. The term "untranslated region" or "UTR" relates to a region in a DNA molecule which is transcribed but is not translated into an amino acid sequence, or to the corresponding region in an RNA molecule, such as an mRNA molecule. An untranslated region (UTR) can be present 5' (upstream) of an open reading frame (5'-UTR) and/or 3' (downstream) of an open reading frame (3'-UTR). A 5'-UTR, if present, is located at the 5'-end, upstream of the start codon of a protein-encoding region. A 5'-UTR is downstream of the 5'-cap (if present), *e.g.*, directly adjacent to the 5'-cap. A 3'-UTR, if present, is located at the 3'-end, downstream of the termination codon of a protein-encoding region, but the term "3'-UTR" does generally not include the poly-A sequence. Thus, the 3'-UTR is upstream of the poly-A sequence (if present), *e.g.*, directly adjacent to the poly-A sequence. Incorporation of a 3'-UTR into the 3'-non translated region of an RNA (preferably mRNA) molecule can result in an enhancement in translation efficiency. A synergistic effect may be achieved by incorporating two or more of such 3'-UTRs (which are preferably arranged in a head-to-tail orientation; cf., *e.g.*, Holtkamp *et al.*, Blood 108, 4009-4017 (2006)). The 3'-UTRs may be autologous or heterologous to the RNA (*e.g.*, mRNA) into which they are introduced. In certain embodiments, the 3'-UTR is derived from a globin gene or mRNA, such as a gene or mRNA of alpha2-globin, alpha1-globin, or beta-globin, *e.g.*, beta-globin, *e.g.*, human beta-globin. For example, the RNA (*e.g.*, mRNA) may be modified by the replacement of the existing 3'-UTR with or the insertion of one or more, *e.g.*, two copies of a 3'-UTR derived from a globin gene, such as alpha2-globin, alpha1-globin, beta-globin, *e.g.*, beta-globin, *e.g.*, human beta-globin.

In some embodiments, a 5'-UTR is or comprises a modified human alpha-globin 5'-UTR. In some embodiments, a 3'-UTR comprises a first sequence from the amino terminal enhancer

of split (AES) messenger RNA and a second sequence from the mitochondrial encoded 12S ribosomal RNA.

The RNA (in particular, mRNA) described herein may have modified ribonucleotides in order to increase its stability and/or decrease immunogenicity and/or decrease cytotoxicity. For example, in some embodiments, uridine in the RNA (in particular, mRNA) described herein is replaced (partially or completely, preferably completely) by a modified nucleoside. In some embodiments, the modified nucleoside is a modified uridine.

In some embodiments, the modified uridine replacing uridine is selected from the group consisting of pseudouridine (ψ), N1-methyl-pseudouridine ($m1\psi$), 5-methyl-uridine ($m5U$), and combinations thereof.

In some embodiments, the modified nucleoside replacing (partially or completely, preferably completely) uridine in the RNA may be any one or more of 3-methyl-uridine ($m3U$), 5-methoxy-uridine ($mo5U$), 5-aza-uridine, 6-aza-uridine, 2-thio-5-aza-uridine, 2-thio-uridine ($s2U$), 4-thio-uridine ($s4U$), 4-thio-pseudouridine, 2-thio-pseudouridine, 5-hydroxy-uridine ($ho5U$), 5-aminoallyl-uridine, 5-halo-uridine (e.g., 5-iodo-uridine or 5-bromo-uridine), uridine 5-oxyacetic acid ($cmo5U$), uridine 5-oxyacetic acid methyl ester ($mcmo5U$), 5-carboxymethyl-uridine ($cm5U$), 1-carboxymethyl-pseudouridine, 5-carboxyhydroxymethyl-uridine ($chm5U$), 5-carboxyhydroxymethyl-uridine methyl ester ($mchm5U$), 5-methoxycarbonylmethyl-uridine ($mcm5U$), 5-methoxycarbonylmethyl-2-thio-uridine ($mcm5s2U$), 5-aminomethyl-2-thio-uridine ($nm5s2U$), 5-methylaminomethyl-uridine ($mnm5U$), 1-ethyl-pseudouridine, 5-methylaminomethyl-2-thio-uridine ($mnm5s2U$), 5-methylaminomethyl-2-seleno-uridine ($mnm5se2U$), 5-carbamoylmethyl-uridine ($ncm5U$), 5-carboxymethylaminomethyl-uridine ($cmnm5U$), 5-carboxymethylaminomethyl-2-thio-uridine ($cmnm5s2U$), 5-propynyl-uridine, 1-propynyl-pseudouridine, 5-aurinomethyl-uridine ($tm5U$), 1-aurinomethyl-pseudouridine, 5-aurinomethyl-2-thio-uridine ($tm5s2U$), 1-aurinomethyl-4-thio-pseudouridine, 5-methyl-2-thio-uridine ($m5s2U$), 1-methyl-4-thio-pseudouridine ($m1s4\psi$), 4-thio-1-methyl-pseudouridine, 3-methyl-pseudouridine ($m3\psi$), 2-thio-1-methyl-pseudouridine, 1-methyl-1-deaza-pseudouridine, 2-thio-1-methyl-1-deaza-pseudouridine, dihydrouridine (D), dihydropseudouridine, 5,6-dihydrouridine, 5-methyl-dihydrouridine ($m5D$), 2-thio-dihydrouridine, 2-thio-dihydropseudouridine, 2-methoxy-uridine, 2-methoxy-4-thio-uridine,

4-methoxy-pseudouridine, 4-methoxy-2-thio-pseudouridine, N1-methyl-pseudouridine, 3-(3-amino-3-carboxypropyl)uridine (acp3U), 1-methyl-3-(3-amino-3-carboxypropyl)pseudouridine (acp3 ψ), 5-(isopentenylaminomethyl)uridine (inm5U), 5-(isopentenylaminomethyl)-2-thio-uridine (inm5s2U), α -thio-uridine, 2'-O-methyl-uridine (Um), 5,2'-O-dimethyl-uridine (m5Um), 2'-O-methyl-pseudouridine (ψ m), 2-thio-2'-O-methyl-uridine (s2Um), 5-methoxycarbonylmethyl-2'-O-methyl-uridine (mcm5Um), 5-carbamoylmethyl-2'-O-methyl-uridine (ncm5Um), 5-carboxymethylaminomethyl-2'-O-methyl-uridine (cmnm5Um), 3,2'-O-dimethyl-uridine (m3Um), 5-(isopentenylaminomethyl)-2'-O-methyl-uridine (inm5Um), 1-thio-uridine, deoxythymidine, 2'-F-ara-uridine, 2'-F-uridine, 2'-OH-ara-uridine, 5-(2-carbomethoxyvinyl) uridine, 5-[3-(1-E-propenylamino)uridine, or any other modified uridine known in the art.

An RNA (preferably mRNA) which is modified by pseudouridine (replacing partially or completely, preferably completely, uridine) is referred to herein as " Ψ -modified", whereas the term "m1 Ψ -modified" means that the RNA (preferably mRNA) contains N(1)-methylpseudouridine (replacing partially or completely, preferably completely, uridine). Furthermore, the term "m5U-modified" means that the RNA (preferably mRNA) contains 5-methyluridine (replacing partially or completely, preferably completely, uridine). Such Ψ - or m1 Ψ - or m5U-modified RNAs usually exhibit decreased immunogenicity compared to their unmodified forms and, thus, are preferred in applications where the induction of an immune response is to be avoided or minimized. In some embodiments, the RNA (preferably mRNA) contains N(1)-methylpseudouridine replacing completely uridine.

The codons of the RNA (in particular, mRNA) described in the present disclosure may further be optimized, *e.g.*, to increase the GC content of the RNA and/or to replace codons which are rare in the cell (or subject) in which the peptide or polypeptide of interest is to be expressed by codons which are synonymous frequent codons in said cell (or subject). In some embodiments, the amino acid sequence encoded by the RNA (in particular, mRNA) described in the present disclosure is encoded by a coding sequence which is codon-optimized and/or the G/C content of which is increased compared to wild type coding sequence. This also includes embodiments, wherein one or more sequence regions of the coding sequence are codon-optimized and/or increased in the G/C content compared to the corresponding

sequence regions of the wild type coding sequence. In some embodiments, the codon-optimization and/or the increase in the G/C content preferably does not change the sequence of the encoded amino acid sequence.

The term "codon-optimized" refers to the alteration of codons in the coding region of a nucleic acid molecule to reflect the typical codon usage of a host organism without preferably altering the amino acid sequence encoded by the nucleic acid molecule. Within the context of the present disclosure, coding regions may be codon-optimized for optimal expression in a subject to be treated using the RNA (in particular, mRNA) described herein. Codon-optimization is based on the finding that the translation efficiency is also determined by a different frequency in the occurrence of tRNAs in cells. Thus, the sequence of RNA (in particular, mRNA) may be modified such that codons for which frequently occurring tRNAs are available are inserted in place of "rare codons".

In some embodiments, the guanosine/cytosine (G/C) content of the coding region of the RNA (in particular, mRNA) described herein is increased compared to the G/C content of the corresponding coding sequence of the wild type RNA, wherein the amino acid sequence encoded by the RNA is preferably not modified compared to the amino acid sequence encoded by the wild type RNA. This modification of the RNA sequence is based on the fact that the sequence of any RNA region to be translated is important for efficient translation of that RNA. Sequences having an increased G (guanosine)/C (cytosine) content are more stable than sequences having an increased A (adenosine)/U (uracil) content. In respect to the fact that several codons code for one and the same amino acid (so-called degeneration of the genetic code), the most favorable codons for the stability can be determined (so-called alternative codon usage). Depending on the amino acid to be encoded by the RNA, there are various possibilities for modification of the RNA sequence, compared to its wild type sequence. In particular, codons which contain A and/or U nucleotides can be modified by substituting these codons by other codons, which code for the same amino acids but contain no A and/or U or contain a lower content of A and/or U nucleotides.

In various embodiments, the G/C content of the coding region of the RNA (in particular, mRNA) described herein is increased by at least 10%, at least 20%, at least 30%, at least 40%,

at least 50%, at least 55%, or even more compared to the G/C content of the coding region of the wild type RNA.

A combination of the above described modifications, *i.e.*, incorporation of a 5'-cap structure, incorporation of a poly-A sequence, unmasking of a poly-A sequence, alteration of the 5'- and/or 3'-UTR (such as incorporation of one or more 3'-UTRs), replacing one or more naturally occurring nucleotides with synthetic nucleotides (*e.g.*, 5-methylcytidine for cytidine and/or pseudouridine (Ψ) or N(1)-methylpseudouridine (m1 Ψ) or 5-methyluridine (m5U) for uridine), and codon optimization, has a synergistic influence on the stability of RNA (preferably mRNA) and increase in translation efficiency. Thus, in some embodiments, the RNA (in particular, mRNA) described in the present disclosure contains a combination of at least two, at least three, at least four or all five of the above-mentioned modifications, *i.e.*, (i) incorporation of a 5'-cap structure, (ii) incorporation of a poly-A sequence, unmasking of a poly-A sequence; (iii) alteration of the 5'- and/or 3'-UTR (such as incorporation of one or more 3'-UTRs); (iv) replacing one or more naturally occurring nucleotides with synthetic nucleotides (*e.g.*, 5-methylcytidine for cytidine and/or pseudouridine (Ψ) or N(1)-methylpseudouridine (m1 Ψ) or 5-methyluridine (m5U) for uridine), and (v) codon optimization.

Nucleic acid particles

A nucleic acid such as RNA may be administered with one or more delivery vehicles that protect the nucleic acid from degradation, maximize delivery to on-target cells and minimize exposure to off-target cells. Such delivery vehicles may complex or encapsulate the nucleic acid and include a range of materials, including polymers, lipids and mixtures thereof. In some embodiments, such delivery vehicles may form particles with the nucleic acid.

In the context of the present disclosure, the term "particle" relates to a structured entity formed by molecules or molecule complexes, in particular particle forming compounds. In some embodiments, a particle is a nucleic acid containing particle such as a particle comprising DNA, RNA or a mixture thereof. In some embodiments, the particle contains an envelope (*e.g.*, one or more layers or lamellas) made of one or more types of amphiphilic substances (*e.g.*, amphiphilic lipids). In this context, the expression "amphiphilic substance" means that the substance possesses both hydrophilic and lipophilic properties. The envelope may also

comprise additional substances (*e.g.*, additional lipids) which do not have to be amphiphilic. Thus, the particle may be a monolamellar or multilamellar structure, wherein the substances constituting the one or more layers or lamellas comprise one or more types of amphiphilic substances (in particular selected from the group consisting of amphiphilic lipids) optionally in combination with additional substances (*e.g.*, additional lipids) which do not have to be amphiphilic. In some embodiments, the term "particle" relates to a micro- or nano-sized structure, such as a micro- or nano-sized compact structure. According to the present disclosure, the term "particle" includes nanoparticles.

The term "nanoparticle" relates to a nano-sized particle comprising at least one particle forming agent, *e.g.*, at least one cationic or cationically ionizable lipid, wherein all three external dimensions of the particle are in the nanoscale, *i.e.*, at least about 1 nm and below about 1000 nm. Preferably, the size of a particle is its diameter.

In some embodiments, the particles described herein have a size (such as a diameter) in the range of about 10 to about 2000 nm, such as in the range of about 20 to about 1500 nm, such as about 30 to about 1200 nm, about 40 to about 1100 nm, about 50 to about 1000 nm, about 60 to about 900 nm, about 70 to about 800 nm, about 80 to about 700 nm, about 90 to about 600 nm, or about 50 to about 500 nm or about 100 to about 500 nm, such as in the range of 10 to 1000 nm, 15 to 500 nm, 20 to 450 nm, 25 to 400 nm, 30 to 350 nm, 40 to 300 nm, 50 to 250 nm, 60 to 200 nm, 70 to 150 nm, or 80 to 150 nm. In some embodiments, the particles described herein have a size (such as a diameter) in the range of from about 40 nm to about 200 nm, such as from about 50 nm to about 180 nm, from about 60 nm to about 160 nm, from about 80 nm to about 150 nm or from about 80 nm to about 120 nm.

Particles described herein may exhibit a polydispersity index (PDI) less than about 0.5, less than about 0.4, less than about 0.3, less than about 0.2, less than about 0.1, or less than about 0.05. By way of example, the particles can exhibit a polydispersity index in a range of about 0.01 to about 0.4 or about 0.1 to about 0.3.

A "nucleic acid particle" can be used to deliver nucleic acid to a target site of interest (*e.g.*, cell, tissue, organ, and the like). A nucleic acid particle may be formed from at least one cationic or cationically ionizable compound such as a polymer or lipid complexing the nucleic acid. Without intending to be bound by any theory, it is believed that the cationic or

cationically ionizable compound combines together with the nucleic acid to form aggregates, and this aggregation results in colloidally stable particles.

In some embodiments, nucleic acid may be noncovalently associated with a particle. In some embodiments, the nucleic acid may be adhered to the outer surface of the particle (surface nucleic acid) and/or may be contained in the particle (encapsulated nucleic acid).

The N/P ratio gives the ratio of the nitrogen groups in the lipid to the number of phosphate groups in the nucleic acid. It is correlated to the charge ratio, as the nitrogen atoms (depending on the pH) are usually positively charged and the phosphate groups are negatively charged. The N/P ratio, where a charge equilibrium exists, depends on the pH. Lipid formulations are frequently formed at N/P ratios larger than four up to twelve, because positively charged nanoparticles are considered favorable for transfection. In that case, nucleic acid is considered to be completely bound to nanoparticles.

The term "particle forming components" or "particle forming agents" relates to any components which form particles, e.g., by associating with nucleic acid. Nucleic acid delivery vehicles such as particle forming agents useful herein include polymers, polymer derivatives, lipids, and mixtures thereof. Such components include any component which can be part of nucleic acid particles, e.g., cationic or cationically ionizable lipids.

A "polymer," as used herein, is given its ordinary meaning, i.e., a molecular structure comprising one or more repeat units (monomers), connected by covalent bonds. The repeat units can all be identical, or in some cases, there can be more than one type of repeat unit present within the polymer. In some cases, the polymer is biologically derived, i.e., a biopolymer such as a protein. In some cases, additional moieties can also be present in the polymer, for example targeting moieties.

If more than one type of repeat unit is present within the polymer, then the polymer is said to be a "copolymer." It is to be understood that the polymer being employed herein can be a copolymer. The repeat units forming the copolymer can be arranged in any fashion. For example, the repeat units can be arranged in a random order, in an alternating order, or as a "block" copolymer, i.e., comprising one or more regions each comprising a first repeat unit (e.g., a first block), and one or more regions each comprising a second repeat unit (e.g., a

second block), etc. Block copolymers can have two (a diblock copolymer), three (a triblock copolymer), or more numbers of distinct blocks.

In certain embodiments, the polymer is biocompatible. Biocompatible polymers are polymers that typically do not result in significant cell death at moderate concentrations. In certain embodiments, the biocompatible polymer is biodegradable, i.e., the polymer is able to degrade, chemically and/or biologically, within a physiological environment, such as within the body.

In certain embodiments, polymer may be protamine or polyalkyleneimine.

The term "protamine" refers to any of various strongly basic proteins of relatively low molecular weight that are rich in arginine and are found associated especially with DNA in place of somatic histones in the sperm cells of various animals (as fish). In particular, the term "protamine" refers to proteins found in fish sperm that are strongly basic, are soluble in water, are not coagulated by heat, and yield chiefly arginine upon hydrolysis. In purified form, they are used in a long-acting formulation of insulin and to neutralize the anticoagulant effects of heparin.

According to the disclosure, the term "protamine" as used herein is meant to comprise any protamine amino acid sequence obtained or derived from natural or biological sources including fragments thereof and multimeric forms of said amino acid sequence or fragment thereof as well as (synthesized) polypeptides which are artificial and specifically designed for specific purposes and cannot be isolated from native or biological sources.

In some embodiments, the polyalkyleneimine comprises polyethylenimine and/or polypropylenimine, preferably polyethylenimine. A preferred polyalkyleneimine is polyethylenimine (PEI). The average molecular weight of PEI is preferably $0.75 \cdot 10^2$ to 10^7 Da, preferably 1000 to 10^5 Da, more preferably 10000 to 40000 Da, more preferably 15000 to 30000 Da, even more preferably 20000 to 25000 Da.

Preferred according to the disclosure is linear polyalkyleneimine such as linear polyethylenimine (PEI).

Cationic polymers (including polycationic polymers) contemplated for use herein include any cationic polymers which are able to electrostatically bind nucleic acid. In some embodiments, cationic polymers contemplated for use herein include any cationic polymers with which

nucleic acid can be associated, e.g. by forming complexes with the nucleic acid or forming vesicles in which the nucleic acid is enclosed or encapsulated.

Particles described herein may also comprise polymers other than cationic polymers, i.e., non-cationic polymers and/or anionic polymers. Collectively, anionic and neutral polymers are referred to herein as non-cationic polymers.

In some embodiments of the present disclosure, the nucleic acid such as RNA described herein may be present in lipoplex particles.

Lipoplexes (LPX) are electrostatic complexes which are generally formed by mixing preformed cationic lipid liposomes with anionic nucleic acid. Formed lipoplexes possess distinct internal arrangements of molecules that arise due to the transformation from liposomal structure into compact nucleic acid lipoplexes.

Generally, lipoplex particles are obtainable by adding nucleic acid to a colloidal liposome dispersion. Using the ethanol injection technique, such colloidal liposome dispersion is, in some embodiments, formed as follows: an ethanol solution comprising lipids, such as cationic or cationically ionizable lipids and additional lipids, is injected into an aqueous solution under stirring.

In some embodiments, liposomes are self-closed unilamellar or multilamellar vesicular particles wherein the lamellae comprise lipid bilayers and the encapsulated lumen comprises an aqueous phase. A prerequisite for using liposomes for nanoparticle formation is that the lipids in the mixture as required are able to form lamellar (bilayer) phases in the applied aqueous environment.

In certain embodiments, the nucleic acid lipoplex particles include both a cationic lipid and an additional lipid. In an exemplary embodiment, the cationic lipid is DOTMA and the additional lipid is DOPE.

In some embodiments, nucleic acid described herein is present in the form of lipid nanoparticles (LNPs).

In general, lipid nanoparticles are obtainable from direct mixing of nucleic acid, e.g., RNA, in an aqueous phase with lipids in a phase comprising an organic solvent, such as ethanol. In that case, lipids or lipid mixtures can be used for particle formation, which do not form lamellar (bilayer) phases in water.

LNPs typically comprise four components: cationically ionizable lipid, neutral lipids such as phospholipids, a steroid such as cholesterol, and a polymer-conjugated lipid such as PEG-lipid. In some embodiments, the LNP comprises from 40 to 60 mol percent, 40 to 55 mol percent, from 45 to 55 mol percent, or from 45 to 50 mol percent of the cationically ionizable lipid.

In some embodiments, the neutral lipid is present in a concentration ranging from 5 to 15 mol percent, from 7 to 13 mol percent, or from 9 to 11 mol percent.

In some embodiments, the steroid is present in a concentration ranging from 30 to 50 mol percent, from 30 to 45 mol percent, from 35 to 45 mol percent or from 35 to 43 mol percent.

In some embodiments, the LNP comprises from 1 to 10 mol percent, from 1 to 5 mol percent, or from 1 to 2.5 mol percent of the polymer-conjugated lipid.

In some embodiments, the LNP comprises from 45 to 55 mol percent of a cationically ionizable lipid; from 5 to 15 mol percent of a neutral lipid; from 30 to 45 mol percent of a steroid; from 1 to 5 mol percent of a polymer-conjugated lipid; and the nucleic acid, encapsulated within or associated with the lipid nanoparticle.

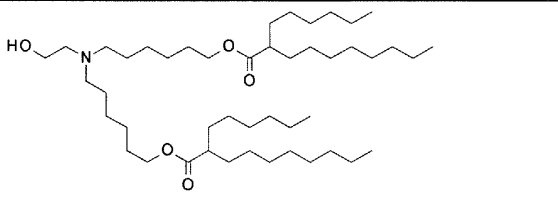
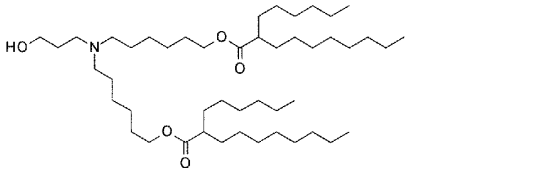
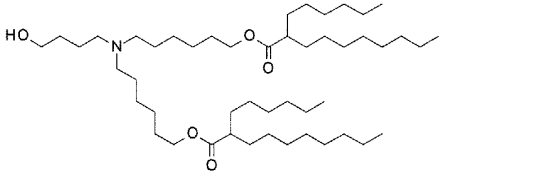
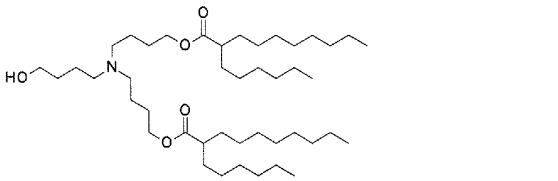
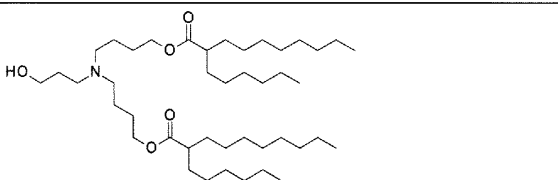
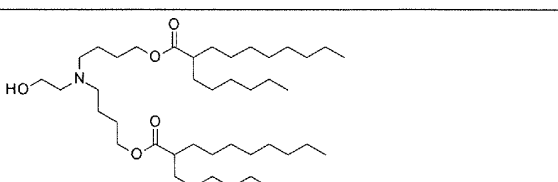
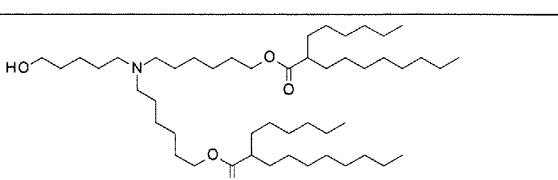
In some embodiments, the mol percent is determined based on total mol of lipid present in the lipid nanoparticle. In some embodiments, the mol percent is determined based on total mol of cationically ionizable lipid, neutral lipid, steroid and polymer-conjugated lipid present in the lipid nanoparticle.

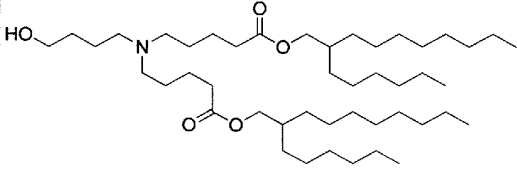
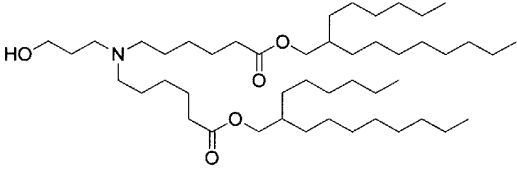
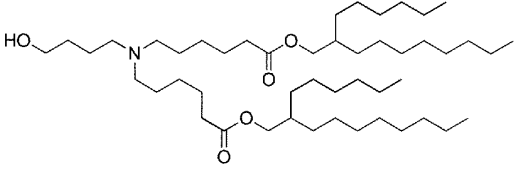
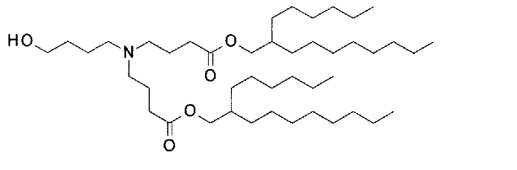
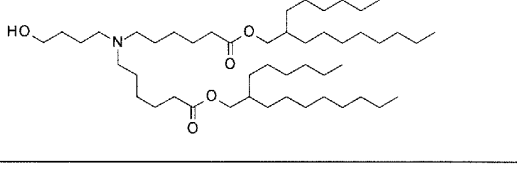
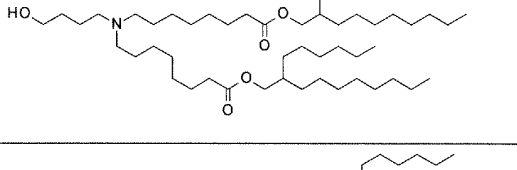
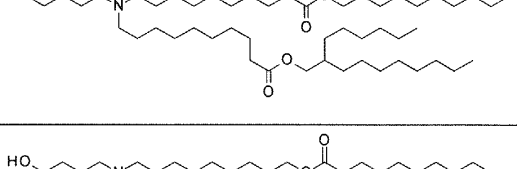
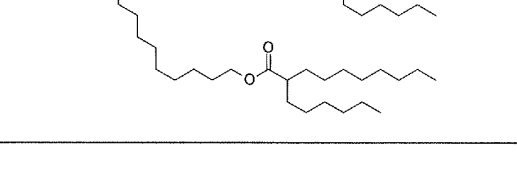
In some embodiments, the neutral lipid is selected from the group consisting of DSPC, DPPC, DMPC, DOPC, POPC, DOPE, DOPG, DPPG, POPE, DPPE, DMPE, DSPE, and SM. In some embodiments, the neutral lipid is selected from the group consisting of DSPC, DPPC, DMPC, DOPC, POPC, DOPE and SM. In some embodiments, the neutral lipid is DSPC.

In some embodiments, the steroid is cholesterol.

In some embodiments, the polymer conjugated lipid is a pegylated lipid.

In some embodiments, the cationically ionizable lipid component of the LNPs has one of the following structures:

No.	Structure
1	
2	
3	
4	
5	
6	
7	

No.	Structure
15	
16	
17	
18	
19	
20	
21	
22	

No.	Structure
23	<chem>CCCCCCCCC(=O)OCCCCN(CCCCCCOC(=O)CCCCCCCC)CCCCCOC(=O)CCCCCCCC</chem>
24	<chem>CCCCC(=O)OCCCCN(CCCCCCOC(=O)CCCC)CCCCCOC(=O)CCCC</chem>
25	<chem>CCCCC(=O)OCCCCN(CCCCCCOC(=O)CCCC)CCCCCOC(=O)CCCC</chem>
26	<chem>CCCCCCCCC(=O)OCCCCN(CCCCCCOC(=O)CCCCCCCC)CCCCCOC(=O)CCCCCCCC</chem>
27	<chem>CCCCCCCCC(=O)OCCCCN(CCCCCCOC(=O)CCCCCCCC)CCCCCOC(=O)CCCCCCCC</chem>
28	<chem>CCCCC(=O)OCCCCN(CCCCCCOC(=O)CCCC)CCCCCOC(=O)CCCC</chem>
29	<chem>CCCCC(=O)OCCCCN(CCCCCCOC(=O)CCCC)CCCCCOC(=O)CCCC</chem>

No.	Structure
30	 <chem>CCCCCCCCCCCC(=O)OCCN(CCCCCCCCCC=O)CCCCCCCCCCCC(=O)OCC</chem>
31	 <chem>CCCCCCCCCCCC(=O)OCCN(CCCCCCCCCC=O)CCCCCCCCCCCC(=O)OCC</chem>
32	 <chem>CCCCCCCCCCCC(=O)OCCN(CCCCCCCCCC=O)CCCCCCCCCCCC(=O)OCC</chem>
33	 <chem>CCCCCCCCCCCC(=O)OCCN(CCCCCCCCCC=O)CCCCCCCCCCCC(=O)OCC</chem>
34	 <chem>CCCCCCCCCCCC(=O)OCCN(CCCCCCCCCC=O)CCCCCCCCCCCC(=O)OCC</chem>
35	 <chem>CCCCCCCCCCCC(=O)OCCN(CCCCCCCCCC=O)CCCCCCCCCCCC(=O)OCC</chem>
36	 <chem>CCCCCCCCCCCC(=O)OCCN(CCCCCCCCCC=O)CCCCCCCCCCCC(=O)OCC</chem>

No.	Structure
37	
38	
39	
40	
41	
42	
43	

No.	Structure
44	
45	
46	
47	
48	
49	

Pharmaceutical compositions

The agents described herein may be administered in pharmaceutical compositions and may be administered in the form of any suitable pharmaceutical composition.

In some embodiments, the agents described herein may be administered in a pharmaceutical composition which may comprise a pharmaceutically acceptable carrier and may optionally comprise one or more stabilizers etc. In some embodiments, the pharmaceutical composition is for therapeutic or prophylactic treatments, e.g., for use in treating or preventing a disease. The term "pharmaceutical composition" relates to a formulation comprising a therapeutically effective agent, preferably together with pharmaceutically acceptable carriers, diluents and/or excipients. Said pharmaceutical composition is useful for treating, preventing, or reducing the severity of a disease or disorder by administration of said pharmaceutical composition to a subject. A pharmaceutical composition is also known in the art as a pharmaceutical formulation.

The pharmaceutical compositions according to the present disclosure are generally applied in a "pharmaceutically effective amount" and in "a pharmaceutically acceptable preparation".

The term "pharmaceutically acceptable" refers to the non-toxicity of a material which does not interact with the action of the active component of the pharmaceutical composition.

The term "pharmaceutically effective amount" or "therapeutically effective amount" refers to the amount which achieves a desired reaction or a desired effect alone or together with further doses and/or agents. In the case of the treatment of a particular disease, the desired reaction preferably relates to inhibition of the course of the disease. This comprises slowing down the progress of the disease and, in particular, interrupting or reversing the progress of the disease. The desired reaction in a treatment of a disease may also be delay of the onset or a prevention of the onset of said disease or said condition. An effective amount of the compositions described herein will depend on the condition to be treated, the severeness of the disease, the individual parameters of the patient, including age, physiological condition, size and weight, the duration of treatment, the type of an accompanying therapy (if present), the specific route of administration and similar factors. Accordingly, the doses administered of the compositions described herein may depend on various of such parameters. In the case that a reaction in a patient is insufficient with an initial dose, higher doses (or effectively higher doses achieved by a different, more localized route of administration) may be used.

The pharmaceutical compositions of the present disclosure may contain salts, buffers, preservatives, and optionally other therapeutic agents. In some embodiments, the

pharmaceutical compositions of the present disclosure comprise one or more pharmaceutically acceptable carriers, diluents and/or excipients.

Suitable preservatives for use in the pharmaceutical compositions of the present disclosure include, without limitation, benzalkonium chloride, chlorobutanol, paraben and thimerosal.

The term "excipient" as used herein refers to a substance which may be present in a pharmaceutical composition of the present disclosure but is not an active ingredient. Examples of excipients, include without limitation, carriers, binders, diluents, lubricants, thickeners, surface active agents, preservatives, stabilizers, emulsifiers, buffers, flavoring agents, or colorants.

The term "diluent" relates a diluting and/or thinning agent. Moreover, the term "diluent" includes any one or more of fluid, liquid or solid suspension and/or mixing media. Examples of suitable diluents include ethanol, glycerol and water.

The term "carrier" refers to a component which may be natural, synthetic, organic, inorganic in which the active component is combined in order to facilitate, enhance or enable administration of the pharmaceutical composition. A carrier as used herein may be one or more compatible solid or liquid fillers, diluents or encapsulating substances, which are suitable for administration to subject. Suitable carriers include, without limitation, sterile water, Ringer, Ringer lactate, sterile sodium chloride solution, isotonic saline, polyalkylene glycols, hydrogenated naphthalenes and, in particular, biocompatible lactide polymers, lactide/glycolide copolymers or polyoxyethylene/polyoxy-propylene copolymers. In some embodiments, the pharmaceutical composition of the present disclosure includes isotonic saline.

Pharmaceutically acceptable carriers, excipients or diluents for therapeutic use are well known in the pharmaceutical art, and are described, for example, in Remington's Pharmaceutical Sciences, Mack Publishing Co. (A. R Gennaro edit. 1985).

Pharmaceutical carriers, excipients or diluents can be selected with regard to the intended route of administration and standard pharmaceutical practice.

In some embodiments, pharmaceutical compositions described herein may be administered intravenously, intraarterially, subcutaneously, intradermally or intramuscularly. In certain embodiments, the pharmaceutical composition is formulated for local administration or

systemic administration. Systemic administration may include enteral administration, which involves absorption through the gastrointestinal tract, or parenteral administration. As used herein, "parenteral administration" refers to the administration in any manner other than through the gastrointestinal tract, such as by intravenous injection. In some embodiments, the pharmaceutical composition is formulated for systemic administration, e.g., for intravenous administration.

Treatments

The agents, compositions and methods described herein can be used to treat a subject with a disease, e.g., a disease characterized by the presence of diseased cells expressing an antigen. The agents, compositions and methods described herein may be used in the therapeutic or prophylactic treatment of various diseases. Particularly preferred diseases are cancer diseases. In some embodiments, the agents, compositions and methods described herein are useful in a prophylactic and/or therapeutic treatment of a disease involving an antigen.

Such antigen may serve as target for a docking compound. For example, if the antigen is derived from a virus, the agents, compositions and methods may be useful in the treatment of a viral disease caused by said virus. If the antigen is a tumor antigen, the agents, compositions and methods may be useful in the treatment of a cancer disease wherein cancer cells express said tumor antigen.

The term "disease" refers to an abnormal condition that affects the body of an individual. A disease is often construed as a medical condition associated with specific symptoms and signs. A disease may be caused by factors originally from an external source, such as infectious disease, or it may be caused by internal dysfunctions, such as autoimmune diseases. In humans, "disease" is often used more broadly to refer to any condition that causes pain, dysfunction, distress, social problems, or death to the individual afflicted, or similar problems for those in contact with the individual. In this broader sense, it sometimes includes injuries, disabilities, disorders, syndromes, infections, isolated symptoms, deviant behaviors, and atypical variations of structure and function, while in other contexts and for other purposes these may be considered distinguishable categories. Diseases usually affect individuals not

only physically, but also emotionally, as contracting and living with many diseases can alter one's perspective on life, and one's personality.

In the present context, the term "treatment", "treating" or "therapeutic intervention" relates to the management and care of a subject for the purpose of combating a condition such as a disease or disorder. The term is intended to include the full spectrum of treatments for a given condition from which the subject is suffering, such as administration of the therapeutically effective compound to alleviate the symptoms or complications, to delay the progression of the disease, disorder or condition, to alleviate or relief the symptoms and complications, and/or to cure or eliminate the disease, disorder or condition as well as to prevent the condition, wherein prevention is to be understood as the management and care of an individual for the purpose of combating the disease, condition or disorder and includes the administration of the active compounds to prevent the onset of the symptoms or complications.

The term "therapeutic treatment" relates to any treatment which improves the health status and/or prolongs (increases) the lifespan of an individual. Said treatment may eliminate the disease in an individual, arrest or slow the development of a disease in an individual, inhibit or slow the development of a disease in an individual, decrease the frequency or severity of symptoms in an individual, and/or decrease the recurrence in an individual who currently has or who previously has had a disease.

The terms "prophylactic treatment" or "preventive treatment" relate to any treatment that is intended to prevent a disease from occurring in an individual. The terms "prophylactic treatment" or "preventive treatment" are used herein interchangeably.

The terms "individual" and "subject" are used herein interchangeably. They refer to a human or another mammal (e.g. mouse, rat, rabbit, dog, cat, cattle, swine, sheep, horse or primate) that can be afflicted with or is susceptible to a disease or disorder (e.g., cancer) but may or may not have the disease or disorder. In many embodiments, the individual is a human being. Unless otherwise stated, the terms "individual" and "subject" do not denote a particular age, and thus encompass adults, elderlies, children, and newborns. In embodiments of the present disclosure, the "individual" or "subject" is a "patient".

The term "patient" means an individual or subject for treatment, in particular a diseased individual or subject.

In some embodiments of the disclosure, the aim is to deliver immune effector cells to diseased cells expressing an antigen such as cancer cells expressing a tumor antigen, and to treat a disease such as a cancer disease involving cells expressing an antigen such as a tumor antigen. In some embodiments, immune effector cells exert one or more immune effector functions on diseased cells, e.g., kill diseased cells by means of a cellular immune response.

The term "disease involving an antigen", "disease involving cells expressing an antigen" or similar terms refer to any disease which implicates an antigen, e.g. a disease which is characterized by the presence of an antigen. The disease involving an antigen can be an infectious disease, or a cancer disease or simply cancer. As mentioned above, the antigen may be a disease-associated antigen, such as a tumor-associated antigen, a viral antigen, or a bacterial antigen. In some embodiments, a disease involving an antigen is a disease involving cells expressing an antigen, preferably on the cell surface.

The term "infectious disease" refers to any disease which can be transmitted from individual to individual or from organism to organism, and is caused by a microbial agent (e.g. common cold). Infectious diseases are known in the art and include, for example, a viral disease, a bacterial disease, or a parasitic disease, which diseases are caused by a virus, a bacterium, and a parasite, respectively. In this regard, the infectious disease can be, for example, hepatitis, sexually transmitted diseases (e.g. chlamydia or gonorrhea), tuberculosis, HIV/acquired immune deficiency syndrome (AIDS), diphtheria, hepatitis B, hepatitis C, cholera, severe acute respiratory syndrome (SARS), the bird flu, and influenza.

The terms "cancer disease" or "cancer" refer to or describe the physiological condition in an individual that is typically characterized by unregulated cell growth. Examples of cancers include, but are not limited to, carcinoma, lymphoma, blastoma, sarcoma, and leukemia. More particularly, examples of such cancers include bone cancer, blood cancer, lung cancer, liver cancer, pancreatic cancer, skin cancer, cancer of the head or neck, cutaneous or intraocular melanoma, uterine cancer, ovarian cancer, rectal cancer, cancer of the anal region, stomach cancer, colon cancer, breast cancer, prostate cancer, uterine cancer, carcinoma of the sexual and reproductive organs, Hodgkin's Disease, cancer of the esophagus, cancer of the

small intestine, cancer of the endocrine system, cancer of the thyroid gland, cancer of the parathyroid gland, cancer of the adrenal gland, sarcoma of soft tissue, cancer of the bladder, cancer of the kidney, renal cell carcinoma, carcinoma of the renal pelvis, neoplasms of the central nervous system (CNS), neuroectodermal cancer, spinal axis tumors, glioma, meningioma, and pituitary adenoma. The term "cancer" according to the disclosure also comprises cancer metastases.

The term "solid tumor" or "solid cancer" as used herein refers to the manifestation of a cancerous mass, as is well known in the art for example in Harrison's Principles of Internal Medicine, 14th edition. Preferably, the term refers to a cancer or carcinoma of body tissues other than blood, preferably other than blood, bone marrow, and lymphoid system. For example, but not by way of limitation, solid tumors include cancers of the prostate, lung cancer, colorectal tissue, bladder, oropharyngeal/laryngeal tissue, kidney, breast, endometrium, ovary, cervix, stomach, pancreas, brain, and central nervous system.

The methods and agents described herein are, in particular, useful for the treatment of cancers, e.g., solid cancers, characterized by diseased cells expressing an antigen a docking compound is directed to.

"Cell-mediated immunity", "cellular immunity", "cellular immune response", or similar terms are meant to include a cellular response directed to cells characterized by expression of an antigen, in particular characterized by presentation of an antigen with class I or class II MHC. The cellular response relates to cells called T cells or T lymphocytes which act as either "helpers" or "killers". The helper T cells (also termed CD4+ T cells) play a central role by regulating the immune response and the killer cells (also termed cytotoxic T cells, cytolytic T cells, CD8+ T cells or CTLs) kill diseased cells such as cancer cells, preventing the production of more diseased cells.

The term "antigen presenting cell" (APC) is a cell of a variety of cells capable of displaying, acquiring, and/or presenting at least one antigen or antigenic fragment on (or at) its cell surface. Antigen-presenting cells can be distinguished in professional antigen presenting cells and non-professional antigen presenting cells.

The term "professional antigen presenting cells" relates to antigen presenting cells which constitutively express the Major Histocompatibility Complex class II (MHC class II) molecules

required for interaction with naive T cells. If a T cell interacts with the MHC class II molecule complex on the membrane of the antigen presenting cell, the antigen presenting cell produces a co-stimulatory molecule inducing activation of the T cell. Professional antigen presenting cells comprise dendritic cells and macrophages.

The term "non-professional antigen presenting cells" relates to antigen presenting cells which do not constitutively express MHC class II molecules, but upon stimulation by certain cytokines such as interferon-gamma. Exemplary, non-professional antigen presenting cells include fibroblasts, thymic epithelial cells, thyroid epithelial cells, glial cells, pancreatic beta cells or vascular endothelial cells.

"Antigen processing" refers to the degradation of an antigen into procession products, which are fragments of said antigen (e.g., the degradation of a protein into peptides) and the association of one or more of these fragments (e.g., via binding) with MHC molecules for presentation by cells, such as antigen presenting cells to specific T cells.

Citation of documents and studies referenced herein is not intended as an admission that any of the foregoing is pertinent prior art. All statements as to the contents of these documents are based on the information available to the applicants and do not constitute any admission as to the correctness of the contents of these documents.

The following description is presented to enable a person of ordinary skill in the art to make and use the various embodiments. Descriptions of specific devices, techniques, and applications are provided only as examples. Various modifications to the examples described herein will be readily apparent to those of ordinary skill in the art, and the general principles defined herein may be applied to other examples and applications without departing from the spirit and scope of the various embodiments. Thus, the various embodiments are not intended to be limited to the examples described herein and shown, but are to be accorded the scope consistent with the claims.

Examples

Example 1: CDR grafting and humanization of NbALFA

In the first step, the CDRs of NbALFA were grafted onto multiple acceptor frameworks. As described herein, CDRs were defined broadly by a combination of Kabat and IMGT numbering schemes. For CDR grafting, seven variants of humanized VHH acceptor frameworks were chosen. Variants 1 to 7 of humanised NbALFA antibody are set forth in SEQ ID NOs: 111 to 117. Genes encoding variants 1 to 7 were cloned into expression vectors and subsequently produced in *E. coli* by procedures well known to a person skilled in the art. Antibody variants were purified with IMAC chromatography and subsequently tested for antigen binding with biolayer interferometry.

Variants 1 and 2 failed in production and variants 3 to 6 showed no binding to the epitope tag of SEQ ID NO: 2. Variant 7 showed binding to the epitope tag of SEQ ID NO: 2, albeit with a significantly impaired antigen recognition, demonstrated by a fast off-rate and low KD (see Table 1).

Table 1: Binding affinities of humanized NbALFA variants

Variant	KD [M]	Ka [Ms ⁻¹]	Kdis [s ⁻¹]
wtNbALFA (SEQ ID NO: 1)	1.60*10 ⁻¹⁰	3.99*10 ⁵	6.37*10 ⁻⁵
Variant 1 (SEQ ID NO: 111)	Failed production		
Variant 2 (SEQ ID NO: 112)	Failed production		
Variant 3 (SEQ ID NO: 113)	Loss of binding		
Variant 4 (SEQ ID NO: 114)	Loss of binding		
Variant 5 (SEQ ID NO: 115)	Loss of binding		
Variant 6 (SEQ ID NO: 116)	Loss of binding		
Variant 7 (SEQ ID NO: 117)	1.86*10 ⁻⁹	1.49*10 ⁵	2.76*10 ⁻⁴

Example 2: Further engineering of humanized variants

Based on extensive structural analysis, 8 further variants of humanized NbALFA were generated. Variants 8 to 15 are set forth in SEQ ID NO: 118 to 125 and their affinity, on/off rates, thermal stability, aggregation behavior and IGHV3-23 framework identity are shown in Table 2. Surprisingly, all variants showed binding to SEQ ID NO: 2 with at comparable binding affinity as compared to wtNbALFA. Strikingly, variants 9, 10, 12, 13, 14, and 15 exhibited wild type affinity, while variant 11 even showed elevated affinity towards SEQ ID NO: 2.

Table 2: affinity optimized variants

Variant #	K _D [M]	K _a [1/Ms]	k _{dis} [1/s]	SEC Area [% @ 280nm]
wtNbALFA (SEQ ID NO: 1)	1.60*10 ⁻¹⁰	3.99*10 ⁵	6.37*10 ⁻⁵	100
Variant 8 (SEQ ID NO: 118)	2.63*10 ⁻¹⁰	1.00*10 ⁶	2.63*10 ⁻⁴	87.16
Variant 9 (SEQ ID NO: 119)	1.82*10 ⁻¹⁰	2.03*10 ⁵	3.69*10 ⁻⁵	100
Variant 10 (SEQ ID NO: 120)	3.06*10 ⁻¹⁰	4.80*10 ⁵	1.47*10 ⁻⁴	56.46
Variant 11 (SEQ ID NO: 121)	5.95*10 ⁻¹¹	6.22*10 ⁵	3.70*10 ⁻⁵	100
Variant 12 (SEQ ID NO: 122)	3.48*10 ⁻¹⁰	6.09*10 ⁵	1.51*10 ⁻⁴	97.56
Variant 13 (SEQ ID NO: 123)	2.17*10 ⁻¹⁰	7.67*10 ⁵	1.67*10 ⁻⁴	100
Variant 14 (SEQ ID NO: 124)	2.18*10 ⁻¹⁰	5.96*10 ⁵	1.30*10 ⁻⁴	100
Variant 15 (SEQ ID NO: 125)	2.35*10 ⁻¹⁰	1.59*10 ⁶	3.02*10 ⁻⁴	100

Table 3: Framework identity and V segment identity of humanized variants

Variant #	Framework identity [%] IGHV3-23, excluding CDRs					V segment identity [%] IGHV3-23
	FR1	FR2	FR3	FR4	Overall framework identity (Weighted average)	
wtNbALFA (SEQ ID NO: 1)	92.0	53.3	55.3	90.9	69.7	59.6
Variant 8 (SEQ ID NO: 118)	100	66.7	78.9	90.9	84.3	72.7
Variant 9 (SEQ ID NO: 119)	100	66.7	78.9	90.9	84.3	72.7
Variant 10 (SEQ ID NO: 120)	100	66.7	78.9	90.9	84.3	73.7
Variant 11 (SEQ ID NO: 121)	100	66.7	81.6	90.9	85.4	73.7
Variant 12 (SEQ ID NO: 122)	100	66.7	81.6	90.9	85.4	74.7
Variant 13 (SEQ ID NO: 123)	100	66.7	86.8	90.9	87.6	76.8
Variant 14 (SEQ ID NO: 124)	100	66.7	92.1	90.9	89.9	77.8
Variant 15 (SEQ ID NO: 125)	100	66.7	97.4	90.9	92.2	80.8

Example 3: Determination of SpA binding

To further investigate the newly generated variants, SpA affinity was determined with the humanized NbALFA. Capacity of SpA binding may provide a straightforward purification process of the humanised antibody under GMP conditions and to circumvent the need to establish a new purification method. SpA binding and the capability to be purified using SpA columns were thus investigated (see Fig. 3, Table 4). The spike experiment revealed a quantitative binding of variant 10, 12, 13 and 15 to the column as the flow through lacked any presence of the VHH. These four variants were not detectable in the two wash fractions and eluted in the two elution steps. However, variants 11 and 14 were bound by the column but remained detectable in the flow through fractions. In the washing steps a clear “bleeding” was detected, and the elution did not lead to an elevated presence of the VHH in the fraction, rendering both variants unfit for SpA-based purification.

Table 4: determination of SpA binding

Variant	SpA binding
wtNbALFA (SEQ ID NO:1)	-
Variant 10 (SEQ ID NO: 120)	+
Variant 11 (SEQ ID NO: 121)	-
Variant 12 (SEQ ID NO: 122)	+
Variant 13 (SEQ ID NO: 123)	+
Variant 14 (SEQ ID NO: 124)	-
Variant 15 (SEQ ID NO: 125)	+

Example 4: Determination of thermal stability

Thermal stability of the humanized variants of NbALFA was determined. T_m values representing thermal stability are shown in Table 5. Strikingly, despite the humanization, variant 8 showed comparable T_m values as wtNbALFA, while variants 9-15 even surprisingly showed higher thermal stability as wtNbALFA.

Table 5: determination of thermal stability

Variant #	T _m [°C]
wtNbALFA (SEQ ID NO: 1)	61.3
Variant 8 (SEQ ID NO: 118)	61.1
Variant 9 (SEQ ID NO: 119)	65.2
Variant 10 (SEQ ID NO: 120)	63.8
Variant 11 (SEQ ID NO: 121)	64.6
Variant 12 (SEQ ID NO: 122)	64.8
Variant 13 (SEQ ID NO: 123)	66.7
Variant 14 (SEQ ID NO: 124)	64.2
Variant 15 (SEQ ID NO: 125)	66.8

Table 6: VHH sequences of humanized NbALFA nanobodies

Variant #	Sequence
Variant 1 (SEQ ID NO: 111)	EVQLLESGGGLVQPGGSLRLSCAASGVTISALNAMAMGWYRQA PGKRREMVSAVSEKGNAMYRESVQGRFTISRDNKNTLYLQMNS LRAEDTAVYYCAKLEDKVDKSFHDYWGQGTLVTVSS
Variant 2 (SEQ ID NO: 112)	EVQLLESGGGLVQPGGSLRLSCAASGVTISALNAMAMGWFRQA PGKEREGVSAVSEKGNAMYRESVQGRFTISRDNKNTLYLQMNS LRAEDTAVYYCAKLEDKVDKSFHDYWGQGTLVTVSS
Variant 3 (SEQ ID NO: 113)	EVQLLESGGGLVQPGGSLRLSCAASGVTISALNAMAMGWFRQA PGKGLEFVAASVSEKGNAMYRESVQGRFTISRDNKNTLYLQMNS LRAEDTAVYYCAKLEDKVDKSFHDYWGQGTLVTVSS
Variant 4 (SEQ ID NO: 114)	EVQLLESGGGLVQPGGSLRLSCAASGVTISALNAMAMGWFRQA PGKGLEFVAASVSEKGNAYYRDSVKGRFTISRDNKNTLYLQMNSL RAEDTAVYYCAKLEDKVDKSFHDYWGQGTLVTVSS

Variant 5 (SEQ ID NO: 115)	EVQLQASGGGFVQPGGSLRLSCAASGVTISALNAMAMGWFRQA PGKEREFSVAISERGNAYYRDSVKGRFTISRDN SKNTVYLQMNSLR AEDTATYYCAALED RVDSFHDYWGQGTQVTVSS
Variant 6 (SEQ ID NO: 116)	EVQLQASGGGFVQPGGSLRLSCAASGVTISALNAMAMGWFRQA PGKEREFSVAISERGNAMYRESVQGRFTISRDN SKNTVYLQMNS LRAEDTATYYCAALED RVDSFHDYWGQGTQVTVSS
Variant 7 (SEQ ID NO: 117)	EVQLVESGGGLVQPGGSLRLSCAASGVTISALNAMAMGWFRQA PGKGRELVAASERGNAMYRESVQGRFTISRDN AKRMVYLQMN SLRAEDTAVYYCHVLED RVDSFHDYWGQGTQVTVSS
Variant 8 (SEQ ID NO: 118)	EVQLVESGGGLVQPGGSLRLSCAASGVTISALNAMAMGWFRQA PGKRREMVAASERGNAMYRESVQGRFTISRDN AKRMVYLQM NSLRAEDTAVYYCHVLED RVDSFHDYWGQGTQVTVSS
Variant 9 (SEQ ID NO: 119)	EVQLVESGGGLVQPGGSLRLSCAASGVTISALNAMAMGWYRQA PGKRREMVAASERGNAMYRESVQGRFTISRDN AKRMVYLQM NSLRAEDTAVYYCHVLED RVDSFHDYWGQGTQVTVSS
Variant 10 (SEQ ID NO: 120)	EVQLVESGGGLVQPGGSLRLSCAASGVTISALNAMAMGWYRQA PGKRREMVAASERGN TMYRESVQGRFTISRDN AKRMVYLQM NSLRAEDTAVYYCHVLED RVDSFHDYWGQGTQVTVSS
Variant 11 (SEQ ID NO: 121)	EVQLVESGGGLVQPGGSLRLSCAASGVTISALNAMAMGWYRQA PGKRREMVAASERGNAMYRESVKGRFTISRDN AKRMVYLQMN SLRAEDTAVYYCHVLED RVDSFHDYWGQGTQVTVSS
Variant 12 (SEQ ID NO: 122)	EVQLVESGGGLVQPGGSLRLSCAASGVTISALNAMAMGWYRQA PGKRREMVAASERGN TMYRESVKGRFTISRDN AKRMVYLQMN SLRAEDTAVYYCHVLED RVDSFHDYWGQGTQVTVSS
Variant 13 (SEQ ID NO: 123)	EVQLVESGGGLVQPGGSLRLSCAASGVTISALNAMAMGWYRQA PGKRREMVAASERGN TYYRDSVKGRFTISRDN AKRMVYLQMN SLRAEDTAVYYCHVLED RVDSFHDYWGQGTQVTVSS
Variant 14 (SEQ ID NO: 124)	EVQLVESGGGLVQPGGSLRLSCAASGVTISALNAMAMGWYRQA PGKRREMVAASERGNAMYRESVKGRFTISRDN SKNTLYLQMNS LRAEDTAVYYCHVLED RVDSFHDYWGQGTQVTVSS

Variant 15 (SEQ ID NO: 125)	EVQLVESGGGLVQPGGSLRLSCAASGVTISALNAMAMGWYRQA PGKRREMVAASVSESGNTYYRDSVKGRFTISRDN SKNTLYLQMNSL RAEDTAVYYCHVLEDRVDSFHDYWGQGTQVTVSS
-----------------------------	--

Table 7: CDR and Framework region sequences of humanized NbALFA nanobodies of the present invention

SEQ ID NO	Region	Applicable to variants	Sequence
SEQ ID NO: 94	CDR1	Variants 8 to 15	GVTISALNAMAMG
SEQ ID NO: 95	CDR2	Variants 8, 9, 11, 14	VSESGNA
SEQ ID NO: 96	CDR2	Variants 10, 12, 13, 15	VSESGNT
SEQ ID NO: 97	CDR2 consensus	Variants 8 to 15	VSESGNX1, wherein X1 is A or T
SEQ ID NO: 98	CDR3	Variants 8 to 15	HVLEDRVDSFHDY
SEQ ID NO: 99	FR1	Variants 8 to 15	EVQLVESGGGLVQPGGSLRLSCAAS
SEQ ID NO: 100	FR2	Variant 8	WFRQAPGKRREMVAAS
SEQ ID NO: 101	FR2	Variants 9 to 15	WYRQAPGKRREMVAAS
SEQ ID NO: 102	FR2 consensus	Variants 8 to 15	WX1RQAPGKRREMVAAS, wherein X1 is F or Y
SEQ ID NO: 103	FR3	Variants 8 to 10	MYRESVQGRFTISRDN AKRMVYLQ MNSLRAEDTAVYYC
SEQ ID NO: 104	FR3	Variants 11, 12	MYRESVKGRFTISRDN AKRMVYLQ MNSLRAEDTAVYYC
SEQ ID NO: 105	FR3	Variant 13	YYRDSVKGRFTISRDN AKRMVYLQM NSLRAEDTAVYYC
SEQ ID NO: 106	FR3	Variant 14	MYRESVKGRFTISRDN SKNTLYLQM NSLRAEDTAVYYC
SEQ ID NO: 107	FR3	Variant 15	YYRDSVKGRFTISRDN SKNTLYLQM NSLRAEDTAVYYC
SEQ ID NO: 108	FR3 consensus	Variants 8 to 15	X1YRX2SVX3GRFTISRDNX4KX5X6 X7YLQMNSLRAEDTAVYYC,

			wherein X1 is M or Y, X2 is D or E, X3 is Q or K, X4 is A or S, X5 is N or R, X6 is M or T, X7 is V or L
SEQ ID NO: 109	FR4	Variants 8 to 15	WGQGTQVTVSS

Example 5: Methods

Method 1: Affinity determination

To determine the affinity of humanized NbALFA variants to the epitope tag of SEQ ID NO: 2, an Octet RED96 from ForteBio was utilized. The measurement principle is based on biolayer interferometry (BLI). SAX2.0 biosensors were soaked for 10 min in kinetic buffer (KB), supplied by the manufacturer of the instrument. At first, a 60 or 120 sec baseline in KB was measured, followed by a loading step of 240 sec with 83 ng/ml of biotinylated peptide of SEQ ID NO: 2. After loading, free streptavidin entities on the biosensor were quenched during a 120 sec incubation step in a 100 µg/ml biocytin solution. Subsequently, a 60 to 600 sec baseline in KB was measured followed by association of the VHH. The VHH was applied with a maximal concentration of 10 nM and was 1:1 serial diluted in KB. As a reference, only KB was measured. Association was performed for 900 sec, followed by a dissociation time of 1500 sec in KB. The signal of the reference well was subtracted from the signal of all other biosensors and the signals of these processed data were aligned to the average of the second baseline. For inter-step correction, the data were aligned to the dissociation step and Savitzky-Golay filtering was applied to all curves. Association and dissociations were globally fitted using a 1:1 Langmuir binding model.

Method 2: Thermal stability analysis

Thermal stability was investigated using the Prometheus Tycho from NanoTemper. 10 µl of a VHH sample was loaded into capillaries. After loading the capillaries into the instrument, a heat ramp of 30°C/min from 35°C to 95°C was applied to all samples. During this process the intrinsic fluorescence of the proteins were measured at 350 nm and 330 nm. The ratio was plotted against the temperature and the first derivative was calculated. Minima and maxima correspond to different T_M values that are observed in the process of denaturation.

Method 3: SEC analysis

Aggregation analysis was performed via SEC utilizing an Agilent Infinity II HPLC and a Biozen 1.8 μm dSEC-2, 200 Å LC column (300 x 4.6 mm). Flowrates were adjusted to 0.25 mL/min, resulting in approximately 255 bar pressure. As mobile phase 0.2 M potassium phosphate, 250 mM KCl, pH 6.2, 5% acetonitrile was used. Each run took 20 min excluding a 2-3 min wash step between each analysis. 10 μL of the VHH proteins were applied and detected by absorption at 280 nm.

Method 4: SpA binding determination

45 μg of the purified VHHs were diluted in 200 μL of CHO cell culture supernatant. SpA HP SpinTrap columns from Cytiva were equilibrated once with 600 μL of PBS and subsequently the 200 μL sample was applied and incubated according to the manufactures protocol. Subsequently two washing steps with 600 μL of PBS were performed followed by two elution steps with 400 μL of 100 mM Glycin pH 2.7. Of the VHH in the CHO supernatant, the flow through the two wash steps and the two elution fractions SDS-PAGE analysis was performed.

Example 6: Testing of humanized NbALFA-CARs

Constructs and method

Two humanized Nanobody (Nb) versions (variant 11/variant 13), specific for the ALFA-Tag (NbALFA), were used to generate modular (Mod) CARs. The humanized NbALFA domains were fused to a CAR scaffold comprising the human CD8 α hinge/transmembrane domain, followed by human 4-1BB co-stimulatory and human CD3 ζ signaling chain with the amino acid sequence of SEQ ID NO: 486. The CARs were cloned into DNA transposon vectors in order to genetically modify T cells by transposition. For this purpose, primary T cells of healthy donors were co-transfected with the transposon vector and transposase-encoding mRNA by electroporation. For *in vitro* functional testing, the humanized as well as the wildtype (NbALFA-CAR) ModCARs were compared with each other. The different ModCAR scaffolds were tested in combination with an anti-CLDN6 targeting adaptor molecule composed of a CLDN6-specific scFv fused to the ALFA-tag to evaluate antigen-specific binding and activation of ModCAR T cells by CLDN6 expressing tumor cell lines.

In order to test the functionality of the humanized ModCARs (variant 11/variant 13) in comparison to the NbALFA-CAR, the generated CAR T cells were co-cultured with CLDN6⁺

tumor cells. For the 2D tumor experiment, T cells of two individual donors and for the 3D tumor experiment T cells of one donor were modified and tested. Adaptor-containing cell culture supernatant, previously generated by adaptor-mRNA-transfection of HEK293T/17 cells, was added in different dilutions to the co-culture.

Example 6.1: 2D Tumor Cell Killing Experiment

The CLDN6⁺ tumor cell monolayer was seeded 24 hours prior to the co-culture with ModCAR T cells. On the day of co-culture, the CAR frequency of ModCAR T cells was adjusted to 50 % for both donors. The final CAR⁺ effector to target cell ratio equated 5:1. Adaptor-containing supernatant was added to the co-culture with a final dilution of 2 %, 0.5 %, 0.125 % and 0.03 %. As control condition, no adaptor-containing supernatant was administered. 24 hours after co-culture, culture supernatants were collected for cytokine analysis (IL-2, GM-CSF, IFN- γ , TNF- α). Moreover, CAR T cells were analyzed for upregulation of activation (CD25, CD69) and degranulation (CD107a) markers by flow cytometry.

Tumor cell lysis of the CLDN6⁺ tumor monolayer was analyzed via xCELLigence over 96 hours. Data were normalized to the Mock (non-genetically modified) T cell control and set into relation to maximum tumor cell lysis achieved by administration of Triton-X-100.

Analysis of the degranulation marker CD107a after 24 hours of co-culture with target cells demonstrated an adaptor dose-dependent increase of CD107a expression on CAR⁺ T cells, while no effect was observed for the Mock control cells. Overall, the CD107a upregulation on CAR⁺ T cells was comparable for all tested ModCARs (figure 5).

Both humanized ModCARs showed slightly elevated upregulation of activation marker (CD25, CD69) after 24 hours of co-culture when compared to the NbALFA-CAR (figure 5). Moreover, variant 13 showed higher secretion of effector cytokines (IL-2, GM-CSF, TNF- α and IFN- γ) at lower adaptor dose (≤ 0.5 %) than variant 11 or NbALFA-CAR (figure 6).

All ModCARs induced efficient and comparable lysis of a CLDN6⁺ tumor monolayer with 2 % adaptor dilution. With decreased adaptor concentration (0.5 %, 0.125 %) humanized ModCAR variant 13 showed superiority in tumor cell lysis over variant 11 or NbALFA-CAR for both donors. None of the ModCARs showed unspecific tumor cell lysis without adaptor supernatant administration (figure 7).

Example 6.2: 3D Tumor Spheroid Killing Experiment

CLDN6⁺ or CLDN6⁻ tumor spheroids were generated 24 hours prior to the co-culture with ModCAR T cells. On the day of co-culture, the CAR frequency was adjusted to 70 %. The final CAR⁺ effector to target cell ratio equated 7:1. Adaptor-containing supernatant was added to the co-culture with a final dilution of 4 %, 1 % and 0.25 % at time point 0 hours. As control condition, no adaptor-containing supernatant was administered. For the co-culture with CLDN6⁺ spheroids, a second and third round of CLDN6⁺ spheroids were added at time points 166 hours and 334 hours, respectively.

Spheroid lysis was followed by Incucyte over 168 hours for the co-culture with CLDN6⁻ spheroids and over 480 hours for the co-culture with CLDN6⁺ spheroids. Analysis of T cell expansion and spheroid lysis was performed at the end of the experiment using the Incucyte software. Data were normalized to the Mock (non-genetically modified) T cell control.

In co-culture with CLDN6⁺ tumor spheroids, all ModCARs induced efficient repetitive spheroid lysis with a high adaptor dose of 4 %. With lower adaptor concentration (1 %, 0.25 % dilution) differences in CAR functionality became evident. Here, ModCAR variant 13 showed superior tumor spheroid lysis with still complete lysis of a third spheroid in the presence of 1 % adaptor-containing supernatant. With 0.25 % adaptor-containing supernatant, variant 13 still induced around 50 % lysis of a third spheroid while ModCAR variant 11 and NbALFA-CAR could not induce any cytotoxicity (figure 8). None of the ModCARs showed unspecific lysis of CLDN6⁺ and CLDN6⁻ tumor spheroids without adaptor supernatant administration (figure 8, figure 9). Moreover, none of the ModCARs showed elevated unspecific T cell expansion on CLDN6⁻ spheroids over the time course of 7 days (figure 9).

These results as well as the results from example 6.1 show that one of the humanized ModCARs (variant 13) demonstrates even superior functionality than the wildtype ModCAR. The CAR T cells expressing humanized ModCAR variant 13 presented a higher functionality at lower adaptor concentrations leading to slightly increased secretion of effector cytokines and induction of stronger cytotoxicity towards CLDN6⁺ tumor cells. This could be confirmed in 2D and 3D tumor models *in vitro*. Further, the humanized ModCARs did not induce unspecific

tumor cell lysis. This could be demonstrated in 2D and 3D tumor models *in vitro* by not adding any adaptor-containing supernatant in a co-culture with CLDN6⁺ or CLDN6⁻ tumor cells.

Example 7: 3D Tumor Spheroid Killing Experiment with various Targets

Tumor spheroids expressing CD19, CLDN6 or CLDN18.2 were generated 24 hours prior to the co-culture with ModCAR T cells. On the day of co-culture, frequencies of NbALFA-CAR and variant 13 were adjusted to 50 %, resulting in a final CAR⁺ effector to target cell ratio of 5:1. Supernatant containing the matching adaptor (specific for CD19, CLDN6 or CLDN18.2) was added to the co-culture with a final dilution of 0.5 % at time point 0 hours. The ModCAR-T cells were repetitively challenged with antigen⁺ spheroids. A second CD19⁺ spheroid was administered at day 3, CLDN6⁺ spheroids at day 3 and day 10, and CLDN18.2⁺ spheroids at day 3, 7 and 10. Spheroid lysis was followed by Incucyte over 136 hours for the co-culture with CD19⁺ spheroids and over 336 hours for the co-culture with CLDN6⁺ and CLDN18.2⁺ spheroids. Analysis of spheroid lysis was performed at the end of the experiment using the Incucyte software. Data were normalized to the Mock (non-genetically modified) T cell control.

In co-culture with CD19⁺ tumor spheroids, both ModCARs (NbALFA-CAR/variant 13) induced similar spheroid lysis of two consecutive spheroids with a 0.5 % dilution of CD19-adaptor-containing supernatant (figure 10A). The first two CLDN6⁺ spheroids were also comparably lysed by both ModCARs in combination with 0.5 % dilution of CLDN6-adaptor supernatant, however, the third CLDN6⁺ spheroid, administered at 244 hours, was only eradicated by variant 13 (figure 10B). In co-culture with CLDN18.2⁺ spheroids and CLDN18.2-adaptor (0.5 % supernatant), variant 13 outperformed NbALFA-CAR at a third and fourth spheroid challenge (figure 10C).

Taken together, the CAR T cells expressing humanized ModCAR variant 13 presented a higher functionality in repetitive tumor spheroid lysis. These results confirm previous data (figure 5-9) demonstrating superior functionality of variant 13 to the wildtype ModCAR.

Example 8: In vivo Tumor Control Experiment

In vivo anti-tumoral activity of two different ModCAR T cells variants was evaluated in a xenograft mouse model. NSG mice were subcutaneously inoculated with CLDN6-positive OV90-SC12 tumor cells. The tumor was engrafted until day 38 before 1×10^6 NbALFA-CAR or variant 13 CAR T cells or corresponding numbers of non-modified T cells were adoptively transferred into mice bearing a mean size of over 277 mm^3 . Treatment with CLDN6-adaptor formulated into lipidnanoparticles (Riboadaptor LNP) in two different doses (30 and $5 \mu\text{g}$) or no LNP was performed weekly for five consecutive weeks. Both, tumor volume and CAR-T cell numbers in peripheral blood, were monitored over the study period. (figure 11).

Tumor outgrowth was observed for the treatment groups with T cells and for both ModCAR T cell variants in the absence of Riboadaptor LNP (figure 12A). In contrast, complete anti-tumor response was achieved with NbALFA-CAR and variant 13 CAR T cells when they were functionalized with one of the two Riboadaptor LNP doses (figure 12B, C). In direct comparison of the different ModCAR variants, variant 13 showed a slightly faster and superior anti-tumor kinetic than NbALFA-CAR and was observed for both LNP dose levels. Furthermore, ModCAR-T expansion in peripheral blood correlated to the anti-tumoral response in the mice (figure 13). At $5 \mu\text{g}$ dose of Riboadaptor LNP functionalization highest CAR T cell numbers were reached for both ModCAR variants with similar mean CAR T cell counts in the groups (figure 13 E). However, at $30 \mu\text{g}$ dose level, variant 13 showed noticeably lower CAR T cell numbers in the periphery than the NbALFA-CAR construct at the analyzed time points (figure C, F).

Taken together, ModCAR T cells showed potent anti-tumoral activity in vivo and correlating peripheral ModCAR T cell expansion. Importantly this was strictly dependent on Riboadaptor mediated functionalization. Finally, variant 13 demonstrated an overall superior tumor clearance compared to NbALFA-CAR, which is in line with previous in vitro results.

Claims

1. A system comprising:
 - (i) an immune effector cell genetically modified to express a chimeric antigen receptor (CAR), or a nucleic acid encoding a CAR useful for genetically modifying an immune effector cell to express the CAR; and
 - (ii) a compound comprising a binding moiety for the CAR and a binding moiety for an antigen on target cells, or a nucleic acid encoding the compound,wherein the CAR and the binding moiety for the CAR comprise a binding system comprising an ALFA-tag and a single-domain antibody binding to the ALFA-tag,
wherein the antibody comprises the CDR1 sequence as shown in SEQ ID NO: 94 or a sequence having 1 or 2 mutations relative to said sequence, the CDR2 sequence as shown in SEQ ID NO: 97 or a sequence having 1 or 2 mutations relative to said sequence, and the CDR3 sequence as shown in SEQ ID NO: 98 or a sequence having 1 or 2 mutations relative to said sequence;
wherein the antibody further comprises at positions corresponding to sequence position 47, 48, and 50 of SEQ ID NO: 1 the following amino acid residues: Arg at position 47, Arg at position 48, Met at position 50;
and wherein the antibody comprises a sequence having at least 90% sequence identity to any one of SEQ ID NOs: 118-125.
2. The system of claim 1, wherein the immune effector cell comprises a T cell.
3. The system of claim 1 or 2, wherein the immune effector cell expresses the CAR on its cell surface.
4. The system of any one of claims 1 to 3, wherein the CAR comprises a moiety binding to the binding moiety for the CAR.
5. The system of any one of claims 1 to 4, wherein the moiety of the CAR binding to the binding moiety for the CAR is comprised by an extracellular moiety of the CAR.

6. The system of any one of claims 1 to 5, wherein the moiety of the CAR binding to the binding moiety for the CAR comprises an ALFA-tag and the binding moiety for the CAR comprises the antibody.
7. The system of any one of claims 1 to 5, wherein the moiety of the CAR binding to the binding moiety for the CAR comprises the antibody and the binding moiety for the CAR comprises an ALFA-tag.
8. The system of any one of claims 1 to 7, wherein the antibody specifically binds to an epitope tag comprising the sequence SRLEEEELRRRLTE (SEQ ID NO: 2).
9. The system of any one of claims 1 to 8, wherein the antibody is humanized and/or wherein the antibody comprises a sequence that has at least about 70% V segment identity to the closest human germline sequence.
10. The system of any one of claims 1 to 9, wherein the antibody has a higher T_m as compared to the antibody of SEQ ID NO: 1.
11. The system of any one of claims 1 to 10, wherein the antibody comprises a His residue at the position corresponding to sequence position 99 of SEQ ID NO: 1, a Val residue at the position corresponding to sequence position 100 of SEQ ID NO: 1, and/or a Tyr residue at the position corresponding to sequence position 40 of SEQ ID NO: 1.
12. The system of any one of claims 1 to 11, wherein the antibody comprises FR1, FR2, FR3, and FR4 sequences, that have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 102, 108, and 109, respectively.
13. The system of any one of claims 1 to 12, wherein the antibody comprises a protein A binding site.

14. The system of any one of claims 1 to 13, wherein the antibody comprises a sequence having at least 95% sequence identity to any one of SEQ ID NOs: 118-125, or wherein the antibody comprises a sequence set forth in any one of SEQ ID NOs: 118-125.
15. The system of any one of claims 1 to 14, wherein the ALFA-tag comprises the amino acid sequence of X1-X2-L-E-X5-E-X7-R-R-R-L-X12-X13 (SEQ ID NO: 3), wherein X1 is G or S or P or T, X2 is R or G or P, X5 is E or Q, X7 is L or I, X12 is S or T or P or A or D or E, and wherein X13 is P or A or S or K or D or E or no amino acid.
16. The system of any one of claims 1 to 15, wherein the binding moiety for the CAR and the binding moiety for an antigen on target cells are connected through a linking moiety.
17. The system of claim 16, wherein the linking moiety comprises a polymer.
18. The system of claim 16 or 17, wherein the linking moiety comprises a non-peptidic moiety.
19. The system of any one of claims 16 to 18, wherein the linking moiety comprises a polymer selected from the group consisting of poly(ethylene glycol) (PEG), polysarcosine (pSar) (poly(N-methylglycine)), polyoxazoline (POX), polyoxazine (POZ), and poly-2-(2-(2-aminoethoxy)ethoxy)acetic acid (pAEEA), derivatives and combinations thereof.
20. The system of claim 16 or 17, wherein the linking moiety comprises a peptidic moiety.
21. The system of any one of claims 1 to 20, wherein the binding moiety for an antigen on target cells comprises a moiety selected from the group consisting of an antibody binding to the antigen, an antibody derivative binding to the antigen and a ligand of the antigen.
22. The system of claim 21, wherein the antibody derivative binding to the antigen comprises an antibody fragment.

23. The system of any one of claims 1 to 22, wherein the binding moiety for an antigen on target cells comprises a peptidic moiety.
24. The system of any one of claims 1 to 17 and 20 to 23, wherein the compound comprises a fusion polypeptide comprising the binding moiety for the CAR and the binding moiety for an antigen on target cells.
25. The system of claim 21, wherein the ligand of the antigen comprises DUPA or FAPI.
26. The system of any one of claims 1 to 25, wherein the target cells are diseased cells.
27. The system of any one of claims 1 to 26, wherein the target cells are cancer cells.
28. The system of any one of claims 1 to 27, wherein the antigen on target cells comprises a cell surface antigen.
29. The system of any one of claims 1 to 28, wherein the antigen on target cells comprises a tumor antigen.
30. The system of any one of claims 1 to 29, wherein the nucleic acid encoding a CAR comprises DNA.
31. The system of any one of claims 1 to 17, 20 to 24, and 26 to 30, wherein the nucleic acid encoding the compound comprises RNA.
32. The system of any one of claims 1 to 31, wherein binding of a complex comprising:
- (i) the immune effector cell genetically modified to express a chimeric antigen receptor (CAR); and
 - (ii) the compound comprising a binding moiety for the CAR and a binding moiety for an antigen on target cells
- to cells expressing the antigen results in killing of the cells expressing the antigen.

33. A method for treating a subject having a disease, disorder or condition characterized by cells expressing an antigen, comprising:

- (i) providing to the subject immune effector cells genetically modified to express a chimeric antigen receptor (CAR); and
- (ii) providing to the subject a compound comprising a binding moiety for the CAR and a binding moiety for the antigen,

wherein the CAR and the binding moiety for the CAR comprise a binding system comprising an ALFA-tag and a single-domain antibody binding to the ALFA-tag,

wherein the antibody comprises the CDR1 sequence as shown in SEQ ID NO: 94 or a sequence having 1 or 2 mutations relative to said sequence, the CDR2 sequence as shown in SEQ ID NO: 97 or a sequence having 1 or 2 mutations relative to said sequence, and the CDR3 sequence as shown in SEQ ID NO: 98 or a sequence having 1 or 2 mutations relative to said sequence; wherein the antibody further comprises at positions corresponding to sequence position 47, 48, and 50 of SEQ ID NO: 1 the following amino acid residues: Arg at position 47, Arg at position 48, Met at position 50;

and wherein the antibody comprises a sequence having at least 90% sequence identity to any one of SEQ ID NOs: 118-125.

34. The method of claim 33, wherein the immune effector cells comprise T cells.

35. The method of claim 33 or 34, wherein the immune effector cells express the CAR on their cell surface.

36. The method of any one of claims 33 to 35, wherein the CAR comprises a moiety binding to the binding moiety for the CAR.

37. The method of any one of claims 33 to 36, wherein the moiety of the CAR binding to the binding moiety for the CAR is comprised by an extracellular moiety of the CAR.

38. The method of any one of claims 33 to 37, wherein the moiety of the CAR binding to the binding moiety for the CAR comprises an ALFA-tag and the binding moiety for the CAR comprises the antibody.

39. The method of any one of claims 33 to 37, wherein the moiety of the CAR binding to the binding moiety for the CAR comprises the antibody and the binding moiety for the CAR comprises an ALFA-tag.

40. The method of any one of claims 33 to 39, wherein the antibody specifically binds to an epitope tag comprising the sequence SRLEELRRRLTE (SEQ ID NO: 2).

41. The method of any one of claims 33 to 40, wherein the antibody is humanized and/or wherein the antibody comprises a sequence that has at least about 70% V segment identity to the closest human germline sequence.

42. The method of any one of claims 33 to 41, wherein the antibody has a higher T_m as compared to the antibody of SEQ ID NO: 1.

43. The method of any one of claims 33 to 42, wherein the antibody comprises a His residue at the position corresponding to sequence position 99 of SEQ ID NO: 1, a Val residue at the position corresponding to sequence position 100 of SEQ ID NO: 1, and/or a Tyr residue at the position corresponding to sequence position 40 of SEQ ID NO: 1.

44. The method of any one of claims 33 to 43, wherein the antibody comprises FR1, FR2, FR3, and FR4 sequences, that have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 102, 108, and 109, respectively.

45. The method of any one of claims 33 to 44, wherein the antibody comprises a protein A binding site.

46. The method of any one of claims 33 to 45, wherein the antibody comprises a sequence having at least 95% sequence identity to any one of SEQ ID NOs: 118-125, or wherein the antibody comprises a sequence set forth in any one of SEQ ID NOs: 118-125.

47. The method of any one of claims 33 to 46, wherein the ALFA-tag comprises the amino acid sequence of X1-X2-L-E-X5-E-X7-R-R-R-L-X12-X13 (SEQ ID NO: 3), wherein X1 is G or S or P or T, X2 is R or G or P, X5 is E or Q, X7 is L or I, X12 is S or T or P or A or D or E, and wherein X13 is P or A or S or K or D or E or no amino acid.
48. The method of any one of claims 33 to 47, wherein the binding moiety for the CAR and the binding moiety for an antigen on target cells are connected through a linking moiety.
49. The method of claim 48, wherein the linking moiety comprises a polymer.
50. The method of claim 48 or 49, wherein the linking moiety comprises a non-peptidic moiety.
51. The method of any one of claims 48 to 50, wherein the linking moiety comprises a polymer selected from the group consisting of poly(ethylene glycol) (PEG), polysarcosine (pSar) (poly(N-methylglycine)), polyoxazoline (POX), polyoxazine (POZ), and poly-2-(2-(2-aminoethoxy)ethoxy)acetic acid (pAEEA), derivatives and combinations thereof.
52. The method of claim 48 or 49, wherein the linking moiety comprises a peptidic moiety.
53. The method of any one of claims 33 to 52, wherein the binding moiety for the antigen comprises a moiety selected from the group consisting of an antibody binding to the antigen, an antibody derivative binding to the antigen and a ligand of the antigen.
54. The method of claim 53, wherein the antibody derivative binding to the antigen comprises an antibody fragment.
55. The method of any one of claims 33 to 54, wherein the binding moiety for the antigen comprises a peptidic moiety.

56. The method of any one of claims 33 to 49 and 52 to 55, wherein the compound comprises a fusion polypeptide comprising the binding moiety for the CAR and the binding moiety for the antigen.

57. The method of claim 53, wherein the ligand of the antigen comprises DUPA or FAPI.

58. The method of any one of claims 33 to 57, wherein the cells expressing an antigen are diseased cells.

59. The method of any one of claims 33 to 58, wherein the cells expressing an antigen are cancer cells.

60. The method of any one of claims 33 to 59, wherein the antigen comprises a cell surface antigen.

61. The method of any one of claims 33 to 60, wherein the antigen comprises a tumor antigen.

62. The method of any one of claims 33 to 61, wherein binding of a complex comprising:
(i) the immune effector cells genetically modified to express a chimeric antigen receptor (CAR); and
(ii) the compound comprising a binding moiety for the CAR and a binding moiety for an antigen
to cells expressing the antigen results in killing of the cells expressing the antigen.

63. The method of any one of claims 33 to 62, wherein the method comprises administering the immune effector cells genetically modified to express a CAR to the subject.

64. The method of any one of claims 33 to 62, wherein the method comprises generating the immune effector cells genetically modified to express a CAR in the subject.

65. The method of any one of claims 33 to 62 and 64, wherein the method comprises administering nucleic acid encoding the CAR to the subject.

66. The method of claim 65, wherein the nucleic acid encoding the CAR comprises DNA.

67. The method of any one of claims 33 to 66, wherein the method comprises administering the compound to the subject.

68. The method of any one of claims 33 to 49, 52 to 56, and 58 to 66, wherein the method comprises generating the compound in the subject.

69. The method of any one of claims 33 to 49, 52 to 56, 58 to 66, and 68, wherein the method comprises administering nucleic acid encoding the compound to the subject.

70. The method of claim 69, wherein the nucleic acid encoding the compound comprises RNA.

71. The method of claim 69 or 70, wherein the method comprises allowing expression of the compound by cells in the subject such that the compound becomes associated with cells expressing the antigen and the binding moiety for the CAR is available for binding by the immune effector cells.

72. The method of claim 71, wherein the cells expressing the compound are transfected with the nucleic acid encoding the compound.

73. The method of claim 71 or 72, wherein the cells expressing the compound secrete the compound.

74. The method of any one of claims 71 to 73, wherein the cells expressing the compound express the compound such that it is released into the bloodstream.

75. The method of any one of claims 65 to 74, wherein the nucleic acid encoding the CAR and/or the nucleic acid encoding the compound is administered as particulate formulation such as formulated as lipid nanoparticles.

76. The method of any one of claims 33 to 75, wherein the disease, disorder or condition is cancer.

77. A chimeric antigen receptor (CAR) comprising a single-domain antibody binding to an ALFA-tag,
wherein the antibody comprises the CDR1 sequence as shown in SEQ ID NO: 94 or a sequence having 1 or 2 mutations relative to said sequence, the CDR2 sequence as shown in SEQ ID NO: 97 or a sequence having 1 or 2 mutations relative to said sequence, and the CDR3 sequence as shown in SEQ ID NO: 98 or a sequence having 1 or 2 mutations relative to said sequence;
wherein the antibody further comprises at positions corresponding to sequence position 47, 48, and 50 of SEQ ID NO: 1 the following amino acid residues: Arg at position 47, Arg at position 48, Met at position 50;
and wherein the antibody comprises a sequence having at least 90% sequence identity to any one of SEQ ID NOs: 118-125.

78. The CAR of claim 77, wherein the antibody specifically binds to an epitope tag comprising the sequence SRLEELRRRLTE (SEQ ID NO: 2).

79. The CAR of claim 77 or 78, wherein the antibody is humanized and/or wherein the antibody comprises a sequence that has at least about 70% V segment identity to the closest human germline sequence.

80. The CAR of any one of claims 77 to 79, wherein the antibody has a higher T_m as compared to the antibody of SEQ ID NO: 1.

81. The CAR of any one of claims 77 to 80, wherein the antibody comprises a His residue at the position corresponding to sequence position 99 of SEQ ID NO: 1, a Val residue at the

position corresponding to sequence position 100 of SEQ ID NO: 1, and/or a Tyr residue at the position corresponding to sequence position 40 of SEQ ID NO: 1.

82. The CAR of any one of claims 77 to 81, wherein the antibody comprises FR1, FR2, FR3, and FR4 sequences, that have at least 80%, at least 85%, at least 90%, or at least 95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 102, 108, and 109, respectively.

83. The CAR of any one of claims 77 to 82, wherein the antibody comprises a protein A binding site.

84. The CAR of any one of claims 77 to 83, wherein the antibody comprises a sequence having at least 95% sequence identity to any one of SEQ ID NOs: 118-125, or wherein the antibody comprises a sequence set forth in any one of SEQ ID NOs: 118-125.

85. The CAR of any one of claims 77 to 84, wherein the antibody is comprised by an extracellular domain of the CAR.

86. A nucleic acid encoding the CAR of any one of claims 77 to 85.

87. The nucleic acid of claim 86 which is DNA.

88. A host cell comprising the nucleic acid of claim 86 or 87 and/or expressing the CAR of any one of claims 77 to 85.

89. The host cell of claim 88 which comprises an immune effector cell.

90. The host cell of claim 89, wherein the immune effector cell comprises a T cell.

91. The host cell of any one of claims 88 to 90, which is genetically modified to comprise the nucleic acid and/or expresses the CAR.

92. The host cell of any one of claims 88 to 91, which expresses the CAR on its cell surface.
93. A compound comprising a single-domain antibody binding to an ALFA-tag and a binding moiety for an antigen on target cells,
wherein the antibody comprises the CDR1 sequence as shown in SEQ ID NO: 94 or a sequence having 1 or 2 mutations relative to said sequence, the CDR2 sequence as shown in SEQ ID NO: 97 or a sequence having 1 or 2 mutations relative to said sequence, and the CDR3 sequence as shown in SEQ ID NO: 98 or a sequence having 1 or 2 mutations relative to said sequence;
wherein the antibody further comprises at positions corresponding to sequence position 47, 48, and 50 of SEQ ID NO: 1 the following amino acid residues: Arg at position 47, Arg at position 48, Met at position 50;
and wherein the antibody comprises a sequence having at least 90% sequence identity to any one of SEQ ID NOs: 118-125.
94. The compound of claim 93, wherein the antibody specifically binds to an epitope tag comprising the sequence SRLEELRRRLTE (SEQ ID NO: 2).
95. The compound of claim 93 or 94, wherein the antibody is humanized and/or wherein the antibody comprises a sequence that has at least about 70% V segment identity to the closest human germline sequence.
96. The compound of any one of claims 93 to 95, wherein the antibody has a higher T_m as compared to the antibody of SEQ ID NO: 1.
97. The compound of any one of claims 93 to 96, wherein the antibody comprises a His residue at the position corresponding to sequence position 99 of SEQ ID NO: 1, a Val residue at the position corresponding to sequence position 100 of SEQ ID NO: 1, and/or a Tyr residue at the position corresponding to sequence position 40 of SEQ ID NO: 1.
98. The compound of any one of claims 93 to 97, wherein the antibody comprises FR1, FR2, FR3, and FR4 sequences, that have at least 80%, at least 85%, at least 90%, or at least

95% sequence identity or are identical to the sequences set forth in SEQ ID NOs: 99, 102, 108, and 109, respectively.

99. The compound of any one of claims 93 to 98, wherein the antibody comprises a protein A binding site.

100. The compound of any one of claims 93 to 99, wherein the antibody comprises a sequence having at least 95% sequence identity to any one of SEQ ID NOs: 118-125, or wherein the antibody comprises a sequence set forth in any one of SEQ ID NOs: 118-125.

101. The compound of any one of claims 93 to 100, wherein the antibody and the binding moiety for an antigen on target cells are connected through a linking moiety.

102. The compound of claim 101, wherein the linking moiety comprises a polymer.

103. The compound of claim 101 or 102, wherein the linking moiety comprises a non-peptidic moiety.

104. The compound of any one of claims 101 to 103, wherein the linking moiety comprises a polymer selected from the group consisting of poly(ethylene glycol) (PEG), polysarcosine (pSar) (poly(N-methylglycine)), polyoxazoline (POX), polyoxazine (POZ), and poly-2-(2-(2-aminoethoxy)ethoxy)acetic acid (pAEEA), derivatives and combinations thereof.

105. The compound of claim 101 or 102, wherein the linking moiety comprises a peptidic moiety.

106. The compound of any one of claims 93 to 105, wherein the binding moiety for an antigen on target cells comprises a moiety selected from the group consisting of an antibody binding to the antigen, an antibody derivative binding to the antigen and a ligand of the antigen.

107. The compound of claim 106, wherein the antibody derivative binding to the antigen comprises an antibody fragment.

108. The compound of any one of claims 93 to 107, wherein the binding moiety for an antigen on target cells comprises a peptidic moiety.

109. The compound of any one of claims 93 to 102 and 105 to 108, wherein the compound comprises a fusion polypeptide comprising the antibody and the binding moiety for an antigen on target cells.

110. The compound of claim 106, wherein the ligand of the antigen comprises DUPA or FAPI.

111. The compound of any one of claims 93 to 110, wherein the target cells are diseased cells.

112. The compound of any one of claims 93 to 111, wherein the target cells are cancer cells.

113. The compound of any one of claims 93 to 112, wherein the antigen on target cells comprises a cell surface antigen.

114. The compound of any one of claims 93 to 113, wherein the antigen on target cells comprises a tumor antigen.

115. A nucleic acid encoding the compound of any one of claims 93 to 102, 105 to 109 and 111 to 114.

116. The nucleic acid of claim 115 which is RNA.

117. A host cell comprising the nucleic acid of claim 115 or 116 and/or expressing the compound of any one of claims 93 to 102, 105 to 109 and 111 to 114.