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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
3 November 2016 (03.11.2016)

WIPO | PCT

(10) International Publication Number
WO 2016/174409 A1

(51) International Patent Classification:
A61K 39/00 (2006.01) *C12N 5/0783* (2010.01)
C07K 16/28 (2006.01)

(21) International Application Number:
PCT/GB2016/051165

(22) International Filing Date:
26 April 2016 (26.04.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
1507108.7 27 April 2015 (27.04.2015) GB

(71) Applicant: **UCL BUSINESS PLC** [GB/GB]; The Network Building, 97 Tottenham Court Road, London W1T 4TP (GB).

(72) Inventors: **PULÉ, Martin**; c/o Autolus Ltd, Forest House, Depot Road, London W12 7RP (GB). **CORDOBA, Shaun**; c/o Autolus Ltd, Forest House, Depot Road, London W12 7RP (GB).

(74) Agent: **WILLIAMS, Aylsa**; D Young & Co LLP, 120 Holborn, London EC1N 2DY (GB).

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

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

Published:

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

— *with sequence listing part of description (Rule 5.2(a))*

WO 2016/174409 A1

(54) Title: NUCLEIC ACID CONSTRUCT

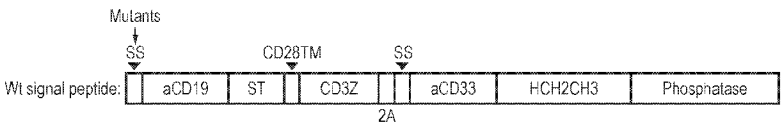

FIG. 3

(57) Abstract: The present invention provides a nucleic acid construct comprising the following structure: A – X – B in which A is nucleic acid sequence encoding a first polypeptide which comprises a first signal peptide; B is nucleic acid sequence encoding a second polypeptide which comprises a second signal peptide and X is a nucleic acid sequence which encodes a cleavage site, wherein the first signal peptide or the second signal peptide comprises one or more mutation(s) such that it has fewer hydrophobic amino acids.

NUCLEIC ACID CONSTRUCT

FIELD OF THE INVENTION

The present invention relates to constructs and approaches for modulating the relative expression of polypeptides co-expressed from a single vector. In particular, the invention relates to modulating the expression of a transmembrane protein co-expressed from a single vector with a second polypeptide.

BACKGROUND TO THE INVENTION

It is often desirable to express different proteins from the same vector since multiple transduction of the same cell is difficult, expensive and unpredictable. Different methods have therefore been developed to allow co-expression of two proteins from a single vector (see Figure 1).

Initial attempts used two different promoters within the same cassette. This results in two separate transcripts each of which code for a separate protein. This is a difficult approach for a number of reasons. A key problem is "promoter interference" whereby one promoter dominates and causes silencing of the second promoter. In addition, different promoters work differently in different cellular contexts and this makes consistent "tuning" of the relative expression of each transgene difficult to achieve.

An alternative approach is to use an Internal Ribosome Entry sequence (IRES). Here, a single transcript is generated. The IRES sequence in the transcript is placed between the open reading frames for the two transgenes and mimics an mRNA cap structure. Hence, the ribosome either initiates translation at the 5' cap or the IRES resulting in expression of two separate proteins. A key limitation with this approach is the inability to control relative expression. The 3' transcript is typically expressed less than the 5' one, but the ratio of expression is difficult to predict and tune.

A further approach has been provided following characterization of the role of foot-and-mouth-disease virus (FMDV) 2A peptide in allowing FMDV (and related viruses) to express multiple proteins from a single open reading frame (ORF) (Donnelly *et al; J. Gen. Virol.;* 82**,** 1027–1041 (2001)). The 2A peptide (and homologs) cleaves at very high efficiency immediately after translation of the ORF, enabling the expression of multiple peptides from a

single ORF. A problem with the use of the 2A peptide to cleave between different peptides in the same ORF is that expression is limited to a 1:1 ratio.

Thus there is a need for alternative methods for expressing more than one protein from a single vector which are not associated with the disadvantages described above.

SUMMARY OF ASPECTS OF THE INVENTION

The present invention is based on the determination that, when two signal-peptide containing proteins are co-expressed as a polyprotein which is cleaved after translation, it is possible to modify the relative "strength" of the signal peptides, and thus control the relative expression of the two proteins. This need not be limited to a pair of transgenes, but may be used to allow control of the relative expression of multiple proteins initially translated as a polyprotein.

As used herein, 'polyprotein' refers to a polypeptide sequence translated from a single nucleic acid construct as a single entity, but which comprises polypeptide sequences which are subsequently separated and which function as discrete entities (e.g. separate proteins).

Thus in a first aspect the present invention provides a nucleic acid construct comprising the following structure:

A – X – B

in which

A is nucleic acid sequence encoding a first polypeptide which comprises a first signal peptide;

B is nucleic acid sequence encoding a second polypeptide which comprises a second signal peptide and

X is a nucleic acid sequence which encodes a cleavage site,

wherein the first signal peptide or the second signal peptide comprises one or more mutation(s) such that it has fewer hydrophobic amino acids.

The mutated signal peptide may have fewer hydrophobic amino acids than the "wild-type" signal peptide sequence from which it is derived.

The mutated signal peptide may have fewer hydrophobic amino acids than the signal peptide of the other polypeptide.

The hydrophobic amino acid(s) may be selected from the group: Alanine (A); Valine (V); Isoleucine (I); Leucine (L); Methionine (M); Phenylalanine (P); Tyrosine (Y); Tryptophan (W).

The first signal peptide and the second signal peptide may be derivable from the same sequence, but one signal peptide may comprise one or more amino acid deletions or substitutions to remove or replace one or more hydrophobic amino acids compared to the other signal peptide.

As explained in more detail below, signal sequences have a tripartite structure, consisting of a hydrophobic core region (h-region) flanked by an n- and c-region. The first signal peptide and the second signal peptide may have identical n- and c- regions, but may differ in the h-region: the h-region of one signal peptide having more hydrophobic amino acids that the other signal peptide.

The hydrophobic amino acid(s) deleted or replaced from the signal peptide may be selected from the group: Alanine (A); Valine (V); Isoleucine (I); Leucine (L); Methionine (M); Phenylalanine (P); Tyrosine (Y); Tryptophan (W).

The hydrophobic amino acid(s) deleted or replaced from the signal peptide may be selected from the group: Valine (V); Isoleucine (I); Leucine (L); and Tryptophan (W).

The signal peptide of one polypeptide may comprise up to five more hydrophobic amino acids than the other signal peptide. The altered signal peptide may have up to 10%, up to 20%, up to 30%, up to 40% or up to 50% of its hydrophobic amino acids replaced or removed.

The first and second polypeptides may be and signal peptide-containing polypeptides, such as secreted, transmembrane or organelle proteins. The present invention enables the relative expression of the proteins to be controlled i.e. the relative amounts of the proteins in the relevant compartment. For secreted proteins, the present invention enables the relative amounts of the proteins produced by a cell to be controlled. For transmembrane proteins, the present invention enables the relative cell surface expression of the two (or more) proteins to be controlled. For organelle proteins, the present invention enables the relative expression of the proteins to be controlled within, or on the membrane of, the organelle in question.

The first and second polypeptides may both be transmembrane proteins, such as T cell receptors or chimeric antigen receptors (CARs).

Where the first and second proteins are transmembrane proteins, the difference between the first and the second signal peptides may be such that, when the nucleic acid construct is expressed in a cell, there is differential relative expression of the first and second polypeptides at the cell surface.

X may be a nucleic acid sequence encoding a self-cleaving peptide, a furin cleavage site or a Tobacco Etch Virus cleavage site.

X may, for example, encode a 2A self-cleaving peptide from an aphtho- or a cardiovirus or a 2A-like peptide.

In a second aspect, the present invention provides a vector comprising a nucleic acid construct according to any preceding claim.

The vector may, for example, be a retroviral vector or a lentiviral vector.

In a third aspect, there is provided a cell comprising a nucleic acid construct according to the first aspect of the invention or a vector according to the second aspect of the invention.

In this third aspect, there is also provided a cell which comprises two chimeric antigen receptors having different signal peptides.

The two signal peptides may have a different number of hydrophobic amino acids.

In this aspect, the chimeric antigen receptor having a signal peptide with the greater number of hydrophobic amino acids may be more highly expressed at the cell surface than the chimeric antigen receptor having a signal peptide with the smaller number of hydrophobic amino acids.

There is also provided a cell which comprises two chimeric antigen receptors having different signal peptides, wherein the signal peptide of one chimeric antigen comprises one or more mutations such that it has fewer hydrophobic amino acids.

The chimeric antigen receptor having a mutated signal peptide may be expressed at a lower level at the cell surface than the chimeric antigen receptor having an unmutated signal peptide.

The cell may be a cytolytic immune cell, such as a T cell or a natural killer (NK) cell.

In a fourth aspect, there is provided a method for making a cell according to the fourth aspect of the invention which comprises the step of introducing a nucleic acid construct according to the first aspect of the invention or a vector according to the second aspect of the invention into a cell.

In a fifth aspect, there is provided a method for modulating the relative cell surface expression of a first signal peptide-containing protein expressed from a single nucleic acid construct with a second signal peptide-containing protein which comprises the step of mutating the nucleic acid sequence which encodes the signal peptide of one protein in order to remove or replace one or more hydrophobic amino acids in comparison with the signal peptide of the other protein.

The removal or replacement of hydrophobic amino acids from the signal peptide of a transmembrane protein reduces the amount of the transmembrane protein expressed on the cell surface, compared to a transmembrane protein having an unmodified signal peptide. As such, the relative expression level of a transmembrane protein derived from a polyprotein including a second polypeptide can be modulated. Where the transmembrane protein is only active at the cell surface (or predominantly active at the cell surface), reducing the relative cell surface expression of the protein also reduces its relative activity.

This invention can be extended to modulate the relative expression of three or more proteins expressed as a concatenated polypeptide, separated by cleavage sites and relative surface expression dictated by signal peptides of differing activity.

DESCRIPTION OF THE FIGURES

Figure 1: Methods utilised to express different proteins from the same vector
(a) Two different promoters within the same cassette result in two different transcripts which each give rise to separate proteins. (b) Use of an Internal Ribosome Entry sequence (IRES) leads to a single transcript which is translated into two separate proteins. (c) Use of the

FMDV 2A peptide results in a single transcript, and a single polyprotein which rapidly cleaves into two separate proteins.

Figure 2: Schematic diagram illustrating the function of signal sequences in protein targeting

Figure 3: Schematic diagram of nucleic acid construct encoding two CARs

Figure 4: Verifying the function of a substituted signal sequence.
PCT/GB2014/053452 describes vector system encoding two chimeric antigen receptors (CARs), one against CD19 and one against CD33. The signal peptide used for the CARs in that study was the signal peptide from the human CD8a signal sequence. For the purposes of this study, this was substituted with the signal peptide from the murine Ig kappa chain V-III region, which has the sequence: METD**TLILWVLLLLV**PGSTG (hydrophobic residues hightlited in bold). In order to establish that the murine Ig kappa chain V-III signal sequence functioned as well as the signal sequence from human CD8a, a comparative study was performed. For both signal sequences, functional expression of the anti-CD33 CAR and the anti-CD19 CAR was observed.

Figure 5: Testing the effect of one amino acid deletion in the murine Ig kappa chain V-III. Mutant 1 kappa chain was created with the following deletion (shown in grey) in the h-region METD**T**L**ILWVLLLLV**PGSTG and the relative expression on the anti-CD33 CAR and the anti-CD19 CAR was observed.

Figure 6: Testing the effect of two amino acid deletions in the murine Ig kappa chain V-III. Mutant 2 kappa chain was created with the following deletions (shown in grey) in the h-region METD**T**L**ILWVLLL**L**V**PGSTG and the relative expression on the anti-CD33 CAR and the anti-CD19 CAR was observed.

Figure 7: Testing the effect of three amino acid deletions in the murine Ig kappa chain V-III. Mutant 2 kappa chain was created with the following deletions (shown in grey) in the h-region METD**T**L**ILWVLLL**LV**PGSTG and the relative expression on the anti-CD33 CAR and the anti-CD19 CAR was observed.

Figure 8: Testing the effect of five amino acid deletions in the murine Ig kappa chain V-III. Mutant 2 kappa chain was created with the following deletions (shown in grey) in the h-region METD**T**L**I**L**W**V**LLL**LV**PGSTG and the relative expression on the anti-CD33 CAR and the anti-CD19 CAR was observed.

DETAILED DESCRIPTION

The present invention provides a nucleic acid construct comprising the following structure:

A – X – B

in which

A is nucleic acid sequence encoding a first polypeptide which comprises a first signal peptide;

B is nucleic acid sequence encoding a second polypeptide which comprises a second signal peptide and

X is a nucleic acid sequence which encodes a cleavage site,

wherein the first signal peptide or the second signal peptide comprises one or more mutation(s) such that it has fewer hydrophobic amino acids.

NUCLEIC ACID CONSTRUCT

POLYPEPTIDE

The polypeptides made by the nucleic acid construct of the invention are signal peptide-containing polypeptides, such as secreted, transmembrane or organelle proteins.

SIGNAL PEPTIDE

The polypeptides A and B (and optionally others, C, D etc) encoded by the nucleic acid construct of the invention each comprise may a signal sequence so that when the polypeptide is expressed inside a cell the nascent protein is directed to the endoplasmic reticulum (ER) (see Figure 2a).

The term “signal peptide” is synonymous with “signal sequence”.

A signal peptide is a short peptide, commonly 5-30 amino acids long, present at the N-terminus of the majority of newly synthesized proteins that are destined towards the secretory pathway. These proteins include those that reside either inside certain organelles (for example, the endoplasmic reticulum, golgi or endosomes), are secreted from the cell, and transmembrane proteins.

Signal peptides commonly contain a core sequence which is a long stretch of hydrophobic amino acids that has a tendency to form a single alpha-helix. The signal peptide may begin with a short positively charged stretch of amino acids, which helps to enforce proper topology of the polypeptide during translocation. At the end of the signal peptide there is typically a stretch of amino acids that is recognized and cleaved by signal peptidase. Signal peptidase may cleave either during or after completion of translocation to generate a free signal peptide and a mature protein. The free signal peptides are then digested by specific proteases.

The signal peptide is commonly positioned at the amino terminus of the molecule, although some carboxy-terminal signal peptides are known.

As mentioned above, signal sequences have a tripartite structure, consisting of a hydrophobic core region (h-region) flanked by an n- and c-region. The latter contains the signal peptidase (SPase) consensus cleavage site. Usually, signal sequences are cleaved off co-translationally, the resulting cleaved signal sequences are termed signal peptides.

In the signal peptide from the murine Ig kappa chain V-III region, which has the sequence: METD**TLILWVLLLLV**PGSTG: the n-region has the sequence METD; the h-region (shown in bold) has the sequence TLILWVLLLLV; and the c-region has the sequence PGSTG.

In the nucleic acid construct of the present invention the signal sequence of the two (or more) polypeptides differ in their h-regions. One polypeptide (which has higher relative expression) has a greater number of hydrophobic amino acids in the h-region that the other polypeptide (which has lower relative expression). The signal peptide of the polypeptide with lower relative expression may comprise one or more amino acid mutations, such as substitutions or deletions, of hydrophobic amino acids in the h-region than the signal peptide of the polypeptide with lower relative expression.

The first signal peptide and the second signal peptide may have substantially the same n- and c- regions, but differ in the h-region as explained above. “Substantially the same” indicates that the n- and c- regions may be identical between the first and second signal peptide or may differ by one, two or three amino acids in the n- or c-chain, without affecting the function of the signal peptide.

The hydrophobic amino acids in the core may, for example be: Alanine (A); Valine (V); Isoleucine (I); Leucine (L); Methionine (M); Phenylalanine (P); Tyrosine (Y); or Tryptophan (W).

The hydrophobic acids mutated in order to alter signal peptide efficiency may be any from the above list, in particular: Valine (V); Isoleucine (I); Leucine (L); and Tryptophan (W).

Of the residues in the h-region, one signal peptide (for example, the altered signal peptide) may comprise at least 10%, 20%, 30%, 40% or 50% fewer hydrophobic amino acids than the other signal peptide (for example, the unaltered signal peptide).

Where the h-region comprises 5-15 amino acids, one signal peptide may comprise 1, 2, 3, 4 or 5 more hydrophobic amino acids than the other signal peptide.

The altered signal peptide may comprise 1, 2, 3, 4 or 5 amino acid deletions or substitutions of hydrophobic amino acids. Hydrophobic amino acids may be replaced with non-hydrophobic amino acids, such as hydrophilic or neutral amino acids.

Signal sequences can be detected or predicted using software techniques (see for example, http://www.predisi.de/).

A very large number of signal sequences are known, and are available in databases. For example, http://www.signalpeptide.de lists 2109 confirmed mammalian signal peptides in its database.

Table 1 provides a list of signal sequences purely for illustrative purposes. The hydrophobic core is highlighted in bold. This includes examples of amino acids which may be substituted or removed for the purposes of the present invention.

Table 1

Accession Number	Entry Name	Protein Name	Length	Signal Sequence (hydrophobic core)
P01730	CD4_HUMAN	T-cell surface glycoprotein CD4	25	MNRGVPFRH**LLLVLQLALL**PAATQG
P08575	CD45_HUMAN	Leukocyte common antigen	23	MYLWLK**LLAFGFAFL**DTEVFVTG
P01732	CD8A_HUMAN	T-cell surface glycoprotein CD8 alpha chain	21	MALP**VTALLLPLALLL**HAARP
P10966	CD8B_HUMAN	T-cell surface glycoprotein CD8 beta chain	21	MRPRLW**LLLAAQLTVL**HGNSV
P06729	CD2_HUMAN	T-cell surface antigen CD2	24	MSFPCK**FVASFLLIF**NVSSKGAVS
P06127	CD5_HUMAN	T-cell surface glycoprotein CD5	24	MPMGSLQP**LATLYLLGML**VASC**L**G

P09564	CD7_HUMAN	T-cell antigen CD7	25	MAGPPR**LLLLPLLLALA**RGLPGALA
P17643	TYRP1_HUMAN	5,6-dihydroxyindole-2-carboxylic acid oxidase	24	MSAPK**LLSLGCIFFPLLLF**QQARA
P00709	LALBA_HUMAN	Alpha-lactalbumin	19	M**RFFVPLFLVGILF**PAILA
P16278	BGAL_HUMAN	Beta-galactosidase	23	MPG**FLVRILPLLLVLLLLG**PTRG
P31358	CD52_HUMAN	CAMPATH-1 antigen	24	M**KRFLFLLLTISLLVMV**QIQTGLS
Q6YHK3	CD109_HUMAN	CD109 antigen	21	MQGPP**LLTAAHLLCVCTAAL**A
P01024	CO3_HUMAN	Complement C3	22	MGPTSGPS**LLLLLLLTHLPLAL**G
P10144	GRAB_HUMAN	Granzyme B	18	MQP**ILLLLAFLLL**PRADA
P04434	KV310_HUMAN	Ig kappa chain V-III region VH	20	MEAPAQ**LLFLLLLWLP**DTTR
P06312	KV401_HUMAN	Ig kappa chain V-IV region	20	MVLQTQ**VFISLLLWIS**GAYG
P06319	LV605_HUMAN	Ig lambda chain V-VI region EB4	19	MAWA**PLLLTLLA**HCTDCWA
P31785	IL2RG_HUMAN	Cytokine receptor common gamma chain	22	MLKPSLPFTS**LLFLQLPLLGV**G
Q8N4F0	BPIL1_HUMAN	Bactericidal/permeability-increasing protein-like 1	20	MAWASR**LGLLLALLLPVV**GA
P55899	FCGRN_HUMAN	IgG receptor FcRn large subunit p51	23	MGVPRPQPW**ALGLLLFLL**PGSLG

The mutated signal peptide comprises one or more mutation(s) such that it has fewer hydrophobic amino acids than the wild-type signal peptide from which it is derived. The term "wild type" means the sequence of the signal peptide which occurs in the natural protein from which it is derived. For example, the signal peptide described in the examples is the signal peptide from the murine Ig kappa chain V-III region, which has the wild-type sequence: METDTLILWVLLLLVPGSTG.

The term "wild-type" also includes signal peptides derived from a naturally occurring protein which comprise one or more amino acid mutations in the n- or c- region. For example it is common to modify a natural signal peptide with a conserved amino acid substitution on the N-terminus to introduce a restriction site. Such modified signal peptide sequences (which do not comprise any mutations in the h-region) are considered "wild-type" for the purposes of the present invention.

The present invention also relates to synthetic signal peptide sequences, which cannot be defined with reference to a wild-type sequence. In this embodiment, the signal peptide of the one polypeptide comprises fewer hydrophobic amino acids than the signal sequence of the other polypeptide. The two signal sequences may be derived from the same synthetic signal peptide sequence, but differ in the number of hydrophobic amino acids in the core region.

TRANSMEMBRANE PROTEIN

The present invention enables modulation of the relative expression of a transmembrane surface protein. The transmembrane surface protein is a protein which is expressed at the cell surface. When expressed at the cell surface at least one domain of the transmembrane protein is exoplasmic (i.e. on the exterior of the cell).

The transmembrane protein may be a single-pass transmembrane protein, i.e. it may comprise a single transmembrane domain or it may comprise multiple transmembrane domains.

Transmembrane proteins may be classified by topology i.e. with reference to the position of the N- and C-terminal domains. Types I, II, and III transmembrane proteins are single-pass molecules, while type IV trans-membrane proteins are multiple-pass molecules. Type I transmembrane proteins are anchored to the lipid membrane with a stop-transfer anchor sequence and have their N-terminal domains targeted to the ER lumen during synthesis (and the extracellular space, when the mature form is located on the plasma membrane). Type II and III are anchored with a signal-anchor sequence, with type II being targeted to the ER lumen with its C-terminal domain, while type III have their N-terminal domains targeted to the ER lumen. Type IV is subdivided into IV-A, with their N-terminal domains targeted to the cytosol and IV-B, with an N-terminal domain targeted to the lumen.

The transmembrane protein(s) made by the nucleic acid construct of the present invention may be any of the types I-IV.

The transmembrane domain may be any protein structure which is thermodynamically stable in a membrane. This is typically an alpha helix comprising of several hydrophobic residues. The transmembrane domain of any transmembrane protein can be used to supply the transmembrane portion. The presence and span of a transmembrane domain of a protein can be determined by those skilled in the art using the TMHMM algorithm (http://www.cbs.dtu.dk/services/TMHMM-2.0/). Further, given that the transmembrane domain of a protein is a relatively simple structure, i.e a polypeptide sequence predicted to form a hydrophobic alpha helix of sufficient length to span the membrane, an artificially designed TM domain may also be used (US7052906 B1 describes synthetic transmembrane components).

The transmembrane domain may be derived from CD28, which gives good stability.

The structure and processing of Type I transmembrane proteins is well known in the art. Such proteins typically comprise an extracellular domain, a transmembrane domain and an intracellular endodomain and are single-pass molecules with a single α-helix passing through the cell membrane.

Type I transmembrane proteins typically have a signal peptide which is quickly recognized by the endoplasmic reticulum (ER) and the protein in translation is therefore quickly re-directed into the ER. A hydrophobic helix locks then anchors the protein in the membrane of the ER.

As mentioned above, Type I transmembrane proteins are anchored to the lipid membrane with a stop-transfer anchor sequence. The stop-transfer sequence halts the further translocation of the polypeptide and acts as a transmembrane anchor.

As used herein, the term Type I transmembrane protein encompasses any protein which comprises a Type I transmembrane domain and a stop-transfer anchor sequence and is, in the absence of an exogenous intracellular retention signal, targeted for expression on the cell surface.

Various type 1 transmembrane proteins which are suitable for use in the present invention are known in the art. Such proteins include, but are not limited to inhibitory receptors, stimulatory receptors, cytokine receptors and G-Proteins.

The transmembrane protein(s) may be a T-cell receptor α or β chain.

The transmembrane protein(s) may be a Chimeric Antigen Receptor (CAR).

CARs are proteins which graft an antigen binding domain to the effector function of a T-cell. Their usual form is that of a type I transmembrane domain protein with an antigen recognizing amino terminus, a spacer, a transmembrane domain all connected to a compound endodomain which transmits T-cell survival and activation signals.

The antigen binding domain may be derived from an antibody or antibody mimetic, or it may be another entity which specifically binds the antigen, such as a ligand.

The most common form of these molecules are fusions of single-chain variable fragments (scFv) derived from monoclonal antibodies which recognize a target antigen, fused via a

spacer and a trans-membrane domain to a signaling endodomain. Such molecules result in activation of the T-cell in response to recognition by the scFv of its target. When T cells express such a CAR, they recognize and kill target cells that express the target antigen. Several CARs have been developed against tumour associated antigens, and adoptive transfer approaches using such CAR-expressing T cells are currently in clinical trial for the treatment of various cancers.

It is also possible for the signalling endodomain to be present on a separate molecule. The term "CAR" in connection with the present invention also encompasses a molecule which comprises an antigen binding domain connected to a transmembrane domain. Such a CAR may be capable of interacting with a separate intracellular signalling domain in order to stimulate T-cell activation.

In the present invention, either of the nucleic acid sequences A or B may be a nucleic acid sequence which encodes a transmembrane protein comprising a signal peptide.

Most transmembrane proteins of interest are only active, or are predominantly active when at the cell membrane. Therefore the use of an inefficient signal peptide reduces the relative expression of the protein at the cell surface and therefore reduces the relative activity of the protein.

POLYPROTEIN

The nucleic acid construct of the present invention may encode a polyprotein, which comprises the first and second polypeptides.

The first and second polypeptides may be chimeric antigen receptors (CARs).

The nucleic acid construct described in the Examples encodes the following polyprotein which comprises the various components in the order they are listed:

Signal peptide derived from Mouse Ig kappa chain V-III region:
METD**TLILWVLLLLV**PGSTG (see below)
scFv aCD19:
DIQMTQTTSSLSASLGDRVTISCRASQDISKYLNWYQQKPDGTVKLLIYHTSRLHSGVPSRF
SGSGSGTDYSLTISNLEQEDIATYFCQQGNTLPYTFGGGTKLEITKAGGGGSGGGGSGGG
GSGGGGSEVKLQESGPGLVAPSQSLSVTCTVSGVSLPDYGVSWIRQPPRKGLEWLGVIW

GSETTYYNSALKSRLTIIKDNSKSQVFLKMNSLQTDDTAIYYCAKHYYYGGSYAMDYWGQG
TSVTVS

Linker:

SD

Human CD8aSTK:

PTTTPAPRPPTPAPTIASQPLSLRPEACRPAAGGAVHTRGLDFACDI

Human CD28TM:

FWVLVVVGGVLACYSLLVTVAFIIFWV

Human CD3zeta intracellular domain:

RRVKFSRSADAPAYQQGQNQLYNELNLGRREEYDVLDKRRGRDPEMGGKPRRKNPQEG
LYNELQKDKMAEAYSEIGMKGERRRGKGHDGLYQGLSTATKDTYDALHMQALPPR

2A peptide:

RAEGRGSLLTCGDVEENPGP

Signal peptide derived from mouse Ig kappa:

MAVPTQVLGLLLLWLTDA

scFv aCD33:

RCDIQMTQSPSSLSASVGDRVTITCRASEDIYFNLVWYQQKPGKAPKLLIYDTNRLADGVPS
RFSGSGSGTQYTLTISSLQPEDFATYYCQHYKNYPLTFGQGTKLEIKRSGGGGSGGGGSG
GGGSGGGGSRSEVQLVESGGGLVQPGGSLRLSCAASGFTLSNYGMHWIRQAPGKGLEW
VSSISLNGGSTYYRDSVKGRFTISRDNAKSTLYLQMNSLRAEDTAVYYCAAQDAYTGGYFD
YWGQGTLVTVSSM

Linker:

DPA

Hinge and Fc derived from human IgG1 with mutations to prevent FcRg association (HCH2CH3pvaa):

EPKSPDKTHTCPPCPAPPVAGPSVFLFPPKPKDTLMIARTPEVTCVVVDVSHEDPEVKFNW
YVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISK
AKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD
SDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

Linker:

KDPK

Human CD148TM:

AVFGCIFGALVIVTVGGFIFW

Human CD148 intracellular domain:

RKKRKDAKNNEVSFSQIKPKKSKLIRVENFEAYFKKQQADSNCGFAEEYEDLKLVGISQPKY
AAELAENRGKNRYNNVLPYDISRVKLSVQTHSTDDYINANYMPGYHSKKDFIATQGPLPNTL
KDFWRMVWEKNVYAIIMLTKCVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRD

FTVKNIQTSESHPLRQFHFTSWPDHGVPDTTDLLINFRYLVRDYMKQSPPESPILVHCSAGV GRTGTFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSK VDLIYQNTTAMTIYENLAPVTTFGKTNGYIA

The signal sequence of from murine Ig kappa chain V-III region, which has the sequence: METD**TLILWVLLLLV**PGSTG (hydrophobic residues highlighted in bold) is altered as described in the Examples. Hydrophobic residues were sequentially deleted from the hydrophobic core in order to reduce the level of expression of the anti-CD19 CAR.

CLEAVAGE SITE

The nucleic acid construct of the first aspect of the invention comprises a sequence encoding a cleavage site positioned between nucleic acid sequences which encode first and second polypeptides, such that first and second polypeptides can be expressed as separate entities.

The cleavage site may be any sequence which enables the polypeptide comprising the first and second polypeptides to become separated.

The term “cleavage” is used herein for convenience, but the cleavage site may cause the first and second polypeptidess to separate into individual entities by a mechanism other than classical cleavage. For example, for the Foot-and-Mouth disease virus (FMDV) 2A self-cleaving peptide (see below), various models have been proposed for to account for the “cleavage” activity: proteolysis by a host-cell proteinase, autoproteolysis or a translational effect (Donnelly et al (2001) J. Gen. Virol. 82:1027-1041). The exact mechanism of such “cleavage” is not important for the purposes of the present invention, as long as the cleavage site, when positioned between nucleic acid sequences which encode first and second polypeptides, causes the first and second polypeptides to be expressed as separate entities.

The cleavage site may be a furin cleavage site.

Furin is an enzyme which belongs to the subtilisin-like proprotein convertase family. The members of this family are proprotein convertases that process latent precursor proteins into their biologically active products. Furin is a calcium-dependent serine endoprotease that can efficiently cleave precursor proteins at their paired basic amino acid processing sites. Examples of furin substrates include proparathyroid hormone, transforming growth factor beta 1 precursor, proalbumin, pro-beta-secretase, membrane type-1 matrix

metalloproteinase, beta subunit of pro-nerve growth factor and von Willebrand factor. Furin cleaves proteins just downstream of a basic amino acid target sequence (canonically, Arg-X-(Arg/Lys)-Arg') and is enriched in the Golgi apparatus.

The cleavage site may be a Tobacco Etch Virus (TEV) cleavage site.

TEV protease is a highly sequence-specific cysteine protease which is chymotrypsin-like proteases. It is very specific for its target cleavage site and is therefore frequently used for the controlled cleavage of fusion proteins both *in vitro* and *in vivo*. The consensus TEV cleavage site is ENLYFQ\S (where '\' denotes the cleaved peptide bond). Mammalian cells, such as human cells, do not express TEV protease. Thus in embodiments in which the present nucleic acid construct comprises a TEV cleavage site and is expressed in a mammalian cell – exogenous TEV protease must also expressed in the mammalian cell.

The cleavage site may encode a self-cleaving peptide.

A 'self-cleaving peptide' refers to a peptide which functions such that when the polypeptide comprising the first and second polypeptides and the self-cleaving peptide is produced, it is immediately "cleaved" or separated into distinct and discrete first and second polypeptides without the need for any external cleavage activity.

The self-cleaving peptide may be a 2A self-cleaving peptide from an aphtho- or a cardiovirus. The primary 2A/2B cleavage of the aptho- and cardioviruses is mediated by 2A "cleaving" at its own C-terminus. In apthoviruses, such as foot-and-mouth disease viruses (FMDV) and equine rhinitis A virus, the 2A region is a short section of about 18 amino acids, which, together with the N-terminal residue of protein 2B (a conserved proline residue) represents an autonomous element capable of mediating "cleavage" at its own C-terminus.

The C-terminal 19 amino acids of the longer cardiovirus protein, together with the N-terminal proline of 2B mediate "cleavage" with an efficiency approximately equal to the apthovirus FMDV 2a sequence. Cardioviruses include encephalomyocarditis virus (EMCV) and Theiler's murine encephalitis virus (TMEV).

Mutational analysis of EMCV and FMDV 2A has revealed that the motif DxExNPGP is intimately involved in "cleavage" activity (Donelly et al (2001) as above).

The cleavage site of the present invention may comprise the amino acid sequence:

Dx_1Ex_2NPGP, where x_1 and x_2 are any amino acid. X_1 may be selected from the following group: I, V, M and S. X_2 may be selected from the following group: T, M, S, L, E, Q and F.

For example, the cleavage site may comprise one of the amino acid sequences shown in Table 2.

Table 2

Motif	Present in:
DIETNPGP (SEQ ID No. 1)	Picornaviruses EMCB, EMCD, EMCPV21
DVETNPGP (SEQ ID No. 2)	Picornaviruses MENGO and TMEBEAN; Insect virus DCV, ABPV
DVEMNPGP (SEQ ID No. 3)	Picornaviruses TMEGD7 and TMEBEAN
DVESNPGP (SEQ ID No. 4)	Picornaviruses FMDA10, FMDA12, FMDC1, FMD01K, FMDSAT3, FMDVSAT2, ERAV; Insect virus CrPV
DMESNPGP (SEQ ID No. 5)	Picornavirus FMDV01G
DVELNPGP (SEQ ID No. 6)	Picornavirus ERBV; Porcine rotavirus
DVEENPGP (SEQ ID No. 7)	Picornavirus PTV-1; Insect virus TaV; Trypanosoma TSR1
DIELNPGP (SEQ ID No. 8)	Bovine Rotavirus, human rotavirus
DIEQNPGP (SEQ ID No. 9)	Trypanosoma AP endonuclease
DSEFNPGP (SEQ ID No. 10)	Bacterial sequence *T. maritima*

The cleavage site, based on a 2A sequence may be, for example 15-22 amino acids in length. The sequence may comprise the C-terminus of a 2A protein, followed by a proline residue (which corresponds to the N-terminal proline of 2B).

Mutational studies have also shown that, in addition to the naturally occurring 2A sequences, some variants are also active. The cleavage site may correspond to a variant sequence from a naturally occurring 2A polypeptide, have one, two or three amino acid substitutions, which retains the capacity to induce the "cleavage" of a polyprotein sequence into two or more separate proteins.

The cleavage sequence may be selected from the following which have all been shown to be active to a certain extent (Donnelly *et al* (2001) as above):

LLNFDLLKLAGDVESNPGP (SEQ ID No. 11)
LLNFDLLKLAGDVQSNPGP (SEQ ID No. 12)
LLNFDLLKLAGDVEINPGP (SEQ ID No. 13)
LLNFDLLKLAGDVEFNPGP (SEQ ID No. 14)
LLNFDLLKLAGDVESHPGP (SEQ ID No. 15)
LLNFDLLKLAGDVESEPGP (SEQ ID No. 16)
LLNFDLLKLAGDVESQPGP (SEQ ID No. 17)
LLNFDLLKLAGDVESNPGG (SEQ ID No. 18)

Based on the sequence of the DxExNPGP "a motif, "2A-like" sequences have been found in picornaviruses other than aptho- or cardioviruses, 'picornavirus-like' insect viruses, type C rotaviruses and repeated sequences within *Trypanosoma spp* and a bacterial sequence (Donnelly et al (2001) as above). The cleavage site may comprise one of these 2A-like sequences, such as:

YHADYYKQRLIHDVEMNPGP (SEQ ID No. 19)
HYAGYFADLLIHDIETNPGP (SEQ ID No. 20)
QCTNYALLKLAGDVESNPGP (SEQ ID No. 21)
ATNFSLLKQAGDVEENPGP (SEQ ID No. 22)
AARQMLLLLSGDVETNPGP (SEQ ID No. 23)
RAEGRGSLLTCGDVEENPGP (SEQ ID No. 24)
TRAEIEDELIRAGIESNPGP (SEQ ID No. 25)
TRAEIEDELIRADIESNPGP (SEQ ID No. 26)
AKFQIDKILISGDVELNPGP (SEQ ID No. 27)

SSIIRTKMLVSGDVEENPGP (SEQ ID No. 28)
CDAQRQKLLLSGDIEQNPGP (SEQ ID No. 29)
YPIDFGGFLVKADSEFNPGP (SEQ ID No. 30)

The cleavage site may comprise the 2A-like sequence shown as SEQ ID No. 24 (RAEGRGSLLTCGDVEENPGP).

It has been shown that including an N-terminal "extension" of between 5 and 39 amino acids can increase activity (Donnelly et al (2001) as above). In particular, the cleavage sequence may comprise one of the following sequences or a variant thereof having, for example, up to 5 amino acid changes which retains cleavage site activity:

VTELLYRMKRAETYCPRPLAIHPTEARHKQKIVAPVKQTLNFDLLKLAGDVESNPGP (SEQ ID No. 31)
LLAIHPTEARHKQKIVAPVKQTLNFDLLKLAGDVESNPGP (SEQ ID No. 32)
EARHKQKIVAPVKQTLNFDLLKLAGDVESNPGP (SEQ ID No. 33)
APVKQTLNFDLLKLAGDVESNPGP (SEQ ID No. 34)

TUNABILITY

The relative expression of one or more protein(s) may be fine tuned using the method of the invention by using signal peptides with different amounts or porportions of hydrophobic amino acids..

The tunability using different signal peptides is especially useful when one considers the expression of multiple proteins, each with their own relative expression. For example, consider a nucleic acid construct having the following structure:

A – X – B – Y - C

in which

A, B and C are nucleic acid sequences encoding polypeptides; and
X and Y are nucleic acid sequences encodes cleavage sites.

The nucleic acid construct will encode three proteins *A*, *B* and *C*, any or all of which may be transmembrane proteins. If it is desired for A, B and C to be expressed such that the relative levels are *A*>*B*>*C*, then the nucleic acid sequence A may have a signal peptide with the most hydrophobic amino acids, the nucleic acid sequence B may have a signal peptide with

a medium amount of hydrophobic amino acids, and the nucleic acid sequence C may have a signal peptide with the least hydrophobic amino acids.

VECTOR

The present invention also provides a vector comprising a nucleic acid construct according to the first aspect of the invention.

Such a vector may be used to introduce the nucleic acid construct into a host cell so that it expresses the first and second polypeptide.

The vector may, for example, be a plasmid or a viral vector, such as a retroviral vector or a lentiviral vector, or a transposon based vector or synthetic mRNA.

The vector may be capable of transfecting or transducing a mammalian cell, for example a T cell.

CELL

The present invention furthers provides a cell comprising a nucleic acid construct or vector of the present invention which expresses the first and second polypeptide encoded by the nucleic acid sequence.

The cell may be any eukaryotic cell capable of expressing a transmembrane protein at the cell surface, such as an immunological cell.

The cell may be a cytolytic immune cell, such as a T cell or natural killer cell.

METHOD

In a further aspect the present invention provides a method for making a cell according to the invention which comprises the step of introducing a nucleic acid construct or a vector of the invention into a cell.

The nucleic acid construct may be introduced by transduction or transfection.

The cell may be a cell isolated from a subject, for example a T cell or an NK cell isolated from a subject.

The present invention also provides a method for modulating the relative cell surface expression of a first protein expressed from a single nucleic acid construct with a second protein which comprises the step of mutating the nucleic acid sequence which encodes the signal peptide of one protein in order to remove or replace on or more hydrophobic amino acids in comparison with the signal peptide of the other protein.

The signal peptide may be altered by techniques known in the art, such as site directed mutagenesis and recombinant techniques.

The invention will now be further described by way of Examples, which are meant to serve to assist one of ordinary skill in the art in carrying out the invention and are not intended in any way to limit the scope of the invention.

EXAMPLES

Example 1 – Using the murine Ig kappa chain V-III signal sequence

PCT/GB2014/053452 describes a vector system encoding two chimeric antigen receptors (CARs), one against CD19 and one against CD33. The signal peptide used for the CARs in that study was the signal peptide from the human CD8a signal sequence. For the purposes of this study, this was substituted with the signal peptide from the murine Ig kappa chain V-III region, which has the sequence: METD**TLILWVLLLLV**PGSTG (hydrophobic residues hightlited in bold). In order to establish that the murine Ig kappa chain V-III signal sequence functioned as well as the signal sequence from human CD8a, a comparative study was performed. For both signal sequences, functional expression of the anti-CD33 CAR and the anti-CD19 CAR was observed. This substituted signal sequence and all subsequent mutations thereof were transiently transfected into 293T cells. Three days after transfection the 293T cells were stained with both soluble chimeric CD19 fused with rabbit Fc chain and soluble chimeric CD33 fused with mouse Fc chain. All cells were then stained with anti-Rabbit Fc-FITC and anti-mouse Fc-APC. Flow cytometry plots show the substituted signal sequence as a comparison with non-transfected (NT) and the construct with Cd8 signal sequences (Figure 4). The murine Ig kappa chain V-III signal sequence was found to function as well as the signal sequence from human CD8a.

Example 2 – Altering relative expression by deleting hydrophobic residues in the signal peptide.

Hydrophobic residues were deleted in a stepwise fashion and the effect on the relative expression of the anti-CD33 CAR and the anti-CD19 CAR was observed. The effect of one, two, three and four amino acid deletions was investigated and the results are shown in Figures 5 to 8 respectively.

All mutant constructs showed a decrease in relative expression of the anti-CD19 CAR compared to the anti-CD33 CAR. The relative decrease of anti-CD19 CAR expression was greater with a greater number of amino acid deletions from 1 to 3, but then plateaued out (four deletions gave a similar decrease in expression as three deletions).

Modification of the signal sequences in a nucleic acid construct encoding two polypeptides can therefore be used to control the relative expression of the two polypeptides.

All publications mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described methods and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology, cell biology or related fields are intended to be within the scope of the following claims.

CLAIMS

1. A nucleic acid construct comprising the following structure:

A – X – B

in which

A is nucleic acid sequence encoding a first polypeptide which comprises a first signal peptide;

B is nucleic acid sequence encoding a second polypeptide which comprises a second signal peptide and

X is a nucleic acid sequence which encodes a cleavage site,

wherein the first signal peptide or the second signal peptide comprises one or more mutation(s) such that it has fewer hydrophobic amino acids.

2. A nucleic acid construct according to claim 1, wherein the first signal peptide and the second signal peptide are derivable from the same sequence, but one signal peptide comprises one or more amino acid deletions/substitutions to remove/replace one or more hydrophobic amino acids compared to the other signal peptide.

3. A nucleic acid construct according to claim 2, wherein the hydrophobic amino acid(s) is/are selected from the group: Alanine (A); Valine (V); Isoleucine (I); Leucine (L); Methionine (M); Phenylalanine (P); Tyrosine (Y); Tryptophan (W).

4. A nucleic acid construct according to claim 3, wherein the hydrophobic amino acid(s) is/are selected from the group: Valine (V); Isoleucine (I); Leucine (L); and Tryptophan (W).

5. A nucleic acid construct according to any preceding claim, wherein one signal peptide comprises up to five more hydrophobic amino acids than the other signal peptide.

6. A nucleic acid construct according to any preceding claim wherein the first and second polypeptides are both transmembrane proteins.

7. A nucleic acid construct according to claim 6, wherein the first and second transmembrane proteins are both chimeric antigen receptors (CARs).

8. A nucleic acid construct according to claim 6 or 7, wherein the difference between the first and the second signal peptides is such that, such that when the nucleic acid construct is expressed in a cell, there is differential relative expression of the first and second polypeptides at the cell surface.

9. A nucleic acid construct according to any preceding claim wherein X is a nucleic acid sequence encoding a self-cleaving peptide, a furin cleavage site or a Tobacco Etch Virus cleavage site.

10. A nucleic acid construct according to claim 9, wherein X encodes a 2A self-cleaving peptide from an aphtho- or a cardiovirus or a 2A-like peptide.

11. A vector comprising a nucleic acid construct according to any preceding claim.

12. A retroviral vector or a lentiviral vector according to claim 11.

13. A cell comprising a nucleic acid construct according to any one of claims 1 to 10 or a vector according to claim 11 or 12.

14. A cell which comprises two chimeric antigen receptors having different signal peptides which have a different number of hydrophobic amino acids, wherein the chimeric antigen receptor having a signal peptide with the greater number of hydrophobic amino acids is more highly expressed at the cell surface than the chimeric antigen receptor having a signal peptide with the smaller number of hydrophobic amino acids.

15. A cell which comprises two chimeric antigen receptors having different signal peptides, wherein the signal peptide of one chimeric antigen comprises one or more mutations such that it has fewer hydrophobic amino acids.

16. A cell according to claim 15, wherein the chimeric antigen receptor having a mutated signal peptide is expressed at a lower level at the cell surface than the chimeric antigen receptor having an unmutated signal peptide.

17. A cell according to any of claims 13 to 16 which is a T cell or a natural killer (NK) cell.

18. A method for making a cell according to any of claims 13 to 17 which comprises the step of introducing a nucleic acid construct according to any of claims 1 to 10 or a vector according to claim 11 or 12 into a cell.

19. A method for modulating the relative cell surface expression of a first protein expressed from a single nucleic acid construct with a second protein which comprises the step of mutating the nucleic acid sequence which encodes the signal peptide of one protein in order to remove or replace one or more hydrophobic amino acids in comparison with the signal peptide of the other protein.

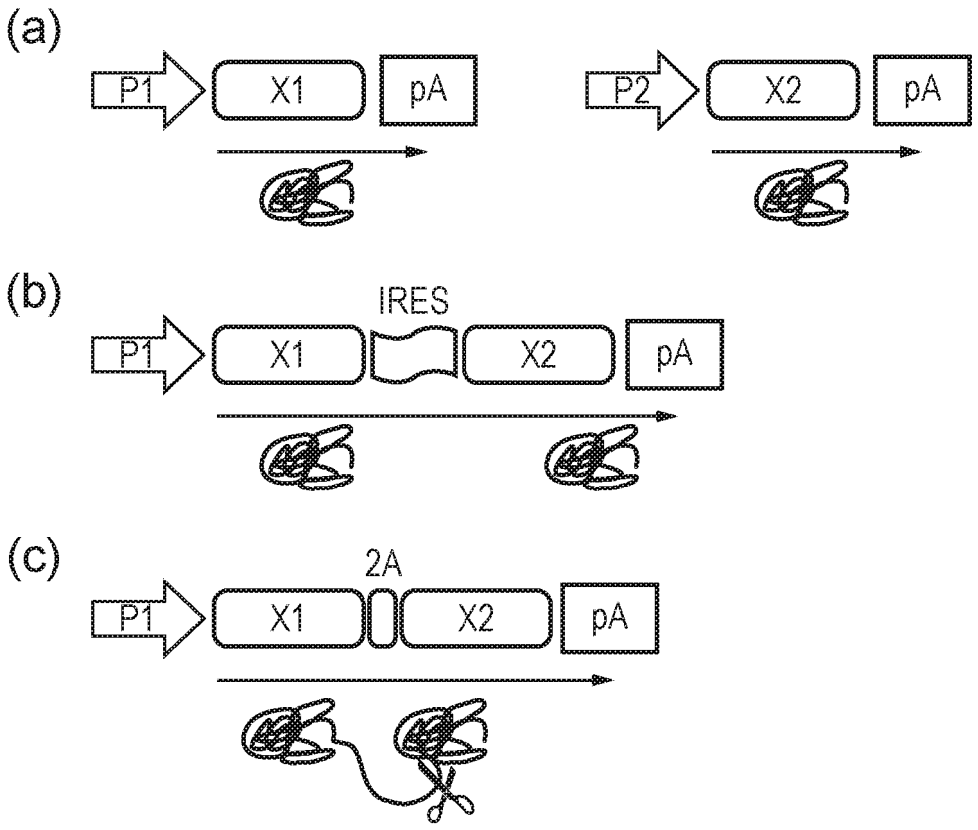

FIG. 1

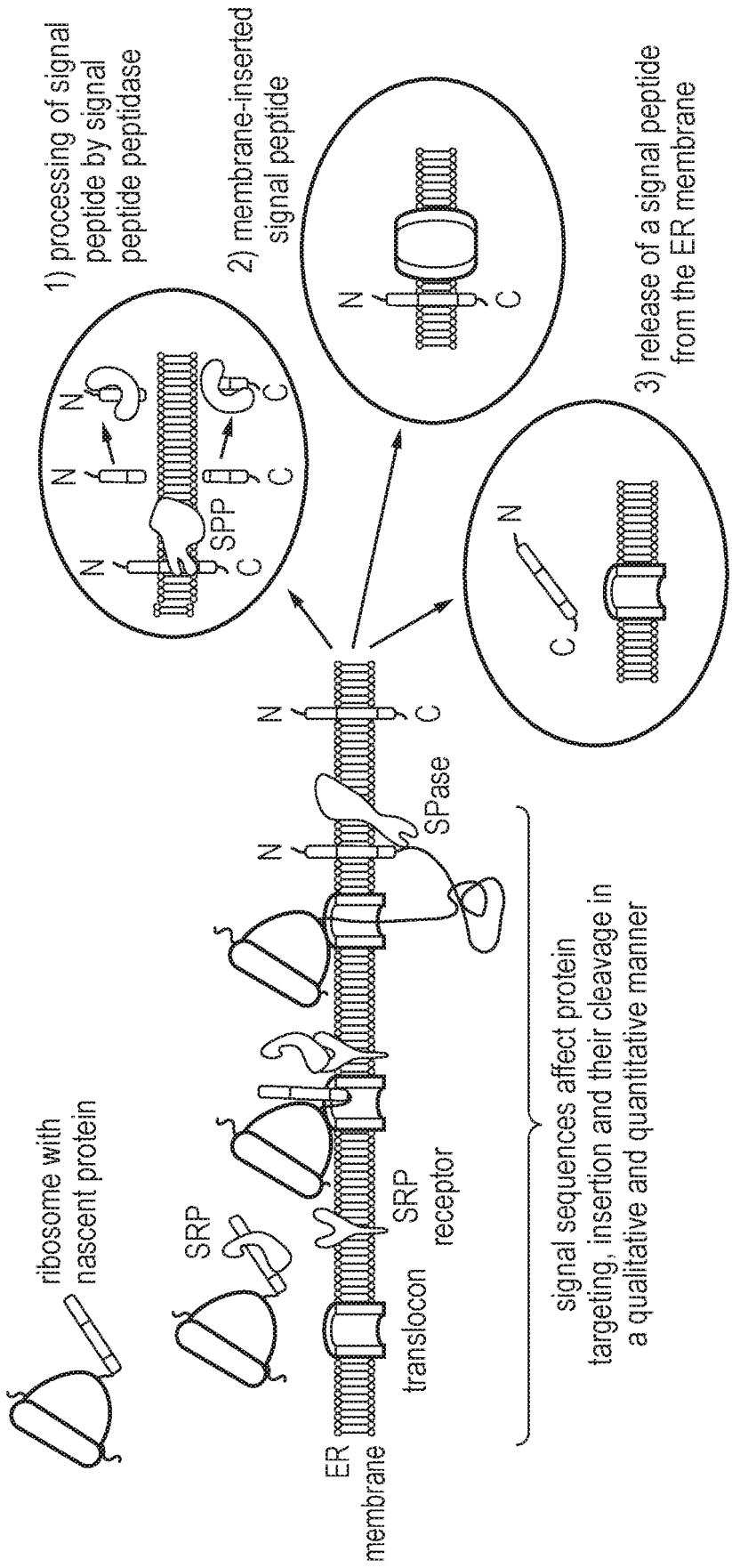
ribosome with nascent protein
SRP
SRP receptor
translocon
ER membrane
SPase
SPP
1) processing of signal peptide by signal peptide peptidase
2) membrane-inserted signal peptide
3) release of a signal peptide from the ER membrane
signal sequences affect protein targeting, insertion and their cleavage in a qualitative and quantitative manner
N
C

FIG. 2

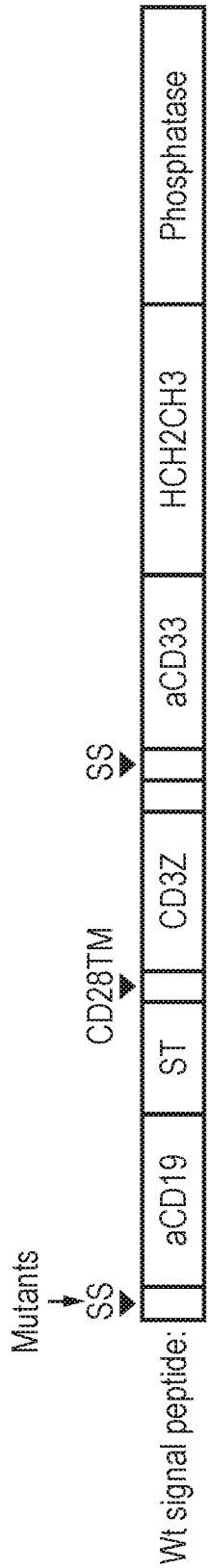
Mutants
SS
CD28TM
SS
Wt signal peptide:
aCD19
ST
CD3Z
aCD33
HCH2CH3
Phosphatase
2A

FIG. 3

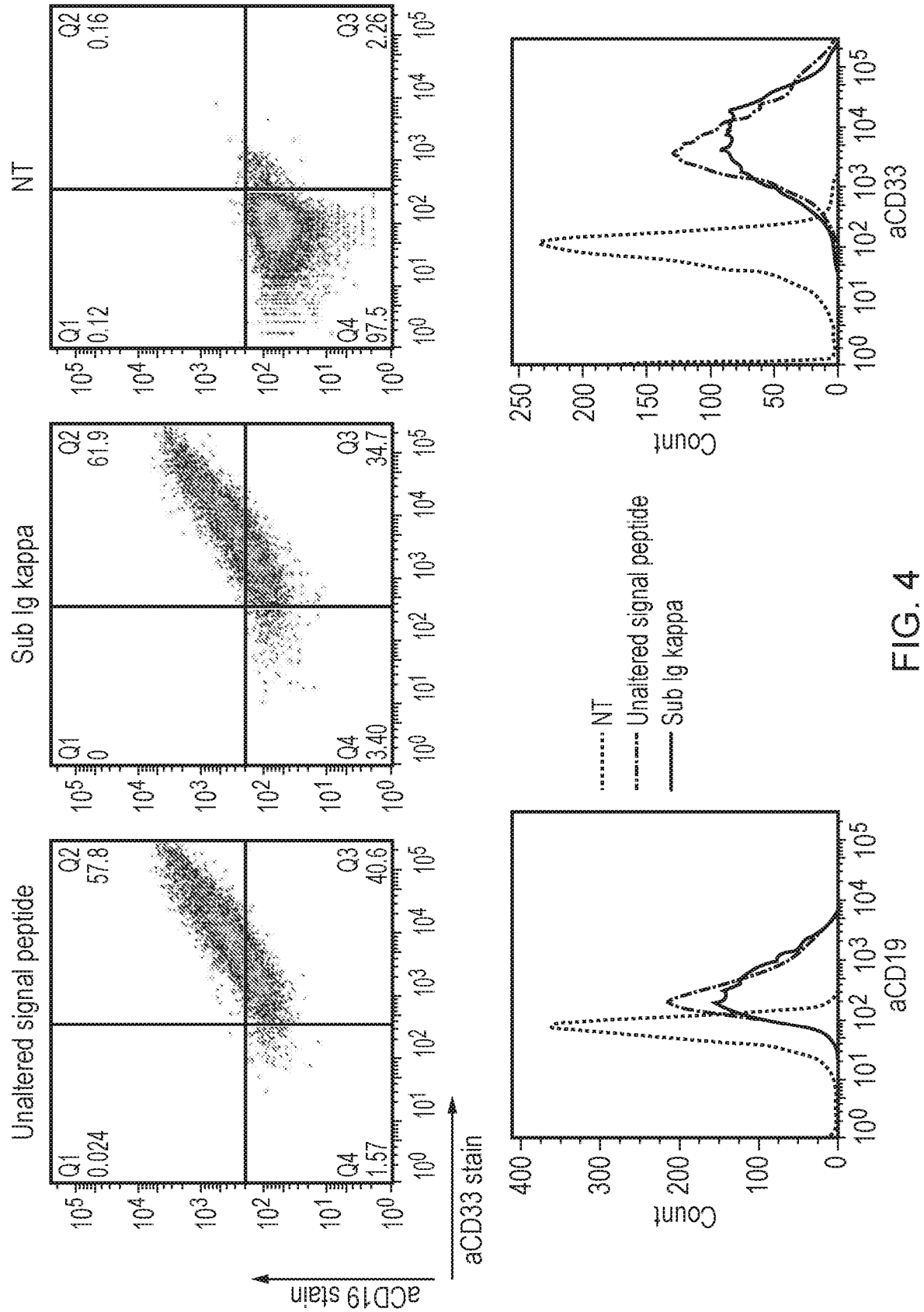
NT
Sub Ig kappa
Unaltered signal peptide
Q1 0.12
Q2 0.16
Q4 97.5
Q3 2.26
Q1 0
Q2 61.9
Q4 3.40
Q3 34.7
Q1 0.024
Q2 57.8
Q4 1.57
Q3 40.6
aCD19 stain
aCD33 stain
NT
Unaltered signal peptide
Sub Ig kappa
Count
aCD33
aCD19

FIG. 4

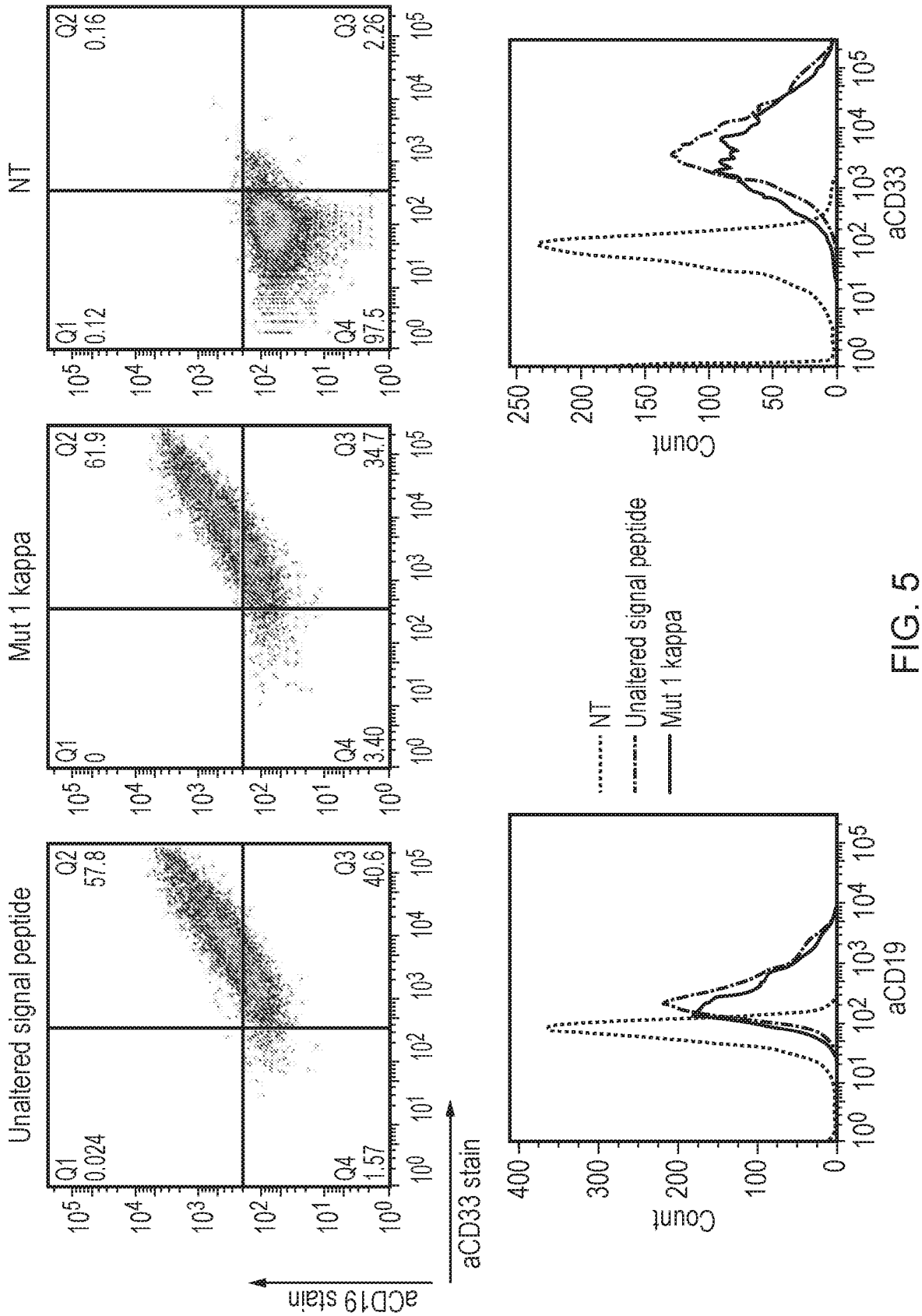
NT
Mut 1 kappa
Unaltered signal peptide
Q1 0.12
Q2 0.16
Q3 2.26
Q4 97.5
Q1 0
Q2 61.9
Q3 34.7
Q4 3.40
Q1 0.024
Q2 57.8
Q3 40.6
Q4 1.57
aCD19 stain
aCD33 stain
NT
Unaltered signal peptide
Mut 1 kappa
Count
aCD33
aCD19

FIG. 5

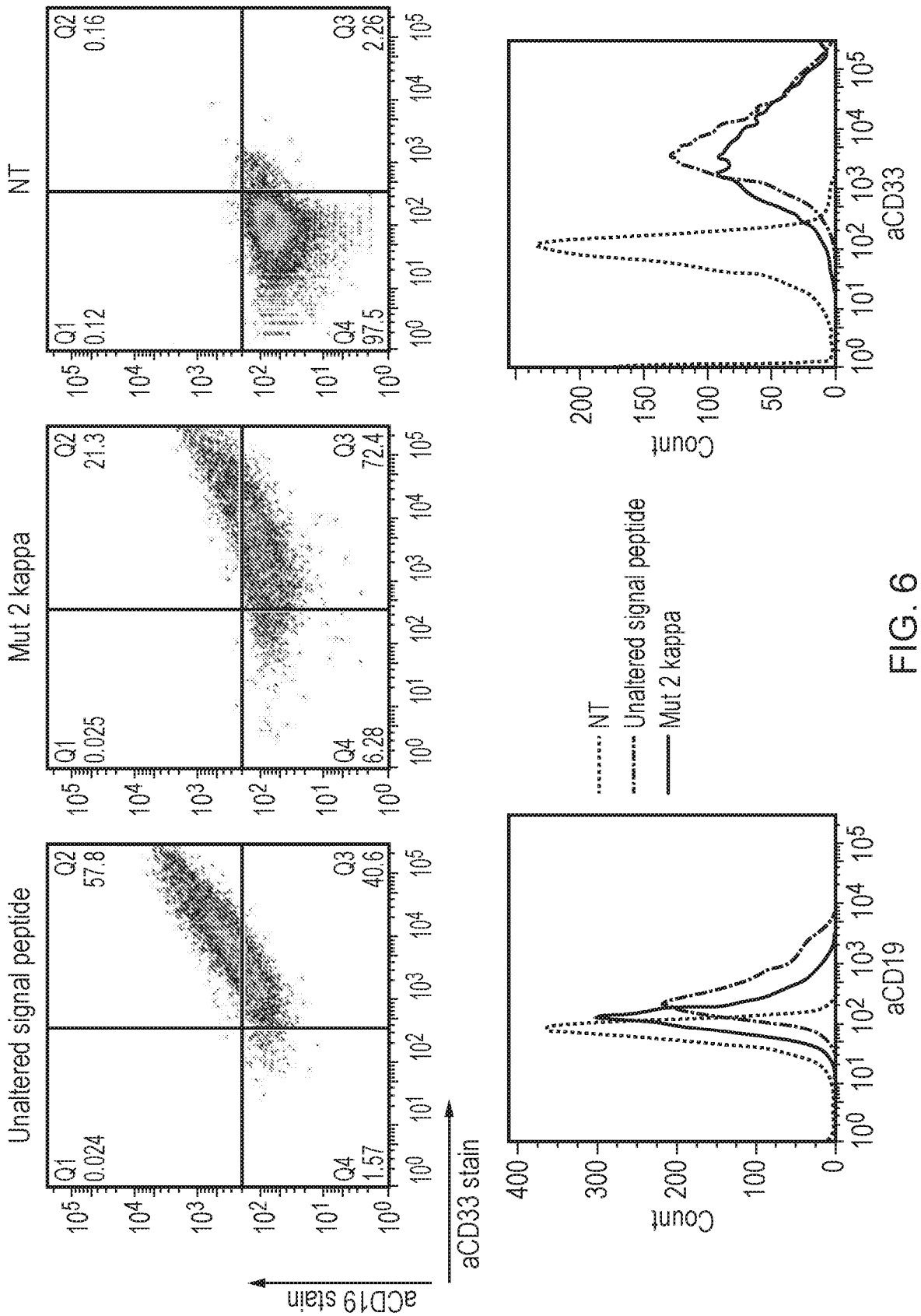
NT
Mut 2 kappa
Unaltered signal peptide
Q1 0.12
Q2 0.16
Q3 2.26
Q4 97.5
Q1 0.025
Q2 21.3
Q3 72.4
Q4 6.28
Q1 0.024
Q2 57.8
Q3 40.6
Q4 1.57
aCD19 stain
aCD33 stain
NT
Unaltered signal peptide
Mut 2 kappa
Count
aCD33
aCD19

FIG. 6

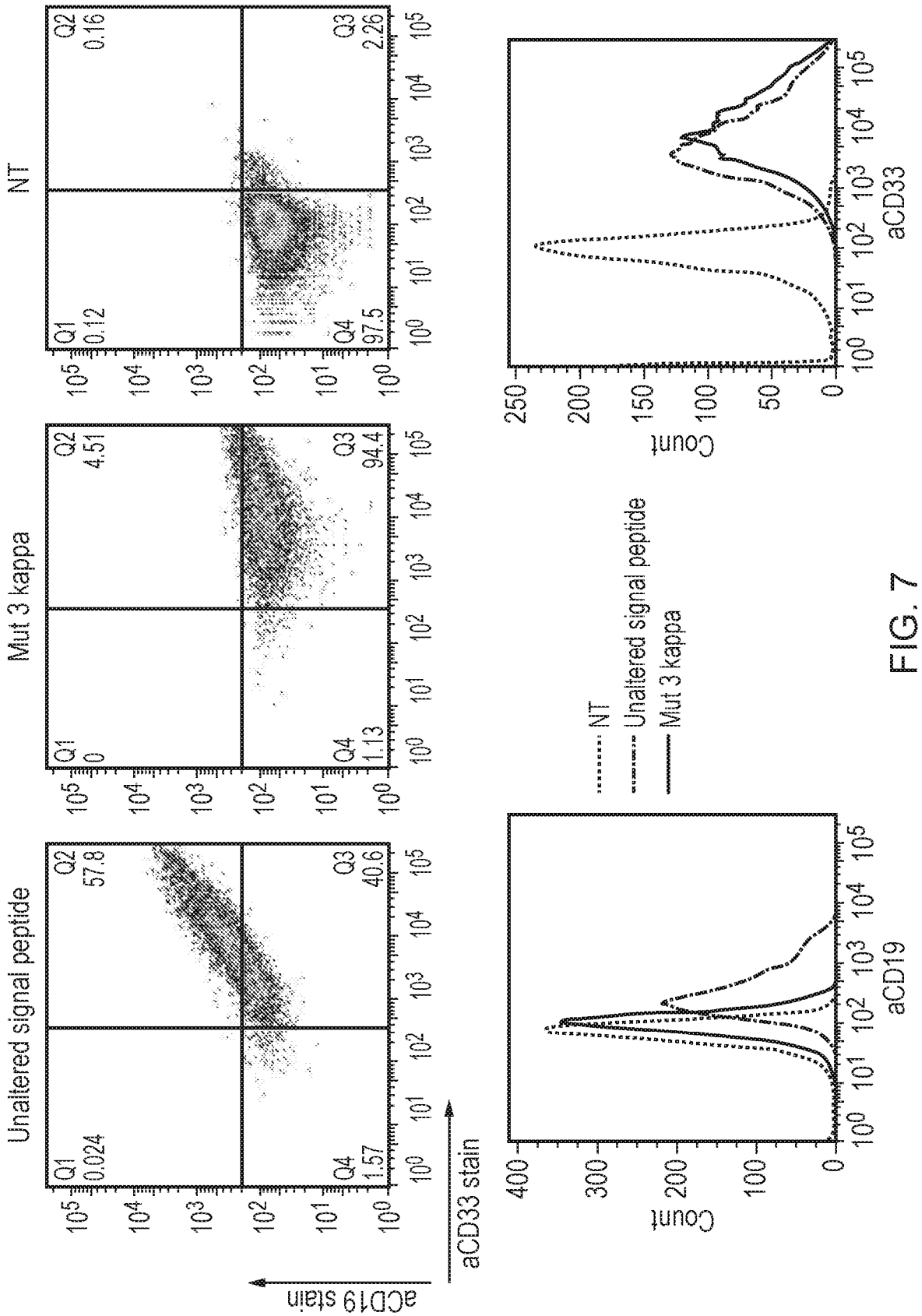
NT
Mut 3 kappa
Unaltered signal peptide
Q1 0.12
Q2 0.16
Q4 97.5
Q3 2.26
Q1 0
Q2 4.51
Q4 1.13
Q3 94.4
Q1 0.024
Q2 57.8
Q4 1.57
Q3 40.6
aCD19 stain
aCD33 stain
NT
Unaltered signal peptide
Mut 3 kappa
Count
aCD33
aCD19

FIG. 7

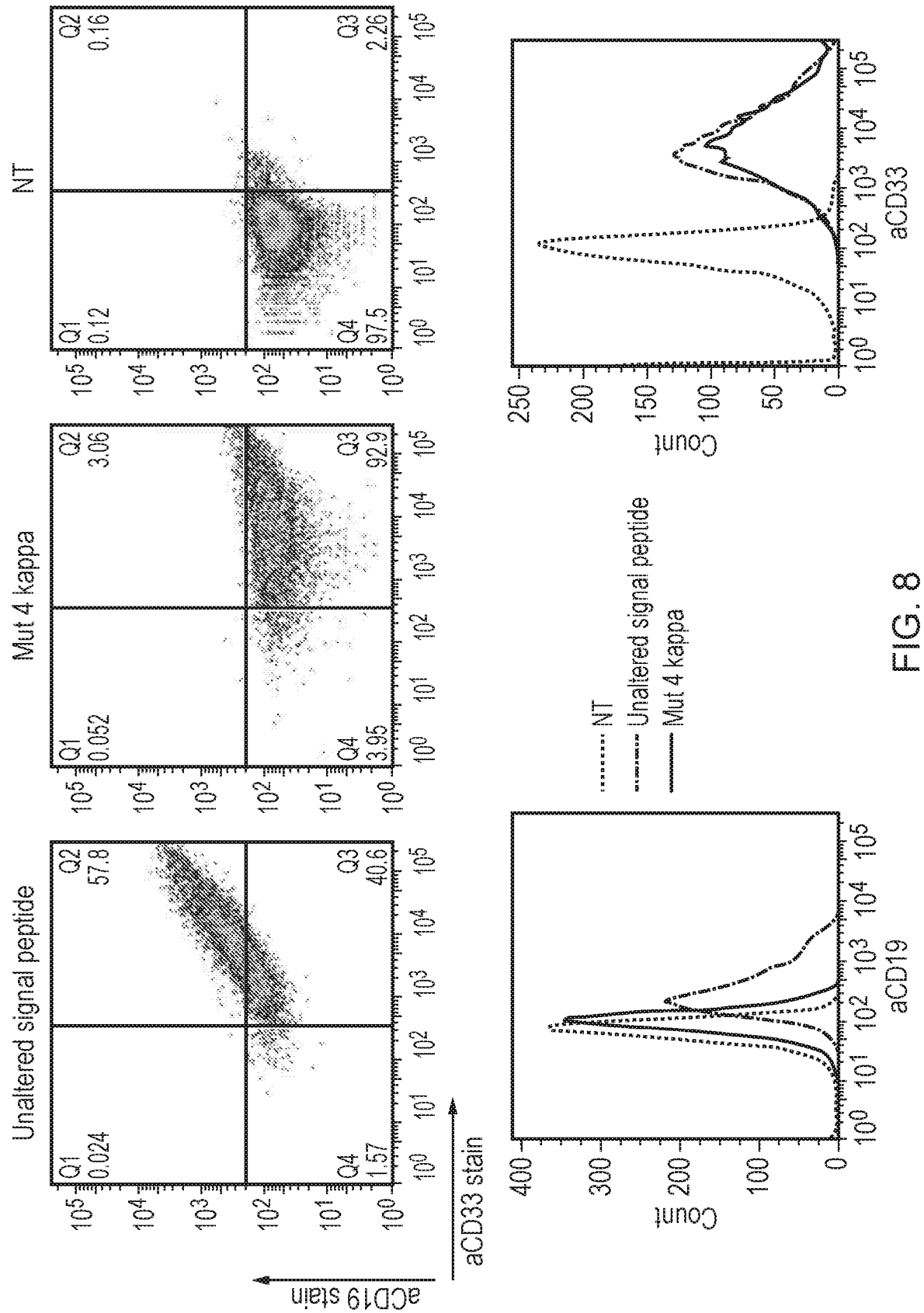
NT
Mut 4 kappa
Unaltered signal peptide
Q1 0.12
Q2 0.16
Q3 2.26
Q4 97.5
Q1 0.052
Q2 3.06
Q3 92.9
Q4 3.95
Q1 0.024
Q2 57.8
Q3 40.6
Q4 1.57
aCD19 stain
aCD33 stain
NT
Unaltered signal peptide
Mut 4 kappa
Count
aCD33
aCD19

FIG. 8

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2016/051165

A. CLASSIFICATION OF SUBJECT MATTER
INV. A61K39/00 C07K16/28 C12N5/0783
ADD.

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
A61K C07K C12N

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, BIOSIS, CHEM ABS Data, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SCOTT WILKIE ET AL: "Dual Targeting of ErbB2 and MUC1 in Breast Cancer Using Chimeric Antigen Receptors Engineered to Provide Complementary Signaling", JOURNAL OF CLINICAL IMMUNOLOGY, KLUWER ACADEMIC PUBLISHERS-PLENUM PUBLISHERS, NE, vol. 32, no. 5, 17 April 2012 (2012-04-17) , pages 1059-1070, XP035113362, ISSN: 1573-2592, DOI: 10.1007/S10875-012-9689-9 figure 1b -----	1-7,9-18
X	WO 2013/185552 A1 (HEPATOBILIARY SURGERY HOSPITAL SECOND MILITARY MEDICAL UNIVERSITY [CN]) 19 December 2013 (2013-12-19) page 11; figures 1, 2; sequences 15, 17 ----- -/--	1-7,9-18

[X] Further documents are listed in the continuation of Box C. [X] See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
28 June 2016	07/07/2016

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Lanzrein, Markus

1

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2016/051165

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2007/014162 A2 (ABBOTT LAB [US]; CARSON GERALD R [US]; GION WENDY [US]; SALFELD JOCHEN) 1 February 2007 (2007-02-01) paragraph [0234] - paragraph [0236]	1-19
A	WO 2015/052538 A1 (UCL BUSINESS PLC [GB]) 16 April 2015 (2015-04-16) page 13, line 20 - page 14, line 14	1-19
X,P	WO 2015/075469 A1 (UCL BUSINESS PLC [GB]) 28 May 2015 (2015-05-28) figure 4 page 34, line 15 - line 33	1-7,9-18

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/GB2016/051165

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2013185552 A1	19-12-2013	CN 103483452 A	01-01-2014
		WO 2013185552 A1	19-12-2013
WO 2007014162 A2	01-02-2007	AU 2006272706 A1	01-02-2007
		BR PI0613784 A2	01-02-2011
		CA 2615983 A1	01-02-2007
		CN 101595228 A	02-12-2009
		EP 1910550 A2	16-04-2008
		EP 2468768 A2	27-06-2012
		EP 2468881 A2	27-06-2012
		EP 2484774 A2	08-08-2012
		JP 2009508470 A	05-03-2009
		JP 2012235782 A	06-12-2012
		KR 20080031024 A	07-04-2008
		SG 164383 A1	29-09-2010
		US 2007065912 A1	22-03-2007
		US 2011034368 A1	10-02-2011
		US 2013243789 A1	19-09-2013
		WO 2007014162 A2	01-02-2007
		ZA 200800207 A	28-01-2009
WO 2015052538 A1	16-04-2015	AU 2014333564 A1	28-04-2016
		CA 2926560 A1	16-04-2015
		SG 11201602414S A	28-04-2016
		WO 2015052538 A1	16-04-2015
WO 2015075469 A1	28-05-2015	AU 2014351557 A1	26-05-2016
		AU 2014351558 A1	26-05-2016
		CA 2929984 A1	28-05-2015
		CA 2930215 A1	28-05-2015
		SG 11201603479T A	30-05-2016
		SG 11201603484P A	30-05-2016
		WO 2015075468 A1	28-05-2015
		WO 2015075469 A1	28-05-2015
		WO 2015075470 A1	28-05-2015

摘要

本發明提供了包含以下結構的核酸構建體：A - X - B其中A是編碼包含第一信號肽的第一多肽的核酸序列; B是編碼包含第二信號肽的第二多肽的核酸序列，並且X是編碼切割位點的核酸序列，其中第一信號肽或第二信號肽包含一個或多個突變以使其具有更少的疏水性氨基酸。