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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
18 November 2021 (18.11.2021)

WIPO | PCT

(10) International Publication Number
WO 2021/229218 A1

(51) International Patent Classification:
A61K 39/00 (2006.01) *C07K 14/54* (2006.01)
A61P 35/00 (2006.01) *C07K 14/725* (2006.01)
C12N 15/63 (2006.01) *C07K 16/30* (2006.01)

(21) International Application Number:
PCT/GB2021/051134

(22) International Filing Date:
12 May 2021 (12.05.2021)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2007044.7 13 May 2020 (13.05.2020) GB

(71) Applicant: **AUTOLUS LIMITED** [GB/GB]; The MediaWorks, 191 Wood Lane, London W12 7FP (GB).

(72) Inventors: **DELLA PERUTA, Marco**; c/o Autolus Limited, The MediaWorks, 191 Wood Lane, London W12 7FP (GB). **RIGHI, Matteo**; c/o Autolus Limited, The MediaWorks, 191 Wood Lane, London W12 7FP (GB). **AGLIARDI, Giulia**; c/o Autolus Limited, The MediaWorks, 191 Wood Lane, London W12 7FP (GB). **PULÉ, Martin**; c/o Autolus Limited, The MediaWorks, 191 Wood Lane, London W12 7FP (GB). **CORDOBA, Shaun**; c/o Autolus Limited, The MediaWorks, 191 Wood Lane, London W12 7FP (GB). **SILLIBOURNE, James**; c/o Autolus Limited, The MediaWorks, 191 Wood Lane, London W12 7FP (GB).

(74) Agent: **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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, 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))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
— *with sequence listing part of description (Rule 5.2(a))*

(54) Title: METHOD

(57) Abstract: The present invention provides a method for treating a solid cancer which comprises the step of administering a cell to a subject, wherein the cell comprises a nucleic acid sequence encoding interleukin 12 (IL-12) downstream of a frame-slip motif (FSM) or a translational readthrough motif (TRM).

WO 2021/229218 A1

METHOD

FIELD OF THE INVENTION

The present invention relates to chimeric antigen receptor (CAR)-expressing cells, such as CAR-T cells, which secrete interleukin-12 (IL-12). The IL-12-encoding sequence is positioned downstream of a frame-slip motif (FSM) or a translational readthrough motif (TRM) so that IL-12 is secreted at very low levels.

BACKGROUND TO THE INVENTION

Tumour heterogeneity describes the observation that different tumour cells can show distinct morphological and phenotypic profiles, including cellular morphology, gene expression, metabolism, motility, proliferation, and metastatic potential.

Heterogeneity occurs between patients, between tumours (inter-tumour heterogeneity) and within tumours (intra-tumour heterogeneity). Multiple types of heterogeneity have been observed between tumour cells, stemming from both genetic and non-genetic variability.

Heterogeneity between tumour cells can be further increased due to heterogeneity in the tumour microenvironment. Regional differences in the tumour (e.g. availability of oxygen) impose different selective pressures on tumour cells, leading to a wider spectrum of dominant subclones in different spatial regions of the tumour. The influence of microenvironment on clonal dominance is also a likely reason for the heterogeneity between primary and metastatic tumours seen in many patients, as well as the inter-tumour heterogeneity observed between patients with the same tumour type.

The heterogeneity of cancer cells introduces significant challenges in designing effective treatment strategies.

For example, heterogeneic tumours may exhibit different sensitivities to cytotoxic drugs among different clonal populations. This is attributed to clonal interactions that may inhibit or alter therapeutic efficacy.

Drug administration in heterogeneic tumours will seldom kill all tumour cells. The initial heterogeneic tumour population may bottleneck, such that few drug resistant

cells (if any) will survive. This allows resistant tumour populations to replicate and grow a new tumour through the branching evolution mechanism. The resulting repopulated tumour is heterogeneic and resistant to the initial drug therapy used. The repopulated tumour may also return in a more aggressive manner.

There is therefore a need for alternative approaches to address the issue of antigen heterogeneity in solid tumours.

Chimeric antigen receptors (CARs)

A number of immunotherapeutic agents have been described for use in cancer treatment, including therapeutic monoclonal antibodies (mAbs), bi-specific T-cell engagers and chimeric antigen receptors (CARs).

Chimeric antigen receptors are proteins which graft the specificity of a monoclonal antibody (mAb) 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 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.

To date, the success of the CAR T cell has largely been in hematological malignancies. A CAR targeted to the B cell antigen CD19 was first used successfully to treat chronic lymphoblastic leukemia (CLL). In August 2017, the FDA approved the use of CART19 (Kymriah) to treat pediatric relapsed or refractory acute lymphoblastic leukemia (ALL) and in October of the same year, another CD19-targeting CAR (Yescarta) was approved by the FDA for adult relapsed or refractory large B cell lymphoma. Additionally, the European Medicines Agency (EMA) also approved the

use of both these drugs in June of 2018. However, despite extensive research, CAR T cell therapy for solid tumours has not been nearly as successful.

Overcoming the hostile tumour microenvironment

One of the key hurdles which need to be overcome for a successful cell-based immunotherapeutic treatment of solid cancers is the tumour microenvironment (TME), which has been extensively characterized as hostile for T cells. This is true for CAR-T cells, but also other immunotherapeutic approaches such as using tumour-infiltrating lymphocyctes (TILs) and T-cells expressing engineered T-cell receptors (TCRs).

The glycolytic metabolism of tumour cells renders the environment hypoxic, acidic, low in nutrients and prone to oxidative stress. In an inflammatory environment, tumours cells often upregulate ligands such as programmed cell death ligand 1 (PD-L1) and Galectin-9 that bind to inhibitory receptors on T cells. The tumour microenvironment also relies on stromal cells like cancer associated fibroblasts (CAFs) and suppressive immune cells, including myeloid-derived suppressor cells (MDSCs), tumour associated macrophages (TAMs), tumour associated neutrophils (TANs), mast cells, and regulatory T cells (Tregs). These cells and tumour cells secrete soluble factors like vascular endothelial growth factor (VEGF) and transforming growth factor β (TGFβ), which contribute to abnormal tumour vasculature, promote anti-inflammatory polarization of TAMs and other immune cells, and are implicated in EMT. They also produce reactive oxygen species (ROS) and molecules like lactate, indoleamine 2,3-dioxygenase (IDO), prostaglandin E2 (PGE2), soluble Fas, and adenosine, which contribute to the suppression of the T cell immune response.

Administration of cytokines to polarize the tumour milieu to be more hospitable to T cells and improve CAR T cell recruitment and functionality has been tested in both preclinical and clinical trials. Local delivery of IL-12, which induces inflammatory immune cell recruitment, has been shown to augment the anti-tumor activity of adoptively transferred anti-VEGFR-2 CAR T cells in mice. CAR T cells which constitutively secrete cytokines such as IL-12, termed “armored” CARs have been created to enhance T cell infiltration and function in solid tumors. However, these approaches have limitations: systemic administration of cytokines can be toxic; constitutive production of cytokines may lead to uncontrolled proliferation and transformation (Nagarkatti et al (1994) PNAS 91:7638-7642; Hassuneh et al (1997) Blood 89:610-620).

There is therefore a need for alternative approaches to help immunotherapeutic cells, such as CAR-T cells, survive and persist in the hostile microenvironment of a solid tumour.

DESCRIPTION OF THE FIGURES

Figure 1 - Schematic diagram showing a classical chimeric antigen receptors
(a) Basic schema of a chimeric antigen receptor; (b) First generation receptors; (c) Second generation receptors; (d) Third generation receptors.

Figure 2: Frame slip motifs and construct design
(A) Diagram illustrating a frame-slip motif where a series of seven uridines are inserted into the transgene sequence, which promotes frame-slip. Frameslipping results in continued transcription/translation of an alternative reading frame, usually in the -1 direction, and the generation of a functional protein. (B) Structure of a construct showing the locations of the transgenes, the frameslip motif (SLIP) and the 2A self-cleaving peptide sequence. (C) Flow cytometric analysis SupT1 cells transduced with constructs containing the RQR8 sort-selection marker followed by either a control sequence (6xU) or a frameslip motif (7xU) and a CD22-CD19 chimeric protein consisting of the ectodomain of CD22 fused to the transmembrane and truncated endodomain of CD19. Introduction of the frameslip motif resulted in dramatically reduced expression of the CD22-CD19 chimera while similar levels of the RQR8 sort selection marker were observed.

Figure 3: Translational readthrough motifs and construct design.
(A) Examples of known translational readthrough motifs. (B-G) Structure of translational readthrough motif constructs. (B) Translational readthrough construct consisting of two transgenes where the translational readthrough motif is placed 3' of the first transgene and 5' of 2A self-cleaving peptide sequence and transgene 2. Expression of transgene 1 will be significantly higher than transgene 2. (B') Double stop translational readthrough construct where two stop codons are incorporated into the translational readthrough motif to reduce expression levels even further than those achieved with a single stop codon. (C) Universal translational readthrough construct consisting of a translational readthrough motif flanked by two 2A self-cleaving peptide sequences. This construct mitigates sequence-dependent effects that can result in unpredictable levels of translational readthrough. (D) Combinatorial

approach where a translational readthrough motif is combined with an attenuated signal peptide sequence to reduce the level of expression of transgene 2 even further. (E and E') Compound translational readthrough motifs where multiple motifs are placed in series to produce a cascade of reduced transgene expression. (E) Multiple transgenes are expressed from a single cassette using translational readthrough motifs and 2A self-cleaving peptide sequences. (E') Self-cleaving peptide sequences are used to separate the translational readthrough motifs, which are 5' of transgene 2.

Figure 4: Schematic diagram illustrating the drastic reduction in IL-12 secretion in cells transduced with a vector in which the IL-12-encoding sequence is placed downstream of a "stop-skip" sequence (SS), i.e. a frame-slip motif or a translational readthrough motif in combination with a stop codon.

Figure 5: ELISA showing concentration of IL-12 in the supernatant of cells either left untransduced (NT) or transduced with: a construct co-expressing a CAR and IL-12 (CAR-2A-IL12); or a construct co-expressing a CAR and IL-12 in which the IL-12 encoding sequence is placed downstream of a "stop-skip" sequence (CAR-SS-IL12).

Figure 6: Graph showing the weight of mice following administration of cells either left untransduced (NT) or transduced with: a construct co-expressing a CAR and IL-12 (CAR-2A-IL12); or a construct co-expressing a CAR and IL-12 in which the IL-12 encoding sequence is placed downstream of a "stop-skip" sequence (CAR-SS-IL12).

Figure 7: Graphs showing tumour size over time in mice following administration of cells either left untransduced (NT) or transduced with: a construct expressing a GD2 CAR alone (CAR); a construct co-expressing a GD2 CAR and IL-12 (CAR-2A); a construct co-expressing a GD2 CAR and IL-12 in which the IL-12 encoding sequence is placed downstream of a "stop-skip" sequence (CAR-SS); or a construct co-expressing a CAR to an irrelevant target antigen and IL-12 in which the IL-12 encoding sequence is placed downstream of a "stop-skip" sequence (MR1-SS).

SUMMARY OF ASPECTS OF THE INVENTION

The present inventors have shown in a mouse model that administration of engineered cells expressing high-level IL-12 is toxic. They have also shown that it is possible to drastically reduce the level of expression of IL-12 by the inclusion of a

motif upstream of the IL-12-encoding sequence. These ultra-low IL-12 expressing cells do not cause significant toxicity in the mouse model. Moreover, in a solid tumour model, ultra-low IL-12 expressing CAR-T cells exhibit more potent anti-tumour activity that cells expressing CAR alone.

Thus, in a first aspect, the present invention provides a method for treating a solid cancer which comprises the step of administering a cell to a subject, wherein the cell comprises a nucleic acid sequence encoding interleukin 12 (IL-12) downstream of a frame-slip motif (FSM) or a translational readthrough motif (TRM).

The method may comprise the following steps:

(i) isolation of a cell sample from a subject;

(ii) transfection or transduction of cells from the cell sample with a nucleic acid sequence encoding interleukin 12 downstream of a frame-slip motif (FSM) or a translational readthrough motif (TRM); and

(iii) administering transduced or transfected cells from step (ii) to the subject.

The nucleic acid sequence encoding IL-12, may encode "flexi-IL-12": a fusion between IL-12α and IL-12β subunits, joined by a linker. Flexi-IL-12 may comprise the sequence shown as SEQ ID No. 1.

Where the nucleic acid sequence encoding interleukin 12 (IL-12) is downstream of a frame-slip motif (FSM), the FSM may comprise a repeat of uracil, thymine or guanine bases. For example, the FSM may comprise the sequence UUUUUUU (SEQ ID No. 2).

The FSM may also comprise a stop codon. For example, the FSM may comprise one of the following sequences:

UUUUUUUGA (SEQ ID NO. 3)

UUUUUUUAG (SEQ ID NO. 4)

UUUUUUUAA (SEQ ID NO. 5).

Where the nucleic acid sequence encoding interleukin 12 (IL-12) is downstream of a translational readthrough motif (TRM), the TRM may comprise the sequence STOP-CUAG or STOP-CAAUUA, in which "STOP" is a stop codon.

The translational readthrough motif may comprise one of the following sequences:

UGA-CUAG (SEQ ID No. 6)
UAG-CUAG (SEQ ID No. 7)
UAA-CUAG (SEQ ID No. 8)
UGA-CAAUUA (SEQ ID No. 9)
UAG-CAAUUA (SEQ ID No. 10)
UAA-CAAUUA (SEQ ID No. 11)

The cell may be a tumour infiltrating immune cell, such as a T-cell, natural killer (NK) cell, NKT cell, cytokine-induced killer (CIK) cell, monocyte, macrophage or tumour-infiltrating lymphocyte (TIL).

The cell may express a chimeric antigen receptor (CAR) or engineered T-cell receptor (TCR), for example, a CAR or engineered TCR which binds to one of the following target antigens: disialoganglioside (GD2); epidermal growth factor receptor (EGFR), Epithelial cell adhesion molecule (EpCAM), Glypican 3 (GPC3), human epidermal growth factor receptor (HER2), L1CAM, Mucin 1 (MUC1), Prostate-specific membrane antigen (PSMA).

.

In one embodiment of the invention the cell expresses a CAR which binds to GD2 and has an antigen-binding domain which comprises
a) a heavy chain variable region (VH) having complementarity determining regions (CDRs) with the following sequences:
CDR1 – SYNIH (SEQ ID No. 12);
CDR2 – VIWAGGSTNYNSALMS (SEQ ID No. 13)
CDR3 – RSDDYSWFAY (SEQ ID No. 14); and
b) a light chain variable region (VL) having CDRs with the following sequences:
CDR1 – RASSSVSSSYLH (SEQ ID No. 15);
CDR2 – STSNLAS (SEQ ID No. 16)
CDR3 – QQYSGYPIT (SEQ ID No. 17).

In another embodiment of the invention, the cell expresses a CAR which binds PSMA and has an antigen-binding domain which comprises
a) a heavy chain variable region (VH) having complementarity determining regions (CDRs) with the following sequences:
CDR1 – SSWMN (SEQ ID No. 18);
CDR2 – RIYPGDGDTNYAQKFQG (SEQ ID No. 19)
CDR3 – GTGYLWYFDV (SEQ ID No. 20); and

b) a light chain variable region (VL) having CDRs with the following sequences:
CDR1 – RASQDINENLA (SEQ ID No. 21);
CDR2 – YTSNRAT (SEQ ID No. 22)
CDR3 – QQYDNLPFT (SEQ ID No. 23).

The cell may also express:

a dominant negative SHP2; and/or a dominant negative TGFβ receptor;
a chimeric cytokine receptor; and/or
one or more additional cytokine(s) or chemokine(s).

The cell may express one or more of the following: IL-7, IL-15, CCL19, CXCL12. In particular, the cell may express IL-7 and CCL19; or IL-15 and/or CXCL12.

The nucleic acid encoding the or each additional cytokine(s) or chemokine(s) may be positioned upstream of the frame-slip motif (FSM) or translational readthrough motif (TRM).

Alternatively, the nucleic acid encoding the or each additional cytokine(s) or chemokine(s) may be positioned downstream of the frame-slip motif (FSM) or translational readthrough motif (TRM).

The method of the invention may be for the treatment of a solid cancer such as: small cell lung cancer (SCLC), melanoma, renal cell cancer (RCC), hepatocellular carcinoma (HCC), ovarian cancer, pancreatic cancer, neuroblastoma or osteosarcoma.

The CAR or engineered TCR may bind a target antigen which has heterogeneous expression on the solid tumour.

In a second aspect, the present invention provides a method for inducing epitope spreading in an anti-tumour immune response in a subject, which comprises the step of administering a plurality of cells as defined in the first aspect of the invention to the subject.

In a third aspect, the present invention provides a method for inducing infiltration of immune cells into a tumour mass in a subject which comprises the step of

administering a plurality of cells as defined in the first aspect of the invention to the subject, such that they induce an anti-tumour immune response.

In a fourth aspect, there is provided a cell for use in treating a solid cancer, the cell comprising a nucleic acid sequence encoding interleukin 12 (IL-12) downstream of a frame-slip motif (FSM) or a translational readthrough motif (TRM).

In a fifth aspect, there is provided the use of a cell in the manufacture of a medicament for treating a solid cancer, wherein the cell comprises a nucleic acid sequence encoding interleukin 12 (IL-12) downstream of a frame-slip motif (FSM) or a translational readthrough motif (TRM).

In a sixth aspect, there is provided a cell which expresses a chimeric antigen receptor (CAR) and comprises;

a nucleic acid sequence encoding interleukin 12 (IL-12) downstream of a frame-slip motif (FSM) or a translational readthrough motif (TRM); and

one or more heterologous nucleic acid sequence(s) encoding one or more additional cytokine(s) or chemokine(s).

The one or more heterologous nucleic acid sequence(s) may encode one or more of the following: IL-7, IL-15, CCL19, CXCL12.

For example, the one or more heterologous nucleic acid sequence(s) may encode IL-7 and CCL19.

Alternatively, the one or more heterologous nucleic acid sequence(s) may encode IL-15 and/or CXCL12.

The nucleic acid sequence encoding an additional cytokine or chemokine may be positioned upstream or downstream of the frame-slip motif (FSM) or translational readthrough motif (TRM). Where there are two or more nucleic acid sequences encoding an additional cytokine(s) or chemokine(s), one or more may be upstream of the FSM or TRM, so it is/they are expressed at a high level; and one or more may be downstream of the FSM or TRM, so it is/they are expressed at a low level.

In a seventh aspect, there is provided a nucleic acid construct comprising:

(i) a nucleic acid sequence encoding a chimeric antigen receptor (CAR); and

(ii) a nucleic acid sequence encoding interleukin 12 (IL-12) downstream of a frame-slip motif (FSM) or a translational readthrough motif (TRM);

(vi) one or more nucleic acid sequence(s) encoding one or more additional cytokine(s) or chemokine(s) as defined above.

In an eighth aspect, there is provided a vector comprising a nucleic acid construct according to the seventh aspect of the invention.

In a ninth aspect, there is provided a kit of vectors, each encoding one or more nucleic acid sequence(s) as defined above.

In a tenth aspect. there is provided a method for making a cell according to the sixth aspect of the invention, which comprises the step of transfecting or transducing a cell with a nucleic acid construct according to the seventh aspect, a vector according to the eighth aspect or a kit of vectors according to the ninth aspect, *ex vivo*.

The present invention provides cells for the treatment of cancers, in particular solid tumours, which secrete IL-12 at very low levels.

Unlike previously developed "armoured" CAR-T cells the level of secretion of IL-12 is reduced by the incorporation of a frame-slip motif or a translational readthrough motif upstream of the IL-12 encoding sequence. The present inventors have shown that the local expression of low-level IL-12 greatly enhances the anti-tumour response but does not cause problematic levels of toxicity.

The local secretion of IL-12 has two effects: firstly, it increases the resistance of the CAR-T cells to the hostile tumour environment, increasing the proliferation and persistence of CAR T cells; secondly, it has an adjuvant effect, causing a general activation of the anti-tumour immune response.

In this respect, the release of IL-12 has various effects during an immune response, including 1) enhancing proliferation of T cells and NK cells, 2) increasing cytolytic activities of T cells, NK cells, and macrophages, 3) activating T helper 1 (Th1) cells, and 4) inducing production of IFN-γ and other cytokines. In addition to activating CAR-T cells, the it is thought that the simultaneous local expression of low-level IL-12 by CAR-T cells will lead to "epitope spreading" i.e. the priming of immune responses against additional target antigens present on tumour cells. The secretion of IL-12 by

CAR-T cells may also induce the infiltration of host immune cells into the tumour mass. The anti-tumour immune response is therefore activated through a combination of direct cell killing by CAR-T cells, infiltration of immune cells into the tumour and the induction of an anti-cancer immune response from existing host immune cells against other tumour-associated antigens.

The induction of epitope spreading is particularly useful for the treatment of solid tumours, which often show heterogeneity in expression of target antigen.

FURTHER ASPECTS

The present invention also provides additional aspects which are summarised in the following numbered paragraphs:

1. A method for treating small cell lung cancer, which comprises the step of administering a cell to a subject, wherein the cell expresses a GD2 CAR, dominant negative SHP2 (dnSHP2) and dominant negative TGFβRII.

2. A method according to paragraph 1, wherein the cell also expresses a chimeric cytokine receptor (CCR).

3. A method according to paragraph 2, where in the CCR has IL-7R endodomains.

4. A method according to paragraph 1, wherein the cell comprises a first nucleic acid construct having the following general structure:

CAR-coexpr1-dnSHP-coexpr2-dnTGFβR,
dnSHP -coexpr1-CAR-coexpr2-dnTGFβR,
CAR-coexpr1-SG-coexpr2-dnSHP-coexpr3-dnTGFβR,
dnSHP -coexpr1-SG-coexpr2-CAR-coexpr2-dnTGFβR,

in which:
dnSHP is a nucleic acid sequence encoding dominant negative SHP-2
"coexpr1", "coexpr2" and "coexpr3" which may be the same or different, are nucleic acid sequences enabling coexpression of each polypeptide as separate entities
"dnTGFβR" is a nucleic acid sequence encoding a dominant negative TGFβ receptor;

"CAR" is a nucleic acid sequence encoding an anti-GD2 chimeric antigen receptor; and

"SG" is a nucleic acid sequence encoding a suicide gene such as RQR8.

5. A method according to paragraph 2 or 3, wherein the cell comprises a second nucleic acid construct having the following general structure:

CAR-coexpr1-CCR, or

CAR-coexpr1-SG-coexpr2-CCR

in which:

"CAR" is a nucleic acid sequence encoding an anti-GD2 chimeric antigen receptor;

"coexpr1" and "coexpr2" and "coexpr3" which may be the same or different, are nucleic acid sequences enabling co-expression of each polypeptide as separate entities;

"CCR" is a nucleic acid sequence encoding a chimeric cytokine receptor; and

"SG" is a nucleic acid sequence encoding a suicide gene such as RQR8.

6. A method according to paragraph 1 which comprises the following steps:

(i) isolation of a cell-containing sample from the subject;

(ii) transduction or transfection of the cells with a vector expressing a first nucleic acid construct as defined in paragraph 4; and

(iii) administration of cells from step (ii) to the subject.

7. A method according to paragraph 2 or 3 which comprises the following steps:

(i) isolation of a cell-containing sample from the subject;

(ii) transduction or transfection of the cells with a first vector expressing a first nucleic acid construct as defined in paragraph 4 and a second vector expressing a second nucleic acid construct as defined in paragraph 5; and

(iii) administration of cells from step (ii) to the subject.

8. A cell for use in treating small cell lung cancer (SCLC), wherein the cell expresses a GD2 CAR, dominant negative SHP2 (dnSHP2) and dominant negative TGFβRII.

9. The use of a cell in the manufacture of a medicament for treating small cell lung cancer (SCLC), wherein the cell expresses a GD2 CAR, dominant negative SHP2 (dnSHP2) and dominant negative TGFβRII.

The following detailed description, as it relates to nucleic acid and polypeptide sequences, polypeptide components, vectors, cells methods etc applies equally to the aspects laid out in the above paragraphs as to the aspects of the invention in the claims.

DETAILED DESCRIPTION

The present invention provides a cell which expresses interleukin 12 (IL-12) at very low levels. The cell comprises a nucleic acid encoding IL-12 downstream of a frame-slip motif (FSM) or a translational readthrough motif (TRM).

IL-12 AND OTHER CYTOKINES/CHEMOKINES

The present invention provides a cell which comprises a nucleic acid sequence encoding interleukin 12 (IL-12) downstream of a frame-slip motif (FSM) or a translational readthrough motif (TRM). The cell may also comprise a heterologous nucleic acid sequence encoding a cytokine other than IL-12 and/or a chemokine.

Interleukin 12 (IL-12) is an interleukin that is naturally produced by dendritic cells, macrophages, neutrophils, and human B-lymphoblastoid cells in response to antigenic stimulation. IL-12 is involved in the differentiation of naive T cells into Th1 cells. It is known as a T cell-stimulating factor, which can stimulate the growth and function of T cells. It stimulates the production of interferon-gamma (IFN-γ) and tumor necrosis factor-alpha (TNF-α) from T cells and natural killer (NK) cells, and reduces IL-4 mediated suppression of IFN-γ.

IL-12 plays an important role in the activities of natural killer cells and T lymphocytes. IL-12 mediates enhancement of the cytotoxic activity of NK cells and CD8+ cytotoxic T lymphocytes.

IL-12 is a potent immunomodulatory cytokine of particular interest for modulating the tumour microenvironment redirecting the immune response against cancer. IL-12 is systemically toxic therefore methods for producing IL-12 locally are of interest.

IL-12 is a heterodimeric cytokine encoded by two separate genes, IL-12A (p35) and IL-12B (p40). The active heterodimer (referred to as 'p70'), is formed following protein synthesis.

The cell of the present invention may comprise a heterologous nucleic acid sequence encoding IL-12A and/or IL-12B. The sequence for human IL-12A is available from Uniprot Accession number P29459. A portion of this sequence, lacking the signal peptide, is shown below as SEQ ID No. 24. The sequence for human IL-12B is available from Uniprot Accession number P29460. A portion of this sequence, lacking the signal peptide, is shown below as SEQ ID No. 25.

SEQ ID No. 24 (human IL-12A)
RNLPVATPDPGMFPCLHHSQNLLRAVSNMLQKARQTLEFYPCTSEEIDHEDITKDKT
STVEACLPLELTKNESCLNSRETSFITNGSCLASRKTSFMMALCLSSIYEDLKMYQVE
FKTMNAKLLMDPKRQIFLDQNMLAVIDELMQALNFNSETVPQKSSLEEPDFYKTKIKL
CILLHAFRIRAVTIDRVMSYLNAS

SEQ ID No. 25 (human IL-12B)
WELKKDVYVVELDWYPDAPGEMVVLTCDTPEEDGITWTLDQSSEVLGSGKTLTIQV
KEFGDAGQYTCHKGGEVLSHSLLLLHKKEDGIWSTDILKDQKEPKNKTFLRCEAKNY
SGRFTCWWLTTISTDLTFSVKSSRGSSDPQGVTCGAATLSAERVRGDNKEYEYSVE
CQEDSACPAAEESLPIEVMVDAVHKLKYENYTSSFFIRDIIKPDPPKNLQLKPLKNSR
QVEVSWEYPDTWSTPHSYFSLTFCVQVQGKSKREKKDRVFTDKTSATVICRKNASI
SVRAQDRYYSSSWSEWASVPC

The heterologous nucleic acid sequence may encode "flexi-IL-12", which is a fusion between the human IL-12α (p35) and IL-12β (p40) subunits, joined by a linker. A suitable flexi-IL-12 sequence is shown below as SEQ ID No. 1.

SEQ ID No. 1 (a flexi-IL-12 sequence)
METDTLLLWVLLLWVPGSTGMWIWELKKDVYVVELDWYPDAPGEMVVLTCDTPEE
DGITWTLDQSSEVLGSGKTLTIQVKEFGDAGQYTCHKGGEVLSHSLLLLHKKEDGIW
STDILKDQKEPKNKTFLRCEAKNYSGRFTCWWLTTISTDLTFSVKSSRGSSDPQGVT
CGAATLSAERVRGDNKEYEYSVECQEDSACPAAEESLPIEVMVDAVHKLKYENYTS
SFFIRDIIKPDPPKNLQLKPLKNSRQVEVSWEYPDTWSTPHSYFSLTFCVQVQGKSK
REKKDRVFTDKTSATVICRKNASISVRAQDRYYSSSWSEWASVPC**SGGGGSGGGG**

SGGGGS*RNLPLATPDPGMFPCLHHSQNLLRAVSNMLQKARQTLEFYPCTSEEIDHE DITKDKTSTVEACLPLELTKNESCLNSRETSFITNGSCLASRKTSFMMALCLSSIYED SKMYQVEFKTMNAKLLMDPKRQIFLDQNMLAVIDELMQALNFNSETVPQKSSLEEP DFYKTKIKLCILLHAFRIRAVTIDRVMSYLNAS*

In SEQ ID No. 1, the signal peptide, which is the signal peptide from Murine kappa chain V-III region MOPC 63 (Uniprot P01661), is shown in bold; and the serine-glycine linker is in bold and underlined. Flexi-IL12 may comprise the IL-12A and B sequences shown in SEQ ID No. 1 but with a different signal peptide and/or linker sequence. It may have the general structure:

SP-IL12A-L-IL12B

where SP is a signal peptide, IL12A is human IL-12A; L is a linker sequence, such as a glycine-serine linker; and IL12B is human IL-12B.

The heterologous nucleic acid sequence may encode one of the sequences shown as SEQ ID No. 1, 24 or 25 or a variant thereof. The variant sequence may have at least 80, 85, 90, 95, 98 or 99% sequence identity to SEQ ID No. 1, 24 or 25, provided that the variant sequence retains IL-12 function when expressed *in vivo*. For example, the variant sequence may retain the capacity to enhance the activity of cytotoxic T cells *in vivo* and/or the variant sequence may stimulate the production of interferon-gamma (IFN-γ) by T cells.

The cell may comprise one or more heterologous nucleic acid(s) encoding one or more cytokines in addition to IL-12. The nucleic acid sequence(s) may, for example, encode a cytokine which enhances the inflammatory response and/or increases the efficacy of CAR-T cell therapy. The cytokine may be selected from the following: IL-7, IL15, IL-17A, IL-18, IL-2, GM-CSF and IL-21. In particular, the additional cytokine may be IL-7.

IL-7

IL-7 is a cytokine important for B and T cell development. IL-7 stimulates the differentiation of multipotent (pluripotent) hematopoietic stem cells into lymphoid progenitor cells and stimulates proliferation of all cells in the lymphoid lineage (B cells, T cells and NK cells.

Il-7 and the hepatocyte growth factor (HGF) form a heterodimer that functions as a pre-pro-B cell growth-stimulating factor. This cytokine is found to be a cofactor for V(D)J rearrangement of the T cell receptor beta (TCRß) during early T cell development. The amino acid sequence of human IL-7 is available from UniProt (Accession No. P13232) and shown below as SEQ ID No. 26.

SEQ ID No. 26 (human IL-7)
MFHVSFRYIFGLPPLILVLLPVASSDCDIEGKDGKQYESVLMVSIDQLLDSMKEIGSN
CLNNEFNFFKRHICDANKEGMFLFRAARKLRQFLKMNSTGDFDLHLLKVSEGTTILL
NCTGQVKGRKPAALGEAQPTKSLEENKSLKEQKKLNDLCFLKRLLQEIKTCWNKILM
GTKEH

IL-15

Interleukin-15 (IL-15) is a cytokine with structural similarity to Interleukin-2 (IL-2). Like IL-2, IL-15 binds to and signals through a complex composed of IL-2/IL-15 receptor beta chain and the common gamma chain. IL-15 is secreted by mononuclear phagocytes and other cells following infection by virus(es) and regulates the activation and proliferation of T and natural killer (NK) cells.

The amino acid sequence of human IL-15 is available from UniProt (Accession No. P40933) and shown below as SEQ ID No. 27.

SEQ ID No. 27 (human IL-15)
MRISKPHLRSISIQCYLCLLLNSHFLTEAGIHVFILGCFSAGLPKTEANWVNVISDLKKI
EDLIQSMHIDATLYTESDVHPSCKVTAMKCFLLELQVISLESGDASIHDTVENLIILANN
SLSSNGNVTESGCKECEELEEKNIKEFLQSFVHIVQMFINTS

The amino acid sequences of IL-17A, IL-18, IL-2, Granulocyte-macrophage colony-stimulating factor (GM-CSF) and IL-21 are available from UniProt, as shown in the table below.

Cytokine	UniProt Accession No.
IL-17A	Q16552
IL-18	Q14116
IL-2	P60568
GM-CSF	P04141

IL-21	Q9HBE4

The one or more heterologous nucleic acid sequence(s) may encode one or more chemokines.

Chemokines are a family of small cytokines secreted by cells. Their name is derived from their ability to induce directed chemotaxis in nearby responsive cells; they are chemotactic cytokines. Chemokines released by infected or damaged cells form a concentration gradient. Attracted cells move through the gradient towards the higher concentration of chemokine.

Chemokines are all approximately 8-10 kilodaltons in mass and have four cysteine residues in conserved locations that are key to forming their 3-dimensional shape. Some chemokines are considered pro-inflammatory and can be induced during an immune response to recruit cells of the immune system to a site of infection. Examples are: CXCL-8, CCL2, CCL3, CCL4, CCL5, CCL11, CXCL10. Other chemokines are considered homeostatic and are involved in controlling the migration of cells during normal processes of tissue maintenance or development. These include: CCL14, CCL19, CCL20, CCL21, CCL25, CCL27, CXCL12 and CXCL13. This classification is not strict; for example, CCL20 can act also as pro-inflammatory chemokine.

Chemokines have been classified into four main subfamilies: CXC, CC, CX3C and XC. All of these proteins exert their biological effects by interacting with G protein-linked transmembrane receptors called chemokine receptors, that are selectively found on the surfaces of their target cells

The cell may express a homeostatic chemokine such as CCL19 or CXCL12.

Chemokine (C-C motif) ligand 19 (CCL19) is a small cytokine belonging to the CC chemokine family that is also known as EBI1 ligand chemokine (ELC) and macrophage inflammatory protein-3-beta (MIP-3-beta). CCL19 elicits its effects on its target cells by binding to the chemokine receptor chemokine receptor CCR7. It attracts certain cells of the immune system, including dendritic cells and antigen-engaged B cells and CCR7+ central-memory T-cells. The amino acid sequence for human CCL19 is available from UniProt (Accession number Q99731) and shown below as SEQ ID No. 28.

SEQ ID No. 28 (CCL19)
MALLLALSLLVLWTSPAPTLSGTNDAEDCCLSVTQKPIPGYIVRNFHYLLIKDGCRVP
AV VFTTLRGRQLCAPPDQPWVERIIQRLQRTSAKMKRRSS

C-X-C motif chemokine 12 (CXCL12), stromal cell-derived factor 1 (SDF1), is produced in two forms, CXCL12a and CXCL12b, by alternate splicing of the same gene. Chemokines are characterized by the presence of four conserved cysteines, which form two disulfide bonds. The CXCL12 proteins belong to the group of CXC chemokines, whose initial pair of cysteines are separated by one intervening amino acid. In addition, the first 8 residues of the CXCL12 N-terminal serve as a receptor binding site, and the RFFESH motif (residues 12-17, SEQ ID NO: 102) in the loop region function as a docking site for CXCL12 receptor binding.

CXCL12 is expressed in many tissues in mice including brain, thymus, heart, lung, liver, kidney, spleen, and bone marrow. CXCL12 is strongly chemotactic for lymphocytes. The amino acid sequence for human CXCL12 is available from UniProt (Accession number P48061) and shown below as SEQ ID No. 29.

SEQ ID No. 29 (CXCL12)
MNAKVVVVLVLVLTALCLSDGKPVSLSYRCPCRFFESHVARANVKHLKILNTPNCAL
QIVARLKNNNRQVCIDPKLKWIQEYLEKALNKRFKM

STOP SKIP MOTIFS

FRAME-SLIP

During transcription, RNA polymerase catalyzes incorporation of nucleotides into growing RNA chains on the basis of complementarity to the DNA template. When RNA polymerase encounters a stretch of repeated bases, however, slippage or 'stuttering' can occur. Transcription slippage is utilized in nature, for example for regulation of the Escherichia coli pyrBI and codBA operons, expression of the P gene in paramyxoviruses and decoding of the cellular dnaX gene of Thermus thermophiles.

When RNA polymerase slips this can result in the synthesis of an mRNA encoding an alternative reading frame because it lacks one (or maybe two) of the repeated bases.

The nucleic acid construct of the present invention may comprise a transcriptional frameslip site, such that transcription of the downstream transgene only occurs if these is slippage of the RNA polymerase resulting in mRNA encoding an alternative reading frame.

Translation of an mRNA sequence to protein is a complex process involving the orchestration of ribosomes, initiation and elongation factors, aminoacyl transfer RNAs (aa-tRNAs), aminoacyl tRNA synthetases and release factor. The initiation of translation begins when a complex of factors binds to the 5' end of the mRNA, which then results in the recruitment of 40S ribosomes and the scanning of the mRNA. When the complex of initiation factors and 40S ribosome encounters a codon (a triplet of nucleotides) encoding the amino acid methionine (AUG codon), the 60S ribosome is recruited and polypeptide synthesis begins with the pairing of cognate tRNAs loaded with the appropriate amino acid. Initiation of translation can occur at alternative start codons other than AUG, but it is the most frequently used start codon. Extension of the polypeptide occurs in a cyclical manner, whereby a tRNA binds to its cognate codon, a peptide bond forms between the newly added amino acid and the extending polypeptide, and the polypeptide translocates to expose the next codon.

During translation, ribosome pausing may occur where the ribosome arrests at a particular codon. Ribosome pausing can promote mRNA degradation by a nucleocytic pathway or it can induce a slip in the -1 or -2 direction. Repetitive sequences within an mRNA, such as UUUUUUA, are known to induce translational frameslipping, and this sequence is present at the 3' end of the group-specific antigen gene (gag) and polymerase (pol) gene of the human immunodeficiency virus (HIV). Translational frameslipping -1 of the HIV gag/pol gene results in the expression of the Gag-Pol polyprotein.

The nucleic acid construct of the present invention may comprise a translational frameslipping site, such that translation of the downstream transcript only occurs if these is a frameslip by the ribosome.

The frame-slip site may comprise a stretch of bases of the same type.

The frame-slip motif may be placed upstream and/or downstream of a cleavage site in the nucleic acid construct. The frame-slip motif is located between first and second transgenes in the nucleic acid construct.

The frame-slip motif may be used alone. In this embodiment, the second transgene may be placed out of frame downstream of the frame-slipping site, such that frame-slip is needed for transcription or translation of the second transgene.

The motif may, for example comprise a stretch of 5, 7, 8, 10 or 11 bases. The site may, for example, comprise a repeat of uracil, adenosine or guanine bases. The site may, for example, comprise the sequence shown as SEQ ID No. 2

UUUUUUU (SEQ ID No. 2).

Alternatively, the frame-slip site may be used in combination with a stop codon. In this embodiment, a stop codon is positioned in frame downstream of the frame-slip site, such that frame-slipping is required for the stop codon to be ignored.

The stop codon may be UGA, UAG or UAA. The frame-slip site may comprise a repeat of uracil, adenosine or guanine bases in multiples of three, for example, 3, 6 or 9 repeats of uracil, adenosine or guanine bases.

Examples of frame-slip site/stop codon combinations are shown as SEQ ID No. 3, 4 and 5.

UUUUUUUGA (SEQ ID NO. 3)
UUUUUUUAG (SEQ ID NO. 4)
UUUUUUUAA (SEQ ID NO. 5).

TRANSLATIONAL READTHROUGH

Translation terminates when the ribosome encounters a UGA, UAG or UAA stop codon. At this point, release factor recognises the stop codon and facilitates dissociation and recycling of the ribosome. Termination of translation usually occurs with high fidelity, with recoding of the stop codon, due to competition between release factor and a near cognate tRNA, and continued extension of the polypeptide only occurring 0.1% of the time. However, an elevated level of stop codon of recoding,

which results in translational readthrough, has been reported in certain genes. In some cases, this has resulted in the generation of a longer polypeptide with additional functional motifs, in a process referred to as functional readthrough.

Translational readthrough occurs when release factor 1 (RF1) fails to recognise a stop codon and a near cognate aa-tRNA inserts an amino acid into the extending polypeptide, thereby suppressing the stop codon. The local concentration of release factor and the aa-tRNAs affects the level of stop codon suppression and translational readthrough, with low concentrations of release factor promoting translational readthrough. In mammals suppression of the UAG stop codon results in the insertion of a tryptophan, arginine or cysteine residue.

One of the earliest discovered examples of stop codon suppression is the rabbit beta globin gene, where translational readthrough results in the addition of 22 amino acids to the C-terminus of the protein 3.

The frequency of translational readthrough depends on a number of factors including: 1) the stop codon used (UGA, UAG or UAA); 2) the immediate sequence flanking the stop codon, with the six nucleotides upstream and downstream of the stop codon being particularly important and; 3) the presence of cis-acting sequences in the 3' end of the mRNA.

The termination efficiency of the three stop codons varies, with UAA being the strongest stop codon and UGA being the weakest and hierarchy of termination efficiency is defined as UAA>UAG>UGA. Consequently, the highest level of translational readthrough is exhibited with the UGA stop codon and the lowest with the UAA stop codon.

Sequence analysis of genes exhibiting translational readthrough has identified at least two different motifs that promote stop codon suppression and sustained translation: STOP-CUAG and STOP-CAAUUA (where stop can be UGA, UAG or UAA). The level of translational readthrough is dependent on the stop codon used: UGA supports the highest level of translational readthrough and decreasing levels of readthrough are obtained from the UAG and UAA stop codons, so the overall hierarchy of readthrough is UGA>UAG>UAA.

The nucleic acid construct of the present invention may comprise one of the sequences shown as SEQ ID No. 6 to 11.

UGACUAG (SEQ ID No. 6)
UAGCUAG (SEQ ID No. 7)
UAACUAG (SEQ ID No. 8)
UGACAAUUA (SEQ ID No. 9)
UAGCAAUUA (SEQ ID No. 10)
UAACAAUUA (SEQ ID No. 11)

The translational readthrough site may be located between first and second transgenes in the nucleic acid construct. The translational readthrough site may be placed upstream and/or downstream of a cleavage site in the nucleic acid construct. The translational readthrough site may be flanked by cleavage sites (Figure 3C). Two or more translational readthrough sites may be used, for example, either positioned next to each other (Figure 3B'), or flanking a cleavage site (Figure E').

To reduce the expression level of a downstream transgene further, multiple stop codons can inserted 5' of the translational readthrough motif (Figure 2B). An example multiple stop translational readthrough motif is shown as SEQ ID NO. 30, where two UGA stop codons are positioned 5' of the CUAG readthrough motif.

UGAUGACUAG (SEQ ID No. 30).

The above sequences are presented as RNA, but in the construct they may be the equivalent DNA sequence which produces these RNA sequences once transcribed, which for SEQ ID No. 6-11 above would be:

TGACTAG (SEQ ID No. 31)
TAGCTAG (SEQ ID No. 32)
TAACTAG (SEQ ID No. 33)
TGACAATTA (SEQ ID No. 34)
TAGCAATTA (SEQ ID No. 35)
TAACAATTA (SEQ ID No. 36)

In particular a DNA sequence encoding a TRM may have the sequence shown as SEQ ID No. 37, which codes for STOP; Leucine; Alanine

TGACTAGCA (SEQ ID No. 37)

Where a nucleic acid construct comprises more than two transgenes, compound translational readthrough motifs may be placed in series, 5' for sequences encoding cleavage sites. This enables multiple transgenes to be expressed at different ratios. With each additional readthrough motif the level of expression should be reduced 10 to 50-fold relative to the upstream transgene.

The nucleic acid sequence encoding IL-12 may be positioned downstream of a frame-slip motif (FSM) or a translational readthrough motif (TRM). The nucleic acid construct may have the general structure:

CAR-FSM/TRM-coexpr-IL12

in which:
"CAR" is a nucleic acid sequence encoding a chimeric antigen receptor
" FSM/TRM" is a frame-slip motif or a translational readthrough motif
"coexpr" is a sequence enabling the co-expression of the CAR and IL-12 as separate polypeptides; and
"IL-12" is a nucleic acid sequence encoding IL-12 or flexi-IL12

In constructs which comprise a nucleic acid sequence encoding another cytokine or chemokine or a chemokine, this may also be placed downstream of the FSM or TRM, such that the level of expression of the other cytokine(s) or chemokine(s) is also reduced (compared to the level of expression which would be obtained in the absence of the FSM or TRM sequence). In this respect, the level of expression of the other cytokine(s) or chemokine(s) may be reduced comparable to that of IL-12. Such a construct may have the general structure:

CAR-FSM/TRM-coexpr1-IL12-coexpr2-CC; or
CAR-FSM/TRM-coexpr1-CC-coexpr2-IL12

in which:
"CAR" is a nucleic acid sequence encoding a chimeric antigen receptor
" FSM/TRM" is a frame-slip motif or a translational readthrough motif

"coexpr1" and "coexpr2", which may be the same or different, are sequence enabling the co-expression of the CAR, IL-12 and the cytokine or chemokine as separate polypeptides;
"IL-12" is a nucleic acid sequence encoding IL-12 or flexi-IL12; and
"CC" is a nucleic acid sequence encoding a cytokine (other than IL-12) or a chemokine.

Alternatively, the nucleic acid sequence encoding the other cytokine or encoding a chemokine may be placed upstream of the FSM or TRM or on a separate construct, such that its expression is unaffected by the FSM or TRM. With such an arrangement, the level of expression of other cytokine(s) or chemokine(s) will exceed the level of expression of Il-12. Such a construct may have the general structure:

CAR-coexpr1-CC-FSM/TRM-coexpr2-IL12; or
CC-coexpr1-CAR-FSM/TRM-coexpr2-IL12

in which:
"CAR" is a nucleic acid sequence encoding a chimeric antigen receptor
" FSM/TRM" is a frame-slip motif or a translational readthrough motif
"coexpr1" and "coexpr2", which may be the same or different, are sequence enabling the co-expression of the CAR, IL-12 and the cytokine or chemokine as separate polypeptides;
"IL-12" is a nucleic acid sequence encoding IL-12 or flexi-IL12; and
"CC" is a nucleic acid sequence encoding a cytokine (other than IL-12) or a chemokine.

A third option is the nucleic acid sequence encoding other cytokine(s) or chemokine(s) may be placed downstream of a FSM or TRM which is different to the one upstream of the IL-12 encoding sequence. This may have "compound" effects. For example, a nucleic acid construct with compound translational readthrough motifs may have the structure:

CAR- FSM/TRM1-coexpr1-CC-FSM/TRM2-coexpr2-IL12

in which:
"CAR" is a nucleic acid sequence encoding a chimeric antigen receptor

" FSM/TRM1" and "FSM/TRM2", which may be the same of different is/are frame-slip motif(s) or translational readthrough motif(s);
"coexpr1" and "coexpr2", which may be the same or different, are sequence enabling the co-expression of the CAR, IL-12 and the cytokine or chemokine as separate polypeptides;
"IL-12" is a nucleic acid sequence encoding IL-12 or flexi-IL12; and
"CC" is a nucleic acid sequence encoding a cytokine (other than IL-12) or a chemokine.

In this arrangement, the level of expression of CAR will be higher than the level of expression of the cytokine/chemokine, which will be higher than the level of expression of IL-12.

ALTERED SIGNAL PEPTIDES

While translational readthrough motifs significantly decrease transgene expression, in certain situations it might be necessary to reduce the expression levels even further. For secreted proteins a further reduction in expression can be achieved by combining a translational readthrough motif with an altered signal peptide sequence. In this case, the altered signal peptide and second transgene are placed 3' of the translational readthrough motif and self-cleaving peptide sequence (Figure 3D).

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.

Altered signal peptides are described in detail in WO2016/174409, which is herein incorporated by reference. The altered signal peptide may comprise one or more mutation(s), such as substitutions or deletions, 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.

Where the nucleic acid construct comprises two transgenes both encoding transmembrane proteins, the protein encoded by the downstream transgene (which has lower relative expression) may comprise fewer hydrophobic amino acids than the protein encoded by the upstream transgene (which has higher relative expression).

The hydrophobic amino acids mutated in order to alter signal peptide efficiency may be: Alanine (A); Valine (V); Isoleucine (I); Leucine (L); Methionine (M); Phenylalanine (P); Tyrosine (Y); or Tryptophan (W).

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.

CLEAVAGE SITE

The invention provides a nucleic acid construct comprising:

(i) a nucleic acid sequence encoding a chimeric antigen receptor (CAR); and

(ii) a nucleic acid sequence encoding interleukin 12 (IL-12) downstream of an FSM/TRM.

The nucleic acid construct may also comprise one or more of the following:

(iii) a nucleic acid sequence encoding a dominant negative SHP2;

(iv) a nucleic acid sequence encoding a dominant negative TGFβ receptor;

(v) a nucleic acid sequence encoding a chimeric cytokine receptor:

(vi) one or more nucleic acid sequence(s) encoding one or more additional cytokine(s) or chemokine(s).

In order that the CAR and IL-12 and optionally one or more additional components may be expressed as separate polypeptides, the construct may comprise a sequence encoding a cleavage site positioned between nucleic acid sequences which encode the two polypeptides.

The cleavage site may be any sequence which enables the polypeptide produced by translation of the nucleic acid construct to split or become separated.

The term “cleavage” is used herein for convenience, but the cleavage site may cause two polypeptides 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' (SEQ ID No. 38)) 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 (SEQ ID No. 39) (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 (SEQ ID No. 40) 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 1.

Table 1

Motif	Present in:
DIETNPGP (SEQ ID No. 41)	Picornaviruses EMCB, EMCD, EMCPV21
DVETNPGP (SEQ ID No. 42)	Picornaviruses MENGO and TMEBEAN; Insect virus DCV, ABPV
DVEMNPGP (SEQ ID No. 43)	Picornaviruses TMEGD7 and TMEBEAN
DVESNPGP (SEQ ID No. 44)	Picornaviruses FMDA10, FMDA12, FMDC1, FMD01K, FMDSAT3, FMDVSAT2, ERAV; Insect virus CrPV
DMESNPGP (SEQ ID No. 45)	Picornavirus FMDV01G
DVELNPGP (SEQ ID No. 103)	Picornavirus ERBV; Porcine rotavirus
DVEENPGP (SEQ ID No. 46)	Picornavirus PTV-1; Insect virus TaV; Trypanosoma TSR1
DIELNPGP (SEQ ID No. 47)	Bovine Rotavirus, human rotavirus
DIEQNPGP (SEQ ID No. 48)	Trypanosoma AP endonuclease
DSEFNPGP (SEQ ID No. 49)	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. 50)
LLNFDLLKLAGDVQSNPGP (SEQ ID No. 51)
LLNFDLLKLAGDVEINPGP (SEQ ID No. 52)
LLNFDLLKLAGDVEFNPGP (SEQ ID No. 53)
LLNFDLLKLAGDVESHPGP (SEQ ID No. 54)
LLNFDLLKLAGDVESEPGP (SEQ ID No. 55)
LLNFDLLKLAGDVESQPGP (SEQ ID No. 56)
LLNFDLLKLAGDVESNPGG (SEQ ID No. 57)

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. 58)
HYAGYFADLLIHDIETNPGP (SEQ ID No. 59)
QCTNYALLKLAGDVESNPGP (SEQ ID No. 60)
ATNFSLLKQAGDVEENPGP (SEQ ID No. 61)
AARQMLLLLSGDVETNPGP (SEQ ID No. 62)
RAEGRGSLLTCGDVEENPGP (SEQ ID No. 63)
TRAEIEDELIRAGIESNPGP (SEQ ID No. 64)
TRAEIEDELIRADIESNPGP (SEQ ID No. 65)
AKFQIDKILISGDVELNPGP (SEQ ID No. 66)
SSIIRTKMLVSGDVEENPGP (SEQ ID No. 67)
CDAQRQKLLLSGDIEQNPGP (SEQ ID No. 68)
YPIDFGGFLVKADSEFNPGP (SEQ ID No. 69)

The cleavage site may comprise the 2A-like sequence shown as SEQ ID No. 63 (RAEGRGSLLTCGDVEENPGP) or 2A peptide from Thosea asigna virus capsid protein, which has the sequence EGRGSLLTCGDVEENPGP (SEQ ID No. 70)

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. 71)
LLAIHPTEARHKQKIVAPVKQTLNFDLLKLAGDVESNPGP (SEQ ID No. 72)
EARHKQKIVAPVKQTLNFDLLKLAGDVESNPGP (SEQ ID No. 73)
APVKQTLNFDLLKLAGDVESNPGP (SEQ ID No. 74)

CHIMERIC ANTIGEN RECEPTORS (CARS)

CARs, which are shown schematically in Figure 1, are chimeric type I trans-membrane proteins which connect an extracellular antigen-recognizing domain (binder) to an intracellular signalling domain (endodomain). The binder is typically a single-chain variable fragment (scFv) derived from a monoclonal antibody (mAb), but it can be based on other formats which comprise an antibody-like antigen binding site. A spacer domain is usually necessary to isolate the binder from the membrane and to allow it a suitable orientation. A common spacer domain used is the Fc of IgG1. More compact spacers can suffice e.g. the stalk from CD8α and even just the IgG1 hinge alone, depending on the antigen. A trans-membrane domain anchors the protein in the cell membrane and connects the spacer to the endodomain.

Early CAR designs had endodomains derived from the intracellular parts of either the γ chain of the FcεR1 or CD3ζ. Consequently, these first generation receptors transmitted immunological signal 1, which was sufficient to trigger T-cell killing of cognate target cells but failed to fully activate the T-cell to proliferate and survive. To overcome this limitation, compound endodomains have been constructed: fusion of the intracellular part of a T-cell co-stimulatory molecule to that of CD3ζ results in second generation receptors which can transmit an activating and co-stimulatory signal simultaneously after antigen recognition. The co-stimulatory domain most

commonly used is that of CD28. This supplies the most potent co-stimulatory signal - namely immunological signal 2, which triggers T-cell proliferation. Some receptors have also been described which include TNF receptor family endodomains, such as the closely related OX40 and 41BB which transmit survival signals. Even more potent third generation CARs have now been described which have endodomains capable of transmitting activation, proliferation and survival signals.

CAR-encoding nucleic acids may be transferred to T cells using, for example, retroviral or lentiviral vectors to generate cancer-specific T cells for adoptive cell transfer. When the CAR binds the target-antigen, this results in the transmission of an activating signal to the T-cell it is expressed on. Thus, the CAR directs the specificity and cytotoxicity of the T cell towards tumour cells expressing the targeted antigen.

Bispecific CARs, known as tandem CARs or TanCARs, have been developed to target two or more cancer specific markers simultaneously. In a TanCAR, the extracellular domain comprises two antigen binding specificities in tandem, joined by a linker. The two binding specificities (scFvs) are thus both linked to a single transmembrane portion: one scFv being juxtaposed to the membrane and the other being in a distal position. When a TanCAR binds either or both of the target antigens, this results in the transmission of an activating signal to the cell on which it is expressed.

ANTIGEN BINDING DOMAIN

The antigen binding domain is the portion of CAR which recognizes antigen. Numerous antigen-binding domains are known in the art, including those based on the antigen binding site of an antibody, antibody mimetics, and T-cell receptors. For example, the antigen-binding domain may comprise: a single-chain variable fragment (scFv) derived from a monoclonal antibody; a natural ligand of the target antigen; a peptide with sufficient affinity for the target; a single domain antibody; an artificial single binder such as a Darpin (designed ankyrin repeat protein); or a single-chain derived from a T-cell receptor.

In a classical CAR, the antigen-binding domain comprises: a single-chain variable fragment (scFv) derived from a monoclonal antibody. CARs have also been produced with domain antibody (dAb) or VHH antigen binding domains or which

comprise a Fab fragment of, for example, a monoclonal antibody. A FabCAR comprises two chains: one having an antibody-like light chain variable region (VL) and constant region (CL); and one having a heavy chain variable region (VH) and constant region (CH). One chain also comprises a transmembrane domain and an intracellular signalling domain. Association between the CL and CH causes assembly of the receptor.

The two chains of a Fab CAR may have the general structure:

VH - CH - spacer - transmembrane domain - intracellular signalling domain; and
VL - CL

or

VL - CL - spacer- transmembrane domain - intracellular signalling domain; and
VH - CH

For Fab-type chimeric receptors, the antigen binding domain is made up of a VH from one polypeptide chain and a VL from another polypeptide chain.

The polypeptide chains may comprise a linker between the VH/VL domain and the CH/CL domains. The linker may be flexible and serve to spatially separate the VH/VL domain from the CH/CL domain.

The antigen-binding domain of the CAR may bind a tumour associated antigen. Various tumour associated antigens (TAA) are known, for example as shown in the following Table 2.

Table 2

Cancer type	**TAA**
Diffuse Large B-cell Lymphoma	CD19, CD20, CD22
Breast cancer	ErbB2, MUC1
AML	CD13, CD33
Neuroblastoma	GD2, NCAM, ALK, GD2
B-CLL	CD19, CD52, CD160
Colorectal cancer	Folate binding protein, CA-125

Chronic Lymphocytic Leukaemia	CD5, CD19
Glioma	EGFR, Vimentin
Multiple myeloma	BCMA, CD138
Renal Cell Carcinoma	Carbonic anhydrase IX, G250
Prostate cancer	PSMA
Bowel cancer	A33

The CAR may bind one of the targets summarised in Table 3 below. In particular, the CAR may bind one of the following target antigens: disialoganglioside (GD2); epidermal growth factor receptor (EGFR), Epithelial cell adhesion molecule (EpCAM), Glypican 3 (GPC3), human epidermal growth factor receptor (HER2), L1CAM, Mucin 1 (MUC1), or Prostate-specific membrane antigen (PSMA).

In one embodiment of the invention the CAR may bind an antigen other than PSMA. The CAR may bind one of the target antigens listed in Table 3 which is not PSMA. For example, the CAR may bind one of the following target antigens: disialoganglioside (GD2); epidermal growth factor receptor (EGFR), Epithelial cell adhesion molecule (EpCAM), Glypican 3 (GPC3), human epidermal growth factor receptor (HER2), L1CAM and Mucin 1 (MUC1).

Disialoganglioside GD2

CARs have been developed which bind disialoganglioside (GD2) a sialic acid-containing glycosphinolipid. Such CARs may, for example, be based on the GD2 binder 14g2a, or huK666 as described in WO2015/132604.

A CAR which binds GD2 may have an antigen-binding domain which comprises:

a) a heavy chain variable region (VH) having complementarity determining regions (CDRs) with the following sequences:

CDR1 – SYNIH (SEQ ID No. 12);

CDR2 – VIWAGGSTNYNSALMS (SEQ ID No. 13)

CDR3 – RSDDYSWFAY (SEQ ID No. 14); and

b) a light chain variable region (VL) having CDRs with the following sequences:

CDR1 – RASSSVSSSYLH (SEQ ID No. 15);

CDR2 – STSNLAS (SEQ ID No. 16)

CDR3 – QQYSGYPIT (SEQ ID No. 17).

The GD2 binding domain may comprise a VH domain having the sequence shown as SEQ ID No. 75; and/or a VL domain having the sequence shown as SEQ ID No 76.

SEQ ID No. 75 (Humanised KM666 VH sequence)
QVQLQESGPGLVKPSQTLSITCTVSGFSLASYNIHWVRQPPGKGLEWLGVIWAGGS
TNYNSALMSRLTISKDNSKNQVFLKMSSLTAADTAVYYCAKRSDDYSWFAYWGQG
TLVTVSS

SEQ ID No. 76 (Humanised KM666 VL sequence)
ENQMTQSPSSLSASVGDRVTMTCRASSSVSSSYLHWYQQKSGKAPKVWIYSTSNL
ASGVPSRFSGSGSGTDYTLTISSLQPEDFATYYCQQYSGYPITFGQGTKVEIK

The antigen-binding domain of the CAR may comprise a variant of SEQ ID NO: 75 or 76 having at least 80, 85, 90, 95, 98 or 99% sequence identity, provided that the variant sequence retains the capacity to bind GD2.

Epidermal growth factor receptor (EGFR)

Epidermal growth factor receptor (EGFR; ErbB-1; HER1 in humans) is a transmembrane protein that is a receptor for members of the epidermal growth factor family (EGF family) of extracellular protein ligands. The amino acid sequence for human EpCAM is available from UniProt, Accession No. P00533. EGFR undergoes a transition from an inactive monomeric form to an active homodimer. EGFR dimerization stimulates its intrinsic intracellular protein-tyrosine kinase activity, causing autophosphorylation of several tyrosine (Y) residues in the C-terminal domain, including Y992, Y1045, Y1068, Y1148 and Y1173. Autophosphorylation elicits downstream activation and signaling by several other proteins that associate with the phosphorylated tyrosines through their own phosphotyrosine-binding SH2 domains. These downstream signaling proteins initiate several signal transduction cascades, principally the MAPK, Akt and JNK pathways, leading to DNA synthesis and cell proliferation.

Over-expression of EGFR is associated with the development of a wide variety of tumours. Mutations that lead to EGFR overexpression (known as upregulation or amplification) have been associated with a number of cancers, including adenocarcinoma of the lung (40% of cases), anal cancers, glioblastoma (50%) and epithelial tumors of the head and neck (80-100%). These somatic mutations involving EGFR lead to its constant activation, which produces uncontrolled cell division. In

glioblastoma a specific mutation of EGFR, called EGFRvIII, is often observed. Mutations, amplifications or misregulations of EGFR or family members are implicated in about 30% of all epithelial cancers.

Several anti-EGFR antibodies are known, for example, cetuximab, panitumumab, zalutumumab, nimotuzumab, and matuzumab.

Epithelial cell adhesion molecule (EpCAM)

Epithelial cell adhesion molecule (EpCAM) is a transmembrane glycoprotein mediating Ca2+-independent homotypic cell–cell adhesion in epithelia. EpCAM is also involved in cell signaling, migration, proliferation, and differentiation. Additionally, EpCAM has oncogenic potential via its capacity to upregulate c-myc, e-fabp, and cyclins A & E. Since EpCAM is expressed exclusively in epithelia and epithelial-derived neoplasms, EpCAM can be used as diagnostic marker for various cancers. It appears to play a role in tumorigenesis and metastasis of carcinomas, so it can also act as a prognostic marker and as a target for immunotherapeutic strategies. EpCAM is often overexpressed in certain carcinomas, including in breast cancer, colon cancer and basal cell carcinoma of the skin.

EpCAM is a glycosylated, 30- to 40-kDa type I membrane protein. The amino acid sequence for human EpCAM is available from UniProt, Accession No. P16422. EpCAM can be cleaved which lends the molecule oncogenic potential. Upon cleavage, the extracellular domain (EpEX) is released into the area surrounding the cell, and the intracellular domain (EpICD) is released into the cytoplasm of the cell. EpICD forms a complex with the proteins FHL2, β-catenin, and Lef inside the nucleus. This complex then binds to DNA and promotes the transcription of various genes. Targets of upregulation include c-myc, e-fabp, and cyclins A & E. This has the effect of promoting tumour growth. Additionally, EpEX that has been cleaved can stimulate the cleavage of additional EpCAM molecules resulting in a positive feedback loop. EpCAM may also play a role in epithelial mesenchymal transition (EMT) in tumours.

Since EpCAM is expressed mostly on the basolateral membrane in normal epithelia, it should be much less accessible to antibodies than EpCAM in cancer tissue, where it is homogeneously distributed on the cancer cell surface. In addition to being overexpressed in many carcinomas, EpCAM is expressed in cancer stem cells, making EpCAM an attractive target for immunotherapy. However, the heterogeneous

expression of EpCAM in carcinomas and the fact that EpCAM is not tumor-specific (i.e., it is found in normal epithelium) has hampered development of immunotherapy approaches directed towards EpCAM to date.

Several anti-EpCAM antibodies are known including murine IgG2a edrecolomab and its murine/human chimeric IgG1 antibody version, and humanized, human-engineered and fully human IgG1 antibodies 3622W94, ING-1, and adecatumumab (MT201), respectively.

Glypican 3 (GPC3)

Glypican-3 plays a role in the control of cell division and growth regulation. The protein core of GPC3 consists of two subunits, where the N-terminal subunit has a size of ~40 kDa and the C-terminal subunit is ~30 kDa. The amino acid sequence for human GPC3 is available from UniProt, Accession No. P51654.

Six glypicans (GPC1-6) have been identified in mammals. Cell surface heparan sulfate proteoglycans are composed of a membrane-associated protein core substituted with a variable number of heparan sulfate chains. Members of the glypican-related integral membrane proteoglycan family (GRIPS) contain a core protein anchored to the cytoplasmic membrane via a glycosyl phosphatidylinositol linkage. GPC3 interacts with both Wnt and frizzled (FZD) to form a complex and trigger downstream signalling.

GPC3 is a promising therapeutic target for treating liver cancer. Several therapeutic anti-GPC3 antibodies have been developed, including GC33, HN3 and YP7. YP7 and its humanised form hYP7 binds the C-lobe of GPC3; the single-domain antibody, HN3, targets the N-lobe of GPC3; whereas the human monoclonal antibody targets the heparan sulfate moiety of GPC3. Chimeric antigen receptor (CAR) T cell immunotherapies based on GC33, hYP7 and HN3 are being developed at various stages for treating liver cancer.

Human epidermal growth factor receptor (HER2)

HER2 (from human epidermal growth factor receptor 2) is also known as HER2/neu, Receptor tyrosine-protein kinase erbB-2, CD340 (cluster of differentiation 340), proto-oncogene Neu, Erbb2 (rodent), and ERBB2 (human.

HER2 is a member of the human epidermal growth factor receptor (HER/EGFR/ERBB) family. Amplification or over-expression of this oncogene has been shown to play an important role in the development and progression of certain aggressive types of breast cancer. In recent years the protein has become an important biomarker and target of therapy for approximately 30% of breast cancer patients.

The ErbB family consists of four plasma membrane-bound receptor tyrosine kinases. One of which is erbB-2, and the other members being epidermal growth factor receptor, erbB-3 (neuregulin-binding; lacks kinase domain), and erbB-4. All four contain an extracellular ligand binding domain, a transmembrane domain, and an intracellular domain that can interact with a multitude of signaling molecules and exhibit both ligand-dependent and ligand-independent activity. HER2 can heterodimerise with any of the other three receptors and is considered to be the preferred dimerisation partner of the other ErbB receptors. Dimerisation results in the autophosphorylation of tyrosine residues within the cytoplasmic domain of the receptors and initiates a variety of signaling pathways. The amino acid sequence for human HER2 is available from UniProt, Accession No. P04626.

HER2 is the target of the monoclonal antibody trastuzumab (marketed as Herceptin). Trastuzumab is effective only in cancers where HER2 is over-expressed. An important downstream effect of trastuzumab binding to HER2 is an increase in p27, a protein that halts cell proliferation. Another monoclonal antibody, Pertuzumab, which inhibits dimerisation of HER2 and HER3 receptors, was approved by the FDA for use in combination with trastuzumab in June 2012.

L1 cell adhesion molecule (L1CAM)

L1CAM, also known as L1, is a transmembrane protein member of the L1 protein family, encoded by the L1CAM gene. This protein, of 200-220 kDa, is formed of six immunoglobulin domains followed by five fibronectin type III domains which are connected to a small intracellular domain by a transmembrane helix. The amino acid sequence for human L1CAM is available from UniProt, Accession No. P32004.

L1CAM is a neuronal cell adhesion molecule with a strong implication in cell migration, adhesion, neurite outgrowth, myelination and neuronal differentiation. L1CAM is located all over the nervous system on the surface of neurons. It is placed along the cellular membrane so that one end of the protein remains inside the nerve

cell while the other end stays on the outer surface of the neurone. This position allows the protein to activate chemical signals which spread through the neurone.

A wide variety of cells express L1CAM, including immature oligodendrocytes and Schwann cells, which are non-neuronal cells that provide support and protection for neurons and form myelin; T cells which are lymphocytes involved in cell-mediated immunity; other types of lymphocytes such as B cells and monocytes. L1CAM is expressed in multiple tumour types, for example melanoma and lung carcinoma cells.

Wolterink *et al* (2010) Cancer Res 15:2504-2515 describe the generation a series of novel monoclonal antibodies (mAb) to the L1CAM ectodomain, including the antibody L1-9.3.

Mucin 1 (MUC1)

Mucin 1, cell surface associated (MUC1), also called polymorphic epithelial mucin (PEM) or epithelial membrane antigen or EMA, is a mucin encoded by the MUC1 gene in humans. MUC1 is a glycoprotein with extensive O-linked glycosylation of its extracellular domain. Mucins line the apical surface of epithelial cells in the lungs, stomach, intestines, eyes and several other organs. Mucins protect the body from infection by pathogen binding to oligosaccharides in the extracellular domain, preventing the pathogen from reaching the cell surface. Overexpression of MUC1 is often associated with colon, breast, ovarian, lung and pancreatic cancers. The amino acid sequence for human MUC1 is available from UniProt, Accession No. P15941.

Numerous anti-mucin 1 (anti-MUC1) antibodies have been developed and characterised, including 1B2 and 12D10 which recognise specific O-glycan structures at the PDT*R motif (the asterisk represents an O-glycosylation site, SEQ ID NO: 104). 1B2 recognizes O-glycans with an unsubstituted O-6 position of the GalNAc residue (Tn, T, and 23ST), whereas 12D10 recognizes Neu5Ac at the same position (STn, 26ST, and dST).

Prostate-specific membrane antigen (PSMA)

Prostate-specific membrane antigen (PSMA), also known as glutamate carboxypeptidase II (GCPII), N-acetyl-L-aspartyl-L-glutamate peptidase I (NAALADase I) or NAAG peptidase, is a zinc metalloenzyme. It is a class II membrane glycoprotein that catalyzes the hydrolysis of N-acetylaspartylglutamate

(NAAG) to glutamate and N-acetylaspartate (NAA). The amino acid sequence for human MUC1 is available from UniProt, Accession No. Q04609.

Human PSMA is highly expressed in the prostate, roughly a hundred times greater than in most other tissues. In some prostate cancers, PSMA is the second-most upregulated gene product, with an 8- to 12-fold increase over levels in noncancerous prostate cells. Because of this high expression, PSMA is being developed as potential biomarker for therapy and imaging of some cancers. In human prostate cancer, the higher expressing tumours are associated with quicker time to progression and a greater percentage of patients suffering relapse. *In vitro* studies using prostate and breast cancer cell lines with decreased PSMA levels showed a significant decrease in the proliferation, migration, invasion, adhesion and survival of the cells.

T-cells expressing CARs specific for prostate-specific membrane antigen (PSMA) are currenty in clinical trial for the treatment of prostate cancer (Junhans et al (2016) Prostate 76:1257-1270).

United Kingdom patent application No. 1919019.8 described various PSMA antigen-binding domains suitable for use in CARs.

A CAR which binds PSMA may have an antigen-binding domain which comprises:
a) a heavy chain variable region (VH) having complementarity determining regions (CDRs) with the following sequences:
CDR1 – SSWMN (SEQ ID No. 18);
CDR2 – RIYPGDGDTNYAQKFQG (SEQ ID No. 19)
CDR3 – GTGYLWYFDV (SEQ ID No. 20); and
b) a light chain variable region (VL) having CDRs with the following sequences:
CDR1 – RASQDINENLA (SEQ ID No. 21);
CDR2 – YTSNRAT (SEQ ID No. 22)
CDR3 – QQYDNLPFT (SEQ ID No. 23).

The PSMA binding domain may comprise a VH domain having the sequence shown as SEQ ID No. 77; and/or a VL domain having the sequence shown as SEQ ID No 78.

SEQ ID No. 77 (Humanised 7A12 VH sequence)

QVQLVQSGAEVKKPGASVKVSCKASGYTFTSSWMNWVRQAPGQGLEWMGRIYPG
DGDTNYAQKFQGRVTMTRDTSISTAYMELSRLRSDDTAVYYCARGTGYLWYFDVW
GQGTMVTVSS

SEQ ID No. 78 (Humanised 7A12 VL sequence)
EIVLTQSPATLSLSPGERATLSCRASQDINENLAWYQQKPGQAPRLLIYYTSNRATGI
PARFSGSGSGRDFTLTISSLEPEDFAVYYCQQYDNLPFTFGQGTKVEIKR

The antigen-binding domain of the CAR may comprise a variant of SEQ ID NO: 77 or 78 having at least 80, 85, 90, 95, 98 or 99% sequence identity, provided that the variant sequence retains the capacity to bind PSMA.

In one embodiment of the invention, however, the CAR binds to an antigen other than PSMA. The CAR may have a non-PSMA specific antigen binding domain.

INTRACELLULAR T CELL SIGNALING DOMAIN (ENDODOMAIN)

The CAR may comprise or associate with an activating endodomain: the signal-transmission portion of the CAR. After antigen recognition, receptors cluster and a signal is transmitted to the cell. The most commonly used endodomain component is that of CD3-zeta which contains 3 ITAMs. This transmits an activation signal to the T cell after antigen is bound. CD3-zeta may not provide a fully competent activation signal and additional co-stimulatory signaling may be needed. For example, chimeric CD28 and OX40 can be used with CD3-Zeta to transmit a proliferative/survival signal, or all three can be used together.

The endodomain of the CAR may comprise the CD28 endodomain and CD3-Zeta endodomain. The sequences for these endodomains are shown below as SEQ ID Nos. 79 and 80 respectively:

SEQ ID No. 79 (CD28 endodomain)
RSKRSRLLHSDYMNMTPRRPGPTRKHYQPYAPPRDFAAYRS

SEQ ID No. 80 (CD3z endodomain)
RVKFSRSADAPAYQQGQNQLYNELNLGRREEYDVLDKRRGRDPEMGGKPRRKNP
QEGLYNELQKDKMAEAYSEIGMKGERRRGKGHDGLYQGLSTATKDTYDALHMQAL
PPR

The endodomain may comprise:

(i) an ITAM-containing endodomain, such as the endodomain from CD3 zeta; and/or

(ii) a co-stimulatory domain, such as the endodomain from CD28; and/or

(iii) a domain which transmits a survival signal, for example a TNF receptor family endodomain such as OX-40 or 4-1BB.

An endodomain which contains an ITAM motif can act as an activation endodomain in this invention. Several proteins are known to contain endodomains with one or more ITAM motifs. Examples of such proteins include the CD3 epsilon chain, the CD3 gamma chain and the CD3 delta chain to name a few. The ITAM motif can be easily recognized as a tyrosine separated from a leucine or isoleucine by any two other amino acids, giving the signature YxxL/I (SeQ ID NO. 81). Typically, but not always, two of these motifs are separated by between 6 and 8 amino acids in the tail of the molecule (YxxL/Ix(6-8)YxxL/I). Hence, one skilled in the art can readily find existing proteins which contain one or more ITAM to transmit an activation signal. Further, given the motif is simple and a complex secondary structure is not required, one skilled in the art can design polypeptides containing artificial ITAMs to transmit an activation signal (see WO 2000/063372, which relates to synthetic signalling molecules).

A number of systems have been described in which the antigen recognition portion of the CAR is on a separate molecule from the signal transmission portion, such as those described in WO015/150771; WO2016/124930 and WO2016/030691. One or more of the viral vectors used in the method of the invention may encode such a "split CAR". Alternatively one vector may comprise a nucleic acid sequence encoding the antigen recognition portion and one vector may comprise a nucleic acid sequence encoding the intracellular signalling domain.

SIGNAL PEPTIDE

One or more nucleic acid sequences in the vector may encode a signal peptide so that when the CAR, cytokine or chemokine is expressed inside a cell, the nascent protein is directed to the endoplasmic reticulum and subsequently to the cell surface, where it is expressed (or secreted in the case of the cytokine or chemokine).

The core of the signal peptide may contain a long stretch of hydrophobic amino acids that tends 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 may be at the amino terminus of the molecule.

A CAR may have the general formula:

Signal peptide – antigen binding domain – spacer domain - transmembrane domain - intracellular T cell signaling domain (endodomain).

SPACER

The CAR may comprise a spacer sequence to connect the antigen binding domain with the transmembrane domain and spatially separate the antigen binding domain from the endodomain. A flexible spacer allows to the antigen binding domain to orient in different directions to enable antigen binding.

The spacer sequence may, for example, comprise an IgG1 Fc region, an IgG1 hinge or a CD8 stalk, or a combination thereof. The spacer may alternatively comprise an alternative sequence which has similar length and/or domain spacing properties as an IgG1 Fc region, an IgG1 hinge or a CD8 stalk.

Where the composition of viral vectors includes more than one vector comprising a nucleic acid sequence encoding a CAR, the CARs may have different spacers.

OR GATES

A cell composition of the present invention may comprise two or more CARs. This may be as a result of transduction with two or more vectors, each comprising a nucleic acid sequence encoding a CAR; or it may be as a result of transduction with a single vector which comprises a nucleic acid construct encoding two or more CARs.

A CAR may be used in a combination with one or more other activatory or inhibitory chimeric antigen receptors. For example, they may be used in combination with one or more other CARs in a "logic-gate", a CAR combination which, when expressed by a cell, such as a T cell, are capable of detecting a particular pattern of expression of at least two target antigens. If the at least two target antigens are arbitrarily denoted as antigen A and antigen B, the three possible options are as follows:

"OR GATE" – T cell triggers when either antigen A or antigen B is present on the target cell
"AND GATE" – T cell triggers only when both antigens A and B are present on the target cell
"AND NOT GATE" – T cell triggers if antigen A is present alone on the target cell, but not if both antigens A and B are present on the target cell

Engineered T cells expressing these CAR combinations can be tailored to be exquisitely specific for cancer cells, based on their particular expression (or lack of expression) of two or more markers.

Such "Logic Gates" are described, for example, in WO2015/075469, WO2015/075470 and WO2015/075470.

An "OR Gate" comprises two or more activatory CARs each directed to a distinct target antigen expressed by a target cell. The advantage of an OR gate is that the effective targetable antigen is increased on the target cell, as it is effectively antigen A + antigen B. This is especially important for antigens expressed at variable or low density on the target cell, as the level of a single antigen may be below the threshold needed for effective targeting by a CAR-T cell. Also, it avoids the phenomenon of antigen escape.

TRANSGENIC T-CELL RECEPTOR (TCR)

The cell of the present invention may express a transgenic TCR.

The T-cell receptor (TCR) is a molecule found on the surface of T cells which is responsible for recognizing fragments of antigen as peptides bound to major histocompatibility complex (MHC) molecules.

The TCR is a heterodimer composed of two different protein chains. In humans, in 95% of T cells the TCR consists of an alpha (α) chain and a beta (β) chain (encoded by TRA and TRB, respectively), whereas in 5% of T cells the TCR consists of gamma and delta (γ/δ) chains (encoded by TRG and TRD, respectively).

When the TCR engages with antigenic peptide and MHC (peptide/MHC), the T lymphocyte is activated through signal transduction.

In contrast to conventional antibody-directed target antigens, antigens recognized by the TCR can include the entire array of potential intracellular proteins, which are processed and delivered to the cell surface as a peptide/MHC complex.

It is possible to engineer cells to express heterologous (i.e. non-native) TCR molecules by artificially introducing the TRA and TRB genes; or TRG and TRD genes into the cell using a vector. For example, the genes for engineered TCRs may be introduced into autologous T cells and transferred back into patients for T cell adoptive therapies. Such 'heterologous' TCRs may also be referred to herein as 'transgenic TCRs' or engineered TCRs.

The cell of the present invention may comprise one or more heterologous nucleic acid sequence(s) encoding a TRA and a TRB gene; or one or more heterologous nucleic acid sequence(s) encoding a TRG and a TRD gene.

DOMINANT NEGATIVE SHP-2

The cell of the present invention may express a dominant negative SHP-2.

WO2016/193696 describes truncated versions of SHP-1 and SHP-2 which block or reduces the inhibition mediated by inhibitory immunoreceptors such as CTLA4, PD-1, LAG-3, 2B4 or BTLA 1. The truncated forms of SHP-1 or SHP-2 comprise one or both SH2 domains, but lacks the phosphatase domain. When expressed in a CAR-T cell, these molecules act as dominant negative versions of wild-type SHP-1 and SHP-2 and compete with the endogenous molecule for binding to phosphorylated ITIMs.

The cell of the present invention may express a truncated protein which comprises an SH2 domain from a protein which binds a phosphorylated immunoreceptor tyrosine-

based inhibition motif (ITIM) but lacks a phosphatase domain. The truncated protein may comprise one or both SHP-1 SH2 domain(s) but lack the SHP-1 phosphatase domain. Alternatively the truncated protein may comprise one or both SHP-2 SH2 domain(s) but lack the SHP-2 phosphatase domain.

SHP-1

Src homology region 2 domain-containing phosphatase-1 (SHP-1) is a member of the protein tyrosine phosphatase family. It is also known as PTPN6.

The N-terminal region of SHP-1 contains two tandem SH2 domains which mediate the interaction of SHP-1 and its substrates. The C-terminal region contains a tyrosine-protein phosphatase domain.

SHP-1 is capable of binding to, and propagating signals from, a number of inhibitory immune receptors or ITIM containing receptors. Examples of such receptors include, but are not limited to, PD1, PDCD1, BTLA4, LILRB1, LAIR1, CTLA4, KIR2DL1, KIR2DL4, KIR2DL5, KIR3DL1 and KIR3DL3.

Human SHP-1 protein has the UniProtKB accession number P29350.

An activity modulator may comprise or consist of the SHP-1 tandem SH2 domain which is shown below as SEQ ID NO: 82.

SHP-1 SH2 complete domain (SEQ ID NO: 82)

MVRWFHRDLSGLDAETLLKGRGVHGSFLARPSRKNQGDFSLSVRVGDQVTHIRIQN
SGDFYDLYGGEKFATLTELVEYYTQQQGVLQDRDGTIIHLKYPLNCSDPTSERWYH
GHMSGGQAETLLQAKGEPWTFLVRESLSQPGDFVLSVLSDQPKAGPGSPLRVTHIK
VMCEGGRYTVGGLETFDSLTDLVEHFKKTGIEEASGAFVYLRQPYY

SHP-1 has two SH2 domains at the N-terminal end of the sequence, at residues 4-100 and 110-213. An activity modulator may comprise one or both of the sequences shown as SEQ ID No. 83 and 84.

SHP-1 SH2 1 (SEQ ID NO: 83)

WFHRDLSGLDAETLLKGRGVHGSFLARPSRKNQGDFSLSVRVGDQVTHIRIQNSGD
FYDLYGGEKFATLTELVEYYTQQQGVLQDRDGTIIHLKYPL

SHP-1 SH2 2 (SEQ ID No. 84)
WYHGHMSGGQAETLLQAKGEPWTFLVRESLSQPGDFVLSVLSDQPKAGPGSPLRV
THIKVMCEGGRYTVGGLETFDSLTDLVEHFKKTGIEEASGAFVYLRQPY

The cell may express a variant of SEQ ID NO: 82, 83 or 84 having at least 80, 85, 90, 95, 98 or 99% sequence identity, provided that the variant sequence is a SH2 domain sequence has the required properties. In other words, the variant sequence should be capable of binding to the phosphorylated tyrosine residues in the cytoplasmic tail of at least one of PD1, PDCD1, BTLA4, LILRB1, LAIR1, CTLA4, KIR2DL1, KIR2DL4, KIR2DL5, KIR3DL1 or KIR3DL3 which allow the recruitment of SHP-1.

SHP-2

SHP-2, also known as PTPN11, PTP-1D and PTP-2C is is a member of the protein tyrosine phosphatase (PTP) family. Like PTPN6, SHP-2 has a domain structure that consists of two tandem SH2 domains in its N-terminus followed by a protein tyrosine phosphatase (PTP) domain. In the inactive state, the N-terminal SH2 domain binds the PTP domain and blocks access of potential substrates to the active site. Thus, SHP-2 is auto-inhibited. Upon binding to target phospho-tyrosyl residues, the N-terminal SH2 domain is released from the PTP domain, catalytically activating the enzyme by relieving the auto-inhibition.

Human SHP-2 has the UniProtKB accession number P35235-1.

An activity modulator may comprise or consist of the SHP-1 tandem SH2 domain which is shown below as SEQ ID NO: 87. SHP-1 has two SH2 domains at the N-terminal end of the sequence, at residues 6-102 and 112-216. An activity modulator may comprise one or both of the sequences shown as SEQ ID No. 85 and 86.

SHP-2 first SH2 domain (SEQ ID NO: 85)
WFHPNITGVEAENLLLTRGVDGSFLARPSKSNPGDFTLSVRRNGAVTHIKIQNTGDY
YDLYGGEKFATLAELVQYYMEHHGQLKEKNGDVIELKYPL

SHP-2 second SH2 domain (SEQ ID No. 86)
WFHGHLSGKEAEKLLTEKGKHGSFLVRESQSHPGDFVLSVRTGDDKGESNDGKSK
VTHVMIRCQELKYDVGGGERFDSLTDLVEHYKKNPMVETLGTVLQLKQPL

SHP-2 both SH2 domains (SEQ ID No. 87)

WFHPNITGVEAENLLLTRGVDGSFLARPSKSNPGDFTLSVRRNGAVTHIKIQNTGDY YDLYGGEKFATLAELVQYYMEHHGQLKEKNGDVIELKYPLNCADPTSERWFHGHLS GKEAEKLLTEKGKHGSFLVRESQSHPGDFVLSVRTGDDKGESNDGKSKVTHVMIRC QELKYDVGGGERFDSLTDLVEHYKKNPMVETLGTVLQLKQPL

The cell may express a variant of SEQ ID NO: 85, 86 or 87 having at least 80, 85, 90, 95, 98 or 99% sequence identity, provided that the variant sequence is a SH2 domain sequence has the required properties. In other words, the variant sequence should be capable of binding to the phosphorylated tyrosine residues in the cytoplasmic tail of at least one of PD1, PDCD1, BTLA4, LILRB1, LAIR1, CTLA4, KIR2DL1, KIR2DL4, KIR2DL5, KIR3DL1 or KIR3DL3 which allow the recruitment of SHP-2.

DOMINANT NEGATIVE TGFβ RECEPTOR

The cell of the present invention may express a dominant negative TGFβ receptor.

Engineered cells face hostile microenvironments which limit adoptive immunotherapy. One of the main inhibitory mechanisms within the tumour microenvironment is transforming growth factor beta (TGFβ). The TGFβ signalling pathway has a pivotal role in the regulatory signalling that controls a variety of cellular processes. TGFβ play also a central role in T cell homeostasis and control of cellular function. Particularly, TGFβ signalling is linked to an immuno-depressed state of the T-cells, with reduced proliferation and activation. TGFβ expression is associated with the immunosuppressive microenvironment of tumour.

A variety of cancerous tumour cells are known to produce TGFβ directly. In addition to the TGFβ production by cancerous cells, TGFβ can be produced by the wide variety of non-cancerous cells present at the tumour site such as tumour-associated T cells, natural killer (NK) cells, macrophages, epithelial cells and stromal cells.

The transforming growth factor beta receptors are a superfamily of serine/threonine kinase receptors. These receptors bind members of the TGFβ superfamily of growth factor and cytokine signalling proteins. There are five type II receptors (which are activatory receptors) and seven type I receptors (which are signalling propagating receptors).

Auxiliary co-receptors (also known as type III receptors) also exist. Each subfamily of the TGFβ superfamily of ligands binds to type I and type II receptors.

The three transforming growth factors have many activities. TGFβ1 and 2 are implicated in cancer, where they may stimulate the cancer stem cell, increase fibrosis /desmoplastic reactions and suppress immune recognition of the tumour.

The active TGFβ receptor (TβR) is a hetero-tetramer, composed by two TGFβ receptor I (TβRI) and two TGFβ receptor II (TβRII). TGFβ1 is secreted in a latent form and is activated by multiple mechanisms. Once activated it forms a complex with the TβRII TβRI that phosphorylates and activates TβRI.

The cell of the invention may express a dominant negative TGFβ receptor. A dominant negative TGFβ receptor may lack the kinase domain.

For example, the dominant negative TGFβ receptor may comprise or consist of the sequence shown as SEQ ID No. 88, which is a monomeric version of TGF receptor II

SEQ ID No. 88 (dn TGFβ RII)
TIPPHVQKSVNNDMIVTDNNGAVKFPQLCKFCDVRFSTCDNQKSCMSNCSITSICEK PQEVCVAVWRKNDENITLETVCHDPKLPYHDFILEDAASPKCIMKEKKKPGETFFMC SCSSDECNDNIIFSEEYNTSNPDLLLVIFQVTGISLLPPLGVAISVIIIFYCYRVNRQQKL SS

A dominant-negative TGF-βRII (dnTGF-βRII) has been reported to enhance PSMA targeted CAR-T cell proliferation, cytokine secretion, resistance to exhaustion, long-term in vivo persistence, and the induction of tumour eradication in aggressive human prostate cancer mouse models (Kloss et al (2018) Mol. Ther.26:1855-1866).

CHIMERIC CYTOKINE RECEPTOR

The cell of the present invention may express a chimeric cytokine receptor.

WO2017/029512 describes chimeric cytokine receptors (CCR) comprising: an exodomain which binds to a non-cytokine ligand; and a cytokine receptor endodomain.

The non-cytokine ligand may be selected from selected from a tumour secreted factor, a chemokine and a cell-surface antigen.

The chimeric cytokine receptor may comprise two polypeptides:

(i) a first polypeptide which comprises:

(a) a first antigen-binding domain which binds a first epitope of the ligand

(b) a first chain of the cytokine receptor endodomain; and

(ii) a second polypeptide which comprises:

(a) a second antigen-binding domain which binds a second epitope of the ligand

(b) a second chain of the cytokine-receptor endodomain.

Alternatively the chimeric cytokine receptor which comprises two polypeptides:

(i) a first polypeptide which comprises:

(a) a heavy chain variable domain (VH)

(b) a first chain of the cytokine receptor endodomain; and

(ii) a second polypeptide which comprises:

(a) a light chain variable domain (VL)

(b) a second chain of the cytokine-receptor endodomain.

For example, the cytokine receptor endodomain may comprise:

(i) IL-2 receptor β-chain endodomain

(ii) IL-7 receptor α-chain endodomain;

(iii) IL-15 receptor α-chain endodomain; or

(iv) common γ-chain receptor endodomain.

The cytokine receptor endodomain may comprise (i), (ii) or (iii); and (iv).

The cytokine receptor endodomain may comprise the α-chain endodomain and the β-chain endodomain from granulocyte-macrophage colony-stimulating factor receptor (GMCSF-R)

The ligand may be a tumour secreted factor, for example a tumour secreted factor selected from: prostate-specific antigen (PSA), carcinoembryonic antigen (CEA), vascular endothelial growth factor (VEGF) and CA125.

The ligand may be a chemokine, for example a chemokine selected from chemokine selected from: CXCL12, CCL2, CCL4, CCL5 and CCL22.

The ligand may be a cell-surface molecule, such as a transmembrane protein. The ligand may be, for example, CD22.

Constitutively active chimeric cytokine receptors

The cell of the invention may express a constitutively active chimeric cytokine receptor. A constitutively active chimeric cytokine receptor may comprise two chains which dimerise, either spontaneously or in the presence of an agent (a chemical inducer of dimerization or CID) bringing together two cytokine receptor endodomains.

The constitutively active chimeric cytokine receptor may therefore comprise a dimerization domain; and a cytokine receptor endodomain.

Dimerisation may occur spontaneously, in which case the chimeric transmembrane protein will be constitutively active. Alternatively, dimerization may occur only in the presence of a chemical inducer of dimerization (CID) in which case the transmembrane protein only causes cytokine-type signalling in the presence of the CID.

Suitable dimerization domains and CIDs are described in WO2015/150771, the contents of which are hereby incorporated by reference.

Where the dimerization domain spontaneously heterodimerizes, it may be based on the dimerization domain of an antibody. In particular it may comprise the dimerization portion of a heavy chain constant domain (CH) and a light chain constant domain (CL). The “dimerization portion” of a constant domain is the part of the sequence which forms the inter-chain disulphide bond.

The chimeric cytokine receptor may comprise the Fab portion of an antibody as exodomain. In this respect, the chimeric antigen may comprise two polypeptides:

(i) a first polypeptide which comprises:

(a) a heavy chain constant domain (CH)

(b) a first chain of the cytokine receptor endodomain; and

(ii) a second polypeptide which comprises:

(a) a light chain constant domain (CL)

(b) a second chain of the cytokine-receptor endodomain.

The cytokine receptor endodomain may comprise:

(i) IL-2 receptor β-chain endodomain

(ii) IL-7 receptor α-chain endodomain; or

(iii) IL-15 receptor α-chain endodomain; and/or

(iv) common γ-chain receptor endodomain.

The cytokine receptor endodomain may comprise the α-chain endodomain and the β-chain endodomain from granulocyte-macrophage colony-stimulating factor receptor (GMCSF-R)

A constitutively active CCR having an IL-2, IL-7 or GM-CSF receptor endodomain may have one of the following structures:

Fab_CCR_IL2: HuLightKappa-IL2RgTM-IL2RgEndo-2A-HuCH1-IL2bTM-IL2RbENDO

Fab_CCR_IL7:
HuLightKappa-IL2RgTM-IL2RgEndo-2A-HuCH1-IL7RaTM-IL7RaENDO

Fab_CCR_GMCSF:
HuLightKappa-GMCSFRbTM-GMCSFRbEndo-2A-HuCH1-GMCSFRaTM-GMCSFRaENDO

In which:

HuLightKappa is a human light kappa chain

IL2RgTM is a transmembrane domain from human IL2R common gamma chain

IL2RgEndo is an endodomain derived from human IL2R common gamma chain

2A is a sequence enabling the co-expesion of the two polypeptides, which may be a self cleaving peptide such as a 2A peptide

HuCH1 is a human CH1

IL2bTM is a transmembrane domain from human IL-2R beta

IL2RbENDO is an endodomain from human IL2R beta

IL7RaTM is a transmembrane domain from human IL-7R alpha

IL7RaENDO is an endodomain from human IL-7R alpha

GMCSFRbTM is a transmembrane domain from Human GM-CSFR common beta chain

GMCSFRbEndo is an endodomain from GM-CSFR common beta chain

GMCSFRaTM is a transmembrane domain from Human GF-CSFR alpha

GMCSFRaENDO is an endodomain Derived from Human GM-CSFR alpha

The sequences for the components for making a constitutively active cytokine receptor as shown below as SEQ ID NO. 89 to 101.

SEQ ID No. 89 (Human Ig Kappa constant domain)
RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESV
TEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC

SEQ ID No. 90 (Human Hinge)
EPKSCDKTHTCPPCP

SEQ ID No. 91 (Human IgG1 CH1 domain)
STKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVL
QSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKRV

SEQ ID No. 92 (Transmembrane domain from human IL2R common gamma chain):
VVISVGSMGLIISLLCVYFWL

SEQ ID No. 93 (Transmembrane domain from human IL-2R beta)
IPWLGHLLVGLSGAFGFIILVYLLI

SEQ ID No. 94 (Transmembrane domain from human IL-7R alpha)
PILLTISILSFFSVALLVILACVLW

SEQ ID No. 95 (Transmembrane domain from Human GF-CSFR alpha)
NLGSVYIYVLLIVGTLVCGIVLGFLF

SEQ ID No. 96 (Transmembrane domain from Human GM-CSFR common beta chain)
VLALIVIFLTIAVLLAL

SEQ ID No. 97 (Endodomain from human IL2R common gamma chain)
ERTMPRIPTLKNLEDLVTEYHGNFSAWSGVSKGLAESLQPDYSERLCLVSEIPPKGG
ALGEGPGASPCNQHSPYWAPPCYTLKPET

SEQ ID No. 98 (Endodomain from human IL-2R beta)

NCRNTGPWLKKVLKCNTPDPSKFFSQLSSEHGGDVQKWLSSPFPSSSFSPGGLAP
EISPLEVLERDKVTQLLLQQDKVPEPASLSSNHSLTSCFTNQGYFFFHLPDALEIEAC
QVYFTYDPYSEEDPDEGVAGAPTGSSPQPLQPLSGEDDAYCTFPSRDDLLLFSPSL
LGGPSPPSTAPGGSGAGEERMPPSLQERVPRDWDPQPLGPPTPGVPDLVDFQPP
PELVLREAGEEVPDAGPREGVSFPWSRPPGQGEFRALNARLPLNTDAYLSLQELQ
GQDPTHLV

SEQ ID No. 99 (Endodomain from human IL-7R alpha)
KKRIKPIVWPSLPDHKKTLEHLCKKPRKNLNVSFNPESFLDCQIHRVDDIQARDEVEG
FLQDTFPQQLEESEKQRLGGDVQSPNCPSEDVVITPESFGRDSSLTCLAGNVSACD
APILSSSRSLDCRESGKNGPHVYQDLLLSLGTTNSTLPPPFSLQSGILTLNPVAQGQ
PILTSLGSNQEEAYVTMSSFYQNQ

SEQ ID No. 100 (Endodomain Derived from Human GM-CSFR alpha)
KRFLRIQRLFPPVPQIKDKLNDNHEVEDEIIWEEFTPEEGKGYREEVLTVKEIT

SEQ ID No. 101 (Endodomain from GM-CSFR common beta chain)
RFCGIYGYRLRRKWEEKIPNPSKSHLFQNGSAELWPPGSMSAFTSGSPPHQGPWG
SRFPELEGVFPVGFGDSEVSPLTIEDPKHVCDPPSGPDTTPAASDLPTEQPPSPQP
GPPAASHTPEKQASSFDFNGPYLGPPHSRSLPDILGQPEPPQEGGSQKSPPPGSLE
YLCLPAGGQVQLVPLAQAMGPGQAVEVERRPSQGAAGSPSLESGGGPAPPALGP
RVGGQDQKDSPVAIPMSSGDTEDPGVASGYVSSADLVFTPNSGASSVSLVPSLGLP
SDQTPSLCPGLASGPPGAPGPVKSGFEGYVELPPIEGRSPRSPRNNPVPPEAKSPV
LNPGERPADVSPTSPQPEGLLVLQQVGDYCFLPGLGPGPLSLRSKPSSPGPGPEIK
NLDQAFQVKKPPGQAVPQVPVIQLFKALKQQDYLSLPPWEVNKPGEVC

The constitutively active CCR may comprise a variant of one or more of SEQ ID NO: 89 to 101 having at least 80, 85, 90, 95, 98 or 99% sequence identity, provided that the variant sequence has the required properties. For example, a variant CH or CL sequence should retain the capacity to dimerise with a CL/CH containing-chain. A variant chain from a cytokine receptor endodomain should retain the capacity to trigger cytokine-mediated signalling when coupled with the reciprocal chain for that cytokine receptor.

NUCLEIC ACID SEQUENCE

The present invention provides a nucleic acid sequence encoding interleukin 12 (IL-12) downstream of a frame-slip motif (FSM) or a translational readthrough motif (TRM).

The nucleic acid sequence may be introduced into a cell such that the cell secretes IL-12 at very low levels.

The nucleic acid sequence may encode a polypeptide comprising a sequence shown as SEQ ID No 1, 24 or 25.

As used herein, the terms “polynucleotide”, “nucleotide”, and “nucleic acid” are intended to be synonymous with each other.

It will be understood by a skilled person that numerous different polynucleotides and nucleic acids can encode the same polypeptide as a result of the degeneracy of the genetic code. In addition, it is to be understood that skilled persons may, using routine techniques, make nucleotide substitutions that do not affect the polypeptide sequence encoded by the polynucleotides described here to reflect the codon usage of any particular host organism in which the polypeptides are to be expressed.

Nucleic acids according to the invention may comprise DNA or RNA. They may be single-stranded or double-stranded. They may also be polynucleotides which include within them synthetic or modified nucleotides. A number of different types of modification to oligonucleotides are known in the art. These include methylphosphonate and phosphorothioate backbones, addition of acridine or polylysine chains at the 3' and/or 5' ends of the molecule. For the purposes of the use as described herein, it is to be understood that the polynucleotides may be modified by any method available in the art. Such modifications may be carried out in order to enhance the in vivo activity or life span of polynucleotides of interest.

The terms “variant”, “homologue” or “derivative” in relation to a nucleotide sequence include any substitution of, variation of, modification of, replacement of, deletion of or addition of one (or more) nucleic acid from or to the sequence.

NUCLEIC ACID CONSTRUCT

The present invention provides a nucleic acid construct comprising the nucleic acid sequence of the present invention. The nucleic acid construct may also comprise a nucleic acid sequence encoding a chimeric antigen receptor (CAR) or transgenic TCR.

The nucleic acid construct according may also comprises one or more of the following:

a nucleic acid sequence encoding a dominant negative SHP2;

a nucleic acid sequence encoding a dominant negative TGFβ receptor;

a nucleic acid sequence encoding a chimeric cytokine receptor:

one or more nucleic acid sequence(s) encoding one or more additional cytokine(s) or chemokine(s).

The present invention provides a nucleic acid construct comprising:

(i) a nucleic acid sequence encoding a chimeric antigen receptor (CAR);

(ii) a nucleic acid sequence encoding interleukin 12 (IL-12) downstream of a frame-slip motif (FSM) or a translational readthrough motif (TRM); and

(iii) a nucleic acid sequence encoding IL-15; and/or

(iv) a nucleic acid sequence encoding CXCL12

The present invention provides a nucleic acid construct comprising:

(i) a nucleic acid sequence encoding a chimeric antigen receptor (CAR);

(ii) a nucleic acid sequence encoding interleukin 12 (IL-12) downstream of a frame-slip motif (FSM) or a translational readthrough motif (TRM); and

(iii) a nucleic acid sequence encoding IL-7; and/or

(iv) a nucleic acid sequence encoding CCL19.

The nucleic acid construct may comprise a co-expression sequence between nucleic acid sequences so that the transgenes upstream and downstream of the co-expression sequence are expressed by the cell as separate polypeptides. The co-expression sequence may encode a self-cleaving peptide, which are described in more detail above.

As explained above, a nucleic acid construct for co-expression of a CAR with low-level IL-12 may have the general structure:

CAR-FSM/TRM-coexpr-IL12

in which:
"CAR" is a nucleic acid sequence encoding a chimeric antigen receptor
" FSM/TRM" is a frame-slip motif or a translational readthrough motif
"coexpr" is a sequence enabling the co-expression of the CAR and IL-12 as separate polypeptides; and
"IL-12" is a nucleic acid sequence encoding IL-12 or flexi-IL12

Constructs which comprise a nucleic acid sequence encoding another cytokine or chemokine or a chemokine, may have the general structure:

CAR-FSM/TRM-coexpr1-IL12-coexpr2-CC; or
CAR-FSM/TRM-coexpr1-CC-coexpr2-IL12;
CAR-coexpr1-CC-FSM/TRM-coexpr2-IL12;
CC-coexpr1-CAR-FSM/TRM-coexpr2-IL12

in which:
"CAR" is a nucleic acid sequence encoding a chimeric antigen receptor
" FSM/TRM" is a frame-slip motif or a translational readthrough motif
"coexpr1" and "coexpr2", which may be the same or different, are sequence enabling the co-expression of the CAR, IL-12 and the cytokine or chemokine as separate polypeptides;
"IL-12" is a nucleic acid sequence encoding IL-12 or flexi-IL12; and
"CC" is a nucleic acid sequence encoding a cytokine (other than IL-12) or a chemokine.

A nucleic acid construct with compound translational readthrough motifs may have the structure:

CAR- FSM/TRM1-coexpr1-CC-FSM/TRM2-coexpr2-IL12

in which:
"CAR" is a nucleic acid sequence encoding a chimeric antigen receptor
" FSM/TRM1" and "FSM/TRM2", which may be the same of different is/are frame-slip motif(s) or translational readthrough motif(s);

"coexpr1" and "coexpr2", which may be the same or different, are sequence enabling the co-expression of the CAR, IL-12 and the cytokine or chemokine as separate polypeptides;

"IL-12" is a nucleic acid sequence encoding IL-12 or flexi-IL12; and

"CC" is a nucleic acid sequence encoding a cytokine (other than IL-12) or a chemokine.

VECTOR

The present invention also provides a vector comprising a nucleic acid sequence or a nucleic acid construct of the invention.

Such a vector may be used to introduce the nucleic acid sequence or construct into a host cell so that it expresses low-level IL-12.

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 or an NK cell.

KIT OF VECTORS

The invention also provides a kit comprising a plurality of vectors, in which at least one vector comprises a nucleic acid sequence encoding interleukin 12 (IL-12) downstream of a frame-slip motif (FSM) or a translational readthrough motif (TRM).

The vector, or other vectors in the kit may comprise one or more of the following:

a nucleic acid sequence encoding a CAR or transgenic TCR

a nucleic acid sequence encoding a dominant negative SHP2;

a nucleic acid sequence encoding a dominant negative TGFβ receptor;

a nucleic acid sequence encoding a chimeric cytokine receptor:

one or more nucleic acid sequence(s) encoding one or more additional cytokine(s) or chemokine(s).

The kit may comprise:

a first vector which comprises a nucleic acid sequence encoding a CAR and a nucleic acid sequence encoding interleukin 12 (IL-12) downstream of a frame-slip motif (FSM) or a translational readthrough motif (TRM); and

a second vector which comprises a nucleic acid sequence encoding a CAR, a nucleic acid sequence encoding a dominant negative SHP2 and a nucleic acid sequence encoding a dominant negative TGFβ.

CELL

The present invention provides a cell which comprises a nucleic acid sequence encoding interleukin 12 (IL-12) downstream of a frame-slip motif (FSM) or a translational readthrough motif (TRM).

The cell may be a is a tumour infiltrating immune cell, such as a T-cell, natural killer (NK) cell, NKT cell, cytokine-induced killer (CIK) cell, monocyte, macrophage, tumour-infiltrating lymphocyte (TIL), dendritic cell, neutrophil, mast cell, eosinophil or basophil.

The cell may be a cytolytic immune cell such as a T cell and/or or NK cell.

T cells or T lymphocytes are a type of lymphocyte that play a central role in cell-mediated immunity. They can be distinguished from other lymphocytes, such as B cells and natural killer cells (NK cells), by the presence of a T-cell receptor (TCR) on the cell surface. There are various types of T cell, as summarised below.

Helper T helper cells (TH cells) assist other white blood cells in immunologic processes, including maturation of B cells into plasma cells and memory B cells, and activation of cytotoxic T cells and macrophages. TH cells express CD4 on their surface. TH cells become activated when they are presented with peptide antigens by MHC class II molecules on the surface of antigen presenting cells (APCs). These cells can differentiate into one of several subtypes, including TH1, TH2, TH3, TH17, Th9, or TFH, which secrete different cytokines to facilitate different types of immune responses.

Cytolytic T cells (TC cells, or CTLs) destroy virally infected cells and tumor cells, and are also implicated in transplant rejection. CTLs express the CD8 at their surface. These cells recognize their targets by binding to antigen associated with MHC class I,

which is present on the surface of all nucleated cells. Through IL-10, adenosine and other molecules secreted by regulatory T cells, the CD8+ cells can be inactivated to an anergic state, which prevent autoimmune diseases such as experimental autoimmune encephalomyelitis.

Memory T cells are a subset of antigen-specific T cells that persist long-term after an infection has resolved. They quickly expand to large numbers of effector T cells upon re-exposure to their cognate antigen, thus providing the immune system with "memory" against past infections. Memory T cells comprise three subtypes: central memory T cells (TCM cells) and two types of effector memory T cells (TEM cells and TEMRA cells). Memory cells may be either CD4+ or CD8+. Memory T cells typically express the cell surface protein CD45RO.

Regulatory T cells (Treg cells), formerly known as suppressor T cells, are crucial for the maintenance of immunological tolerance. Their major role is to shut down T cell-mediated immunity toward the end of an immune reaction and to suppress auto-reactive T cells that escaped the process of negative selection in the thymus.

Two major classes of CD4+ Treg cells have been described — naturally occurring Treg cells and adaptive Treg cells.

Naturally occurring Treg cells (also known as CD4+CD25+FoxP3+ Treg cells) arise in the thymus and have been linked to interactions between developing T cells with both myeloid (CD11c+) and plasmacytoid (CD123+) dendritic cells that have been activated with TSLP. Naturally occurring Treg cells can be distinguished from other T cells by the presence of an intracellular molecule called FoxP3. Mutations of the FOXP3 gene can prevent regulatory T cell development, causing the fatal autoimmune disease IPEX.

Adaptive Treg cells (also known as Tr1 cells or Th3 cells) may originate during a normal immune response.

Natural Killer cells (or NK cells) form part of the innate immune system. NK cells provide rapid responses to innate signals from virally infected cells in an MHC independent manner

NK cells (belonging to the group of innate lymphoid cells) are defined as large granular lymphocytes (LGL) and constitute the third kind of cells differentiated from the common lymphoid progenitor generating B and T lymphocytes. NK cells are known to differentiate and mature in the bone marrow, lymph node, spleen, tonsils and thymus where they then enter into the circulation.

Natural killer T (NKT) cells are a heterogeneous group of T cells that share properties of both T cells and natural killer cells. Many of these cells recognize the non-polymorphic CD1d molecule, an antigen-presenting molecule that binds self and foreign lipids and glycolipids. They constitute only approximately 0.1% of all peripheral blood T cells.

NKT cells are a subset of T cells that coexpress an αβ T-cell receptor, but also express a variety of molecular markers that are typically associated with NK cells, such as NK1.1. The best-known NKT cells differ from conventional αβ T cells in that their T-cell receptors are far more limited in diversity ('invariant' or 'type 1' NKT). They and other CD1d-restricted T cells ('type 2' NKT) recognize lipids and glycolipids presented by CD1d molecules, a member of the CD1 family of antigen-presenting molecules, rather than peptide-major histocompatibility complexes (MHCs). As such, NKT cells are important in recognizing glycolipids from organisms such as Mycobacterium, which causes tuberculosis.

NKT cells include both NK1.1+ and NK1.1−, as well as CD4+, CD4−, CD8+ and CD8− cells. Natural killer T cells can also share other features with NK cells, such as CD16 and CD56 expression and granzyme production.

Cytokine-induced killer cells (CIK) cells are a group of immune effector cells featuring a mixed T- and natural killer (NK) cell-like phenotype. They are generated by *ex vivo* incubation of human peripheral blood mononuclear cells (PBMC) or cord blood mononuclear cells with interferon-gamma (IFN-γ), anti-CD3 antibody, recombinant human interleukin (IL-) 1 and recombinant human interleukin (IL)-2.

Typically, immune cells detect major histocompatibility complex (MHC) presented on infected cell surfaces, triggering cytokine release, causing lysis or apoptosis. However, CIK cells have the ability to recognize infected or even malignant cells in the absence of antibodies and MHC, allowing for a fast and unbiased immune reaction. This is of particular importance as harmful cells that are missing MHC markers cannot be tracked and attacked by other immune cells, such as T-

lymphocytes. As a special feature, terminally differentiated CD3+CD56+ CIK cells possess the capacity for both MHC-restricted and MHC-unrestricted anti-tumour cytotoxicity. These properties, inter alia, rendered CIK cells attractive as a potential therapy for cancer and viral infections.

A new subclass of NK cells have been created both *in vitro* and *in vivo*. These NK cells referred to as cytokine induced memory-like natural killer cells are induced using cytokines, most commonly a mix of IL-12, IL-15, and IL-18. These NK cells are activated by these cytokines to stimulate an infection and induce an adaptive immune response. If cocultured with target cells such as tumour targets, these NK cells have memory-like abilities and are more adapt and effective at mounting a defence.

The cell of the invention may be any of the cell types mentioned above.

The cell may be derived from a blood sample, for example from a leukapheresate. The cells may be or comprise peripheral blood mononuclear cells (PBMCs).

Cells may either be created *ex vivo* either from a patient's own peripheral blood (1st party), or in the setting of a haematopoietic stem cell transplant from donor peripheral blood (2nd party), or peripheral blood from an unconnected donor (3rd party).

Alternatively, cells may be derived from *ex vivo* differentiation of inducible progenitor cells or embryonic progenitor cells to, for example, T or NK cells. Alternatively, an immortalized T-cell line which retains its lytic function and could act as a therapeutic may be used.

The cells may be activated and/or expanded prior to being transduced with nucleic acid encoding the molecules providing the chimeric polypeptide according to the first aspect of the invention, for example by treatment with an anti-CD3 monoclonal antibody.

PHARMACEUTICAL COMPOSITION

The cell of the present inventionmay be administered to a patient as a pharmaceutical composition.

The pharmaceutical composition may additionally comprise a pharmaceutically acceptable carrier, diluent or excipient. The pharmaceutical composition may optionally comprise one or more further pharmaceutically active polypeptides and/or compounds. Such a formulation may, for example, be in a form suitable for intravenous infusion.

METHOD FOR MAKING CELL

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 sequence, a nucleic acid construct, a vector or a kit of vectors of the invention into a cell.

The nucleic acid sequence, nucleic acid construct, vector or kit of vectors may, for example, be introduced by transduction or transfection *in vitro* or *ex vivo*.

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

The cell may be autologous or allogeneic.

METHOD OF TREATMENT

The present invention provides a method for treating a disease which comprises the step of administering a cell composition of the present invention (for example in a pharmaceutical composition as described above) to a subject.

A method for treating a disease relates to the therapeutic use of the cell composition of the present invention. The cell composition may be administered to a subject having an existing disease or condition in order to lessen, reduce or improve at least one symptom associated with the disease and/or to slow down, reduce or block the progression of the disease.

The method for preventing a disease relates to the prophylactic use of the cell composition of the present invention. The cell composition may be administered to a subject who has not yet contracted the disease and/or who is not showing any symptoms of the disease to prevent or impair the cause of the disease or to reduce or

prevent development of at least one symptom associated with the disease. The subject may have a predisposition for, or be thought to be at risk of developing, the disease.

The method may involve the steps of:

(i) isolating a cell-containing sample;

(ii) transducing the such cells with a mixture of at least two viral vectors;

(iii) administering the cells from (ii) to a subject.

The present invention also provides a cell composition of the present invention for use in treating and/or preventing a disease.

The invention also relates to the use of a cell composition of the present invention in the manufacture of a medicament for the treatment of a disease.

The disease to be treated by the methods of the present invention may be a cancerous disease, such as bladder cancer, breast cancer, colon cancer, endometrial cancer, kidney cancer (renal cell), leukaemia, lung cancer, melanoma, non-Hodgkin lymphoma, pancreatic cancer, prostate cancer and thyroid cancer.

The cells of the composition of the present invention may be capable of killing target cells, such as cancer cells. The target cell may be characterised by the presence of a tumour secreted ligand or chemokine ligand in the vicinity of the target cell. The target cell may be characterised by the presence of a soluble ligand together with the expression of a tumour-associated antigen (TAA) at the target cell surface.

In particular, the cells of the invention may be used to treat a solid cancer. A solid cancer is a malignancy that forms a discrete tumour mass, for example of brain, breast, prostate, colorectum, kidney; sarcoma; melanoma, in contrast to lymphoproliferative malignancies leukemia, which may diffusely infiltrate a tissue without forming a mass.

The cells of the present invention may be used to treat one of the cancers listed in Table 3.

Table 3 - Exemplary solid cancer targets and their associated indications

Target	Indications

CD44v6	Metastatic cancers, colorectal cancers, soft-tissue sarcoma
CAIX	Renal cell cancer (RCC), glioblastoma (GBM)
CEA	Ovarian cancer, colorectal cancer, cancers of the gastrointestinal tract, hepatocellular carcinoma (HCC)
CD133	Ovarian cancer, GBM, HCC
Claudin-6	Ovarian cancer
Claudin-18.2	Gastric cancer, pancreatic cancer
c-Met	Breast cancer, melanoma, HCC
EGFR	Non-small cell lung cancer (NSCLC), GBM, gliomas, medulloblastoma, soft-tissue sarcoma (STS), osteosarcoma, Ewing's mesothelioma
EGFRvIII	GBM, colorectal cancer, STS, pancreatic cancer
EpCAM	HCC, NSCLC, ovarian cancer, breast cancer, gastric cancer, pancreatic cancer
EpA2	GBM, gliomas
Fetal ACho receptor	Osteosarcoma, STS
Folate receptor	Ovarian cancer, urothelial cancer
GD2	Neuroblastoma, melanoma, osteosarcoma, STS, Ewing's sarcoma
GPC3	HCC, SCC
GUCY2C	Colorectal cancer
HER1	NSCLC, prostate cancer
HER2	Breast cancer, ovarian cancer, osteosarcoma, STS, GBM, pediatric CNS cancer, medulloblastoma, gastric cancer, mesothelioma
ICAM-1	Thyroid
IL13Ra2	GBM, gliomas
IL11Ra	Osteosarcoma
L1CAM	Ovarian cancer
Mesothelin	Pancreatic cancer, ovarian cancer, lung adenocarcinoma (LUAD), GBM, mesothelioma
MUC1	HCC, NSCLC, pancreatic cancer, breast cancer, glioma, colorectal, gastric cancer
MUC16	Ovarian cancer
NKG2D	Ovarian cancer, osteosarcoma, Ewing's sarcoma
NY-ES01	STS, neuroblastoma, ovarian cancer, breast cancer, GBM,

	NSCLC
PSA	Prostate cancer, pancreatic cancer
PSMA	Prostate cancer, cervical cancer
WT1	Ovarian cancer

In particular, the cells of the present invention may be used to treat small cell lung cancer (SCLC), melanoma, renal cell cancer (RCC), hepatocellular carcinoma (HCC), ovarian cancer, pancreatic cancer, neuroblastoma, osteosarcoma. In these embodiments, the cells may express a GD2-specific CAR.

CANCER HETEROGENEITY AND EPITOPE SPREADING

The cancer may be heterogeneous. In this respect, the level of expression of the CAR-target antigen may vary between cancer cells. The proportion of cells in the tumour which express the CAR-target antigen may be less than 100%, For example, less than about 95, 90, 80, 70, 60, 50, 40, 30, 20 or 10% of the cells in the tumour may express the target antigen at detectable levels, or at levels recognisable to a CAR-T cell. Some of the cells of the solid cancer may be antigen-negative or antigen-dim, i.e. express low levels of target antigen. An "antigen-dim" target cell may express fewer than about 1000, 750, 500 or 250 copies of the target antigen per cell, on average.

The cells of the present invention are particularly well suited to target heterogenous tumours because of the adjuvant effect of IL-12. IL-12 activates an anti-tumour immune response by activating naïve T cells and macrophages. It has been shown to 1) enhance proliferation of T cells and NK cells, 2) increase cytolytic activities of T cells, NK cells, and macrophages, 3) activate T helper 1 (Th1) cells, and 4) induce production of IFN-γ and other cytokines. IL-12 activates neighbouring CAR-T cells and enhances their cytotoxicity, enhancing their capacity to kill tumour cells with low levels of target antigen

The mechanism of CAR T cell therapy is through direct killing of cells expressing target antigen by T cells expressing the CAR. Without wishing to be bound by theory, the present inventors believe that the simultaneous local expression of IL-12 by CAR-T cells will lead to "epitope spreading" i.e. the priming of immune responses against additional target antigens present on tumour cells. The secretion of IL-12 by CAR-T cells may also induce the infiltration of host immune cells into the tumour mass

IL-12 secreted from T cells thus not only overcomes immune suppression mediated by tumour cells but also provides a conducive environment for epitope spreading, where immune cells diversify to attack multiple tumour-associated targets in addition to the original antigen. The mechanism of action of tumour eradication is through a combination of direct cell killing by CAR-T cells, infiltration of immune cells into the tumour and the induction of an anti-cancer immune response from existing host immune cells against other tumour-associated antigens.

Epitope spreading is important for the treatment of a heterogeneous tumour as it enables the killing of cells which are dim or negative for target antigen. Moreover, it helps prevent immune escape of tumours from CAR-T cell therapy preventing where malignancies may otherwise recur as antigen-negative populations.

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 - Expression of IL-12 by CAR-T cells

In a construct encoding two transgenes, the inclusion of a self-cleaving peptide such as an FMDV-derived 2A peptide sequence, leads to approximately equal expression of the two polypeptides. Where the construct co-expresses a CAR and IL-12 this will lead to secretion of high levels of IL-12 by the CAR-T cell (illustrated schematically in Figure 4A, left-hand diagram).

However, incorporation of a "stop-skip" sequence upstream of the IL-12 encoding sequence should drastically reduce the amount of IL-12 produced by the CAR-T cell (illustrated schematically in Figure 4A, left-hand diagram).

In order to demonstrate the experimentally, peripheral blood mononuclear cells (PBMCs) were transduced with viruses encoding the following constructs:

CAR-2A-IL-12: comprising a CAR encoding sequence, a 2A self-cleaving peptide, and a sequence encoding flexi-IL-12 (having the sequence shown as SEQ ID No. 1), and

CAR-SS-IL-12: comprising a CAR encoding sequence, a stop-skip sequence (having the sequence shown as SEQ ID No. 37), and a sequence encoding flexi-IL-12 positioned downstream of the stop-skip sequence

Untransduced cells (NT) and PBMCs transduced with either construct, were plated in a 96 well plate at $1x10E^6$ cells/ml and after 48 hours, medium was collected for quantification of released IL-12 by ELISA. The results are shown in Figure 4B.

Production of IL-12 by the transduced cells was drastically reduced, but still detectable, when the sequence encoding IL-12 was placed downstream of a stop-skip sequence.

Example 2 - Investigating anti-tumour activity of CAR-T cell expressing IL-12 *in vivo*

Next, the functionality of the SS-IL12 module *in vivo* was investigated using an immunocompetent mouse model.

CAR-expressing cells were created by transduction of murine PBMCs with vectors expressing the following constructs:

NAME	CONSTRUCT
GD2 CAR	SFGmR.Thy1.1-2A-aGD2_muK666-muCD8STK-muCD28Z
CAR-2A-IL-12	SFGmR.Thy1.1-2A-aGD2_muK666-muCD8STK-muCD28Z-2A-mflL12
CAR-SS-IL-12	SFGmR.Thy1.1-2A-aGD2_muK666-muCD8STK-muCD28Z-SKIP_TGACAATTA-2A-MUflL12
EGFRvIII-SS-IL-12	SFGmR.THY1.1-2A-aEGFRvIII_MR1-muCD8STK-muCD28Z-SKIP_TGACAATTA-2A-MUflL12

in which "Thy1.1" is a transduction marker; "aGD2_muK666-muCD8STK-muCD28Z" is a second generation murine GD2 CAR; "mflL-12" is murine flexi-IL-12; "SKIP_TGACAATTA" is a translational readthrough motif placed upstream of the murine flexi-IL-12-encoding sequence; "aEGFRvIII_MR1-muCD8STK-muCD28Z" is a

second generation murine CAR against aEGFRvIII, which is used as a negative control.

On the right side of the slide you can see in the chart the drastic weight loss in the CAR-2A-IL12 cohort, due to IL-12 systemic toxicity and these mice had to be euthanized. With regards to tumor growth control; on the right side we can see that tumor growth was not controlled at all in the CAR alone cohort, while mice injected with CAR-SS-IL12 had a drastic reduction an control of the tumor growth, without the toxicity showed by the CAR-2A-IL12 mice cohort.

We can conclude that SS-IL12 module prevents toxicity whilst maintaining potent anti-tumor activity in a challenging murine model

An in vivo assay was used to investigate the anti-tumour activity of T cells expressing a GD2 with or without IL-12 in the murine B16 melanoma model.

B16.F10.GD2 ($0.1x10^6$) were administered subcutaneously to mice on day 0. After 7 days $3x10^6$ CAR-T cells were administered intravenously and tumour growth was monitored for the following 14 days.

As shown in Figure 6, the CAR-2A-IL12 cohort, i.e. mice receiving cells expressing the CAR in combination with IL-12 <u>without</u> a stop-skip sequence, showed drastic weight loss due to IL-12 systemic toxicity. These mice had to be euthanized before day 12. The CAR-SS-IL12 cohort, i.e. mice receiving cells expressing the CAR in combination with ultra-low level IL-12 did not show any significant weight loss or IL-12-related toxicity.

The tumour growth assay results are shown in Figure 7. Intravenous delivery of CAR T cells expressing a simple GD2 CAR alone had no significant effect on tumour growth (Figure 7 "CAR"). The introduction of CAR-T cells expressing IL-12, without a translational readthrough motif to control the level of expression, was fatally toxic to the mice (Figure 7 "CAR-2A"). However, intravenous delivery of CAR T cells co-expressing ultra low-level IL-12, in which the IL-12 encoding sequence was placed downstream of a translational readthrough motif, exhibited potent anti-tumour activity and extended survival of the mice (Figure 7 "CAR-SS").

Example 3 - <u>Investigating the capacity for ultra-low IL-12 expressing CAR-T cells to induce epitope spreading during an immune response</u>

To examine the ability of the ssIL12 module to induce an epitope spreading response *in vivo*, a target cell line was generated from the murine melanoma cell line B16.F10 transduced to express the two precursor enzymes (GD2 and GD3 synthase) required to synthesise GD2 on the plasmid membrane. This target cell line is 100% positive for GD2 expression and is used to make a heterogenous population of target cells by the addition of non-transduced B16.F10 cells at fixed ratios.

Heterogeneous target populations generated are 10%, 30%, 50%, 70% or 90% GD2-positive and are compared with a homogeneous control of 100% GD2-positive cells. $1E^5$ target cells are subcutaneously injected into an autologous recipient and allowed to engraft for 6 days before the administration of 5 Gray total body irradiation.

The following day, $3E^6$ CAR T cells are injected intravenously. Mice are monitored regularly for signs of toxicity (weight, serum cytokine and general conditioning) and efficacy (calliper measurement of tumour size).

At the time of take down, spleens from the mice are taken for analysis of the proportion of CAR cells, CD4+ T cells, CD8+ T cells, B cells, NK cells, and monocytes. Any remaining tumour at the time of take down is collected for both immunohistochemical analysis (to assess immune infiltrate and tumour morphology) and flow cytometry (to assess the immune infiltrate and proportion and expression levels of GD2 positive tumour).

The degree of epitope spreading is determined by the ability of the module to resolve heterogeneous targets from the recipient mouse and/or the ability of the module to induce immune infiltrate in the tumour mass.

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 or related fields are intended to be within the scope of the following claims.

CLAIMS

1. A method for treating a solid cancer which comprises the step of administering a cell to a subject, wherein the cell comprises a nucleic acid sequence encoding interleukin 12 (IL-12) downstream of a frame-slip motif (FSM) or a translational readthrough motif (TRM).

2. A method according to claim 1, which comprises the following steps:

(i) isolation of a cell sample from a subject;

(ii) transfection or transduction of cells from the cell sample with a nucleic acid sequence encoding interleukin 12 downstream of a frame-slip motif (FSM) or a translational readthrough motif (TRM); and

(iii) administering transduced or transfected cells from step (ii) to the subject.

3. A method according to claim 1 or 2, wherein the nucleic acid sequence encoding IL-12, encodes "flexi-IL-12": a fusion between IL-12α and IL-12β subunits, joined by a linker.

4. A method according to claim 3, wherein flexi-IL-12 comprises the sequence shown as SEQ ID No. 1.

5. A method according to any preceding claim, wherein the nucleic acid sequence encoding interleukin 12 (IL-12) is downstream of a frame-slip motif (FSM), and the FSM comprises a repeat of uracil, thymine or guanine bases.

6. A method according to claim 5, wherein the FSM comprises the sequence UUUUUUU (SEQ ID No. 2).

7. A method according to claim 5 or 6, wherein the FSM also comprises a stop codon.

8. A method according to claim 7, wherein the FSM comprises one of the following sequences:

UUUUUUUGA (SEQ ID NO. 3)

UUUUUUUAG (SEQ ID NO. 4)

UUUUUUUAA (SEQ ID NO. 5).

9. A method according to any of claims 1 to 4, wherein the nucleic acid sequence encoding interleukin 12 (IL-12) is downstream of a translational readthrough motif (TRM)

which comprises the sequence STOP-CUAG or STOP-CAAUUA, in which "STOP" is a stop codon.

10. A method according to claim 9, wherein the translational readthrough motif comprises one of the following sequences:

UGA-CUAG (SEQ ID No. 6)
UAG-CUAG (SEQ ID No. 7)
UAA-CUAG (SEQ ID No. 8)
UGA-CAAUUA (SEQ ID No. 9)
UAG-CAAUUA (SEQ ID No. 10)
UAA-CAAUUA (SEQ ID No. 11)

11. A method according to any preceding claim wherein the cell is a tumour infiltrating immune cell.

12. A method according to claim 11, wherein the cell is T-cell, natural killer (NK) cell, NKT cell, cytokine-induced killer (CIK) cell, monocyte, macrophage or tumour-infiltrating lymphocyte (TIL).

13. A method according to any preceding claim wherein the cell expresses a chimeric antigen receptor (CAR) or engineered T-cell receptor (TCR)

14. A method according to claim 12 or 13, wherein the CAR or engineered TCR binds to one of the following target antigens: disialoganglioside (GD2); epidermal growth factor receptor (EGFR), Epithelial cell adhesion molecule (EpCAM), Glypican 3 (GPC3), human epidermal growth factor receptor (HER2), L1CAM, Mucin 1 (MUC1), Prostate-specific membrane antigen (PSMA).

.

15. A method according to claim 14, wherein the cell expresses a CAR which binds to GD2 and has an antigen-binding domain which comprises

a) a heavy chain variable region (VH) having complementarity determining regions (CDRs) with the following sequences:

CDR1 – SYNIH (SEQ ID No. 12);
CDR2 – VIWAGGSTNYNSALMS (SEQ ID No. 13)
CDR3 – RSDDYSWFAY (SEQ ID No. 14); and

b) a light chain variable region (VL) having CDRs with the following sequences:

CDR1 – RASSSVSSSYLH (SEQ ID No. 15);

CDR2 – STSNLAS (SEQ ID No. 16)
CDR3 – QQYSGYPIT (SEQ ID No. 17).

16. A method according to claim 14, wherein the cell expresses a CAR which binds PSMA and has an antigen-binding domain which comprises
a) a heavy chain variable region (VH) having complementarity determining regions (CDRs) with the following sequences:
CDR1 – SSWMN (SEQ ID No. 18);
CDR2 – RIYPGDGDTNYAQKFQG (SEQ ID No. 19)
CDR3 – GTGYLWYFDV (SEQ ID No. 20); and
b) a light chain variable region (VL) having CDRs with the following sequences:
CDR1 – RASQDINENLA (SEQ ID No. 21);
CDR2 – YTSNRAT (SEQ ID No. 22)
CDR3 – QQYDNLPFT (SEQ ID No. 23).

17. A method according any preceding claim, wherein the cell also expresses a dominant negative SHP2; and/or a dominant negative TGFβ receptor.

18. A method according to any preceding claim, wherein the cell also expresses a chimeric cytokine receptor.

19. A method according to any preceding claim, wherein the cell also expresses one or more additional cytokine(s) or chemokine(s).

20. A method according to claim 19, wherein the cell expresses one or more of the following: IL-7, IL-15, CCL19, CXCL12.

21. A method according to claim 20, wherein the cell expresses IL-7 and CCL19.

22. A method according to claim 20, wherein the cell expresses IL-15 and/or CXCL12.

23. A method according to any of claims 19 to 22, wherein the nucleic acid encoding the or each additional cytokine(s) or chemokine(s) is positioned upstream of the frame-slip motif (FSM) or translational readthrough motif (TRM).

24. A method according to any of claims 19 to 22, wherein the nucleic acid encoding the or each additional cytokine(s) or chemokine(s) is positioned downstream of the frame-slip motif (FSM) or translational readthrough motif (TRM).

25. A method according to any preceding claim for the treatment of small cell lung cancer (SCLC), melanoma, renal cell cancer (RCC), hepatocellular carcinoma (HCC), ovarian cancer, pancreatic cancer, neuroblastoma, osteosarcoma.

26. A method according to any preceding claim wherein the cell expresses a CAR or engineered TCR which binds a target antigen and expression of the target antigen on the solid tumour is heterogeneous.

27. A method for inducing epitope spreading in an anti-tumour immune response in a subject, which comprises the step of administering a plurality of cells as defined in any preceding claim to the subject.

28. A method for inducing infiltration of immune cells into a tumour mass in a subject which comprises the step of administering a plurality of cells as defined in any preceding claim to the subject, such that they induce an anti-tumour immune response.

29. A cell for use in treating a solid cancer, the cell comprising a nucleic acid sequence encoding interleukin 12 (IL-12) downstream of a frame-slip motif (FSM) or a translational readthrough motif (TRM).

30. The use of a cell in the manufacture of a medicament for treating a solid cancer, wherein the cell comprises a nucleic acid sequence encoding interleukin 12 (IL-12) downstream of a frame-slip motif (FSM) or a translational readthrough motif (TRM).

31. A cell which expresses a chimeric antigen receptor (CAR) and comprises;

a nucleic acid sequence encoding interleukin 12 (IL-12) downstream of a frame-slip motif (FSM) or a translational readthrough motif (TRM); and

one or more heterologous nucleic acid sequence(s) encoding one or more additional cytokine(s) or chemokine(s).

32. A cell according to claim 25, wherein the one or more heterologous nucleic acid sequence(s) encode one or more of the following: IL-7, IL-15, CCL19, CXCL12.

33. A cell according to claim 25, wherein the one or more heterologous nucleic acid sequence(s) encode IL-7 and CCL19.

34. A cell according to claim 25, wherein the one or more heterologous nucleic acid sequence(s) encode IL-15 and/or CXCL12.

35. A cell according to any of claims 31 to 34, wherein one or more of the heterologous nucleic acid sequence(s) encoding one or more additional cytokine(s) or chemokine(s) is/are positioned upstream of the frame-slip motif (FSM) or translational readthrough motif (TRM).

36. A cell according to any of claims 31 to 34, wherein one or more of the heterologous nucleic acid sequence(s) encoding one or more additional cytokine(s) or chemokine(s) is/are positioned downstream of the frame-slip motif (FSM) or translational readthrough motif (TRM).

37. A nucleic acid construct comprising:

(i) a nucleic acid sequence encoding a chimeric antigen receptor (CAR); and

(ii) a nucleic acid sequence encoding interleukin 12 (IL-12) downstream of a frame-slip motif (FSM) or a translational readthrough motif (TRM);

(vi) one or more nucleic acid sequence(s) encoding one or more additional cytokine(s) or chemokine(s) as defined in any of claims 32 to 36.

38. A vector comprising a nucleic acid construct according to claim 37.

39. A kit of vectors, each encoding one or more nucleic acid sequence(s) as defined in claim 37.

40. A method for making a cell according to any of claims 31 to 36, which comprises the step of transfecting or transducing a cell with a nucleic acid construct according to claim 37, a vector according to claim 38 or a kit of vectors according to claim 39, *ex vivo*.

FIGURE 1

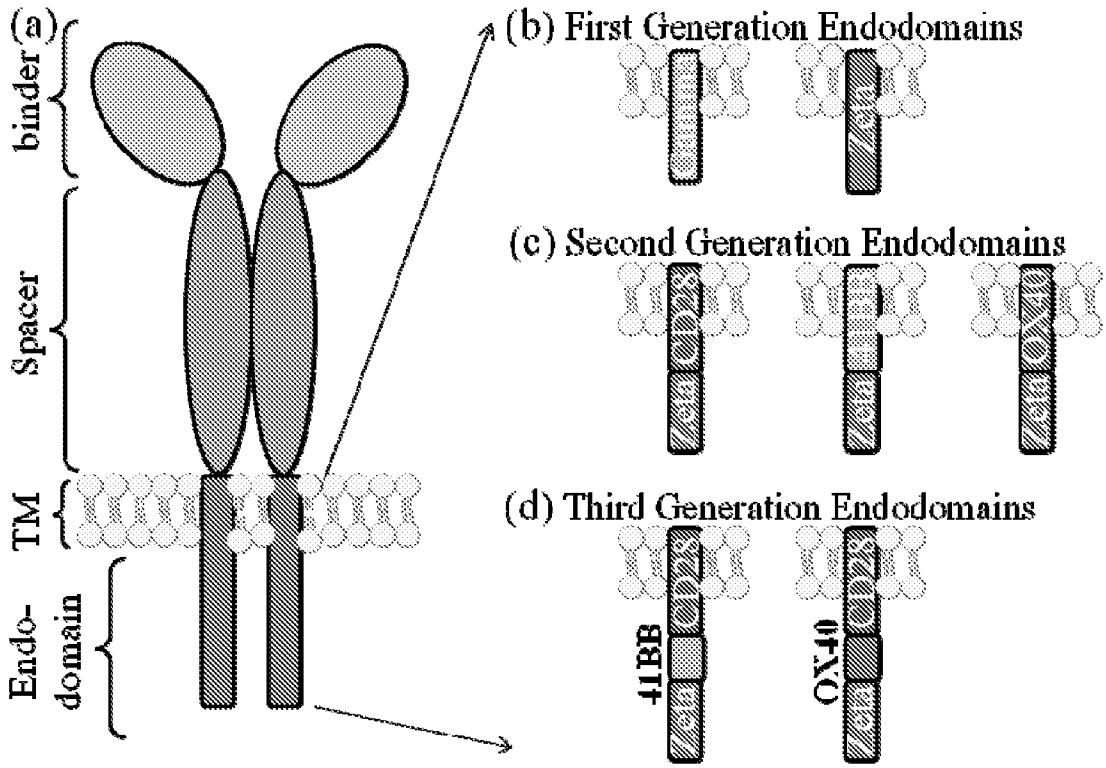

(a)
binder
Spacer
TM
Endo-domain
(b) First Generation Endodomains
(c) Second Generation Endodomains
(d) Third Generation Endodomains
41BB
OX40

FIGURE 2

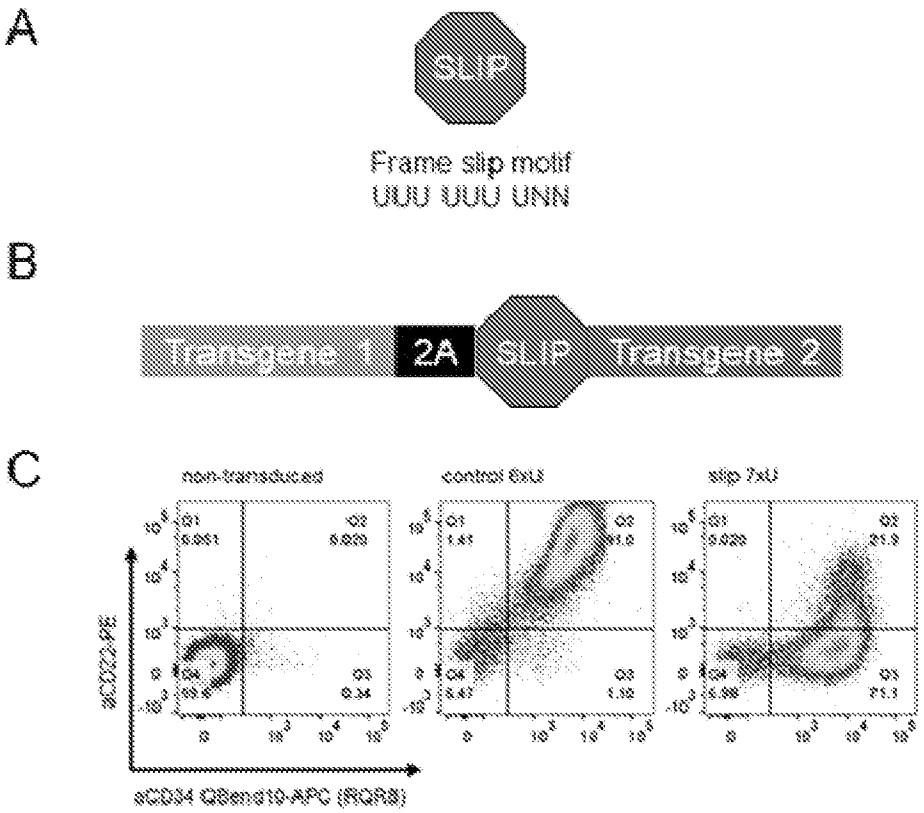

A
SLIP
Frame slip motif
UUU UUU UNN
B
Transgene 1
2A
SLIP
Transgene 2
C
non-transduced
control 6xU
slip 7xU
αCD22-PE
αCD34 QBend10-APC (RORB)

FIGURE 3

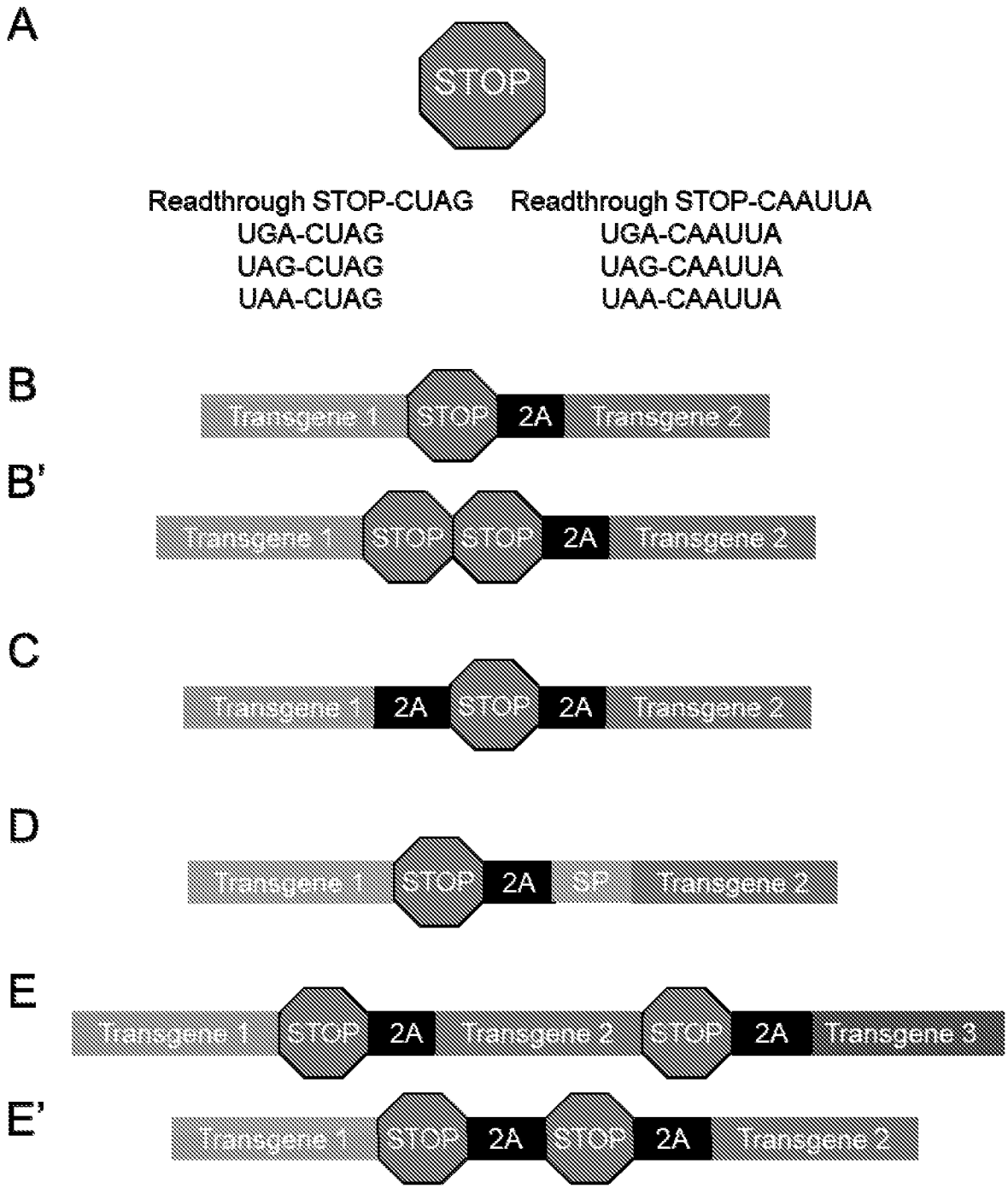

A
STOP
Readthrough STOP-CUAG
UGA-CUAG
UAG-CUAG
UAA-CUAG
Readthrough STOP-CAAUUA
UGA-CAAUUA
UAG-CAAUUA
UAA-CAAUUA
B
Transgene 1
STOP
2A
Transgene 2
B'
Transgene 1
STOP
STOP
2A
Transgene 2
C
Transgene 1
2A
STOP
2A
Transgene 2
D
Transgene 1
STOP
2A
SP
Transgene 2
E
Transgene 1
STOP
2A
Transgene 2
STOP
2A
Transgene 3
E'
Transgene 1
STOP
2A
STOP
2A
Transgene 2

FIGURE 4

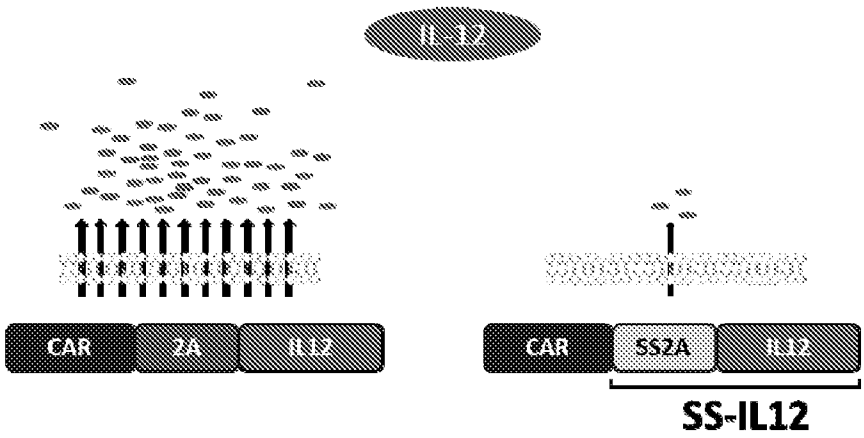


FIGURE 5

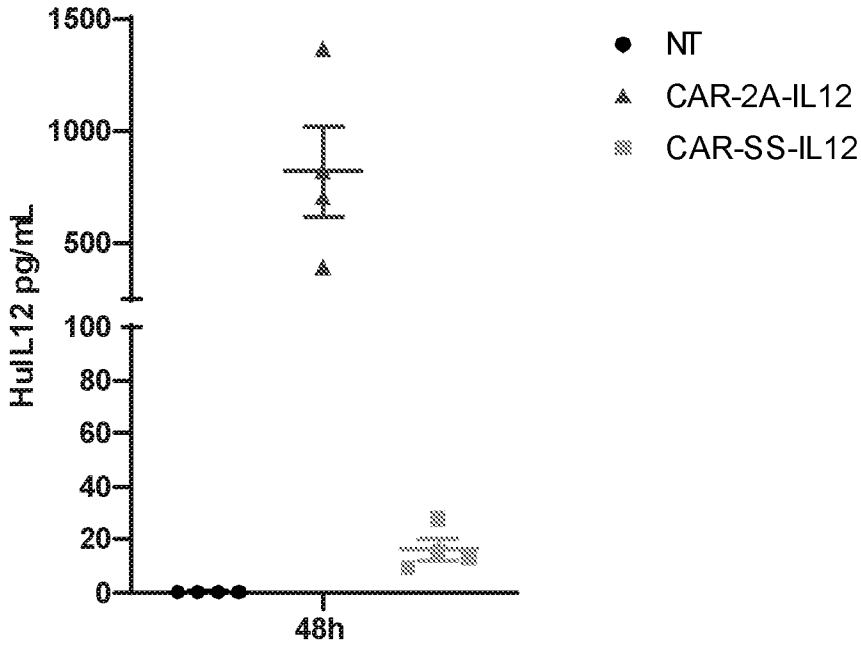


FIGURE 6

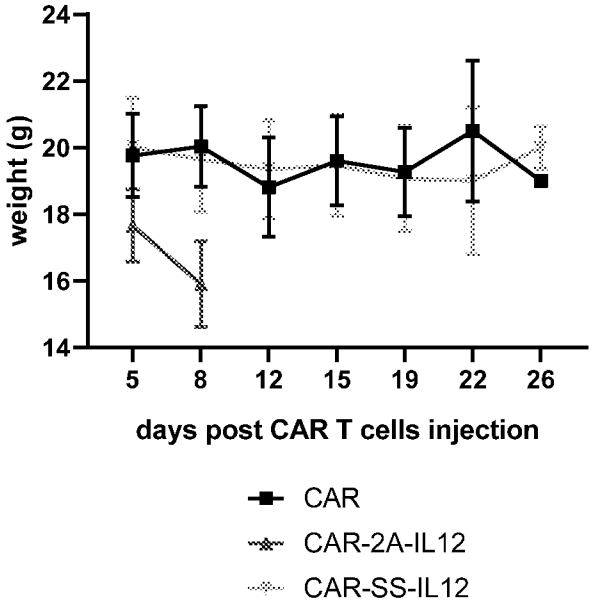
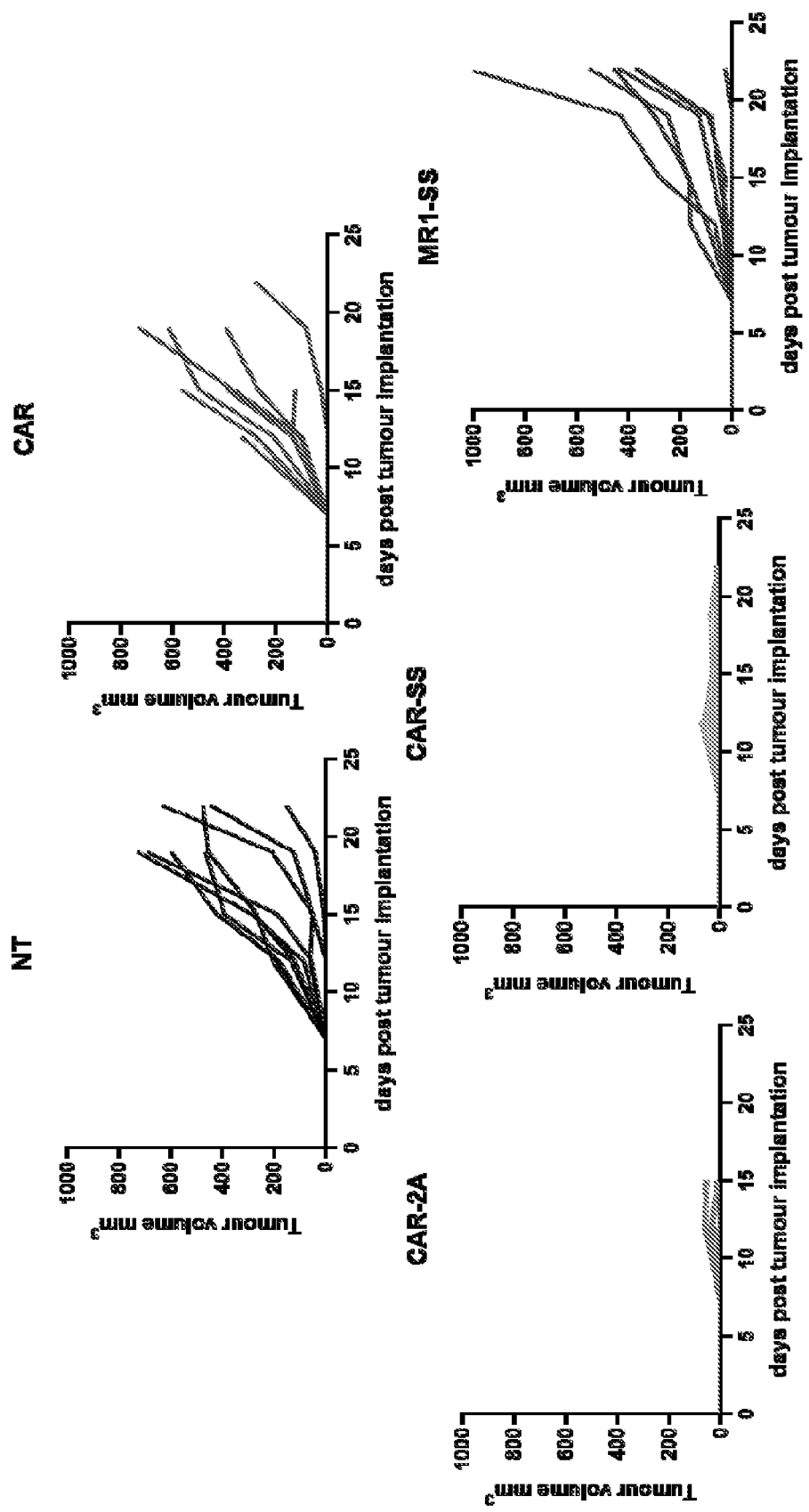

24
22
20
18
16
14
weight (g)
5
8
12
15
19
22
26
days post CAR T cells injection
CAR
CAR-2A-IL12
CAR-SS-IL12

FIGURE 7


NT
CAR
CAR-2A
CAR-SS
MR1-SS
Tumour volume mm3
days post tumour implantation
0
200
400
600
800
1000
0
5
10
15
20
25

SEQUENCE LISTING

<110> Autolus Limited

<120> METHOD

<130> P120285PCT

<150> GB 2007044.7
<151> 2020-05-13

<160> 104

<170> PatentIn version 3.5

<210> 1
<211> 540
<212> PRT
<213> Artificial Sequence

<220>
<223> flexi-IL-12 sequence

<400> 1

```
Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
1               5                  10                  15


Gly Ser Thr Gly Met Trp Ile Trp Glu Leu Lys Lys Asp Val Tyr Val
            20                  25                  30


Val Glu Leu Asp Trp Tyr Pro Asp Ala Pro Gly Glu Met Val Val Leu
        35                  40                  45


Thr Cys Asp Thr Pro Glu Glu Asp Gly Ile Thr Trp Thr Leu Asp Gln
    50                  55                  60


Ser Ser Glu Val Leu Gly Ser Gly Lys Thr Leu Thr Ile Gln Val Lys
65                  70                  75                  80


Glu Phe Gly Asp Ala Gly Gln Tyr Thr Cys His Lys Gly Gly Glu Val
                85                  90                  95


Leu Ser His Ser Leu Leu Leu Leu His Lys Lys Glu Asp Gly Ile Trp
            100                 105                 110


Ser Thr Asp Ile Leu Lys Asp Gln Lys Glu Pro Lys Asn Lys Thr Phe
```

```
        115                 120                 125

Leu Arg Cys Glu Ala Lys Asn Tyr Ser Gly Arg Phe Thr Cys Trp Trp
    130                 135                 140

Leu Thr Thr Ile Ser Thr Asp Leu Thr Phe Ser Val Lys Ser Ser Arg
145                 150                 155                 160

Gly Ser Ser Asp Pro Gln Gly Val Thr Cys Gly Ala Ala Thr Leu Ser
                165                 170                 175

Ala Glu Arg Val Arg Gly Asp Asn Lys Glu Tyr Glu Tyr Ser Val Glu
            180                 185                 190

Cys Gln Glu Asp Ser Ala Cys Pro Ala Ala Glu Glu Ser Leu Pro Ile
        195                 200                 205

Glu Val Met Val Asp Ala Val His Lys Leu Lys Tyr Glu Asn Tyr Thr
    210                 215                 220

Ser Ser Phe Phe Ile Arg Asp Ile Ile Lys Pro Asp Pro Pro Lys Asn
225                 230                 235                 240

Leu Gln Leu Lys Pro Leu Lys Asn Ser Arg Gln Val Glu Val Ser Trp
                245                 250                 255

Glu Tyr Pro Asp Thr Trp Ser Thr Pro His Ser Tyr Phe Ser Leu Thr
            260                 265                 270

Phe Cys Val Gln Val Gln Gly Lys Ser Lys Arg Glu Lys Lys Asp Arg
        275                 280                 285

Val Phe Thr Asp Lys Thr Ser Ala Thr Val Ile Cys Arg Lys Asn Ala
    290                 295                 300

Ser Ile Ser Val Arg Ala Gln Asp Arg Tyr Tyr Ser Ser Ser Trp Ser
305                 310                 315                 320

Glu Trp Ala Ser Val Pro Cys Ser Gly Gly Gly Gly Ser Gly Gly Gly
                325                 330                 335
```

```
Gly Ser Gly Gly Gly Gly Ser Arg Asn Leu Pro Leu Ala Thr Pro Asp
            340                 345                 350

Pro Gly Met Phe Pro Cys Leu His His Ser Gln Asn Leu Leu Arg Ala
        355                 360                 365

Val Ser Asn Met Leu Gln Lys Ala Arg Gln Thr Leu Glu Phe Tyr Pro
    370                 375                 380

Cys Thr Ser Glu Glu Ile Asp His Glu Asp Ile Thr Lys Asp Lys Thr
385                 390                 395                 400

Ser Thr Val Glu Ala Cys Leu Pro Leu Glu Leu Thr Lys Asn Glu Ser
                405                 410                 415

Cys Leu Asn Ser Arg Glu Thr Ser Phe Ile Thr Asn Gly Ser Cys Leu
            420                 425                 430

Ala Ser Arg Lys Thr Ser Phe Met Met Ala Leu Cys Leu Ser Ser Ile
        435                 440                 445

Tyr Glu Asp Ser Lys Met Tyr Gln Val Glu Phe Lys Thr Met Asn Ala
    450                 455                 460

Lys Leu Leu Met Asp Pro Lys Arg Gln Ile Phe Leu Asp Gln Asn Met
465                 470                 475                 480

Leu Ala Val Ile Asp Glu Leu Met Gln Ala Leu Asn Phe Asn Ser Glu
                485                 490                 495

Thr Val Pro Gln Lys Ser Ser Leu Glu Glu Pro Asp Phe Tyr Lys Thr
            500                 505                 510

Lys Ile Lys Leu Cys Ile Leu Leu His Ala Phe Arg Ile Arg Ala Val
        515                 520                 525

Thr Ile Asp Arg Val Met Ser Tyr Leu Asn Ala Ser
    530                 535                 540
```

<210> 2

<211> 7
<212> RNA
<213> Artificial Sequence

<220>
<223> frame-slip motif (FSM) sequence

<400> 2
uuuuuuu 7

<210> 3
<211> 9
<212> RNA
<213> Artificial Sequence

<220>
<223> FSM sequence

<400> 3
uuuuuuuga 9

<210> 4
<211> 9
<212> RNA
<213> Artificial Sequence

<220>
<223> FSM sequence

<400> 4
uuuuuuuag 9

<210> 5
<211> 9
<212> RNA
<213> Artificial Sequence

<220>
<223> FSM sequence

<400> 5
uuuuuuuaa 9

<210> 6
<211> 7
<212> RNA
<213> Artificial Sequence

<220>
<223> translational readthrough motif (TRM) sequence

<400> 6
ugacuag 7

<210> 7
<211> 7
<212> RNA
<213> Artificial Sequence

<220>
<223> TRM sequence

<400> 7
uagcuag 7

<210> 8
<211> 7
<212> RNA
<213> Artificial Sequence

<220>
<223> TRM sequence

<400> 8
uaacuag 7

<210> 9
<211> 9
<212> RNA
<213> Artificial Sequence

<220>
<223> TRM sequence

<400> 9
ugacaauua 9

<210> 10
<211> 9
<212> RNA
<213> Artificial Sequence

<220>
<223> TRM sequence

<400> 10
uagcaauua 9

<210> 11

<211> 9
<212> RNA
<213> Artificial Sequence

<220>
<223> TRM sequence

<400> 11
uaacaauua 9

<210> 12
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> heavy chain variable region (VH) complementarity determining region (CDR), CDR1

<400> 12

Ser Tyr Asn Ile His
1 5

<210> 13
<211> 16
<212> PRT
<213> Artificial Sequence

<220>
<223> VH CDR, CDR2

<400> 13

Val Ile Trp Ala Gly Gly Ser Thr Asn Tyr Asn Ser Ala Leu Met Ser
1 5 10 15

<210> 14
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> VH CDR, CDR3

<400> 14

Arg Ser Asp Asp Tyr Ser Trp Phe Ala Tyr
1 5 10

<210> 15
<211> 12
<212> PRT
<213> Artificial Sequence

<220>
<223> light chain variable region (VL) CDR, CDR1

<400> 15

Arg Ala Ser Ser Ser Val Ser Ser Ser Tyr Leu His
1 5 10

<210> 16
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> VL CDR, CDR2

<400> 16

Ser Thr Ser Asn Leu Ala Ser
1 5

<210> 17
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> VL CDR, CDR3

<400> 17

Gln Gln Tyr Ser Gly Tyr Pro Ile Thr
1 5

<210> 18
<211> 5
<212> PRT
<213> Artificial Sequence

<220>
<223> VH CDR, CDR1

<400> 18

Ser Ser Trp Met Asn
1 5

<210> 19
<211> 17
<212> PRT
<213> Artificial Sequence

<220>
<223> VH CDR, CDR2

<400> 19

Arg Ile Tyr Pro Gly Asp Gly Asp Thr Asn Tyr Ala Gln Lys Phe Gln
1 5 10 15

Gly

<210> 20
<211> 10
<212> PRT
<213> Artificial Sequence

<220>
<223> VH CDR, CDR3

<400> 20

Gly Thr Gly Tyr Leu Trp Tyr Phe Asp Val
1 5 10

<210> 21
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> VL CDR, CDR1

<400> 21

Arg Ala Ser Gln Asp Ile Asn Glu Asn Leu Ala
1 5 10

<210> 22
<211> 7
<212> PRT
<213> Artificial Sequence

<220>

<223> VL CDR, CDR2

<400> 22

Tyr Thr Ser Asn Arg Ala Thr
1   5

<210> 23
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> VL CDR, CDR3

<400> 23

Gln Gln Tyr Asp Asn Leu Pro Phe Thr
1   5

<210> 24
<211> 197
<212> PRT
<213> Homo sapiens

<400> 24

Arg Asn Leu Pro Val Ala Thr Pro Asp Pro Gly Met Phe Pro Cys Leu
1   5   10   15

His His Ser Gln Asn Leu Leu Arg Ala Val Ser Asn Met Leu Gln Lys
            20   25   30

Ala Arg Gln Thr Leu Glu Phe Tyr Pro Cys Thr Ser Glu Glu Ile Asp
        35   40   45

His Glu Asp Ile Thr Lys Asp Lys Thr Ser Thr Val Glu Ala Cys Leu
    50   55   60

Pro Leu Glu Leu Thr Lys Asn Glu Ser Cys Leu Asn Ser Arg Glu Thr
65   70   75   80

Ser Phe Ile Thr Asn Gly Ser Cys Leu Ala Ser Arg Lys Thr Ser Phe
                85   90   95

Met Met Ala Leu Cys Leu Ser Ser Ile Tyr Glu Asp Leu Lys Met Tyr

```
            100                 105                 110

Gln Val Glu Phe Lys Thr Met Asn Ala Lys Leu Leu Met Asp Pro Lys
        115                 120                 125

Arg Gln Ile Phe Leu Asp Gln Asn Met Leu Ala Val Ile Asp Glu Leu
    130                 135                 140

Met Gln Ala Leu Asn Phe Asn Ser Glu Thr Val Pro Gln Lys Ser Ser
145                 150                 155                 160

Leu Glu Glu Pro Asp Phe Tyr Lys Thr Lys Ile Lys Leu Cys Ile Leu
                165                 170                 175

Leu His Ala Phe Arg Ile Arg Ala Val Thr Ile Asp Arg Val Met Ser
            180                 185                 190

Tyr Leu Asn Ala Ser
        195


<210>  25
<211>  304
<212>  PRT
<213>  Homo sapiens

<400>  25

Trp Glu Leu Lys Lys Asp Val Tyr Val Val Glu Leu Asp Trp Tyr Pro
1               5                   10                  15

Asp Ala Pro Gly Glu Met Val Val Leu Thr Cys Asp Thr Pro Glu Glu
            20                  25                  30

Asp Gly Ile Thr Trp Thr Leu Asp Gln Ser Ser Glu Val Leu Gly Ser
        35                  40                  45

Gly Lys Thr Leu Thr Ile Gln Val Lys Glu Phe Gly Asp Ala Gly Gln
    50                  55                  60

Tyr Thr Cys His Lys Gly Gly Glu Val Leu Ser His Ser Leu Leu Leu
65                  70                  75                  80
```

```
Leu His Lys Lys Glu Asp Gly Ile Trp Ser Thr Asp Ile Leu Lys Asp
                85                  90                  95

Gln Lys Glu Pro Lys Asn Lys Thr Phe Leu Arg Cys Glu Ala Lys Asn
            100                 105                 110

Tyr Ser Gly Arg Phe Thr Cys Trp Trp Leu Thr Thr Ile Ser Thr Asp
        115                 120                 125

Leu Thr Phe Ser Val Lys Ser Ser Arg Gly Ser Ser Asp Pro Gln Gly
    130                 135                 140

Val Thr Cys Gly Ala Ala Thr Leu Ser Ala Glu Arg Val Arg Gly Asp
145                 150                 155                 160

Asn Lys Glu Tyr Glu Tyr Ser Val Glu Cys Gln Glu Asp Ser Ala Cys
                165                 170                 175

Pro Ala Ala Glu Glu Ser Leu Pro Ile Glu Val Met Val Asp Ala Val
            180                 185                 190

His Lys Leu Lys Tyr Glu Asn Tyr Thr Ser Ser Phe Phe Ile Arg Asp
        195                 200                 205

Ile Ile Lys Pro Asp Pro Pro Lys Asn Leu Gln Leu Lys Pro Leu Lys
    210                 215                 220

Asn Ser Arg Gln Val Glu Val Ser Trp Glu Tyr Pro Asp Thr Trp Ser
225                 230                 235                 240

Thr Pro His Ser Tyr Phe Ser Leu Thr Phe Cys Val Gln Val Gln Gly
                245                 250                 255

Lys Ser Lys Arg Glu Lys Lys Asp Arg Val Phe Thr Asp Lys Thr Ser
            260                 265                 270

Ala Thr Val Ile Cys Arg Lys Asn Ala Ser Ile Ser Val Arg Ala Gln
        275                 280                 285

Asp Arg Tyr Tyr Ser Ser Ser Trp Ser Glu Trp Ala Ser Val Pro Cys
    290                 295                 300
```

<210> 26
<211> 177
<212> PRT
<213> Homo sapiens

<400> 26

```
Met Phe His Val Ser Phe Arg Tyr Ile Phe Gly Leu Pro Pro Leu Ile
1               5                  10                  15


Leu Val Leu Leu Pro Val Ala Ser Ser Asp Cys Asp Ile Glu Gly Lys
            20                  25                  30


Asp Gly Lys Gln Tyr Glu Ser Val Leu Met Val Ser Ile Asp Gln Leu
        35                  40                  45


Leu Asp Ser Met Lys Glu Ile Gly Ser Asn Cys Leu Asn Asn Glu Phe
    50                  55                  60


Asn Phe Phe Lys Arg His Ile Cys Asp Ala Asn Lys Glu Gly Met Phe
65                  70                  75                  80


Leu Phe Arg Ala Ala Arg Lys Leu Arg Gln Phe Leu Lys Met Asn Ser
                85                  90                  95


Thr Gly Asp Phe Asp Leu His Leu Leu Lys Val Ser Glu Gly Thr Thr
            100                 105                 110


Ile Leu Leu Asn Cys Thr Gly Gln Val Lys Gly Arg Lys Pro Ala Ala
        115                 120                 125


Leu Gly Glu Ala Gln Pro Thr Lys Ser Leu Glu Glu Asn Lys Ser Leu
    130                 135                 140


Lys Glu Gln Lys Lys Leu Asn Asp Leu Cys Phe Leu Lys Arg Leu Leu
145                 150                 155                 160


Gln Glu Ile Lys Thr Cys Trp Asn Lys Ile Leu Met Gly Thr Lys Glu
                165                 170                 175


His
```

<210> 27
<211> 162
<212> PRT
<213> Homo sapiens

<400> 27

```
Met Arg Ile Ser Lys Pro His Leu Arg Ser Ile Ser Ile Gln Cys Tyr
1               5                  10                  15


Leu Cys Leu Leu Leu Asn Ser His Phe Leu Thr Glu Ala Gly Ile His
            20                  25                  30


Val Phe Ile Leu Gly Cys Phe Ser Ala Gly Leu Pro Lys Thr Glu Ala
        35                  40                  45


Asn Trp Val Asn Val Ile Ser Asp Leu Lys Lys Ile Glu Asp Leu Ile
    50                  55                  60


Gln Ser Met His Ile Asp Ala Thr Leu Tyr Thr Glu Ser Asp Val His
65                  70                  75                  80


Pro Ser Cys Lys Val Thr Ala Met Lys Cys Phe Leu Leu Glu Leu Gln
                85                  90                  95


Val Ile Ser Leu Glu Ser Gly Asp Ala Ser Ile His Asp Thr Val Glu
            100                 105                 110


Asn Leu Ile Ile Leu Ala Asn Asn Ser Leu Ser Ser Asn Gly Asn Val
        115                 120                 125


Thr Glu Ser Gly Cys Lys Glu Cys Glu Glu Leu Glu Glu Lys Asn Ile
    130                 135                 140


Lys Glu Phe Leu Gln Ser Phe Val His Ile Val Gln Met Phe Ile Asn
145                 150                 155                 160


Thr Ser
```

<210> 28
<211> 98
<212> PRT
<213> Homo sapiens

<400> 28

Met Ala Leu Leu Leu Ala Leu Ser Leu Leu Val Leu Trp Thr Ser Pro
1 5 10 15

Ala Pro Thr Leu Ser Gly Thr Asn Asp Ala Glu Asp Cys Cys Leu Ser
20 25 30

Val Thr Gln Lys Pro Ile Pro Gly Tyr Ile Val Arg Asn Phe His Tyr
35 40 45

Leu Leu Ile Lys Asp Gly Cys Arg Val Pro Ala Val Val Phe Thr Thr
50 55 60

Leu Arg Gly Arg Gln Leu Cys Ala Pro Pro Asp Gln Pro Trp Val Glu
65 70 75 80

Arg Ile Ile Gln Arg Leu Gln Arg Thr Ser Ala Lys Met Lys Arg Arg
85 90 95

Ser Ser

<210> 29
<211> 93
<212> PRT
<213> Homo sapiens

<400> 29

Met Asn Ala Lys Val Val Val Val Leu Val Leu Val Leu Thr Ala Leu
1 5 10 15

Cys Leu Ser Asp Gly Lys Pro Val Ser Leu Ser Tyr Arg Cys Pro Cys
20 25 30

Arg Phe Phe Glu Ser His Val Ala Arg Ala Asn Val Lys His Leu Lys
35 40 45

Ile Leu Asn Thr Pro Asn Cys Ala Leu Gln Ile Val Ala Arg Leu Lys
50 55 60

Asn Asn Asn Arg Gln Val Cys Ile Asp Pro Lys Leu Lys Trp Ile Gln
65 70 75 80

Glu Tyr Leu Glu Lys Ala Leu Asn Lys Arg Phe Lys Met
85 90

<210> 30
<211> 10
<212> RNA
<213> Artificial Sequence

<220>
<223> multiple stop translational readthrough motif

<400> 30
ugaugacuag 10

<210> 31
<211> 7
<212> DNA
<213> Artificial Sequence

<220>
<223> construct DNA sequence encoding a TRM

<400> 31
tgactag 7

<210> 32
<211> 7
<212> DNA
<213> Artificial Sequence

<220>
<223> construct DNA sequence encoding a TRM

<400> 32
tagctag 7

<210> 33
<211> 7
<212> DNA
<213> Artificial Sequence

<220>

<223> construct DNA sequence encoding a TRM

<400> 33
taactag 7

<210> 34
<211> 9
<212> DNA
<213> Artificial Sequence

<220>
<223> construct DNA sequence encoding a TRM

<400> 34
tgacaatta 9

<210> 35
<211> 9
<212> DNA
<213> Artificial Sequence

<220>
<223> construct DNA sequence encoding a TRM

<400> 35
tagcaatta 9

<210> 36
<211> 9
<212> DNA
<213> Artificial Sequence

<220>
<223> construct DNA sequence encoding a TRM

<400> 36
taacaatta 9

<210> 37
<211> 9
<212> DNA
<213> Artificial Sequence

<220>
<223> construct DNA sequence encoding a TRM

<400> 37
tgactagca 9

<210> 38
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> basic amino acid furin target sequence

<220>
<221> misc_feature
<222> (2)..(2)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> MISC_FEATURE
<222> (3)..(3)
<223> Xaa may be Arg or Lys

<400> 38

Arg Xaa Xaa Arg
1

<210> 39
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> consensus Tobacco Etch Virus (TEV) cleavage site

<400> 39

Glu Asn Leu Tyr Phe Gln Ser
1 5

<210> 40
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> motif involved in cleavage activity

<220>
<221> misc_feature
<222> (2)..(2)
<223> Xaa can be any naturally occurring amino acid

<220>

<221> misc_feature
<222> (4)..(4)
<223> Xaa can be any naturally occurring amino acid

<400> 40

Asp Xaa Glu Xaa Asn Pro Gly Pro
1 5

<210> 41
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage site motif

<400> 41

Asp Ile Glu Thr Asn Pro Gly Pro
1 5

<210> 42
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage site motif

<400> 42

Asp Val Glu Thr Asn Pro Gly Pro
1 5

<210> 43
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage site motif

<400> 43

Asp Val Glu Met Asn Pro Gly Pro
1 5

<210> 44
<211> 8

<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage site motif

<400> 44

Asp Val Glu Ser Asn Pro Gly Pro
1   5

<210> 45
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage site motif

<400> 45

Asp Met Glu Ser Asn Pro Gly Pro
1   5

<210> 46
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage site motif

<400> 46

Asp Val Glu Glu Asn Pro Gly Pro
1   5

<210> 47
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage site motif

<400> 47

Asp Ile Glu Leu Asn Pro Gly Pro
1   5

<210> 48
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage site motif

<400> 48

Asp Ile Glu Gln Asn Pro Gly Pro
1 5

<210> 49
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage site motif

<400> 49

Asp Ser Glu Phe Asn Pro Gly Pro
1 5

<210> 50
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage sequence

<400> 50

Leu Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val Glu Ser Asn
1 5 10 15

Pro Gly Pro

<210> 51
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage sequence

<400> 51

Leu Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val Gln Ser Asn
1 5 10 15

Pro Gly Pro

<210> 52
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage sequence

<400> 52

Leu Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val Glu Ile Asn
1 5 10 15

Pro Gly Pro

<210> 53
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage sequence

<400> 53

Leu Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val Glu Phe Asn
1 5 10 15

Pro Gly Pro

<210> 54
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage sequence

<400> 54

Leu Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val Glu Ser His
1 5 10 15

Pro Gly Pro

<210> 55
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage sequence

<400> 55

Leu Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val Glu Ser Glu
1 5 10 15

Pro Gly Pro

<210> 56
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage sequence

<400> 56

Leu Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val Glu Ser Gln
1 5 10 15

Pro Gly Pro

<210> 57
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage sequence

<400> 57

Leu Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val Glu Ser Asn
1 5 10 15

Pro Gly Gly

<210> 58
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage site, 2A-like sequence

<400> 58

Tyr His Ala Asp Tyr Tyr Lys Gln Arg Leu Ile His Asp Val Glu Met
1 5 10 15

Asn Pro Gly Pro
20

<210> 59
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage site, 2A-like sequence

<400> 59

His Tyr Ala Gly Tyr Phe Ala Asp Leu Leu Ile His Asp Ile Glu Thr
1 5 10 15

Asn Pro Gly Pro
20

<210> 60
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage site, 2A-like sequence

<400> 60

Gln Cys Thr Asn Tyr Ala Leu Leu Lys Leu Ala Gly Asp Val Glu Ser
1 5 10 15

Asn Pro Gly Pro
20

<210> 61
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage site, 2A-like sequence

<400> 61

Ala Thr Asn Phe Ser Leu Leu Lys Gln Ala Gly Asp Val Glu Glu Asn
1 5 10 15

Pro Gly Pro

<210> 62
<211> 19
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage site, 2A-like sequence

<400> 62

Ala Ala Arg Gln Met Leu Leu Leu Leu Ser Gly Asp Val Glu Thr Asn
1 5 10 15

Pro Gly Pro

<210> 63
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage site, 2A-like sequence

<400> 63

Arg Ala Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu
1 5 10 15

Asn Pro Gly Pro
20

<210> 64
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage site, 2A-like sequence

<400> 64

Thr Arg Ala Glu Ile Glu Asp Glu Leu Ile Arg Ala Gly Ile Glu Ser
1 5 10 15

Asn Pro Gly Pro
20

<210> 65
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage site, 2A-like sequence

<400> 65

Thr Arg Ala Glu Ile Glu Asp Glu Leu Ile Arg Ala Asp Ile Glu Ser
1 5 10 15

Asn Pro Gly Pro
20

<210> 66
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage site, 2A-like sequence

<400> 66

Ala Lys Phe Gln Ile Asp Lys Ile Leu Ile Ser Gly Asp Val Glu Leu
1 5 10 15

Asn Pro Gly Pro
20

<210> 67
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage site, 2A-like sequence

<400> 67

Ser Ser Ile Ile Arg Thr Lys Met Leu Val Ser Gly Asp Val Glu Glu
1 5 10 15

Asn Pro Gly Pro
20

<210> 68
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage site, 2A-like sequence

<400> 68

Cys Asp Ala Gln Arg Gln Lys Leu Leu Leu Ser Gly Asp Ile Glu Gln
1 5 10 15

Asn Pro Gly Pro
20

<210> 69
<211> 20
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage site, 2A-like sequence

<400> 69

Tyr Pro Ile Asp Phe Gly Gly Phe Leu Val Lys Ala Asp Ser Glu Phe
1 5 10 15

Asn Pro Gly Pro
20

<210> 70
<211> 18
<212> PRT
<213> Artificial Sequence

<220>
<223> 2A peptide from Thosea asigna virus capsid protein

<400> 70

Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu Asn Pro
1 5 10 15

Gly Pro

<210> 71
<211> 57
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage sequence

<400> 71

Val Thr Glu Leu Leu Tyr Arg Met Lys Arg Ala Glu Thr Tyr Cys Pro
1 5 10 15

Arg Pro Leu Ala Ile His Pro Thr Glu Ala Arg His Lys Gln Lys Ile
20 25 30

Val Ala Pro Val Lys Gln Thr Leu Asn Phe Asp Leu Leu Lys Leu Ala
35 40 45

Gly Asp Val Glu Ser Asn Pro Gly Pro
50 55

<210> 72
<211> 40
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage sequence

<400> 72

Leu Leu Ala Ile His Pro Thr Glu Ala Arg His Lys Gln Lys Ile Val
1 5 10 15

Ala Pro Val Lys Gln Thr Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly
20 25 30

Asp Val Glu Ser Asn Pro Gly Pro
35 40

<210> 73
<211> 33
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage sequence

<400> 73

Glu Ala Arg His Lys Gln Lys Ile Val Ala Pro Val Lys Gln Thr Leu
1 5 10 15

Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val Glu Ser Asn Pro Gly
20 25 30

Pro

<210> 74
<211> 24
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage sequence

<400> 74

```
Ala Pro Val Lys Gln Thr Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly
1               5                   10                  15


Asp Val Glu Ser Asn Pro Gly Pro
            20
```

<210> 75
<211> 118
<212> PRT
<213> Artificial Sequence

<220>
<223> Humanised KM666 VH sequence

<400> 75

```
Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Lys Pro Ser Gln
1               5                   10                  15


Thr Leu Ser Ile Thr Cys Thr Val Ser Gly Phe Ser Leu Ala Ser Tyr
            20                  25                  30


Asn Ile His Trp Val Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp Leu
        35                  40                  45


Gly Val Ile Trp Ala Gly Gly Ser Thr Asn Tyr Asn Ser Ala Leu Met
    50                  55                  60


Ser Arg Leu Thr Ile Ser Lys Asp Asn Ser Lys Asn Gln Val Phe Leu
65                  70                  75                  80


Lys Met Ser Ser Leu Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala
                85                  90                  95


Lys Arg Ser Asp Asp Tyr Ser Trp Phe Ala Tyr Trp Gly Gln Gly Thr
            100                 105                 110


Leu Val Thr Val Ser Ser
        115
```

<210> 76
<211> 108
<212> PRT
<213> Artificial Sequence

<220>
<223> Humanised KM666 VL sequence

<400> 76

```
Glu Asn Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
1               5                  10                  15


Asp Arg Val Thr Met Thr Cys Arg Ala Ser Ser Ser Val Ser Ser Ser
            20                  25                  30


Tyr Leu His Trp Tyr Gln Gln Lys Ser Gly Lys Ala Pro Lys Val Trp
        35                  40                  45


Ile Tyr Ser Thr Ser Asn Leu Ala Ser Gly Val Pro Ser Arg Phe Ser
    50                  55                  60


Gly Ser Gly Ser Gly Thr Asp Tyr Thr Leu Thr Ile Ser Ser Leu Gln
65                  70                  75                  80


Pro Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Tyr Ser Gly Tyr Pro
                85                  90                  95


Ile Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys
            100                 105
```

<210> 77
<211> 119
<212> PRT
<213> Artificial Sequence

<220>
<223> Humanised 7A12 VH sequence

<400> 77

```
Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ala
1               5                  10                  15


Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Ser
            20                  25                  30


Trp Met Asn Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met
        35                  40                  45
```

Gly Arg Ile Tyr Pro Gly Asp Gly Asp Thr Asn Tyr Ala Gln Lys Phe
    50   55   60

Gln Gly Arg Val Thr Met Thr Arg Asp Thr Ser Ile Ser Thr Ala Tyr
65   70   75   80

Met Glu Leu Ser Arg Leu Arg Ser Asp Asp Thr Ala Val Tyr Tyr Cys
                85   90   95

Ala Arg Gly Thr Gly Tyr Leu Trp Tyr Phe Asp Val Trp Gly Gln Gly
            100   105   110

Thr Met Val Thr Val Ser Ser
        115

<210> 78
<211> 108
<212> PRT
<213> Artificial Sequence

<220>
<223> Humanised 7A12 VL sequence

<400> 78

Glu Ile Val Leu Thr Gln Ser Pro Ala Thr Leu Ser Leu Ser Pro Gly
1   5   10   15

Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Asp Ile Asn Glu Asn
            20   25   30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu Ile
        35   40   45

Tyr Tyr Thr Ser Asn Arg Ala Thr Gly Ile Pro Ala Arg Phe Ser Gly
    50   55   60

Ser Gly Ser Gly Arg Asp Phe Thr Leu Thr Ile Ser Ser Leu Glu Pro
65   70   75   80

Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Tyr Asp Asn Leu Pro Phe
                85   90   95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg
100 105

<210> 79
<211> 41
<212> PRT
<213> Artificial Sequence

<220>
<223> CD28 endodomain

<400> 79

Arg Ser Lys Arg Ser Arg Leu Leu His Ser Asp Tyr Met Asn Met Thr
1 5 10 15

Pro Arg Arg Pro Gly Pro Thr Arg Lys His Tyr Gln Pro Tyr Ala Pro
20 25 30

Pro Arg Asp Phe Ala Ala Tyr Arg Ser
35 40

<210> 80
<211> 112
<212> PRT
<213> Artificial Sequence

<220>
<223> CD3z endodomain

<400> 80

Arg Val Lys Phe Ser Arg Ser Ala Asp Ala Pro Ala Tyr Gln Gln Gly
1 5 10 15

Gln Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr
20 25 30

Asp Val Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys
35 40 45

Pro Arg Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys
50 55 60

```
Asp Lys Met Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg
65                  70                  75                  80


Arg Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala
                85                  90                  95


Thr Lys Asp Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg
            100                 105                 110


<210>  81
<211>  4
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  ITAM (immunoreceptor tyrosine-based activation motif)


<220>
<221>  misc_feature
<222>  (2)..(3)
<223>  Xaa can be any naturally occurring amino acid

<220>
<221>  MISC_FEATURE
<222>  (4)..(4)
<223>  Xaa may be Leu or Ile

<400>  81

Tyr Xaa Xaa Xaa
1


<210>  82
<211>  214
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  SHP-1 SH2 complete domain

<400>  82

Met Val Arg Trp Phe His Arg Asp Leu Ser Gly Leu Asp Ala Glu Thr
1               5                   10                  15


Leu Leu Lys Gly Arg Gly Val His Gly Ser Phe Leu Ala Arg Pro Ser
            20                  25                  30
```

```
Arg Lys Asn Gln Gly Asp Phe Ser Leu Ser Val Arg Val Gly Asp Gln
        35                  40                  45

Val Thr His Ile Arg Ile Gln Asn Ser Gly Asp Phe Tyr Asp Leu Tyr
    50                  55                  60

Gly Gly Glu Lys Phe Ala Thr Leu Thr Glu Leu Val Glu Tyr Tyr Thr
65                  70                  75                  80

Gln Gln Gln Gly Val Leu Gln Asp Arg Asp Gly Thr Ile Ile His Leu
                85                  90                  95

Lys Tyr Pro Leu Asn Cys Ser Asp Pro Thr Ser Glu Arg Trp Tyr His
            100                 105                 110

Gly His Met Ser Gly Gly Gln Ala Glu Thr Leu Leu Gln Ala Lys Gly
        115                 120                 125

Glu Pro Trp Thr Phe Leu Val Arg Glu Ser Leu Ser Gln Pro Gly Asp
    130                 135                 140

Phe Val Leu Ser Val Leu Ser Asp Gln Pro Lys Ala Gly Pro Gly Ser
145                 150                 155                 160

Pro Leu Arg Val Thr His Ile Lys Val Met Cys Glu Gly Gly Arg Tyr
                165                 170                 175

Thr Val Gly Gly Leu Glu Thr Phe Asp Ser Leu Thr Asp Leu Val Glu
            180                 185                 190

His Phe Lys Lys Thr Gly Ile Glu Glu Ala Ser Gly Ala Phe Val Tyr
        195                 200                 205

Leu Arg Gln Pro Tyr Tyr
    210
```

<210> 83
<211> 97
<212> PRT
<213> Artificial Sequence

<220>
<223> SHP-1 SH2 domain 1

<400> 83

Trp Phe His Arg Asp Leu Ser Gly Leu Asp Ala Glu Thr Leu Leu Lys
1   5   10   15

Gly Arg Gly Val His Gly Ser Phe Leu Ala Arg Pro Ser Arg Lys Asn
            20   25   30

Gln Gly Asp Phe Ser Leu Ser Val Arg Val Gly Asp Gln Val Thr His
        35   40   45

Ile Arg Ile Gln Asn Ser Gly Asp Phe Tyr Asp Leu Tyr Gly Gly Glu
    50   55   60

Lys Phe Ala Thr Leu Thr Glu Leu Val Glu Tyr Tyr Thr Gln Gln Gln
65   70   75   80

Gly Val Leu Gln Asp Arg Asp Gly Thr Ile Ile His Leu Lys Tyr Pro
                85   90   95

Leu

<210> 84
<211> 104
<212> PRT
<213> Artificial Sequence

<220>
<223> SHP-1 SH2 domain 2

<400> 84

Trp Tyr His Gly His Met Ser Gly Gly Gln Ala Glu Thr Leu Leu Gln
1   5   10   15

Ala Lys Gly Glu Pro Trp Thr Phe Leu Val Arg Glu Ser Leu Ser Gln
            20   25   30

Pro Gly Asp Phe Val Leu Ser Val Leu Ser Asp Gln Pro Lys Ala Gly
        35   40   45

Pro Gly Ser Pro Leu Arg Val Thr His Ile Lys Val Met Cys Glu Gly
    50   55   60

Gly Arg Tyr Thr Val Gly Gly Leu Glu Thr Phe Asp Ser Leu Thr Asp
65   70   75   80

Leu Val Glu His Phe Lys Lys Thr Gly Ile Glu Glu Ala Ser Gly Ala
                85   90   95

Phe Val Tyr Leu Arg Gln Pro Tyr
            100

<210> 85
<211> 97
<212> PRT
<213> Artificial Sequence

<220>
<223> SHP-2 first SH2 domain

<400> 85

Trp Phe His Pro Asn Ile Thr Gly Val Glu Ala Glu Asn Leu Leu Leu
1   5   10   15

Thr Arg Gly Val Asp Gly Ser Phe Leu Ala Arg Pro Ser Lys Ser Asn
            20   25   30

Pro Gly Asp Phe Thr Leu Ser Val Arg Arg Asn Gly Ala Val Thr His
        35   40   45

Ile Lys Ile Gln Asn Thr Gly Asp Tyr Tyr Asp Leu Tyr Gly Gly Glu
    50   55   60

Lys Phe Ala Thr Leu Ala Glu Leu Val Gln Tyr Tyr Met Glu His His
65   70   75   80

Gly Gln Leu Lys Glu Lys Asn Gly Asp Val Ile Glu Leu Lys Tyr Pro
                85   90   95

Leu

<210> 86
<211> 105
<212> PRT
<213> Artificial Sequence

<220>
<223> SHP-2 second SH2 domain

<400> 86

Trp Phe His Gly His Leu Ser Gly Lys Glu Ala Glu Lys Leu Leu Thr
1 5 10 15

Glu Lys Gly Lys His Gly Ser Phe Leu Val Arg Glu Ser Gln Ser His
20 25 30

Pro Gly Asp Phe Val Leu Ser Val Arg Thr Gly Asp Asp Lys Gly Glu
35 40 45

Ser Asn Asp Gly Lys Ser Lys Val Thr His Val Met Ile Arg Cys Gln
50 55 60

Glu Leu Lys Tyr Asp Val Gly Gly Gly Glu Arg Phe Asp Ser Leu Thr
65 70 75 80

Asp Leu Val Glu His Tyr Lys Lys Asn Pro Met Val Glu Thr Leu Gly
85 90 95

Thr Val Leu Gln Leu Lys Gln Pro Leu
100 105

<210> 87
<211> 211
<212> PRT
<213> Artificial Sequence

<220>
<223> SHP-2 tandem SH2 domains

<400> 87

Trp Phe His Pro Asn Ile Thr Gly Val Glu Ala Glu Asn Leu Leu Leu
1 5 10 15

Thr Arg Gly Val Asp Gly Ser Phe Leu Ala Arg Pro Ser Lys Ser Asn

```
            20                  25                  30

Pro Gly Asp Phe Thr Leu Ser Val Arg Arg Asn Gly Ala Val Thr His
        35                  40                  45

Ile Lys Ile Gln Asn Thr Gly Asp Tyr Tyr Asp Leu Tyr Gly Gly Glu
    50                  55                  60

Lys Phe Ala Thr Leu Ala Glu Leu Val Gln Tyr Tyr Met Glu His His
65                  70                  75                  80

Gly Gln Leu Lys Glu Lys Asn Gly Asp Val Ile Glu Leu Lys Tyr Pro
                85                  90                  95

Leu Asn Cys Ala Asp Pro Thr Ser Glu Arg Trp Phe His Gly His Leu
            100                 105                 110

Ser Gly Lys Glu Ala Glu Lys Leu Leu Thr Glu Lys Gly Lys His Gly
        115                 120                 125

Ser Phe Leu Val Arg Glu Ser Gln Ser His Pro Gly Asp Phe Val Leu
    130                 135                 140

Ser Val Arg Thr Gly Asp Asp Lys Gly Glu Ser Asn Asp Gly Lys Ser
145                 150                 155                 160

Lys Val Thr His Val Met Ile Arg Cys Gln Glu Leu Lys Tyr Asp Val
                165                 170                 175

Gly Gly Gly Glu Arg Phe Asp Ser Leu Thr Asp Leu Val Glu His Tyr
            180                 185                 190

Lys Lys Asn Pro Met Val Glu Thr Leu Gly Thr Val Leu Gln Leu Lys
        195                 200                 205

Gln Pro Leu
    210
```

<210> 88
<211> 177
<212> PRT

<213> Artificial Sequence

<220>
<223> dominant negative TGFbeta receptor II

<400> 88

Thr Ile Pro Pro His Val Gln Lys Ser Val Asn Asn Asp Met Ile Val
1 5 10 15

Thr Asp Asn Asn Gly Ala Val Lys Phe Pro Gln Leu Cys Lys Phe Cys
20 25 30

Asp Val Arg Phe Ser Thr Cys Asp Asn Gln Lys Ser Cys Met Ser Asn
35 40 45

Cys Ser Ile Thr Ser Ile Cys Glu Lys Pro Gln Glu Val Cys Val Ala
50 55 60

Val Trp Arg Lys Asn Asp Glu Asn Ile Thr Leu Glu Thr Val Cys His
65 70 75 80

Asp Pro Lys Leu Pro Tyr His Asp Phe Ile Leu Glu Asp Ala Ala Ser
85 90 95

Pro Lys Cys Ile Met Lys Glu Lys Lys Lys Pro Gly Glu Thr Phe Phe
100 105 110

Met Cys Ser Cys Ser Ser Asp Glu Cys Asn Asp Asn Ile Ile Phe Ser
115 120 125

Glu Glu Tyr Asn Thr Ser Asn Pro Asp Leu Leu Leu Val Ile Phe Gln
130 135 140

Val Thr Gly Ile Ser Leu Leu Pro Pro Leu Gly Val Ala Ile Ser Val
145 150 155 160

Ile Ile Ile Phe Tyr Cys Tyr Arg Val Asn Arg Gln Gln Lys Leu Ser
165 170 175

Ser

<210> 89
<211> 107
<212> PRT
<213> Artificial Sequence

<220>
<223> Human Ig Kappa constant domain

<400> 89

Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu
1 5 10 15

Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe
20 25 30

Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln
35 40 45

Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser
50 55 60

Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu
65 70 75 80

Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser
85 90 95

Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
100 105

<210> 90
<211> 15
<212> PRT
<213> Artificial Sequence

<220>
<223> Human Hinge

<400> 90

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro
1 5 10 15

<210> 91

<211> 97
<212> PRT
<213> Artificial Sequence

<220>
<223> Human IgG1 CH1 domain

<400> 91

Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys Ser
1 5 10 15

Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe
20 25 30

Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly
35 40 45

Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu
50 55 60

Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr
65 70 75 80

Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys Arg
85 90 95

Val

<210> 92
<211> 21
<212> PRT
<213> Artificial Sequence

<220>
<223> Transmembrane domain from human IL2R common gamma chain

<400> 92

Val Val Ile Ser Val Gly Ser Met Gly Leu Ile Ile Ser Leu Leu Cys
1 5 10 15

Val Tyr Phe Trp Leu
20

<210> 93
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Transmembrane domain from human IL-2R beta

<400> 93

Ile Pro Trp Leu Gly His Leu Leu Val Gly Leu Ser Gly Ala Phe Gly
1 5 10 15

Phe Ile Ile Leu Val Tyr Leu Leu Ile
20 25

<210> 94
<211> 25
<212> PRT
<213> Artificial Sequence

<220>
<223> Transmembrane domain from human IL-7R alpha

<400> 94

Pro Ile Leu Leu Thr Ile Ser Ile Leu Ser Phe Phe Ser Val Ala Leu
1 5 10 15

Leu Val Ile Leu Ala Cys Val Leu Trp
20 25

<210> 95
<211> 26
<212> PRT
<213> Artificial Sequence

<220>
<223> Transmembrane domain from Human GF-CSFR alpha

<400> 95

Asn Leu Gly Ser Val Tyr Ile Tyr Val Leu Leu Ile Val Gly Thr Leu
1 5 10 15

Val Cys Gly Ile Val Leu Gly Phe Leu Phe
20 25

<210> 96
<211> 17
<212> PRT
<213> Artificial Sequence

<220>
<223> Transmembrane domain from Human GM-CSFR common beta chain

<400> 96

Val Leu Ala Leu Ile Val Ile Phe Leu Thr Ile Ala Val Leu Leu Ala
1   5   10   15

Leu

<210> 97
<211> 86
<212> PRT
<213> Artificial Sequence

<220>
<223> Endodomain from human IL2R common gamma chain

<400> 97

Glu Arg Thr Met Pro Arg Ile Pro Thr Leu Lys Asn Leu Glu Asp Leu
1   5   10   15

Val Thr Glu Tyr His Gly Asn Phe Ser Ala Trp Ser Gly Val Ser Lys
            20   25   30

Gly Leu Ala Glu Ser Leu Gln Pro Asp Tyr Ser Glu Arg Leu Cys Leu
        35   40   45

Val Ser Glu Ile Pro Pro Lys Gly Gly Ala Leu Gly Glu Gly Pro Gly
    50   55   60

Ala Ser Pro Cys Asn Gln His Ser Pro Tyr Trp Ala Pro Pro Cys Tyr
65   70   75   80

Thr Leu Lys Pro Glu Thr
                85

<210> 98

<211> 286
<212> PRT
<213> Artificial Sequence

<220>
<223> Endodomain from human IL-2R beta

<400> 98

Asn Cys Arg Asn Thr Gly Pro Trp Leu Lys Lys Val Leu Lys Cys Asn
1 5 10 15

Thr Pro Asp Pro Ser Lys Phe Phe Ser Gln Leu Ser Ser Glu His Gly
20 25 30

Gly Asp Val Gln Lys Trp Leu Ser Ser Pro Phe Pro Ser Ser Ser Phe
35 40 45

Ser Pro Gly Gly Leu Ala Pro Glu Ile Ser Pro Leu Glu Val Leu Glu
50 55 60

Arg Asp Lys Val Thr Gln Leu Leu Leu Gln Gln Asp Lys Val Pro Glu
65 70 75 80

Pro Ala Ser Leu Ser Ser Asn His Ser Leu Thr Ser Cys Phe Thr Asn
85 90 95

Gln Gly Tyr Phe Phe Phe His Leu Pro Asp Ala Leu Glu Ile Glu Ala
100 105 110

Cys Gln Val Tyr Phe Thr Tyr Asp Pro Tyr Ser Glu Glu Asp Pro Asp
115 120 125

Glu Gly Val Ala Gly Ala Pro Thr Gly Ser Ser Pro Gln Pro Leu Gln
130 135 140

Pro Leu Ser Gly Glu Asp Asp Ala Tyr Cys Thr Phe Pro Ser Arg Asp
145 150 155 160

Asp Leu Leu Leu Phe Ser Pro Ser Leu Leu Gly Gly Pro Ser Pro Pro
165 170 175

Ser Thr Ala Pro Gly Gly Ser Gly Ala Gly Glu Glu Arg Met Pro Pro

180 185 190

Ser Leu Gln Glu Arg Val Pro Arg Asp Trp Asp Pro Gln Pro Leu Gly
195 200 205

Pro Pro Thr Pro Gly Val Pro Asp Leu Val Asp Phe Gln Pro Pro Pro
210 215 220

Glu Leu Val Leu Arg Glu Ala Gly Glu Glu Val Pro Asp Ala Gly Pro
225 230 235 240

Arg Glu Gly Val Ser Phe Pro Trp Ser Arg Pro Pro Gly Gln Gly Glu
245 250 255

Phe Arg Ala Leu Asn Ala Arg Leu Pro Leu Asn Thr Asp Ala Tyr Leu
260 265 270

Ser Leu Gln Glu Leu Gln Gly Gln Asp Pro Thr His Leu Val
275 280 285

<210> 99
<211> 195
<212> PRT
<213> Artificial Sequence

<220>
<223> Endodomain from human IL-7R alpha

<400> 99

Lys Lys Arg Ile Lys Pro Ile Val Trp Pro Ser Leu Pro Asp His Lys
1 5 10 15

Lys Thr Leu Glu His Leu Cys Lys Lys Pro Arg Lys Asn Leu Asn Val
20 25 30

Ser Phe Asn Pro Glu Ser Phe Leu Asp Cys Gln Ile His Arg Val Asp
35 40 45

Asp Ile Gln Ala Arg Asp Glu Val Glu Gly Phe Leu Gln Asp Thr Phe
50 55 60

Pro Gln Gln Leu Glu Glu Ser Glu Lys Gln Arg Leu Gly Gly Asp Val

65 70 75 80

Gln Ser Pro Asn Cys Pro Ser Glu Asp Val Val Ile Thr Pro Glu Ser
85 90 95

Phe Gly Arg Asp Ser Ser Leu Thr Cys Leu Ala Gly Asn Val Ser Ala
100 105 110

Cys Asp Ala Pro Ile Leu Ser Ser Ser Arg Ser Leu Asp Cys Arg Glu
115 120 125

Ser Gly Lys Asn Gly Pro His Val Tyr Gln Asp Leu Leu Leu Ser Leu
130 135 140

Gly Thr Thr Asn Ser Thr Leu Pro Pro Pro Phe Ser Leu Gln Ser Gly
145 150 155 160

Ile Leu Thr Leu Asn Pro Val Ala Gln Gly Gln Pro Ile Leu Thr Ser
165 170 175

Leu Gly Ser Asn Gln Glu Glu Ala Tyr Val Thr Met Ser Ser Phe Tyr
180 185 190

Gln Asn Gln
195

<210> 100
<211> 54
<212> PRT
<213> Artificial Sequence

<220>
<223> Endodomain Derived from Human GM-CSFR alpha

<400> 100

Lys Arg Phe Leu Arg Ile Gln Arg Leu Phe Pro Pro Val Pro Gln Ile
1 5 10 15

Lys Asp Lys Leu Asn Asp Asn His Glu Val Glu Asp Glu Ile Ile Trp
20 25 30

Glu Glu Phe Thr Pro Glu Glu Gly Lys Gly Tyr Arg Glu Glu Val Leu

35 40 45

Thr Val Lys Glu Ile Thr
50

<210> 101
<211> 437
<212> PRT
<213> Artificial Sequence

<220>
<223> Endodomain from GM-CSFR common beta chain

<400> 101

Arg Phe Cys Gly Ile Tyr Gly Tyr Arg Leu Arg Arg Lys Trp Glu Glu
1 5 10 15

Lys Ile Pro Asn Pro Ser Lys Ser His Leu Phe Gln Asn Gly Ser Ala
20 25 30

Glu Leu Trp Pro Pro Gly Ser Met Ser Ala Phe Thr Ser Gly Ser Pro
35 40 45

Pro His Gln Gly Pro Trp Gly Ser Arg Phe Pro Glu Leu Glu Gly Val
50 55 60

Phe Pro Val Gly Phe Gly Asp Ser Glu Val Ser Pro Leu Thr Ile Glu
65 70 75 80

Asp Pro Lys His Val Cys Asp Pro Pro Ser Gly Pro Asp Thr Thr Pro
85 90 95

Ala Ala Ser Asp Leu Pro Thr Glu Gln Pro Pro Ser Pro Gln Pro Gly
100 105 110

Pro Pro Ala Ala Ser His Thr Pro Glu Lys Gln Ala Ser Ser Phe Asp
115 120 125

Phe Asn Gly Pro Tyr Leu Gly Pro Pro His Ser Arg Ser Leu Pro Asp
130 135 140

Ile Leu Gly Gln Pro Glu Pro Pro Gln Glu Gly Gly Ser Gln Lys Ser

```
145                 150                 155                 160

Pro Pro Pro Gly Ser Leu Glu Tyr Leu Cys Leu Pro Ala Gly Gly Gln
                165                 170                 175

Val Gln Leu Val Pro Leu Ala Gln Ala Met Gly Pro Gly Gln Ala Val
            180                 185                 190

Glu Val Glu Arg Arg Pro Ser Gln Gly Ala Ala Gly Ser Pro Ser Leu
        195                 200                 205

Glu Ser Gly Gly Gly Pro Ala Pro Pro Ala Leu Gly Pro Arg Val Gly
    210                 215                 220

Gly Gln Asp Gln Lys Asp Ser Pro Val Ala Ile Pro Met Ser Ser Gly
225                 230                 235                 240

Asp Thr Glu Asp Pro Gly Val Ala Ser Gly Tyr Val Ser Ser Ala Asp
                245                 250                 255

Leu Val Phe Thr Pro Asn Ser Gly Ala Ser Ser Val Ser Leu Val Pro
            260                 265                 270

Ser Leu Gly Leu Pro Ser Asp Gln Thr Pro Ser Leu Cys Pro Gly Leu
        275                 280                 285

Ala Ser Gly Pro Pro Gly Ala Pro Gly Pro Val Lys Ser Gly Phe Glu
    290                 295                 300

Gly Tyr Val Glu Leu Pro Pro Ile Glu Gly Arg Ser Pro Arg Ser Pro
305                 310                 315                 320

Arg Asn Asn Pro Val Pro Pro Glu Ala Lys Ser Pro Val Leu Asn Pro
                325                 330                 335

Gly Glu Arg Pro Ala Asp Val Ser Pro Thr Ser Pro Gln Pro Glu Gly
            340                 345                 350

Leu Leu Val Leu Gln Gln Val Gly Asp Tyr Cys Phe Leu Pro Gly Leu
        355                 360                 365
```

Gly Pro Gly Pro Leu Ser Leu Arg Ser Lys Pro Ser Ser Pro Gly Pro
370 375 380

Gly Pro Glu Ile Lys Asn Leu Asp Gln Ala Phe Gln Val Lys Lys Pro
385 390 395 400

Pro Gly Gln Ala Val Pro Gln Val Pro Val Ile Gln Leu Phe Lys Ala
405 410 415

Leu Lys Gln Gln Asp Tyr Leu Ser Leu Pro Pro Trp Glu Val Asn Lys
420 425 430

Pro Gly Glu Val Cys
435

<210> 102
<211> 6
<212> PRT
<213> Artificial Sequence

<220>
<223> CXCL12 loop region motif

<400> 102

Arg Phe Phe Glu Ser His
1 5

<210> 103
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> cleavage site motif

<400> 103

Asp Val Glu Leu Asn Pro Gly Pro
1 5

<210> 104
<211> 4
<212> PRT
<213> Artificial Sequence

<220>
<223> Mucin 1 motif

<400> 104

Pro Asp Thr Arg
1