



(51) International Patent Classification:

C12N 15/63 (2006.01) C12N 15/09 (2006.01)
C07K 14/575 (2006.01) C12N 15/11 (2006.01)
C07K 14/605 (2006.01) C12N 15/13 (2006.01)
C07K 16/00 (2006.01) C12N 15/62 (2006.01)
C07K 19/00 (2006.01) C12N 15/85 (2006.01)

(21) International Application Number:

PCT/CA2013/050204

(22) International Filing Date:

14 March 2013 (14.03.2013)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/610,797 14 March 2012 (14.03.2012) US
61/731,988 30 November 2012 (30.11.2012) US

(71) Applicant: INNOVATIVE TARGETING SOLUTIONS INC. [CA/CA]; 309 - 4475 Wayburne Drive, Burnaby, British Columbia V5G 4X4 (CA).

(72) Inventors: GALLO, Michael; 2650 Colwood Drive, North Vancouver, British Columbia V7R 2R1 (CA). KANG, Jaspal Singh; 16727 109 Avenue, Surrey, British Columbia V4N 5H6 (CA). PIGOTT, Craig Robin; 210-

1775 West 11th Avenue, Vancouver, British Columbia V6J 2C1 (CA).

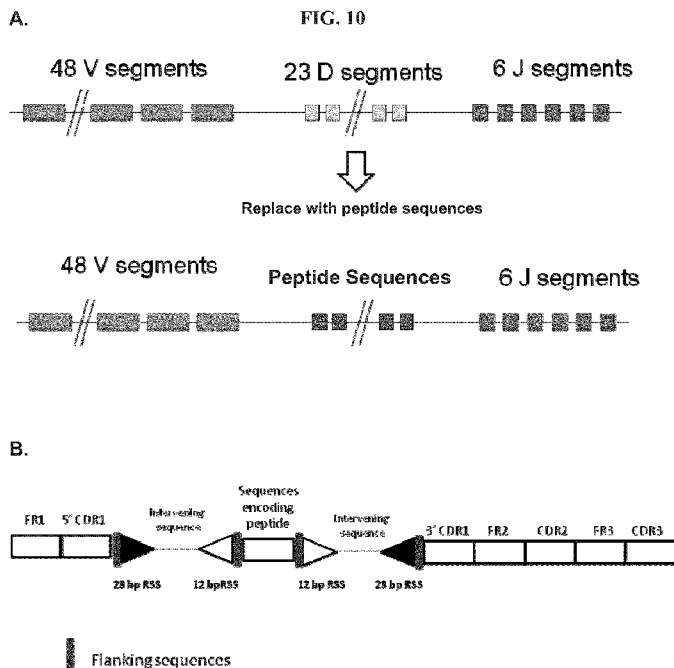
(74) Agents: SECHLEY, Konrad et al.; Gowling Lafleur Henderson LLP, 550 Burrard Street, Suite 2300, Vancouver, British Columbia V6C 2B5 (CA).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KR, KZ, LA, LC, LK, LR, LS, LT, 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, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, 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,

[Continued on next page]

(54) Title: GENERATING TARGETED SEQUENCE DIVERSITY IN FUSION PROTEINS



(57) Abstract: Methods of generating fusion protein variants are provided that comprise introducing sequence diversity at the junction region or regions in the fusion and allows for the generation of variants having a desired activity. Examples include immunoglobulins comprising a domain or polypeptide inserted into, or replacing, a CDR. Also provided are polynucleotides encoding a fusion protein and comprising two or more RSSs, and compositions and host cells comprising same, as well as fusion proteins variants produced by the described methods.





TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, — *with sequence listing part of description (Rule 5.2(a))*
ML, MR, NE, SN, TD, TG).

Published:

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

GENERATING TARGETED SEQUENCE DIVERSITY IN FUSION PROTEINS

FIELD OF THE INVENTION

[0001] The present invention relates to the field of protein engineering and, in particular, to methods of generating optimized fusion proteins.

BACKGROUND OF THE INVENTION

[0002] A number of successful protein therapeutics are recombinant fusion proteins consisting of two proteins or protein domains fused together through a linker, or a protein scaffold into which one or more domains from a second protein have been grafted. Typically, such fusion proteins are designed to leverage beneficial properties of each member of the fusion.

[0003] For example, cytokines or growth factors have been fused with the Fc portion of IgG1 or immunotoxin and expressed as single polypeptides with dual biological activities. Examples of therapeutic fusion proteins that have been developed using cytokines or growth factors and the Fc portion of IgG1 include Enbrel® (TNF-RIFs-IgG1), Ontak® (IL-2/diphtheria toxin), Orencia® (CTLA-4/Fc-IgG1) and Amevive® (LFA-3/Fc-IgG1).

[0004] Protein engineering has been used extensively to introduce novel binding specificities into protein scaffolds. Both rational and combinatorial approaches have been used with a variety of structurally diverse scaffolds (see Binz *et al.*, 2005, *Nature Biotechnology*, 23(10):1257-1268; Nygren & Skerra (2004, *J Immunol. Methods*, 290:3-28) and Gebauer & Skerra (2009, *Curr. Op. Chem. Biol.*, 13:245-255). Antibodies are perhaps the best studied of all protein scaffolds and affinity transfer by loop swapping has become routine. The technique of loop swapping was first described by Jones *et al.* (1986, *Nature* 321(6069):522-525), who substituted the CDRs from the heavy chain variable region of a mouse antibody, which binds to the hapten 4-hydroxy-3-nitrophenacetyl caproic acid (NP-cap), for the corresponding CDRs of a human myeloma antibody. It is now quite common to transfer the complementarity

determining region (CDR) loops from a non-human antibody to the scaffold of a human antibody to increase its therapeutic potential (Jones *et al.*, 1986, *ibid*; Riechmann *et al.*, 1988, *Nature* 332(6162):323-327; Verhoeyen *et al.*, 1988, *Science* 239(4847):1534-1536).

5 [0005] Affinity transfer by CDR replacement has also been successful with non-immunoglobulin scaffolds. Nicaise *et al.* (2004, *Protein Sci* 13(7):1882-1891) grafted the CDR3 of a lysozyme-specific camel antibody onto neocarzinostatin (NCS). Novel binding properties have also been generated by transferring CDR-like loops from proteins other than antibodies, for example, van den Beucken *et al.* (2001, *J Mol Biol*
10 310(3):591-601) made a V_L library with a constant CDR3-like sequence from the protein CLTA-4, and selected variants with specificity for its receptor B7.1 and demonstrated that the flanking conformational context is important in maintaining functional binding properties of the transferred domain. Several non-antibody scaffolds are also being evaluated for use as potential therapeutics including fibronectin (Hackel
15 *et al.*, 2008, *J Mol Biol* 381(5):1238-1252; Lipovsek *et al.*, 2007, *J Mol Biol* 368(4):1024-1041), lipocalins, avimers, adnectins and ankyrins. Zeytun *et al.* (2003, *Nat Biotechnol* 21(12):1473-1479), introduced diverse CDR-H3 sequences into four surface loops of an optimised GFP scaffold to create “fluorobodies.”

[0006] Various methods have been used to introduce diversity into these scaffolds including error prone PCR approaches, degenerate oligo or peptide synthesis or a
20 variety of DNA/CDR shuffling and CDR walking strategies (Bernath *et al.*, 2005, *J Mol Biol* 345(5):1015-1026; Nord *et al.*, 1997, *Nat Biotechnol* 15(8):772-777; Colas *et al.*, 1996, *Nature* 380(6574):548-550).

[0007] The first report of peptide being placed into the CDR of an antibody was by
25 Sallazzo (1990). Placing peptides into a CDR of antibody and maintaining peptide function is often compromised because the peptide is no longer unconstrained or is constrained in an inappropriate confirmation. Successful insertion of RGD peptides into the CDR3 of the antibody heavy chain has been reported (Zanetti *et al.*, 1993, *EMBO J*, 12(11):4375-4384). Simon *et al.* (2005, *Arch Biochem Biophys*, 440(2):148-
30 157) describes the insertion of the somatostatin peptide into the CDRs of the kappa light chain using PCR mediated gene splicing by overlap extension. The points of

insertion were identified through alignment of kappa light chain variable region amino acid sequences and X-ray crystal structures. The authors confirmed that somatostatin peptides inserted into the predicted regions of kappa CDR-1 and CDR-2 were able to bind to membranes containing somatostatin receptor 5.

5 [0008] A TPO agonist antibody has also been described that utilized insertion of two copies of an active peptide into CDR loops of an antibody fragment (Fab) (Frederickson *et al.*, 2006, *PNAS USA*, 103(39):14307-14312). The group reported that the amino acids flanking the peptide required optimization for proper presentation of the peptide in the context of the antibody scaffold. Using phage display, two amino
10 acids on either side of the peptide were randomized and inserted in to CDR3 of the heavy chain and subsequent panning identified binders. Several of the identified binders also showed agonist activity.

[0009] V(D)J recombination is the process responsible of the assembly of antibody gene segments (V, D and J; or V and J in the case of the light chain) and as part of the
15 assembly process creates the CDR3 of the respective antibody chain. V(D)J recombination can be considered conceptually as a segment shuffler for antibodies, *i.e.* it brings together the different VH segments, D segments and JH segments to create an antibody (similarly V(D)J recombination at the light chain assembles different combinations of light chain V and J segments at either the kappa or lambda locus). The
20 recombination event results in large chromosomal deletions in order to bring the required segments together. V(D)J recombination is targeted by the presence of specific DNA sequences called the recombination signal sequences (RSSs). The recombination reaction involves the recombination proteins RAG-1 and RAG-2 and follows a 12/23 rule where an RSS with a 23bp spacer is paired only with an RSS with
25 12 bp spacer and adjacent sequences are subsequently joined by double-stranded break repair proteins.

[0010] U.S. Patent No. 8,012,714 describes compositions and methods for generating sequence diversity in the CDR3 region of *de novo* generated immunoglobulins *in vitro*. The methods comprise constructing nucleic acid molecules that comprise
30 polynucleotide sequences encoding immunoglobulin V, D, J and C regions, together

with recombination signal sequences (RSS), and subsequently introducing these nucleic acid molecules into suitable recombination-competent host cells.

[0011] This background information is provided for the purpose of making known information believed by the applicant to be of possible relevance to the present invention. No admission is necessarily intended, nor should be construed, that any of the preceding information constitutes prior art against the present invention.

SUMMARY OF THE INVENTION

[0012] An object of the present invention is to provide methods and compositions for generating targeted sequence diversity in fusion proteins. One aspect of the invention relates to a method of generating variants of a fusion protein comprising the steps of: providing a first nucleic acid sequence comprising a first coding sequence encoding a first portion of the fusion protein and further comprising a first recombination signal sequence (RSS); providing a second nucleic acid sequence comprising a second coding sequence encoding a second portion of the fusion protein and further comprising a second RSS capable of functional recombination with the first RSS; introducing the first and second nucleic acid sequence into a recombination-competent host cell, and culturing the host cell *in vitro* under conditions allowing (a) recombination of the first and second RSS to generate a chimeric polynucleotide comprising the first and second coding sequences and (b) expression of the chimeric polynucleotide, thereby generating variants of the fusion protein.

[0013] Certain embodiments of the invention relate to a method as described above in which the second nucleic acid sequence further comprises a third RSS and the method further comprises the steps of: providing a third nucleic acid sequence comprising a third coding sequence encoding a third portion of the fusion protein and further comprising a fourth RSS capable of functional recombination with the third RSS, and introducing the third nucleic acid sequence into the recombination-competent host cell, and in which culturing the host cell further allows for recombination of the third and fourth RSS and the chimeric polynucleotide comprises the first, second and third coding sequences.

[0014] Another aspect of the invention relates to a polynucleotide comprising a first nucleic acid sequence comprising a first coding sequence encoding a first portion of a fusion protein and further comprising a first recombination signal sequence (RSS) and a second nucleic acid sequence comprising a second coding sequence encoding a second portion of the fusion protein and further comprising a second RSS capable of functional recombination with the first RSS.

[0015] Certain embodiments of the invention relate to a polynucleotide as described above in which the second nucleic acid sequence further comprises a third RSS and the polynucleotide further comprises a third nucleic acid sequence comprising a third coding sequence encoding a third portion of the fusion protein and further comprising a fourth RSS capable of functional recombination with the third RSS.

[0016] Another aspect of the invention relates to an isolated host cell comprising a polynucleotide as described herein.

[0017] Another aspect of the invention relates to a variant fusion protein produced by the methods described herein.

[0018] Another aspect of the invention relates to a peptide-grafted immunoglobulin comprising an immunoglobulin scaffold and a heterologous polypeptide inserted into at least one CDR of the immunoglobulin scaffold, wherein the heterologous polypeptide comprises a peptide sequence capable of binding to a GPCR, an upstream flanking sequence comprising between about 1 and about 20 amino acids and a downstream flanking sequence comprising between about 1 and about 20 amino acids.

BRIEF DESCRIPTION OF THE DRAWINGS

[0019] These and other features of the invention will become more apparent in the following detailed description in which reference is made to the appended drawings.

[0020] **Figure 1** presents (A) a schematic representation of a peptide grafting acceptor vector to generate antibody variants, and (B) the nucleotide sequence of the vector [SEQ ID NO:28].

[0021] **Figure 2** presents a modified fibronectin sequence [SEQ ID NO:35] that includes an RGD peptide encoding sequence in all three reading frames [SEQ ID NOs:9-11].

[0022] **Figure 3** presents (A) a schematic representation of a cassette for insertion into the fusion protein grafting acceptor vector shown in Figure 1(A) that includes amino acids 7-37 of GLP-1, (B) the nucleotide sequence [SEQ ID NO:36] with coding sequence in bold and (C) the amino acid sequence [SEQ ID NO:38] of the cassette shown in (A), (D) a schematic representation of a cassette that includes amino acids 13-33 of GLP-1, (E) the nucleotide sequence [SEQ ID NO:37] with coding sequence in bold and (F) the amino acid sequence [SEQ ID NO:39] of the cassette shown in (D).

[0023] **Figure 4** presents (A) a schematic representation of a cassette for insertion into the fusion protein grafting acceptor vector shown in Figure 1(A) that includes amino acids 1-39 of exendin-4, (B) the nucleotide sequence [SEQ ID NO:40] and (C) the amino acid sequence [SEQ ID NO:43] of the cassette shown in (A), (D) a schematic representation of a cassette that includes amino acids 9-39 of exendin-4, (E) the nucleotide sequence [SEQ ID NO:41] and (F) the amino acid sequence [SEQ ID NO:44] of the cassette shown in (D), (G) a schematic representation of a cassette that includes amino acids 15-27 of exendin-4, (H) the nucleotide sequence [SEQ ID NO:42] and (F) the amino acid sequence [SEQ ID NO:45] of the cassette shown in (G).

[0024] **Figure 5** presents (A) the IL-8 nucleotide and amino acid sequences [SEQ ID NOs:46 and 47, respectively], (B) the IL-8 nucleotide sequence for peptide grafting [SEQ ID NO:12], (C) the amino acid and nucleotide sequences of the binding domain of Gro-alpha [SEQ ID NOs:48 and 49, respectively], (D) the nucleotide sequence of Gro-alpha for peptide grafting [SEQ ID NO:49], and (E) the nucleotide sequence of LL-37 with a silent substitution of (G>A) that places a stop codon in reading frame #2 [SEQ ID NO:50].

[0025] **Figure 6** presents an alignment of the amino acid sequences of the receptor dimerization arm for various ErbB proteins [SEQ ID NOs:13-16].

[0026] **Figure 7** presents (A) a schematic representation of an acceptor vector for grafting peptides into the 10Fn3 loop, and (B) the nucleotide sequence of the vector [SEQ ID NO:63].

5 [0027] **Figure 8** presents a schematic representation of a generalised cassette for peptide grafting in accordance with one embodiment of the invention.

[0028] **Figure 9** presents (A) a schematic representation of a cassette for generating in-frame selection of a secreted protein (shown is Ig Kappa) showing from constant region to poly(A), and (B) the nucleotide sequence of the cassette [SEQ ID NO:64] with the furin cleavage site in bold.

10 [0029] **Figure 10** presents (A) a schematic overview of a method of grafting peptides in accordance with one embodiment of the invention in which immunoglobulin D segments are replaced with peptide sequences, and (B) a schematic of a recombination substrate for grafting peptides in accordance with another embodiment of the invention in which peptide sequences are grafted into other CDRs of immunoglobulin heavy or
15 light chains.

[0030] **Figure 11** presents (A) a cassette comprising a 5' RSS [SEQ ID NO:17], 5' and 3' flanking sequences and a 3' RSS [SEQ ID NO:18] for peptide grafting of an anti-TPO receptor peptide (encoded by nucleotide sequence [SEQ ID NO:65]), (B) a cassette comprising a 5' RSS [SEQ ID NO:17], 5' and 3' flanking sequences and a 3'
20 RSS [SEQ ID NO:18] for peptide grafting of an anti-GLP-1 receptor peptide, and (C) nucleotide sequences encoding exemplary anti-GLP-1 receptor peptides [SEQ ID NOs:66-70] for incorporation into the cassette shown in (B). N = any nucleotide, and K = T or G.

[0031] **Figure 12** presents the results of FACS scanning analysis of peptide-grafted
25 immunoglobulins in which the peptide has replaced the D segment in the heavy chain CDR3: (A) binding of immunoglobulins grafted with an anti-TPO receptor peptide to the TPO receptor, and (B) binding of immunoglobulins grafted with an anti-GLP-1 receptor peptide to the GLP-1 receptor.

DETAILED DESCRIPTION OF THE INVENTION

[0032] As illustrated herein, the use of components of the antibody V(D)J recombination system can be expanded outside their natural role of mediating assembly of antibody gene segments and may be used to generate sequence diversity in a variety of contexts and, specifically, at the junction region(s) of fusion proteins.

5 [0033] Fusion proteins in the context of the present invention include, in certain embodiments, fusions comprising two full length proteins and fusions comprising two or more protein domains or polypeptides, each associated with a desired function or activity. The protein domains or polypeptides may be derived from the same protein or they may be derived from different proteins, or one or more of the protein domains or
10 polypeptides may be a non-naturally occurring polypeptide. Thus, in certain embodiments, such polypeptides may, for example, be sequences representing all or a portion of a known protein having the desired function or activity, or they may be non-naturally occurring sequences that have been shown to have a desired function or activity. In the simplest case, fusion proteins can be the result of the amino acid
15 sequences corresponding to each protein, domain or polypeptide, being brought together to form the fusion protein. Fusion proteins may also comprise additional amino acids separating each protein, domain or polypeptide, that are not necessarily part of the protein, domain or polypeptide. These additional amino acids sequences joining the components of the fusion protein are referred to as a linker. Accordingly, in
20 certain embodiments of the invention, a fusion protein may be defined as a protein in which two or more amino acid sequences that are not normally contiguous have been joined together. In certain embodiments, the fusion protein can comprise a small functional region of one protein inserted into a region of a second protein, protein domain or protein scaffold - the process used to generate such fusion proteins is often
25 referred to as "peptide grafting."

[0034] One challenge with joining proteins, protein domains or polypeptides to produce a fusion protein is that the context and conformation of the fusion may not be appropriate for the functioning of the proteins, domains or polypeptides. The present invention recognizes that the natural V(D)J reaction has inherent characteristics,
30 specifically the imprecise junctions generated during the joining process, that make it

useful as a general means to generate sequence diversity and thus may be employed in the context of fusion proteins to generate a large repertoire of junctions between the component portions of the fusion protein.

5 [0035] Certain embodiments of the present invention thus relate to methods of generating fusion protein variants by introducing sequence diversity at the junction region or regions of the fusion protein and allows for the identification of a variant which preferably retains the optimal activity of the protein, or domain or polypeptide, of interest.

10 [0036] Certain embodiments of the invention relate to peptide-grafted immunoglobulins which comprise one or more peptides grafted into a CDR of an immunoglobulin scaffold. In some embodiments, the peptide is targeted to a membrane-bound receptor, such as a G-protein coupled receptor (GPCR) or ion channel.

15 [0037] In certain embodiments, the methods of the invention comprise generating fusion protein variants by introducing sequence diversity at a junction region between two proteins, domains or polypeptides. In some embodiments, therefore, the methods make use of a “bipartite” reaction that involves a single pair of RSSs, which may be used with or without flanking sequences, as described in more detail below.

20 [0038] In some embodiments, the methods comprise generating sequence diversity at two or more junctions between proteins, domains or polypeptides. For example, in certain embodiments, the methods are used for peptide grafting in which a protein domain or polypeptide having a desired function is integrated into a structural framework of a stably folded protein with suitable properties for the desired purpose. Examples of such frameworks include antibody scaffolds and other protein scaffolds
25 consisting of a stably folded non-Ig protein. In some embodiments, therefore, the methods make use of a “tripartite” reaction that involves a RSS flanked donor cassette sequence (*i.e.* two pairs of RSSs) and diversity is generated at each junction. In certain embodiments, diversity at both junctions can be accomplished by two sequential bipartite reactions.

[0039] Accordingly, certain embodiments of the invention relate to methods of generating sequence diversity at a junction between proteins, protein domains or polypeptides comprised by a fusion protein by providing polynucleotides comprising coding sequences for the proteins, domains or polypeptides, and further comprising recombination signal sequences (RSSs) and subsequent introduction of the polynucleotides into a recombination-competent host cell, specifically a host cell that is capable of expressing at least RAG-1 and RAG-2 or functional fragments thereof, resulting in the generation and expression of variant fusion proteins. In certain embodiments, the present invention also relates to polynucleotides for generating variant fusion proteins comprising coding sequences for the constituent proteins, or domains or polypeptides, and further comprising recombination signal sequences (RSSs), as well as compositions comprising same. In some embodiments, the invention relates to fusion proteins generated from recombination of such polynucleotides and compositions comprising same.

15 *Definitions*

[0040] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs.

[0041] The term “domain,” as used herein with respect to a protein refers to a portion of the protein that has, or is predicted to have, a desired function. Proteins may comprise more than one domain as distinct, non-contiguous regions of the protein. A domain can vary in size from a few amino acids to several hundred amino acids in length. As such, a domain may comprise substantially all of the protein from which it is derived, or it may be a fragment of the protein. In this context, a fragment is generally considered to be a polypeptide that has an amino-terminal and/or carboxy-terminal deletion compared to a full-length protein. Fragments typically are between about 3 and about 300 amino acids in length. In certain embodiments, a fragment is at least 3, 4, 5, 6, 7, 8, 9 or 10 amino acids long, and no more than 200 amino acids long, for example, between about 5 and 200, between about 5 and 190, between about 5 and 180, between about 5 and 170, between about 5 and 160, between about 5 and 150, between about 5 and 140, between about 5 and 130, between about 5 and 120, between about 5 and 110,

between about 5 and 100, between about 5 and 90, between about 5 and 80, between about 5 and 70, between about 5 and 60 and between about 5 and 50 amino acids long. The term “domain” also encompasses variants of the naturally-occurring domain provided that the variants retain at least partial functionality, for example, 10%, 20%,
5 30%, 40%, 50% or more of the activity of the naturally occurring domain. Variants may be constructed by, for example, substituting or deleting residues not needed for functionality or by inserting residues that will not adversely affect functionality.

[0042] The term “polypeptide,” as used herein refers broadly to an amino acid chain that may have various lengths, including a chain length shorter than 50 amino acids. A
10 polypeptide may therefore range from about 2 to about 3000 amino acids in length, for example, between about 2 and about 1500 amino acids, between about 2 and about 1000 amino acids, between about 2 and about 500 amino acids, between about 2 and about 300 amino acids in length. The term as used herein encompasses analogs and mimetics as known in the art that mimic structural and thus biological function.

15 [0043] “Naturally occurring,” as used herein with reference to an object, refers to the fact that the object can be found in nature. For example, an organism, or a polypeptide or polynucleotide sequence that is present in an organism that can be isolated from a source in nature and which has not been intentionally modified by man in the laboratory is naturally occurring.

20 [0044] The term “isolated,” as used herein with reference to a material, means that the material is removed from its original environment (for example, the natural environment if it is naturally occurring). For example, a naturally occurring polynucleotide or polypeptide present in a living animal is not isolated, but the same polynucleotide or polypeptide separated from some or all of the co-existing materials in
25 the natural system, is isolated. Such polynucleotides could be part of a vector and/or such polynucleotides or polypeptides could be part of a composition, and still be isolated in that such vector or composition is not part of its natural environment.

[0045] The term “deletion” as used herein with reference to a polynucleotide, polypeptide or protein has its common meaning as understood by those familiar with
30 the art and may refer to molecules that lack one or more of a portion of a sequence from either terminus or from a non-terminal region, relative to a corresponding full length

molecule. For example, in certain embodiments, a deletion may be a deletion of between 1 and about 1500 contiguous nucleotide or amino acid residues from the full length sequence.

[0046] The term “expression vector,” as used herein, refers to a vehicle used in a recombinant expression system for the purpose of expressing a polynucleotide sequence constitutively or inducibly in a host cell, including prokaryotic, yeast, fungal, plant, insect or mammalian host cells, either *in vitro* or *in vivo*. The term includes both linear and circular expression systems. The term includes expression systems that remain episomal and expression systems that integrate into the host cell genome. The expression systems can have the ability to self-replicate or they may not (for example, they may drive only transient expression in a cell).

[0047] The term “antigen-binding domain,” as used herein, refers to one or more fragments of an antibody that retain the ability to specifically bind to an antigen. Non-limiting examples of antibody fragments comprising antigen-binding domains include, but are not limited to, (i) a Fab fragment, a monovalent fragment consisting of the V_L , V_H , C_L and C_{H1} domains; (ii) a $F(ab')_2$ fragment, a bivalent fragment comprising two Fab fragments linked by a disulphide bridge at the hinge region; (iii) a Fd fragment consisting of the V_H and C_{H1} domains; (iv) a Fv fragment consisting of the V_L and V_H domains of a single arm of an antibody, (v) a dAb fragment, which consists of a V_H domain; and (vi) an isolated complementarity determining region (CDR). The term also encompasses single chain Fv (scFv) fragments, which comprise the two domains of the Fv fragment, V_L and V_H , joined using recombinant methods by a synthetic linker that enables them to be made as a single protein chain in which the V_L and V_H regions pair to form monovalent molecules.

[0048] The term “bipartite reaction,” as used herein, refers to a recombination reaction that involves a single pair of RSSs (12 bp and 23 bp, or 23 bp and 12 bp). When V(D)J recombination occurs it generates a double-stranded break in the nucleic acid sequence containing the RSSs. The double-stranded break is targeted as a result of the RSSs in that a 12 bp and 23 bp RSS are assembled with the RAG proteins to initiate the reaction. The ends of the DNA that will be subsequently rejoined will comprise the coding joint (or junction). An example of a bipartite reaction is *in vivo*

immunoglobulin light chain recombination, which joins the Variable to the Joining segment - these two segments comprise the “substrates” for the bipartite reaction. The bipartite reaction can occur in the presence or absence of TdT.

[0049] The term “tripartite reaction,” as used herein, refers to a recombination
5 reaction that involves two pairs of RSSs (each 12 bp and 23 bp, or 23 bp and 12 bp). An example of a tripartite reaction is *in vivo* immunoglobulin heavy chain recombination, which joins the V, the D and the J gene segments. A tripartite reaction generates two independent coding junctions. Two sequential bipartite reactions can be considered to be a tripartite reaction in that a tripartite reaction may comprise two
10 bipartite reactions occurring in the same substrate, usually (but not always) in close temporal time. The tripartite reaction can occur in the presence or absence of TdT.

[0050] The term “recombination-competent” when used herein with reference to a host cell means that the host cell is capable of mediating RAG-1/RAG-2 recombination. The host cell may, therefore, express RAG-1 and RAG-2, or functional
15 fragments thereof, or may be modified (for example, transformed or transfected with appropriate genetic constructs) such that it expresses RAG-1 and RAG-2, or functional fragments thereof. The expression of one or both of RAG-1 and RAG-2 in the recombination-competent host cell may be constitutive or it may be inducible. A recombination-competent host cell may optionally further express TdT, or a functional
20 fragment thereof.

[0051] As used herein, the term “about” refers to an approximately +/-10% variation from a given value. It is to be understood that such a variation is always included in any given value provided herein, whether or not it is specifically referred to.

[0052] The term “plurality” as used herein means more than one, for example, two or
25 more, three or more, four or more, and the like.

METHODS OF GENERATING VARIANT FUSION PROTEINS

[0053] The methods according to the present invention comprise utilizing the V(D)J recombination system to generate sequence diversity in fusion proteins.

[0054] The use of V(D)J recombination as a method to modify an existing protein sequence as opposed to assembly of a protein from gene segments can present a number of challenges, including a number of features of the reaction that are under-appreciated in the art.

5 [0055] For example, the V(D)J recombination reaction is known to bring together different DNA sequences and result in large chromosomal deletions, which suggests that its utility to introduce sequence diversity would be limited to extended stretches of nucleic acid sequence that permit such large deletions. As demonstrated herein, however, the components of the V(D)J recombination system can be manipulated to
10 allow the utility of this reaction to be extended to include targeted sequences within a restricted size of protein sequence, such as a small loop.

[0056] In addition, although the involvement of the enzyme TdT, which is responsible for non-template nucleotide additions (N-additions), is central to the reaction, the net size of the product following gene segment assembly is frequently less
15 than would be predicted if no deletions or additions were to occur, *i.e.* the V(D)J reaction often results in a net loss of sequence. For example, the average size of the assembled germline V, D and J segments, without any additions or deletions, is 15 amino acids and yet the average CDR3 reported in humans is 12-13 amino acids, which includes N additions from TdT (Rock *et al.*, 1994, *J Exp Med*, 179:323-328).

20 [0057] Another feature of V(D)J recombination that is under-appreciated is that the additions introduced by TdT are small. *In vivo* and *in vitro* TdT additions have been reported to be typically an average of 2-4 nucleotides (Kallenbach *et al.*, 1992, *PNAS USA*, 89:2799-2903; Bentolila *et al.*, 1997, *J Immunol.*, 158:715-723). A larger number of amino acid changes per variant is generally preferred for mutagenesis techniques in
25 order to allow for a greater amount of diversity to be sampled.

[0058] The above-noted features of V(D)J recombination can represent challenges to the application of V(D)J recombination to a non-antibody scaffold. The methods provided by the present invention, however, allow for this random process to be used as a valuable tool for semi-rational protein engineering.

[0059] In some embodiments, for example, the methods employ flanking sequences adjacent to one or more of the RSSs to allow for incorporation of additional sequences into the junction region(s) of the fusion protein to minimise any net deletion effect of the V(D)J recombination reaction and/or to introduce additional functionality by way of addition of specific amino acid residues. By way of example, when the targeted location is within a small loop of a protein, such as the CDR3 of an antibody, flanking sequences may be used in conjunction with the RSSs to ensure that the loop retains a minimal length once sequence diversification has taken place.

[0060] In certain embodiments, the methods of the present invention allow for the generation of both composition and length diversity simultaneously. In some embodiments, the methods are entirely cell-mediated thus eliminating the requirement for cloning of variants and their subsequent introduction into cells as is required by other methods.

[0061] The methods according to the present invention generally comprise the steps of providing polynucleotides comprising coding sequences for the components of a fusion protein (*i.e.* the constituent proteins, or domains or polypeptides) and further comprising recombination signal sequences (RSSs), introducing the polynucleotides into a recombination-competent host cell, and culturing the cell *in vitro* under conditions that allow for recombination and expression of the polynucleotides, thus generating a variant fusion protein. In certain embodiments, the methods further comprise screening the variant fusion protein for defined functional characteristics.

[0062] The host cell may constitutively express RAG-1 and RAG-2, and optionally TdT, or one or more of these proteins may be under inducible control. In certain embodiments, expression of one or more of RAG-1 and RAG-2, and optionally TdT, in the host cell is under inducible control allowing, for example, for expansion of the host cell prior to the induction of sequence diversity generation. Accordingly, in some embodiments, the method comprises the steps of: providing polynucleotides comprising coding sequences for the components of a fusion protein (*i.e.* the constituent proteins, domains or polypeptides) and further comprising recombination signal sequences (RSSs), introducing the polynucleotides into a recombination-competent host cell, wherein expression of one or more of RAG-1, RAG-2 and TdT is

under inducible control, culturing the host cell under conditions allowing expansion of the host cell, inducing expression of one or more of RAG-1, RAG-2 and TdT, culturing the expanded host cells under conditions allowing recombination and expression of the polynucleotides, thereby generating a fusion protein variant.

5 [0063] In certain embodiments, the methods are used to generate a library of fusion protein variants that can be subsequently screened for variants having defined functional characteristics, and comprise the steps of: providing polynucleotides comprising coding sequences for the components of a fusion protein (*i.e.* the constituent proteins, domains or polypeptides) and further comprising recombination
10 signal sequences (RSSs), introducing the polynucleotides into recombination-competent host cells, and culturing the cells under conditions that allow for recombination and expression of the polynucleotides, thus generating a library of variant fusion proteins. In certain embodiments, the methods further comprise screening the library of variant fusion proteins for variants having the defined
15 functional characteristics.

[0064] In certain embodiments relating to peptide grafting in which the protein scaffold is an immunoglobulin, the methods may be used to generate a library of peptide-grafted variants by replacing the D segment-encoding region of a heavy chain with a cassette comprising a peptide encoding sequence flanked by RSS sequences (see
20 Figure 10A). The library of peptide-grafted variants thus generated will represent a wide variety of possible V segment-peptide-J segment combinations that can be screened for variants having the required functional characteristics.

[0065] Light chains or other heavy chain CDRs can also be similarly grafted by using a RSS flanked peptide sequence encoding cassette as an artificial D segment in a light
25 chain or heavy chain CDR recombination substrate (see Figure 10B). The library of peptide-grafted variants thus generated will represent a variety of antibodies with the peptide encoded DNA sequences and appropriate flanking sequences inserted into the targeted CDR.

[0066] The polynucleotides may be introduced into the host cell by way of a suitable vector or vectors and may be, for example, stably integrated into the genome of the cell,
30 stably maintained exogenously to the genome or transiently expressed.

[0067] In some embodiments, the coding sequence for the protein, domain or polypeptide comprised by the polynucleotide is operably linked to a regulatable promoter, for example, an inducible promoter, such that expression of the encoded sequence can be controlled.

5 [0068] In some embodiments, the polynucleotides may also comprise additional coding sequences that encode a polypeptide that provides additional functionality to the fusion protein. For example, the polypeptide may localize the fusion protein to the cell membrane, nucleus or other organelle; provide for secretion of the fusion protein from the cell; introduce a detectable label, or the like.

10 [0069] In certain embodiments, the recombination is controlled. In some embodiments, the host cell is capable of cell divisions without recombination. As described herein, these and related embodiments permit expansion of the host cell population prior to the initiation of recombination events that give rise to sequence diversity in the fusion protein. Control of recombination in such host cells may be
15 achieved, for example, through the use of an operably linked recombination control element (such as an inducible recombination control element, which may be a tightly regulated inducible recombination control element), and/or through the use of one or more low efficiency RSSs in the nucleic acid composition(s) (as described in more detail below), and/or through the use of low host cell expression levels of one or more
20 of RAG-1 or RAG-2, and/or through design of the polynucleotide to integrate at a chromosomal integration site offering poor accessibility to host cell recombination mechanisms (for example, RAG-1 and/or RAG-2).

[0070] In some embodiments, the methods further comprise selecting a variant having the desired functional characteristics, and subjecting the variant to one or more
25 additional rounds of sequence diversity generation in order to obtain further variants having optimised functional characteristics.

Fusion Protein Components

[0071] The methods of the present invention may be used to generate sequence diversity in fusion proteins comprised of a wide variety of proteins, protein domains or
30 polypeptides. The methods may be used, for example, to generate variants of a known

fusion protein having improved activity, or they may be used to generate new fusion proteins with new activities or combinations of activities.

[0072] In general, the components for fusion proteins are selected on the basis that they have an activity or function that renders them useful for a given application, for example, therapeutic, diagnostic, nutraceutical, agricultural, or industrial application, or otherwise impart desirable characteristics to the fusion protein, such as improved stability, improved pharmacokinetics, decreased antigenicity, and the like.

[0073] For example, components may be selected that have activities/functions such as protein-ligand interaction, protein-protein interaction, enzymatic activity, light capture and emission, antigenic activity, and the like.

[0074] The components of the fusion protein may be derived from naturally occurring proteins or polypeptides, or they may be non-naturally occurring polypeptides known or demonstrated to have a desired activity or function.

[0075] Examples of naturally occurring proteins and polypeptides of interest which may be used in their entirety, or as a source of a domain having a desired function, include, but are not limited to, antibodies (mAbs such as IgG, IgM, IgA, and the like), hormones, protease inhibitors, antibiotics, antimicrobials, HIV entry inhibitors, collagen, human lactoferrin, cytokines, receptors, growth factors, toxins, protein and peptide antigens, enzymes involved in primary and secondary intracellular signaling and metabolic pathways (such as enterokinase, beta-glucuronidase (GUS), phytase, carbonic anhydrase, and the like), industrial enzymes (such as hydrolases, glycosidases, cellulases, oxido-reductases, and the like) and fluorescent proteins (such as green fluorescent protein (GFP), enhanced cyan fluorescent protein (ECFP), red fluorescent protein (DsRed) and the like).

[0076] In certain embodiments, the fusion protein is comprised of a polypeptide with a desired activity that has been “grafted” into a protein scaffold. Examples of polypeptides that may be used for such “peptide grafting” include polypeptides derived from various ligands, toxins, antigens, protein domains involved in protein-protein interactions, and the like. Certain embodiments of the invention contemplate peptide

grafting using peptides that target a membrane bound receptor, such as a GPCR, ion channel, a member of of the hematopoietic receptor superfamily or an integrin.

- [0077] Various protein scaffolds are known. For example, immunoglobulins such as antibodies or antibody fragments that comprise an antigen-binding domain are suitable for use as protein scaffolds. Examples include, but are not limited to, IgA, IgA2, IgD, IgE, IgGs (*i.e.* IgG1, IgG2, IgG3 and/or IgG4) and IgM antibodies; camelid antibodies; shark antibodies; antibody fragments such as Fab, Fab', F(ab')₂, Fd, Fv and single-chain Fv (scFv) antibody fragments; diabodies, nanobodies and fluorobodies. Certain embodiments of the invention relate to immunoglobulin scaffolds.
- [0078] Non-immunoglobulin protein scaffolds are also known and include various stably folded non-Ig proteins as described in Binz, *et al.* (2005, *Nature Biotechnology*, 23(10):1257-1268), Nygren & Skerra (2004, *J Immunol. Methods*, 290:3-28) and Gebauer & Skerra (2009, *Curr. Op. Chem. Biol.*, 13:245-255). Examples of such protein scaffolds include, but are not limited to, cytotoxic lymphocyte-associated antigen-4 (CTLA-4), Tendamistat, 10th fibronectin type 3 domain (¹⁰FN3), carbohydrate-binding module 4 of family 2 of xylanase of *Rhodothermus marinus* (CBM4-2), lipocalins (“anticalins”), T-cell receptor, Protein A domain (protein Z), immunity protein 9 (Im9), designed ankyrin repeat proteins (DARPs), designed tetratricopeptide repeat (TPR) proteins, zinc finger proteins, protein VIII of filamentous bacteriophage (pVIII), avian pancreatic polypeptide, general control nonderepressible (yeast transcription factor) (GCN4), WW domain, Src homology domain 3 (SH3), Src homology domain 2 (SH2), PDZ domains, TEM-1, β-lactamase, green fluorescent protein (GFP), thioredoxin, staphylococcal nuclease, plant homeodomain finger protein (PHD-finger), chymotrypsin inhibitor 2 (CI-2), bovine pancreatic trypsin inhibitor (BPTI), Alzheimer amyloid β-protein precursor inhibitor (APPI), human pancreatic secretory trypsin inhibitor (hPSTI), ecotin, human lipoprotein-associated coagulation inhibitor domain 1 (LACI-D1), leech-derived trypsin inhibitor (LDTI), MTI-II, scorpion toxins, insect defensin A peptide, *Ecballium elaterium* trypsin inhibitor II (EETI-II), Min-23, cellulose-binding domain (CBD), periplasmic binding proteins (PBP), cytochrome b₅₆₂, low density lipoprotein (ldl) receptor domain A, γ-crystallin, ubiquitin, transferrin and C-type lectin-like domain. T-cell receptors are also useful protein scaffolds in certain embodiments.

[0079] Protein scaffolds can be considered as falling into two groups: a first group consisting of loop presenting scaffolds (which includes scaffolds presenting a single loop and scaffolds presenting a plurality of loops), and a second group consisting of interface presenting scaffolds, in which the binding site is presented on a secondary structure element. Examples of scaffolds in the first group include, but are not limited to, Kunitz domain inhibitors, hPSTI, APPI, LACI-D1, ecotin, members of the knottin family of proteins (such as EETI-II), thioredoxin, staphylococcal nuclease, immunoglobulins, CTLA-4, FN3, Tendamistat, GFP, members of the lipocalin family of proteins, and bilin binding protein (BBP) from *Pieris brassicae*. Examples in the second group include, but are not limited to, the immunoglobulin binding domain of Staphylococcal protein A (SPA) (“affibodies”), DARPin, leucine-rich repeat polypeptides, PDZ domains, cellulose binding domains (CBD), members of the lipocalin family of proteins, γ -crystallins, and Cys₂His₂ zinc-finger polypeptides. The binding domains of both of these groups of proteins have been studied and regions suitable for modification have been identified (see review by Nygren & Skerra, *ibid.*).

Polynucleotides

[0080] The methods of the present invention employ polynucleotides that comprise a coding sequence, *i.e.* a nucleic acid sequence encoding the protein, domain or polypeptide of interest. The polynucleotides may be in the form of RNA or in the form of DNA, which DNA includes cDNA, genomic DNA, and synthetic DNA. The DNA may be double-stranded or single-stranded, and if single stranded may be the coding strand or non-coding (anti-sense) strand. A nucleic acid sequence which encodes a protein, domain or polypeptide for use in the methods of the present invention may be identical to the coding sequence known in the art for the protein, domain or polypeptide or may be a different coding sequence, which, as a result of the redundancy or degeneracy of the genetic code, encodes the same protein.

[0081] The polynucleotides may include only the coding sequence for the protein, domain or polypeptide; the coding sequence and additional coding sequence (for example, encoding a polypeptide providing additional functionality to the final fusion protein); the coding sequence (and optionally additional coding sequence) and non-coding sequence, such as introns or non-coding sequences 5' and/or 3' of the coding

sequence. The coding sequence may be in the form of one or more exons, which may be contiguous or may be interspersed with one or more introns. The non-coding sequences may include, for example, one or more regulatory nucleic acid sequences that may be a regulated or regulatable promoter, enhancer, other transcription
5 regulatory sequence, repressor binding sequence, translation regulatory sequence or other regulatory nucleic acid sequence.

[0082] The coding sequence for various proteins, domains or polypeptides that may be used in the methods of the invention may be a known sequence that can be obtained, for example, from public databases such as GenBank. Many proteins have been cloned
10 and polynucleotides comprising the coding sequences for these proteins may be obtained from commercial sources. Alternatively, coding sequences can be obtained from an appropriate source, or otherwise generated or synthesized, using standard molecular biology techniques, such as those described in *Molecular Cloning: A Laboratory Manual* (Third Edition) (Sambrook, *et al.*, 2001, Cold Spring Harbour
15 Laboratory Press, NY) and *Current Protocols in Molecular Biology* (Ausubel *et al.* (Ed.), 1987 & Updates, J. Wiley & Sons, Inc., Hoboken, NJ). In addition, many companies offer custom gene synthesis and may be used as a source of coding sequences.

[0083] In certain embodiments, the polynucleotide may be codon-optimized
20 according to standard codon usage preference tables, such that its expression in the chosen host cell is optimized.

[0084] Certain embodiments of the invention encompass the use of variant polynucleotides in the present methods, for example, polynucleotides that encode analogs and/or derivatives of a protein (or a protein domain or polypeptide). The
25 polynucleotide variants may be, for example, naturally-occurring allelic variants of the polynucleotide or non-naturally occurring variants. As is known in the art, an allelic variant is an alternate form of a nucleic acid sequence which may have at least one of a substitution, a deletion or an addition of one or more nucleotides, any of which does not substantially alter the function of the encoded protein or polypeptide. Non-naturally
30 occurring polynucleotide variants may be accomplished by a number of conventional methods. For example, mutations can be introduced at particular loci by synthesizing

oligonucleotides containing a mutant sequence, flanked by restriction sites enabling ligation to fragments of the native sequence. Following ligation, the resulting reconstructed sequence encodes an analog having the desired amino acid insertion(s), substitution(s), or deletion(s). Alternatively, oligonucleotide-directed site-specific mutagenesis procedures can be employed to provide an altered gene wherein predetermined codons can be altered by substitution, deletion or insertion. Exemplary methods of making such alterations are described, for example, in *Molecular Cloning: A Laboratory Manual* (Third Edition) (Sambrook, *et al.*, 2001, Cold Spring Harbour Laboratory Press, NY) and *Current Protocols in Molecular Biology* (Ausubel *et al.* (Ed.), 1987 & Updates, J. Wiley & Sons, Inc., Hoboken, NJ).

[0085] In certain embodiments of the invention, for example, those relating to peptide grafting and/or those that involve tripartite reactions, polynucleotides may be provided in the form of a cassette comprising a sequence encoding a peptide flanked by pairs of RSS sequences. Such cassettes may be inserted into a larger polynucleotide encoding the protein scaffold, which is then transfected into an appropriate host cell to allow recombination to occur. In some embodiments, such cassettes may comprise a peptide-encoding sequence flanked by pairs of RSS sequences with degenerate nucleotide sequences inserted between the RSS sequences and the peptide encoding sequences. The degenerate sequences allow for introduction of additional sequence diversity over that provided by the V(D)J recombination and are typically about 3 and about 50 nucleotides in length, for example, between about 3 and about 40 nucleotides, between about 3 and about 30 nucleotides, between about 3 and about 20 nucleotides, between about 3 and about 18 nucleotides or between about 3 and about 15 nucleotides in length. Non-limiting examples of cassettes comprising degenerate sequences are shown in Figure 11.

Recombination Signal Sequences (RSSs)

[0086] The polynucleotides employed in the methods of the invention comprise recombination signal sequences (RSSs). The RSS in accordance with the present invention preferably consist of two conserved sequences (for example, heptamer, 5'-CACAGTG-3', and nonamer, 5'-ACAAAAACC-3'), separated by a spacer of either 12 +/- 1 bp (a "12-signal" RSS) or 23 +/- 1 bp (a "23-signal" RSS). Within the host cell,

two RSSs (one 12-signal RSS and one 23-signal RSS) are selected and rearranged under the “12/23 rule.” Recombination does not occur between two RSS signals with the same size spacer. As would be appreciated by one of skill in the art, the orientation of the RSS determines if recombination results in a deletion or inversion of the
5 intervening sequence.

[0087] As a result of extensive investigations of RSS processes, it is known in the art which nucleotide positions within RSSs cannot be varied without compromising RSS functional activity in genetic recombination mechanisms, which nucleotide positions within RSSs can be varied to alter (for example, increase or decrease in a statistically
10 significant manner) the efficiency of RSS functional activity in genetic recombination mechanisms, and which positions within RSSs can be varied without having any significant effect on RSS functional activity in genetic recombination mechanisms (see, for example, Ramsden *et al.*, 1994, *Proc Natl Acad Sci USA* 88(23): 10721-10725; Akamatsu *et al.*, 1994, *J Immunol* 153:4520; Hesse *et al.*, 1989, *Genes Dev* 3:1053;
15 Fanning *et al.*, 1996, *Immunogenetics* 44(2):146-150; Larijani *et al.*, 1999, *Nucleic Acids Res* 27(11):2304-2309; Nadel *et al.*, 1998, *J Exp Med* 187:1495; Lee *et al.*, 2003, *PLoS Biol* 1:E1; and Cowell *et al.*, 2004, *Immunol. Rev.* 200:57).

[0088] In certain embodiments, the invention makes use of an RSS that is known in the art. Also contemplated in some embodiments are sequence variants of known RSSs
20 that comprise one or more nucleotide substitutions (for example, about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18 or more substitutions) relative to a known RSS sequence and which, by virtue of such substitutions, predictably have low efficiency (for example, about 1% or less, relative to a high efficiency RSS), medium efficiency (for example, about 10% to about 20%, relative to a high efficiency RSS) or high
25 efficiency. Also contemplated in some embodiments are those RSS variants for which one or more nucleotide substitutions relative to a known RSS sequence will have no significant effect on the recombination efficiency of the RSS (for example, the success rate of the RSS in promoting formation of a recombination product, as known in the art).

30 [0089] Examples of RSS sequences known to the art, including their characterization as high, medium or low efficiency RSSs, are presented in Table 1A & B.

Table 1A. EXEMPLARY RECOMBINATION SIGNAL SEQUENCES (12 NUCLEOTIDE SPACER)

	Heptamer	Spacer	Nonamer
	H12	S12	N12
Part I. Efficiency: HIGH			
1	CACAGTG	ATACAGACCTTA [SEQ ID NO:1]	ACAAAAACC
2	CACAGTG	CTACAGACTGGA [SEQ ID NO:2]	ACAAAAACC
3	CACAGTG	CTCCAGGGCTGA [SEQ ID NO:3]	ACAAAAACC
4	CACAGTG	CTACAGACTGGA [SEQ ID NO:2]	ACAAAAACC
5	CACAGTG	CTACAGACTGGA [SEQ ID NO:2]	ACAAAAACC
6	CACAGTG	CTACAGACTGGA [SEQ ID NO:2]	ACAAAAACC
7	CACAGTG	GTACAGACCAAT [SEQ ID NO:4]	ACAGAAACC
Part II. Efficiency: MEDIUM (~10-20% of High)			
8	CACGGTG	CTACAGACTGGA [SEQ ID NO:2]	ACAAAAACC
9	CACAATG	CTACAGACTGGA [SEQ ID NO:2]	ACAAAAACC
10	CACAGCG	CTACAGACTGGA [SEQ ID NO:2]	ACAAAAACC
11	CACAGTG	CTACAGACTGGA [SEQ ID NO:2]	ACAAAAACC
12	CACAGTG	CTACAGACTGGA [SEQ ID NO:2]	ACAAAAACC
13	CACAGTG	CTACAGACTGGA [SEQ ID NO:2]	ACAAAAACC
14	CACAGTG	CTACAGACTGGA [SEQ ID NO:2]	ACAAAAACC
15	CACAGTG	CTACAGACTGGA [SEQ ID NO:2]	ACAAAAACC
16	CACAGTG	CTACAGACTGGA [SEQ ID NO:2]	ACAAAAACC
17	CACAGTG	CTACAGACTGGA [SEQ ID NO:2]	CAAAAACCC
18	CACAGTG	CTACAGACTGGA [SEQ ID NO:2]	ACAAAAACC
19	CACAATG	CTACAGACTGGA [SEQ ID NO:2]	ACAAAAACC
20	CACAGCG	CTACAGACTGGA [SEQ ID NO:2]	ACAAAAACC
Part III. Efficiency: LOW (~1% or less of High)			
21	TACAGTG	CTACAGACTGGA [SEQ ID NO:2]	ACAAAAACC

	Heptamer	Spacer	Nonamer
	H12	S12	N12
22	GACAGTG	CTACAGACTGGA [SEQ ID NO:2]	ACAAAAACC
23	CATAGTG	CTACAGACTGGA [SEQ ID NO:2]	ACAAAAACC
24	CACAATG	CTACAGACTGGA [SEQ ID NO:2]	ACAAAAACC
25	CACAGTG	CTACAGACTGGA [SEQ ID NO:2]	ACAAAAACC
26	CAGAGTG	CTCCAGGGCTGA [SEQ ID NO:3]	ACAAAAACC
27	CACAGTG	CTCCAGGGCTGA [SEQ ID NO:3]	AAAAAAACC
28	CTCAGTG	CTCCAGGGCTGA [SEQ ID NO:3]	ACAAAAACC

Table 1B. EXEMPLARY RECOMBINATION SIGNAL SEQUENCES (23 NUCLEOTIDE SPACER)

	Heptamer	Spacer	Nonamer	Ref.*
	H23	S23	N23	
Part I. Efficiency: HIGH				
1	CACAGTG	GTAGTACTCCACTGTCTGGCTGT [SEQ ID NO:5]	ACAAAAACC	4
2	CACAGTG	GTAGTACTCCACTGTCTGGCTGT [SEQ ID NO:5]	ACAAAAACC	3
3	CACAGTG	GTAGTACTCCACTGTCTGGGTGT [SEQ ID NO:5]	ACAAAAACC	1
4	CACAGTG	TTGCAACCACATCCTGAGTGTGT [SEQ ID NO:6]	ACAAAAACC	2
5	CACAGTG	GTAGTACTCCACTGTCTGGCTGT [SEQ ID NO:5]	ACAAAAACC	2
6	CACAGTG	ACGGAGATAAAGGAGGAAGCAGG [SEQ ID NO:7]	ACAAAAACC	2
7	CACAGTG	GCCGGGCCCGCGGCCCGCGGC [SEQ ID NO:8]	ACAAAAACC	5
Part II. Efficiency: MEDIUM (~10-20% of High)				

	Heptamer	Spacer	Nonamer	Ref.*
	H23	S23	N23	
8	CACAGTG	GTAGTACTCCACTGTCTGGCTGT [SEQ ID NO:5]	ACAAAAACC	3
9	CACAGTG	GTAGTACTCCACTGTCTGGCTGT [SEQ ID NO:5]	ACAAAAACC	3
10	CACAGTG	GTAGTACTCCACTGTCTGGCTGT [SEQ ID NO:5]	ACAAAAACC	3
11	CACAATG	GTAGTACTCCACTGTCTGGCTGT [SEQ ID NO:5]	ACAAAAACC	3
12	CACAGCG	GTAGTACTCCACTGTCTGGCTGT [SEQ ID NO:5]	ACAAAAACC	3
13	CACAGTA	GTAGTACTCCACTGTCTGGCTGT [SEQ ID NO:5]	ACAAAAACC	3
14	CACAGTG	GTAGTACTCCACTGTCTGGCTGT [SEQ ID NO:5]	ACAATAACC	3
15	CACAGTG	GTAGTACTCCACTGTCTGGCTGT [SEQ ID NO:5]	ACAAGAACC	3
16	CACAGTG	GTAGTACTCCACTGTCTGGCTGT [SEQ ID NO:5]	ACACGAACC	3
17	CACAGTG	GTAGTACTCCACTGTCTGGCTGT [SEQ ID NO:5]	ACAAAAACC	3
18	CACAGTG	GTAGTACTCCACTGTCTGGCTGT [SEQ ID NO:5]	ACACGAACC	3
19	CACAATG	GTAGTACTCCACTGTCTGGCTGT [SEQ ID NO:5]	ACAAAAACC	3
20	CACAGCG	GTAGTACTCCACTGTCTGGCTGT [SEQ ID NO:5]	ACAAAAACC	3
Part III. Efficiency: LOW (~1% or less of High)				
21	CACAGTA	GTAGTACTCCACTGTCTGGCTGT [SEQ ID NO:5]	ACAAAAACC	3
22	CACAGTG	GTAGTACTCCACTGTCTGGCTGT [SEQ ID NO:5]	ACAAAAACC	3

	Heptamer	Spacer	Nonamer	Ref.*
	H23	S23	N23	
23	CACAATG	GTAGTACTCCACTGTCTGGCTGT [SEQ ID NO:5]	ACAAAAACC	3
24	CATAGTG	GTAGTACTCCACTGTCTGGCTGT [SEQ ID NO:5]	ACAAAAACC	3
25	CACAGTG	GTAGTACTCCACTGTCTGGCTGT [SEQ ID NO:5]	TGTCTCTGA	3
26	CACAGTG	GTAGTACTCCACTGTCTGGGTGT [SEQ ID NO:5]	ACAAAAACC	1
27	CACAGTG	GTAGTACTCCACTGTCTGGGTGT [SEQ ID NO:5]	ACAAAAACC	1
28	CACAGTG	GTAGTACTCCACTGTCTGGGTGT [SEQ ID NO:5]	ACAAAAACC	1

*(1) Akamatsu, 1994, *ibid*; (2) Cowell, 2004, *ibid*; (3) Hesse,1989 *ibid*; (4) Lee, 2003, *ibid*; (5) Nadel,1998, *ibid*.

[0090] In accordance with certain embodiments of the invention, RSSs are used in pairs, in which the first RSS of the pair is capable of functional recombination with the second RSS of the pair (*i.e.* “complementary pairs”). It is to be understood that when a first RSS (for example present in a first polynucleotide or nucleic acid sequence) is described as being capable of functional recombination with a second RSS (for example present in a second polynucleotide or nucleic acid sequence), such capability includes compliance with the above-noted 12/23 rule for RSS spacers, such that if the first RSS comprises a 12-nucleotide spacer then the second RSS will comprise a 23-nucleotide spacer, and similarly if the first RSS comprises a 23-nucleotide spacer then the second RSS will comprise a 12-nucleotide spacer.

[0091] Complementary pairs of RSSs are generally separated by an intervening nucleotide sequence of about 100bp or more in length. The actual nucleotide sequence of this intervening sequence is not critical to the invention and can accommodate a wide variety of sequences, including for example some selectable markers, some promoters and other regulatory elements such as polyadenylation signals, but preferably does not include insulator-like elements as exemplified by *chs4* and *AAV1*.

[0092] In certain embodiments, the intervening sequence comprises an expression cassette, for example containing a promoter and optionally poly(A) sequences that drive expression of a marker such as GFP or a cell surface marker such that recombination can be monitored, or a selectable marker such as a drug resistance gene
5 such that the cell can be maintained in the un-recombined state via drug selection.

[0093] Regardless of the composition of the intervening sequence, it is preferably selected to be at least 100bp in length, for example, at least 110bp, at least 120bp, at least 130bp, at least 140bp, at least 150bp, but may range up to several kilobases in size, for example up to about 5kb. One skilled in the art will understand that the exact
10 upper limit for the intervening sequence will be dictated by the limitation of the vector system used. In certain embodiments, the intervening sequence is selected to be between about 100bp and 5kb, for example, between about 150bp and 5kb, between about 180bp and 5kb, between about 180bp and 4kb, between about 180bp and 3kb or between about 180bp and 2kb. In some embodiments, the intervening sequence is
15 selected to be between about 100bp and 1.5kb, for example, between about 110bp and 1.5kb, between about 120bp and 1.5kb, between about 130bp and 1.5kb, between 140bp and 1.5kb, or between 150bp and 1.5kb. In some embodiments, the intervening sequence is selected to be between about 180bp and 1.9kb, for example, between about 180bp and 1.8kb, between about 180bp and 1.7kb, between about 180bp and 1.6kb, or
20 between 180bp and 1.5kb. Other exemplary embodiments include intervening sequences of between about 190bp and 1.5kb, between about 200bp and 1.5kb, between about 210bp and 1.5kb, between about 220bp and 1.5kb, between about 230bp and 1.5kb, between about 240bp and 1.5kb, and between about 250bp and 1.5kb.

[0094] In certain embodiments, flanking sequences are included adjacent to the
25 heptamer of the RSS. In accordance with this embodiment, the flanking sequences may be chosen to have a defined sequence (for example, to specifically encode one or more amino acids) or they may have a random sequence. In some embodiments, the flanking sequences may be selected to introduce certain characteristics at the site of insertion, for example, through the addition of one or more charged amino acids, histidine
30 residues or cysteine residues. In certain embodiments, the flanking sequence may comprise a duplication of a part of the sequence into which the RSSs are to be introduced. In some embodiments, the position and length of the flanking sequences are

selected to bias diversification towards one side of the insertion point, or to provide a larger loop size prior to diversification.

[0095] When used, the length of the flanking sequence is selected such that it does not interfere with the structural integrity of the target protein. In certain embodiments, the flanking sequences are between about 3 and about 300 bp, for example between about 3 and about 250 bp, between about 3 and about 200 bp, between about 3 and about 150 bp, between about 3 and about 100 bp, between about 3 and about 50 bp, or any amount therebetween.

[0096] The RSSs can be introduced into the polynucleotide by standard genetic engineering techniques such as those described in *Molecular Cloning: A Laboratory Manual* (Third Edition) (Sambrook, *et al.*, 2001, Cold Spring Harbour Laboratory Press, NY) and *Current Protocols in Molecular Biology* (Ausubel *et al.* (Ed.), 1987 & Updates, J. Wiley & Sons, Inc., Hoboken, NJ).

Additional Coding Sequences

[0097] In accordance with certain embodiments of the invention, the polynucleotide may comprise additional coding sequences encoding a polypeptide that provides additional functionality to the fusion protein. Examples of polypeptides that provide additional functionality include, but are not limited to, secretory signal sequences, leader sequences, plasma membrane anchor domain polypeptides such as hydrophobic transmembrane domains (see, for example, Heuck *et al.*, 2002, *Cell Biochem. Biophys.* 36:89; Sadlish *et al.*, 2002, *Biochem J.* 364:777; Phoenix *et al.*, 2002, *Mol. Membr. Biol.* 19:1; Minke *et al.*, 2002, *Physiol. Rev.* 82:429) or glycosylphosphatidylinositol attachment sites (“glypiation” sites) (see, for example, Chatterjee *et al.*, 2001, *Cell Mol. Life Sci.* 58:1969; Hooper, 2001, *Proteomics* 1:748, and Spiro, 2002, *Glycobiol.* 12:43R), and other structural features that assist in localizing the fusion protein to the cell surface such as protein-protein association domains, lipid association domains, glycolipid association domains and proteoglycan association domains, for example, cell surface receptor binding domains, extracellular matrix binding domains, and lipid raft-associating domains (see, for example, Browman *et al.*, 2007, *Trends Cell Biol* 17:394-402; Harder, T., 2004, *Curr Opin Immunol* 16:353-9; Hayashi, T. and Su, T. P., 2005,

Life Sci 77:1612-24; Holowka, D. and Baird, B., 2001, *Semin Immunol* 13:99-105, and Wollscheid *et al.*, 2004, *Subcell Biochem* 37:121- 52).

[0098] Other examples of additional coding sequences that may be employed in some embodiments include intracellular targeting sequences, such as nuclear localization
5 sequences and other sequences that target the protein to an intracellular location.

[0099] In some embodiments, the additional coding sequences may encode a “tag” to facilitate downstream screening and/or purification of the fusion protein. Examples of such sequences include, but are not limited to, affinity tags such as metal-affinity tags, histidine tags, protein A, glutathione S transferase, Glu-Glu affinity tag, substance P,
10 FLAG peptide (Hopp *et al.*, 1988, *Biotechnology* 6:1204), streptavidin binding peptide, or other antigenic epitopes or binding domains (see, in general, Ford *et al.*, 1991, *Protein Expression and Purification* 2:95).

[00100] In some embodiments, the polynucleotide comprises additional coding sequences that encode a plasma membrane anchor domain. For example, a
15 transmembrane polypeptide domain typically comprising a membrane spanning domain (such as an $[\alpha]$ -helical domain) which includes a hydrophobic region capable of energetically favorable interaction with the phospholipid fatty acyl tails that form the interior of the plasma membrane bilayer, or a membrane-inserting domain polypeptide typically comprising a membrane-inserting domain which includes a hydrophobic
20 region capable of energetically favorable interaction with the phospholipid fatty acyl tails that form the interior of the plasma membrane bilayer but that may not span the entire membrane. Well known examples of transmembrane proteins having one or more transmembrane polypeptide domains include members of the integrin family, CD44, glycophorin, MHC Class I and II glycoproteins, EGF receptor, G protein coupled
25 receptor (GPCR) family, receptor tyrosine kinases (such as insulin-like growth factor 1 receptor (IGFR) and platelet-derived growth factor receptor (PDGFR)), porin family and other transmembrane proteins. Certain embodiments of the invention contemplate using a portion of a transmembrane polypeptide domain such as a truncated polypeptide having membrane-inserting characteristics as may be determined according to standard
30 and well known methodologies.

[00101] In some embodiments of the invention, the polynucleotide comprises additional coding sequences that encode a specific protein-protein association domain, for example a protein-protein association domain that is capable of specifically associating with an extracellularly disposed region of a cell surface protein or glycoprotein. In certain embodiments, the protein-protein association domain may result in an association that is initiated intracellularly, for instance, concomitant with the synthesis, processing, folding, assembly, transport and/or export to the cell surface of a cell surface protein. In some embodiments, the protein-protein association domain is known to associate with another cell surface protein that is membrane anchored and exteriorly disposed on a cell surface. Non-limiting examples of such domains include, RGD-containing polypeptides including those that are capable of integrin binding (see, for example, Heckmann, D. and Kessler, H., 2007, *Methods Enzymol* 426:463-503 and Takada *et al.*, 2007, *Genome Biol* 8:215).

[00102] In some embodiments, the polynucleotide comprises a secretory signal sequence that encodes a secretory peptide. A secretory peptide is an amino acid sequence that acts to direct the secretion of a mature polypeptide or protein from a cell and is generally characterized by a core of hydrophobic amino acids. Secretory peptides are typically, but not exclusively, positioned at the amino termini of newly synthesized proteins. The secretory peptide may be cleaved from the mature protein during secretion and may, therefore, contain processing sites that allow cleavage of the signal peptide from the mature protein as it passes through the secretory pathway. Examples of secretory peptides are known in the art and include, but are not limited to, alpha mating factor leader sequence, the secretory pre-peptide of IL-15, the tissue Plasminogen Activator (tPA) secretory leader peptide, transferrin (Tf) signal sequence, IgE secretory peptides, IgHV and IgKV signal peptides and GM-CSF secretory peptides.

[00103] In certain embodiments, sequences encoding transmembrane domain are included in the polynucleotide to provide surface expression of the fusion protein. In some embodiments, the fusion protein is cloned in-frame with a selectable marker to allow for the selection of productive in-frame products. In some embodiments, the polynucleotide comprises sequences encoding transmembrane domain, a selectable

marker and an enzyme cleavage site prior to the selectable marker to allow for cleavage of the fusion protein from the transmembrane domain.

[00104] Additional sequences, when used, can be included in the polynucleotide by standard genetic engineering techniques such as those described in *Molecular Cloning: A Laboratory Manual* (Third Edition) (Sambrook, *et al.*, *ibid.*) and *Current Protocols in Molecular Biology* (Ausubel *et al.* (Ed.), *ibid.*).

Vectors

[00105] Certain embodiments of the invention require the use of vectors as cloning and/or expression vehicles. A wide variety of suitable vectors are known in the art and may be employed as described or according to conventional procedures, including modifications, as described for example in Sambrook *et al.*, *ibid.*; Ausubel *et al.*, *ibid.*, and elsewhere.

[00106] One skilled in the art will appreciate that the precise vector used is not critical to the instant invention and suitable vectors can be readily selected by the skilled person. Examples of expression vectors and cloning vehicles include, but are not limited to, viral particles, baculovirus, phage, plasmids, phagemids, cosmids, fosmids, bacterial artificial chromosomes, retrovirus vectors, viral DNA (for example, vaccinia, adenovirus, fowl pox virus, pseudorabies and derivatives of SV40), P1-based artificial chromosomes, yeast plasmids, yeast artificial chromosomes, and other known vectors specific for specific host cells of interest.

[00107] Large numbers of suitable vectors are known to those of skill in the art, and are many commercially available. Exemplary commercially available vectors include the bacterial vectors: pcDNA (Invitrogen), pQE vectors (Qiagen), pBLUESCRIPT™ plasmids, pNH vectors, lambda-ZAP vectors (Stratagene); ptrc99a, pKK223-3, pDR540, and pRIT2T (Pharmacia); and the eukaryotic vectors: pXT1, pSGS (Stratagene), pSVK3, pBPV, pMSG and pSVLSV40 (Pharmacia). Other vectors include, for example, adenovirus (Ad) vectors (such as, non-replicating Ad5 vectors or replication-competent Ad4 and Ad7 vectors), adeno-associated virus (AAV) vectors (such as, AAV type 5), alphavirus vectors (such as, Venezuelan equine encephalitis virus (VEE), sindbis virus (SIN), semliki forest virus (SFV), and VEE-SIN chimeras),

herpes virus vectors, measles virus vectors, pox virus vectors (such as, vaccinia virus, modified vaccinia virus Ankara (MVA), NYVAC (derived from the Copenhagen strain of vaccinia), and avipox vectors: canarypox (ALVAC) and fowlpox (FPV) vectors), and vesicular stomatitis virus vectors. Other suitable plasmids and vectors are known in the art and can readily be selected by the skilled worker. In accordance with various embodiments of the invention, either low copy number or high copy number vectors may be employed.

[00108] One skilled in the art will understand that the vector may further include regulatory elements, such as transcriptional elements, required for efficient transcription of the DNA sequence encoding the fusion protein. Examples of regulatory elements that can be incorporated into the vector include, but are not limited to, promoters, enhancers, terminators, alpha-factors, ribosome binding sites and polyadenylation signals.

[00109] One skilled in the art will appreciate that selection of suitable regulatory elements is dependent on the host cell chosen for expression of the encoded protein and that such regulatory elements may be derived from a variety of sources, including bacterial, fungal, viral, mammalian or insect genes.

[00110] Mammalian expression vectors, for example, may comprise one or more of an origin of replication, any necessary ribosome binding sites, a polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking non-transcribed sequences. DNA sequences derived from the SV40 splice and polyadenylation sites, for example, may be used to provide the required non-transcribed genetic elements. Eukaryotic expression vectors may also contain one or more enhancers to increase expression levels of the protein. Enhancers are cis-acting elements of DNA, usually from about 10 to about 300 bp in length that act on a promoter to increase its transcription. Examples include, but are not limited to, the SV40 enhancer on the late side of the replication origin by 100 to 270, the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin and the adenovirus enhancers.

[00111] Examples of typical promoters include, but are not limited to, the bacterial promoters: *lacZ*, *T3*, *T7*, *gpt*, λP_R , P_L and *trp*; and the eukaryotic promoters:

CMV immediate early, HSV thymidine kinase, early SV40, late SV40, LTRs from retrovirus and mouse metallothionein-I. Promoter regions can also be selected from a desired gene using chloramphenicol transferase (CAT) vectors or other vectors with selectable markers.

5 [00112] In certain embodiments the vector comprises an expression control sequence which is a “regulated promoter,” which may be a promoter as provided herein or may be a repressor binding site, an activator binding site or other regulatory sequence that controls expression of a nucleic acid sequence. In some embodiments, the vector comprises a tightly regulated promoter that is specifically inducible and that permits
10 little or no transcription of nucleic acid sequences under its control in the absence of an induction signal. Examples of such tightly regulated promoters are known in the art and described, for example, in Guzman *et al.* (1995, *J. Bacteriol.* 177:4121), Carra *et al.* (1993, *EMBO J.* 12:35), Mayer (1995, *Gene* 163:41), Haldimann *et al.* (1998, *J. Bacteriol.* 180:1277), Lutz *et al.* (1997, *NAR.* 25:1203), Allgood *et al.* (1997, *Curr. Opin. Biotechnol.* 8:474) and Makrides (1996, *Microbiol. Rev.* 60:512). In other
15 embodiments of the invention, the vector comprises a regulated promoter that is inducible but that may not be tightly regulated. Inducible systems that include regulated promoters include, for example, the Tet system or other similar expression-regulating components, such as the Tet/on and Tet/off system (Clontech Inc., Palo Alto, CA), the
20 Regulated Mammalian Expression system (Promega, Madison, WI), and the GeneSwitch System (Invitrogen Life Technologies, Carlsbad, CA).

[00113] In certain embodiments, the vector comprises a promoter that is not a regulated promoter; such a promoter may include, for example, a constitutive promoter such as an insect polyhedrin promoter.

25 [00114] In addition, vectors may contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells. Such selectable markers include for example genes encoding dihydrofolate reductase or genes conferring neomycin resistance in eukaryotic host cells, genes conferring ampicillin, chloramphenicol, erythromycin, kanamycin, neomycin or tetracycline resistance in
30 bacterial host cells, and the *S. cerevisiae* TRP1 gene. Promoter regions can be selected from a desired gene using chloramphenicol transferase (CAT) vectors or other vectors

with selectable markers. Selectable markers can also include biosynthetic genes, such as those in the histidine, tryptophan and leucine biosynthetic pathways.

5 [00115] In certain embodiments, the vector can have two replication systems to allow it to be maintained in two organisms, for example in mammalian or insect cells for expression and in a prokaryotic host for cloning and amplification.

[00116] Also contemplated are replicating and non-replicating episomal vectors for transient expression. The replicating vectors containing origin sequences that promote plasmid replication in the presence of the appropriate trans factors. The SV40 and polyoma origins and respective T-antigens are examples. Also contemplated are stably
10 maintained episomal expression vectors. Episomal plasmids are usually based on sequences from DNA viruses, such as BK virus, bovine papilloma virus 1 and Epstein-Barr virus (see, for example, Van Craenenbroeck, K., *et al.*, 2000, *Eur. J. Biochem.* 267:5665-5678). These vectors contain a viral origin of DNA replication and a viral early gene(s), the product of which activates the viral origin and thus allows the
15 episome to reside in the transfected host cell line in a well-controlled manner. Episomal vectors are plasmid constructions that replicate in both eukaryotic and prokaryotic cells and can therefore also be “shuttled” from one host cell system to another.

[00117] In some embodiments the plasmid can be integrated into the host chromosome. Integration can occur by random methods or can be targeted. In some
20 embodiments in which integrating expression vectors are used, the expression vector can contain at least one sequence homologous to the host cell genome, for example, two homologous sequences which flank the expression construct. The integrating vector can thus be directed to a specific locus in the host cell by selecting the appropriate homologous sequence for inclusion in the vector. Constructs and methods
25 for integrating vectors are well known in the art. Alternatively, the use of recombination systems like Cre/Lox and Flp/Frt can be used to target integration. Other methods utilizing zinc-finger proteins as developed by Sangamo Biosciences, Inc. (Richmond, CA) provide another approach to targeting vector integration.

[00118] In certain embodiments, the methods described herein employ a vector or
30 recombination system that allows for stable integration of the polynucleotide into the host cell genome. In some embodiments, the methods described herein employ a vector

or recombination system that allows for stable integration of the polynucleotide into the host cell genome as a single copy.

[00119] In certain embodiments of the invention, the vector employed is a viral vector such as a retroviral vector. For example, retroviruses from which the retroviral plasmid vectors may be derived include, but are not limited to, Moloney Murine Leukemia Virus, spleen necrosis virus, retroviruses such as Rous Sarcoma Virus, Harvey Sarcoma virus, avian leukosis virus, gibbon ape leukemia virus, human immunodeficiency virus, adenovirus, Myeloproliferative Sarcoma Virus, and mammary tumour virus. Suitable promoters for inclusion in viral vectors include, but are not limited to, the retroviral LTR; the SV40 promoter; and the human cytomegalovirus (CMV) promoter described in Miller, *et al.* (1989, *Biotechniques* 7:980-990), or other suitable promoter (for example, cellular promoters such as eukaryotic cellular promoters including, but not limited to, the histone, pol III, and β -actin promoters). Other viral promoters which may be employed include, but are not limited to, adenovirus promoters, thymidine kinase (TK) promoters, and B19 parvovirus promoters. The selection of a suitable promoter will be apparent to those skilled in the art, and may be from among either regulated promoters or promoters as described above.

[00120] In those embodiments that employ a retroviral plasmid vector, the vector is used to transduce packaging cell lines to form producer cell lines. Examples of packaging cells which may be transfected include, but are not limited to, the PE501, PA317, [psi]-2, [psi]-AM, PA12, T19-14X, VT-19-17-H2, [psi]CRE, [psi]CRIP, GP+E-86, GP+envAm12, and DAN cell lines as described in Miller (1990, *Human Gene Therapy*, 7:5-14). The packaging cells may be transduced with the vector using various means known in the art such as, for example, electroporation, the use of liposomes, and CaPO₄ precipitation. The producer cell line generates infectious retroviral vector particles which include the polynucleotide encoding the protein. Such retroviral vector particles then may be employed to transduce eukaryotic cells, either *in vitro* or *in vivo*, and the transduced eukaryotic cells will express the polynucleotide encoding the protein. Eukaryotic cells which may be transduced include, but are not limited to, embryonic stem cells, embryonic carcinoma cells, hematopoietic stem cells, hepatocytes, fibroblasts, myoblasts, keratinocytes, endothelial cells, and bronchial epithelial cells.

[00121] The appropriate DNA or polynucleotide sequences can be inserted into the vector by a variety of procedures known in the art. In general, the DNA sequence is inserted into an appropriate restriction endonuclease site(s) by procedures known in the art. Standard techniques for cloning, DNA isolation, amplification and purification, for enzymatic reactions involving DNA ligase, DNA polymerase, restriction endonucleases and the like, and various separation techniques are those known and commonly employed by those skilled in the art. A number of standard techniques are described, for example, in Sambrook et al., *ibid.*; Ausubel *et al.*, *ibid.*, and elsewhere.

[00122] The vector can be introduced into a suitable host cell by one of a variety of methods known in the art. Such methods can be found generally described in Ausubel *et al.* (*ibid.*) and include, for example, stable or transient transfection, lipofection, electroporation, and infection with recombinant viral vectors. One skilled in the art will understand that selection of the appropriate host cell for expression will be dependent upon the vector chosen. The polynucleotide may stably integrate into the genome of the host cell (for example, with retroviral introduction) or may exist either transiently or stably in the cytoplasm (for example, through the use of traditional plasmids or vectors, utilizing standard regulatory sequences, selection markers, and the like, as described above).

Host cells

[00123] In accordance with the present invention, the host cell employed in the methods described herein is a host cell capable of utilizing recombination signals and undergoing RAG-1/RAG-2 mediated recombination. Accordingly, host cells suitable for use in the methods described herein express or can be engineered to express at least RAG-1 and RAG-2 or functional fragments thereof that allow the host cell to utilize recombination signals and undergo RAG-1/RAG-2 mediated recombination.

[00124] In certain embodiments, cell lines to be used as host cells may additionally contain a functional TdT gene. TdT is encoded by a single gene and expresses a nuclear enzyme whose expression *in vivo* is restricted to lymphoid progenitor cells. TdT has, however, been expressed in non-lymphoid cells and shown to participate in V(D)J recombination using retroviral and transient recombination substrates. TdT has been shown to be expressed as a number of different splice variants, including long form and

short form. Certain embodiments of the invention contemplate the use of different isoforms of TdT.

[00125] TdT has also been shown to have a 3' to 5' exonuclease activity and the different isoforms of TdT have been shown to have different amounts of exonuclease activity. TdT exonuclease activity can be modulated by substitutions at the conserved aspartic acid residue in the exonuclease motif. In addition, expression of both isoforms was shown to modulate nuclease activity. TdT is highly conserved among species. While mice have two isoforms both human and bovine have three isoforms. In certain embodiments, TdT activity in the host cell can be modulated by altering the levels of TdT in the cell. In some embodiments, mutant forms of TdT or different combinations of isoforms may be used in the host cell to generate coding joints with different extents of deletion and addition.

[00126] Recombination-competent host cells may in certain embodiments be pre-B cells or pre-T cells that express RAG-1, RAG-2 and TdT proteins. Such pre-B and pre-T cells may be capable of being induced to express RAG-1, RAG-2 and TdT, or alternatively, may constitutively express RAG-1, RAG-2 and TdT but can be modified to substantially impair the expression of one, two or all three of these enzymes.

[00127] In some embodiments, the recombination-competent host cells are non-immune cells that have been transformed with genes encoding each of RAG-1, RAG-2 and TdT. One skilled in the art can readily select an appropriate non-immune host cell. Examples of host cells include, but are not limited to, yeast and mammalian cells. Specific non-limiting examples include green African monkey kidney (COS) cells, NIH 3T3 cells, Chinese hamster ovary (CHO) cells, BHK cells, human embryonic kidney (HEK 293) cells, Huh7.5 human hepatoma cells, Hep G2 human hepatoma cells, Hep 3B human hepatoma cells, HeLa cells and the like.

[00128] These and other recombination-competent host cells may be used according to contemplated embodiments of the present invention. For example, expression of RAG-1 and/or RAG-2 has been observed in mature B-cells *in vivo* and *in vitro* (Maes *et al.*, 2000, *J Immunol.* 165:703; Hikida *et al.*, 1998, *J Exp Med.* 187:795; Casillas *et al.*, 1995, *Mol Immunol.* 32:167; Rathbun *et al.*, 1993, *Int Immunol.* 5:997; Hikida *et al.*, 1996, *Science* 274:2092).

[00129] RAG-1 and RAG-2 have also been shown to be expressed in mature T-cell lines including Jurkat T-cells. CEM cells have been shown to have V(D)J recombination activity using extrachromosomal substrates (Gauss *et al.* 1998, *Eur J Immunol.* 28:351). Treatment of wild-type Jurkat T cells with chemical inhibitors of signaling components revealed that inhibition of Src family kinases using PP2, FK506, and the like, overcame the repression of RAG-1 and resulted in increased RAG-1 expression. Mature T-cells have also been shown to reactivate recombination with treatment of anti-CD3/IL7 (Lantelme *et al.*, 2008, *Mol Immunol.* 45:328).

[00130] Tumor cells of non-lymphoid origin have also been shown to express RAG-1 and RAG-2 (Zheng *et al.*, 2007, *Mol Immunol.* 44: 2221, Chen *et al.*, 2007, *Faseb J.* 21:2931). Accordingly, in certain embodiments, these cells may also be suitable for use as recombination-competent host cells in the presently described methods. According to other embodiments that are contemplated herein, reactivation of V(D)J recombination would provide another approach to generating a suitable host cell with inducible recombinase expression.

[00131] Alternatively, only one of the RAG-1 or RAG-2 genes may be stably integrated into a host cell, and the other gene can be introduced by transformation to regulate recombination. For example, a cell line that is stably transformed with TdT and RAG-2 would be recombinationally silent. Upon transient transfection with RAG-1, or viral infection with RAG-1, the cell lines would become recombinationally active. The skilled person will appreciate from these illustrative examples that other similar approaches may be used to control the onset of recombination in a host cell.

[00132] Substantial impairment of the expression of one or more recombination control elements (for example, one or more of a RAG-1 gene, RAG-2 gene or TdT gene) may be achieved by a variety of methods that are well known in the art for blocking specific gene expression, including antisense inhibition of gene expression, ribozyme mediated inhibition of gene expression, siRNA mediated inhibition of gene expression, and Cre recombinase regulation of expression control elements using the Cre/Lox system. As used herein, expression of a gene encoding a recombination control element is substantially impaired by such methods for inhibiting when host cells are substantially but not necessarily completely depleted of functional DNA or functional

mRNA encoding the recombination control element, or of the relevant polypeptide. In certain embodiments, recombination control element expression is substantially impaired when cells are at least about 50% depleted of DNA or mRNA encoding the endogenous polypeptide (as detected using high stringency hybridization, for example) or at least about 50% depleted of detectable polypeptide (as measured by Western immunoblot, for example); for example, at least 75% depleted or at least 90% depleted.

[00133] Cell lines can also include added genetic elements giving them useful functionality. Invitrogen provides a flp-in system in which the Frt recombination signal is integrated into different host cell lines (3T3, BHK, CHO, CV-1, 293). Equivalent cell lines incorporating LoxP sites or other sites for targeting integration can be used. A tet inducible system (for example, T-Rex from Invitrogen, Carlsbad, CA) for 293 or HeLa cell lines or other available inducible systems may also be used.

Screening Assays

[00134] The methods according to the present invention may optionally include one or more screening steps, for example, to screen for expression of variant fusion proteins by the host cells and/or to screen for variant fusion proteins having a desired functionality.

[00135] In certain embodiments, the methods of the invention comprise screening transformed host cells for expression of variant fusion proteins. Various protein expression assays are known in the art and include the use of UV/VS spectrophotometry, fluorescence spectrophotometry, mass spectrometry and the like. As noted above, in some embodiments, the variant fusion proteins may comprise an additional polypeptide sequence to facilitate detection, for example, by localizing the protein to the cell surface or by incorporating a detectable label.

[00136] In certain embodiments in which the variant fusion proteins are not localized to the cell surface or secreted, the expression assay may further comprise a cell lysis step or the protein may be assayed directly within the cell for function.

[00137] In certain embodiments, the methods of the invention comprise submitting the variant fusion proteins to a functional assay to identify those variants having a desired

functionality. The specific assay used will be dependent on the functionality being assessed. Various functional assays are known in the art and appropriate assays can be readily selected by the skilled worker. Commonly used assays include, for example, ELISA- and FACS-based assays.

- 5 [00138] The functionality of the variant fusion proteins may be assessed by assaying the cells expressing the variants or the variants may be isolated from the host cells and assayed as isolated proteins.

[00139] In some embodiments of the invention, the methods generate high numbers of variant fusion proteins and in such embodiments high throughput screening approaches
10 are generally preferred. Many high throughput screening approaches are well known in the art and can be readily applied to identify and select variant fusion proteins with a desired functionality.

POLYNUCLEOTIDE COMPOSITIONS

- [00140] In certain embodiments, the invention provides for polynucleotides capable of
15 undergoing RSS-mediated recombination when introduced into a recombination-competent host cell, and compositions comprising same.

[00141] In some embodiments, the polynucleotide is a “bipartite recombination substrate” and preferably comprises: a first nucleic acid sequence including a first coding sequence encoding a first portion of a fusion protein and a first recombination
20 signal sequence (RSS), and a second nucleic acid sequence including a second coding sequence encoding a second portion of the fusion protein and a second RSS capable of functional recombination with the first RSS.

[00142] The coding sequences comprised by the bipartite recombination substrate as described above may encode portions of the same protein, for example different
25 domains of the protein, or each may encode all or a portion of a different protein. When the coding sequences encode portions of the same protein, they may when taken together encode the whole protein, or they may encode a truncated or rearranged version of the protein.

[00143] In some embodiments, the bipartite recombination substrate may further comprise a linker sequence between the first and second coding sequence with the first and second RSSs positioned within or proximal to the linker sequence, such that when the polynucleotide undergoes RSS-mediated recombination, sequence diversity is introduced into the linker sequence.

[00144] In some embodiments, the polynucleotide is a “tripartite recombination substrate” and preferably comprises: a first nucleic acid sequence as described above; a second nucleic acid sequence including a second coding sequence encoding a second portion of the fusion protein, a second RSS capable of functional recombination with the first RSS, and a third RSS; and a third nucleic acid sequence including a third coding sequence encoding a third portion of the fusion protein and a fourth RSS capable of functional recombination with the third RSS.

[00145] In certain embodiments, the polynucleotide is a tripartite recombination substrate as described above in which the first and third coding sequences encode portions of the same protein, and the second coding sequence encodes a heterologous sequence that is inserted into the protein, with sequence diversity being generated at both junctions. In some embodiments, the tripartite recombination substrate may further comprise a linker sequence between the first and second coding sequence with the first and second RSSs positioned within or proximal to the linker sequence and/or a linker sequence between the second and third coding sequence with the third and fourth RSSs positioned within or proximal to the linker sequence, such that when the polynucleotide undergoes RSS-mediated recombination, sequence diversity is introduced into the linker sequence(s).

[00146] In certain embodiments, the polynucleotide is a tripartite recombination substrate as described above in which all three coding sequences encode portions of the same protein, for example different domains of the protein. In this case, the three coding sequences taken together may encode the whole protein, or may encode a truncated or rearranged version of the protein. In some embodiments, the polynucleotide is a tripartite recombination substrate as described above in which two of the three coding sequences encode portions of the same protein, for example different domains of the protein. Embodiments in which the tripartite recombination

substrate comprises three coding sequences each from a different protein, for example each encoding a domain or polypeptide, are also contemplated.

[00147] In some embodiments, the polynucleotide is a tripartite recombination substrate as described above in which the first and third coding sequences encode portions of an antibody variable region, and the second coding sequence encodes a heterologous sequence that is inserted into the antibody variable region, for example into CDR1, CDR2 or CDR3, with sequence diversity being generated at both junctions. In some embodiments, the polynucleotide is a tripartite recombination substrate as described above in which the first and third coding sequences encode portions of a non-Ig protein, and the second coding sequence encodes a heterologous sequence that is inserted into the protein.

[00148] In some embodiments, the polynucleotide comprises RSSs that are accompanied by flanking sequences adjacent to one or both of the heptamers of the RSS. In some embodiments, the polynucleotide comprises RSSs that are accompanied by flanking sequences that encode a specific amino acid, or amino acids, or peptide sequence.

[00149] The polynucleotide compositions may be provided as isolated polynucleotides or they may be provided as part of a vector, in which case they may be operatively linked to one or more regulatory elements, such as, promoters, enhancers, terminators, alpha-factors, ribosome binding sites, polyadenylation signals and the like, as described above. The present invention also contemplates that the compositions may be provided as host cells that have been transformed with the polynucleotide or a vector comprising the polynucleotide. Examples of suitable host cells include those described above.

APPLICATIONS

[00150] In accordance with one aspect of the present invention, the methods can be used to generate variants of a fusion protein, for example fusion proteins having a desired functionality or in which one or more of the components of the fusion protein have an improved or optimized functionality. In certain embodiments, the methods are employed to generate a large number of variants of the fusion protein for subsequent screening for a desired or improved functionality.

[00151] In certain embodiments, the methods are used to generate modified protein scaffolds that include a heterologous amino acid sequence that provides a new functionality to the protein scaffold, for example, a ligand-binding functionality. In some embodiments, the methods are used to generate variants of a fusion protein that
5 comprise two different proteins or protein domains joined by a linker in which sequence diversity is introduced into the linker in order to optimize the functionality of one or both of the components of the fusion protein.

[00152] In some embodiments, the methods are used to graft non-antibody sequences (for example, a protein domain or polypeptide) into an antibody CDR and identify the
10 appropriate sequence context (length and composition) that allows the protein domain or polypeptide to remain functional within the context of the antibody scaffold.

[00153] In certain embodiments, the methods of the invention are used to insert large protein domains into a heterologous coding sequence, such as a protein scaffold, and retain biological function. In some embodiments, the methods use flanking sequences
15 next to the heptamer of one or more of the RSSs such that an inserted protein domain is allowed to maintain an appropriate conformation for functionality within the heterologous protein scaffold.

[00154] In certain embodiments, the invention provides for the use of the methods for peptide grafting to generate fusion proteins with ligand-binding properties (for
20 example, modified antibodies, avimers, adnectins, or other antibody mimetics) for therapeutic purposes, for diagnostic purposes, for drug targeting (for example, through the use of a ligand-binding protein that targets a protein on a particular cell or tissue type as a targeting moiety for attachment to a therapeutic or diagnostic compound), or for research applications (such as screening assays, chromatography and the like).

25 ***PEPTIDE-GRAFTED IMMUNOGLOBULINS***

[00155] Certain embodiments of the invention relate to peptide-grafted immunoglobulins in which one or more peptides of interest having optimized flanking sequences have been grafted into one or more CDRs of an immunoglobulin. In certain
30 embodiments, the peptide(s) comprised by the peptide-grafted immunoglobulins are targeted to a receptor. In some embodiments, the peptide(s) comprised by the peptide-

grafted immunoglobulins are targeted to a receptor from a clinically relevant receptor class, such as a GPCR or ion channel. Such receptors have historically been difficult to target. As demonstrated herein, it is possible to graft peptides with reactivity to a GPCR into the CDRs of a full length human IgG scaffold and retain the ability of the peptide(s) to bind their target thus demonstrating that peptide-grafted immunoglobulins can be used successfully to target these complex membrane proteins.

[00156] Certain embodiments of the invention thus relate to peptide-grafted immunoglobulins that comprise one or more peptides targeted to a GPCR. GPCRs are classified into six families: the rhodopsin family (A), the secretin-receptor family (B), the metabotropic glutamate receptor family (C), fungal pheromone P- and α -factor receptors (D), fungal pheromone A- and M-factor receptors (E) and cyclic-AMP receptors from *Dictyostelium* (F). Peptide-grafted immunoglobulins that comprise one or more peptides targeted to a GPCR from any one of these families are contemplated in various embodiments of the invention. In certain embodiments, the peptide-grafted immunoglobulins comprise one or more peptides targeted to a Family B GPCR.

[00157] Suitable peptides for targeting a GPCR may be derived from, for example, a known natural or synthetic ligand. Peptides may also be derived from snake venom peptides, or toxic peptides from other organisms, which are small and contain a loop structure, and are thus suitable for CDR grafting.

[00158] In some embodiments, the invention relates to peptide-grafted immunoglobulins comprising one or more peptides targeted to the GLP-1 receptor. Suitable peptides include those derived from GLP-1 and from exendin-4. Non-limiting examples of appropriate targeting peptides and flanking sequences are provided in Example 10 (see Tables 5 and 6).

[00159] The immunoglobulin scaffold may be a full-length immunoglobulin (such as a full-length IgA, IgA2, IgD, IgE, IgGs (*i.e.* IgG1, IgG2, IgG3 or IgG4) or IgM) or an immunoglobulin fragment (such as a Fab, Fab', F(ab')₂, Fd, Fv and single-chain Fv (scFv) fragment). Certain embodiments of the invention relate to peptide-grafted immunoglobulins in which the immunoglobulin scaffold is a full-length IgG immunoglobulin.

[00160] Immunoglobulins suitable for use in the methods described herein may be derived from a variety of sources and technologies including, but not limited to, mammals including mice, transgenic mice and humans, phage display or yeast display, or they may be synthetically derived immunoglobulins or fragments thereof.

- 5 [00161] Certain embodiments of the invention contemplate the use of immunoglobulin scaffolds from camelid antibodies; HCAns; single chain antibodies; shark antibodies; diabodies; nanobodies and fluorobodies.

[00162] The peptide(s) may be grafted into a heavy chain CDR or a light chain CDR or both. In some embodiments, the peptide(s) may be grafted into a heavy chain CDR3 or
10 a light chain CDR3 or both. In certain embodiments, the peptide is grafted into at least a heavy chain CDR. In some embodiments, the peptide replaces the D segment in the heavy chain CDR3.

[00163] Optimization of the flanking sequences may be achieved using the methods described above, or by other methods described in the art. For example, gene synthesis
15 can be used to synthesize a V gene segment utilizing degenerate nucleotides in selected positions flanking the peptide encoding sequence. The synthesis of V gene segments with different lengths of flanking sequences in combination with degenerate nucleotides results in both sequence length and composition differences flanking the peptide. These peptide-grafted immunoglobulin variable gene sequences can be cloned
20 and manipulated by any of a variety of methods known in the art for screening or selection, such as phage display, yeast display, or transfection or infection of mammalian cell lines.

[00164] Another method for generating libraries of peptide-grafted immunoglobulin variants includes cloning utilizing oligonucleotides. The V and/or J gene sequences
25 and peptide encoding DNA sequences can be cloned together utilizing different oligonucleotide adapters that contain different amino acid sequences and are designed to ligate and join the variable sequences to the peptide encoding DNA sequences. The oligonucleotides can also differ in length and a large set of oligonucleotides can be generated that represent different lengths and compositions of amino acids. Cloning
30 these pools of oligonucleotides in between the VH gene segment and the DNA

sequences encoding the peptide will generate novel fusions differing in both length and composition. Techniques utilizing trinucleotide mutagenesis have also been described.

[00165] Another method to generate libraries of peptide-grafted immunoglobulin variants utilizes PCR. PCR-based cloning can also be employed to generate amino acid
5 diversity of length and composition between the V and/or the J gene segment. Primers annealing to the peptide encoding sequence can be designed to include degenerate oligonucleotides. The use of degenerate flanking sequences on both the forward and reverse primers will generate a mixture of PCR fragments containing the peptide encoding sequences and a diversity of flanking sequences. PCR primers can be
10 designed with different lengths of flanking sequences. The PCR products are then cloned in between a variable and joining gene segment to generate a library of peptide grafted antibody variable chain variants.

[00166] These above techniques can also be combined. Other techniques are known in the art, including site-directed mutagenesis that would target diversity to the flanking
15 nucleotide sequences, or the use of error-prone PCR.

[00167] Typically the peptide will comprise both an upstream (*i.e.* C-terminal) and a downstream (*i.e.* N-terminal) flanking sequence, although embodiments in which only one of an upstream or a downstream flanking sequence is present are also contemplated. Flanking sequences may be between about 1 and about 30 amino acids in
20 length, for example, between about 2 and about 30 amino acids in length, between about 1 and about 25 amino acids, between about 2 and about 25 amino acids, between about 1 and about 20 amino acids, between about 2 and about 20 amino acids, between about 1 and about 15 amino acids in length, between about 2 and about 15 amino acids in length, or any amount therebetween, for example 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or more
25 amino acids in length. The upstream and downstream flanking sequences may be the same length or may be different lengths.

[00168] In some embodiments, the peptide-grafted immunoglobulins are initially identified by employing the methods of the present invention to graft the peptide into the immunoglobulin scaffold and to optimize the flanking sequences. Certain
30 embodiments of the invention thus relate to the use of the methods described herein that employ components of the V(D)J recombination system to generate the above-

described peptide-grafted immunoglobulins. Once an active peptide-grafted immunoglobulin has been thus identified, it may be sequenced and additional quantities of the molecule prepared by standard molecular biology and protein purification techniques. Further optimization of the peptide-grafted immunoglobulin using the methods described herein is also contemplated in certain embodiments.

KITS

[00169] Certain embodiments of the invention provide for kits comprising a polynucleotide capable of undergoing RSS-mediated recombination when introduced into a recombination-competent host cell, or a composition comprising a polynucleotide capable of undergoing RSS-mediated recombination when introduced into such a host cell, as described above.

[00170] When the kit comprises a composition, the composition may comprise an isolated polynucleotide, a polynucleotide comprised by a vector (in which case the polynucleotide may be operatively linked to one or more regulatory elements, such as, promoters, enhancers, terminators, alpha-factors, ribosome binding sites, polyadenylation signals and the like), or a host cell that has been transformed with the polynucleotide or a vector comprising the polynucleotide.

[00171] When the kit comprises an isolated polynucleotide, the kit may further comprise a vector suitable for expression of the polynucleotide and/or a recombination-competent host cell.

[00172] The kit may further comprise vectors encoding one or more of RAG-1, RAG-2 and TdT that are suitable for transforming a host cell such that the host cell expresses, or is capable of expressing, RAG-1, RAG-2 and/or TdT.

[00173] The kit may further comprise one or more additional components to assist with cloning the polynucleotide and/or transformation of host cells, such as buffers, enzymes, selection reagents, growth media and the like.

[00174] One or more of the components of the kit may optionally be lyophilised and the kit may further comprise reagents suitable for the reconstitution of the lyophilised components. Individual components of the kit would be packaged in separate

containers and, associated with such containers, can be instructions for use. The instructions for use may be provided in paper form or in computer-readable form, such as a disc, CD, DVD or the like.

5 [00175] To gain a better understanding of the invention described herein, the following examples are set forth. It will be understood that these examples are intended to describe illustrative embodiments of the invention and are not intended to limit the scope of the invention in any way.

EXAMPLES

EXAMPLE 1: Construction Of A Fusion Protein Grafting Acceptor Vector

10 [00176] A vector was designed to allow a selected amino acid sequence to be grafted into the variable region of an antibody heavy chain. A schematic of the vector is shown in Figure 1A and the nucleotide sequence [SEQ ID NO:28] is provided in Figure 1(B).

15 [00177] The vector comprises a stuffer sequence between two BsmBI restriction sites, with an upstream V_H sequence and 23bp RSS and a downstream 23bp RSS and J_H sequence operably linked to the IgG1 constant region. The vector is designed to accept a nucleic acid sequence encoding the selected amino acid sequence flanked by 12 bp RSSs. The locations of the various components of the vector are provided in Table 2 below.

20 [00178] Additional details are provided in Example 9, together with an exemplary method for transfection of a recombination substrate into a recombination-competent host cell, expansion of the host cell and recombination and expression of the substrate.

Table 2: Location Of Components Of The Fusion Protein Grafting Acceptor Vector Within SEQ ID NO:28

CMV promoter	1-621
VH3-33	663-1015
Flanking sequences (optional)	1016
23bp-RSS	1016-1054
Intervening sequence #1	1055-1608
BsmBI site #1	1609-1614
Stuffer sequences	1615-3556
BsmBI site #2	3557-3562
Intervening sequences #2	3563-4159
23bp-RSS	4160-4198
Flanking sequences (optional)	4199
JH4	4199-4246
Spice donor-intron-splice acceptor	4247-4541
CH1-hinge-CH2CH3	4542-5530
Transmembrane sequence	5531-5596
Cytoplasmic sequence	5597-5698

EXAMPLE 2: Grafting RGD From Fibronectin Type III 10 Into An Antibody

5 Scaffold

[00179] The grafting of RGD from fibronectin type III 10 into an antibody scaffold using the V(D)J *in vitro* system will be conducted as follows.

[00180] The RGD sequence is derived from fibronectin and is flanked by 12 bp RSSs and BsmBI sites (SEQ ID NO:29, below) and placed into the acceptor vector described
10 in Example 1 to generate a recombination substrate for generating a library of variant fusion proteins.

[00181] Nucleotide sequence of RGD peptide #1 (61bp; 20 amino acids; in capitals) flanked by 12 bp RSSs and BsmBI sites. Nucleotides encoding “RGD” in bold.

cgtctctccaagtgcaaaggacaggaggttttgtaaggctgtatcactgtgTATACCATCACTGTGTAT
15 GCTGTCACTGGCC**GTGGAGAC**AGCCCCGCAAGCAGCAAGCCAATTTCCAT
Tcacagtatacagcccttaacaaaaaccctactgcaacctggcgtaagagacg [SEQ ID NO:29]

[00182] Nucleotide sequence of RGD peptide #1 (61bp)

TATACCATCACTGTGTATGCTGTCACTGGCCGTGGAGACAGCCCCGCAAGC
AGCAAGCCAATTTCCATT [SEQ ID NO:30]

[00183] Amino acid sequence of RGD peptide #1 (20 amino acids) [SEQ ID NO:31].

YTITVYAVT**GRGD**SPASSKPISI

- 5 [00184] Nucleotide sequence of RGD peptide #2 (39bp; 13 amino acids; in capitals) flanked by 12 bp RSSs and BsmBI sites. Nucleotides encoding “RGD” in bold.

cgtctctccaagtgcaaaggacaggagggttttgaagggtgtatcactgtg**TATGCTGTCACTGGCCGT
GGAGACAGCCCCGCAAGCAGC**cacagtgatacagcccttaacaaaaccctactgcaacctggcggg
aagagacg [SEQ ID NO:32]

- 10 [00185] Nucleotide sequence of RGD peptide #2 (39bp)

TATGCTGTCACTGGCCGTGGAGACAGCCCCGCAAGCAGC [SEQ ID NO:33]

[00186] Amino acid sequence of RGD peptide #2 (13 amino acids)

YAVT**GRGD**SPASS [SEQ ID NO:34]

- 15 [00187] The fibronectin sequences can be modified so that an RGD peptide sequence is created in all three reading frames. The modified fibronectin sequence is shown below.

[00188] TATACCATCACGTGGAGACCTGTGTATGCTGTCACTGGCCGTGGAG
ACAGCCCCGCGGAGACAAGCAGCAAGCCAATTTCCATT [SEQ ID NO:35]

- 20 [00189] The three reading frames with the RGD sequences are shown in Figure 2 [SEQ ID NOs:9-11].

- [00190] The acceptor vector with a selected RGD peptide construct is co-integrated into a cell line with a kappa light chain. The cell line is selected for light chain expression and the ability to recombine the integrated substrate. The cell line is then expanded in the unrecombined state and V(D)J recombination is induced to generate a
25 library of greater than 10 million cells each expressing a unique fusion protein on the cell surface. The cell library is then incubated with soluble form of biotinylated

alphaVBeta5 integrin and avidin-conjugated fluorochrome, and FACS sorted to isolate fusion proteins with integrin binding properties.

EXAMPLE 3: Grafting G-Protein Coupled Receptor Ligands Into An Antibody Scaffold

5 [00191] G protein-coupled receptors (GPCRs) are a family of integral transmembrane proteins thought to have the same molecular architecture, consisting of seven transmembrane domains (7TM), three extracellular loops (EC1, EC2, EC3), three intracellular loops (IC1, IC2, and IC3), an amino-terminal extracellular domain and an intracellular carboxyl terminus. This topology is predicted from the analysis of
10 hydrophathy profiles and from a limited amount of experimental evidence, most importantly from the crystal structure rhodopsin. GPCRs were classified into six families: the rhodopsin family (A), the secretin-receptor family (B), the metabotropic glutamate receptor family (C), fungal pheromone P- and α -factor receptors (D), fungal pheromone A- and M-factor receptors (E) and cyclic-AMP receptors from
15 *Dictyostelium* (F). Although many anti-GPCR antibodies have been generated it is generally appreciated that GPCRs are difficult targets to generate neutralizing antibodies. This Example will utilize domain grafting of sequences derived from GPCR ligands to engineer anti-GPCR specificity and antibodies with desired activities.

[00192] Peptide sequences derived from glucagon-like peptide-1 (GLP-1) and
20 Exendin-4 are used to generate fusion proteins targeting GLP-1R, a receptor belonging to the B1 family of seven-transmembrane G protein-coupled receptors. GLP-1 is a peptide hormone generated in intestinal L-cells that binds to GLP-1R on pancreatic beta-cells and potentiates the synthesis and release of insulin in a glucose-dependent manner. Exensin-4 is a peptide hormone found in the saliva of the Gila monster. It is a
25 GLP-1R agonist that binds to the receptor with an affinity and potency similar to GLP-1.

[00193] Two versions of GLP-1 have been selected to generate fusion proteins. The first version is based on the complete hormone sequence (amino acids 7-37). A second version is based on a truncated form of GLP-1 that spans amino acids 13-33. This
30 region has been found to assume an alpha-helical confirmation when bound to GLP-1R

and includes several residues important for receptor binding. Schematic representations of a cassette comprising each sequence together with appropriate 12bp RSS sequences, as well as the respective nucleotide sequences [SEQ ID NOs:36 and 37] and amino acid sequences [SEQ ID NOs:38 and 39] for the cassettes are provided in Figure 3.

5 [00194] Fusion proteins will be generated from three forms of exendin-4. The first version is based on all 39 amino acids of the natural peptide. The second version is based on amino acids 9-39. This truncated form of exendin-4 is a competitive antagonist that binds to GLP-1R with high affinity. The third version was based on residues 15-27. This region is believed to form the most critical interactions with GLP-
10 1R. Schematic representations of a cassette comprising each sequence together with appropriate 12bp RSS sequences, as well as the respective nucleotide sequences [SEQ ID NOs:40, 41 and 42] and amino acid sequences [SEQ ID NOs:43, 44 and 45] for the cassettes are provided in Figure 4.

[00195] The nucleic acid sequences encoding the peptides will be codon optimized and stop codons introduced into the other non-relevant reading frames. The bolded
15 sequences in Figures 3 and 4 represent the specific nucleic acid sequences encoding the appropriate peptide. The sequences are gene synthesized as a cassette that includes flanking BsmBI sites and 12 bp RSS sequences as shown. The BsmBI sites are used to clone the RSS- peptide- RSS cassette into the acceptor vector described in Example 1.

20 **EXAMPLE 4: Peptide Grafting To Generate An Anti-CXCR1 Binding Antibody**

[00196] This example utilizes sequences from IL-8, Gro-alpha and LL-37 (which are all ligands that bind CXCR1) to generate an antibody with anti-CXCR1 specificity. Each of the peptides described below can be generated as a cassette including flanking BsmBI sites and 12 bp RSS sequences as described above for the RGD and GLP-1
25 sequences.

[00197] The canonical chemokine CXCL8 (IL-8) (Nucleotide Accession: NM_000584) is a member of the CXC chemokine family. This chemokine is one of the major mediators of the inflammatory response. This chemokine is secreted by several cell types. It functions as a chemoattractant, and is also a potent angiogenic factor. IL-8

binds with high affinity to two highly homologous chemokine receptors CXCR1 and CXCR2, which mediate pleiotropic responses including the onset of inflammation, angiogenesis, tumorigenesis and wound healing. The CXCR1 and CXCR2 receptors are GPCRs. The chemokines are folded into three anti-parallel β -sheets and a helix on the top, with an unstructured N-terminus containing the ELR triad, and the CXC motif which connects the ELR to the N-loop and the 30 s loop. On the basis of mutagenesis and structural studies of chemokines and their cognate receptors, a two-site model is postulated for the interactions of chemokines with their cognate receptors. Site 1 includes the receptor N-terminus, which recognizes the N-loop of chemokines, and site 2 includes extracellular loops of the receptor for binding to the N-terminus of chemokines to trigger receptor activation.

[00198] The amino acids involved in IL-8 binding to its receptor have been identified. Antibodies that neutralize IL-8 activity have been mapped to the ELRCXC sequences in the IL-8 protein. This Example utilizes sequences from this region of the IL-8 molecule as a peptide for grafting into an Ig scaffold to generate an anti-CXCR1 specific antibody. The use of sequences from the N-loop would also serve as potential sources for IL-8 based domain grafting. Other ligands that bind to CXCR1, such as Gro-alpha and LL-37, are also sources for peptide grafting. In some cases the sequences are engineered so that stop codons are introduced into the other reading frames.

[00199] The IL-8 nucleotide and amino acid sequences [SEQ ID NOs:46 and 47, respectively] are provided in Figure 5(A). The sequence that will be used for peptide grafting is shown in Figure 5(B) [SEQ ID NO:12].

[00200] CXCL1, Gro-alpha, is also a ligand for CXCR2. The binding domain has been identified as: ATELRCQCLQTLQGIHPKNIQSV [SEQ ID NO:48] (also shown in Figure 5(C)).

[00201] The sequence that will be used for peptide grafting is shown in Figure 5(D) [SEQ ID NO:49].

[00202] LL-37 has also been shown to bind to CXCR2. A silent substitution of (G>A) places a stop codon in reading frame #2:

CTGCTAGGTGATTTCTTCCGGAAATCTAAA [SEQ ID NO:50] (also see Figure 5(E)).

EXAMPLE 5: ErbB Receptor Peptides For Peptide Grafting

5 [00203] The source of sequences for peptide grafting is not limited to ligands. This Example identifies regions of the ErbB receptor (a member of the EGF receptor family) that are suitable for grafting into a protein scaffold using the methods of the present invention.

10 [00204] ErbB receptor extracellular regions all contain four distinct domains Domains I, II, III and IV. Domains II and IV are referred to as cysteine rich domains (CRI and CRII). During receptor dimerization all intermolecular contacts are mediated by the receptors. Current models for receptor dimerization suggest that ErbB family members adopt a tethered closed configuration in which loops from CRI and CRII form a structure in which the dimerization arm in CRI is occluded. Upon ligand binding the receptor forms an open configuration in which the dimerization arm in CRI is now
15 exposed and free to engage the dimerization loop on an adjacent receptor thus forming an ErbB dimer. ErbB2 does not have a ligand and its structure is found to be in an untethered form constitutively. Alignment of various ErbB sequences showing the dimerization arm are shown in Figure 6 [SEQ ID NOs:13-16].

20 [00205] As a way to target HER2 and EGFR receptors, receptor-receptor interacting domains can be used to direct the binding of a fusion protein. In the case of HER2, domains can be generated using the receptor dimerization arm. For EGFR receptor the dimerization arm as well as the inhibition domain in Cysteine Rich Domain 2 (CR2) of the Extracellular Domain (CRIIVc and CRIVd) can be used. Given that EGFR forms a closed “tethered” structure, the domains for the EGFR may only bind receptor when
25 the receptor is in the untethered form, which may allow the untethered form of EGFR to be specifically targeted, which in turn could provide the antibody with a unique clinically useful specificity in targeting diseases associated with over-expression of EGFR and its variants.

[00206] Sequences of peptides from ErbB2 selected for grafting into the vector described in Example 1 are provided below. The BsmBI sites and RSSs are not shown but would be generated as described in the previous Examples as gene synthesized cassettes for cloning into the appropriate acceptor vector.

5 **Peptide #1** (V981)

[00207] This peptide contains the ErbB2 dimerization arm. Stop codons are present in alternative forward frames. Stops in an inverted orientation are present in one out of the 3 frames.

[00208] Nucleotide Sequence:

10 TGCCAGCCCTGGTAACCTACAACACAGACACGTTTGAGTCCATGCCCAAT
CCCGAGGGCCGGTATAACATTCGGCGCCAGCTGT [SEQ ID NO:51]

[00209] Amino Acid Sequence:

CPALVTYNTDTFESMPNPEGRYTFGASC [SEQ ID NO:52]

Peptide #2 (V982)

15 [00210] This peptide contains the ErbB2 dimerization arm with additional endogenous sequences flanking the cysteine residues as a means to ensure that a larger portion of the fusion proteins will contain these residues. Stop codons are present in alternative forward frames. Stops in an inverted orientation are present in 1 out of the 3 frames.

[00211] Nucleotide Sequence:

20 CTGCACTGCCAGCCCTGGTAACCTACAACACAGACACGTTTGAGTCCATG
CCCAATCCCGAGGGCCGGTATAACATTCGGCGCCAGCTGTGTGACT [SEQ ID
NO:53]

[00212] Amino Acid Sequence:

LHCPALVTYNTDTFESMPNPEGRYTFGASCVT [SEQ ID NO:54]

25

Peptide #3 (V983)

[00213] This peptide contains the ErbB1 dimerization arm. Stop codons are present in one of the alternative forward frames. Stops in an inverted orientation are present in 2 out of the 3 frames.

5 [00214] Nucleotide Sequence:

TGCCCCCACTCATGCTCTACAACCCACCACGTACCAGATGGATGTGAAC
CCCGAGGGCAAATACAGCTTTGGTGCCACCTGC [SEQ ID NO:55]

[00215] Amino Acid Sequence:

CPPLMLYNPTTYQMDVNPEGKYSFGATC [SEQ ID NO:56]

10 **Peptide #4** (V984)

[00216] This peptide contains two loops from the CRII. These domains bind to the dimerization loop in CRI in the tethered configuration. Upon ligand binding, these loops potentially may form intermolecular contacts with the same loop in another untethered ErbB receptor. Fusion proteins containing this domain could thus
15 potentially bind to CRI or CRII domains of untethered ErbB receptors.

[00217] Stops are present in this peptide in alternative forward frames and in 2 of the 3 reverse frames.

[00218] Nucleotide Sequence:

TGTGCCCACTACATTGACGGCCCCCACTGCGTGAAGACCTGCCCGGCAGGA
20 GTCATGGGTGAAAACAACACCCTGGTCTGGAAGTACGCAGACGCCGGCCAT
GTGTGC [SEQ ID NO:57]

[00219] Amino Acid Sequence:

CAHYIDGPHCVKTCPAGVMGENNLTWVKYADAGHVC [SEQ ID NO:58]

25

Peptide #5 (V985)

[00220] This domain contains the 5' loop in CRII (first loop of peptide #4). As with peptide #4, this domain binds to the dimerization loop in CRI in the tethered configuration and fusion proteins containing the domain represented by peptide #5 could thus also potentially bind to CRI or CRII domains of untethered ErbB receptors.

[00221] Nucleotide Sequence:

TGTGCCCACTACATTGACGGCCCCCACTGC [SEQ ID NO:59]

[00222] Amino Acid Sequence:

CAHYIDGPHC [SEQ ID NO:60]

10 **Peptide #6** (V986)

[00223] This domain contains the 3' loop in CRII (second loop of peptide #4). As with peptide #4, this domain binds to the dimerization loop in CRI in the tethered configuration and fusion proteins containing the domain represented by peptide #6 could thus also potentially bind to CRI or CRII domains of untethered ErbB receptors.

15 [00224] Nucleotide Sequence:

TGCCCGGCAGGAGTCATGGGAGAAAATAACACCCTAGTCTGGAAGTACGC
AGACGCCGCCATGTGTGC [SEQ ID NO:61]

[00225] Amino Acid Sequence:

CPAGVMGENNTLVWKYADAGHVC [SEQ ID NO:62]

20 **EXAMPLE 6: Preparation Of Constructs For Introducing Sequence Diversity Into A Fibronectin Domain**

[00226] The methods of the present invention can also be used to graft peptides into non-Ig scaffolds or to join two non-Ig sequences or a non-Ig sequence and an Ig sequence. This Example demonstrates the principle of using V(D)J in a non-Ig context.
25 The same peptides identified in the preceding Examples 2-5 can be introduced into the

10Fn3 loop. The same method described below is used for Ig and non-Ig peptide grafting using V(D)J. An exemplary acceptor vector for grafting peptides into the 10Fn3 loop is shown in Figure 7. Both this vector and the acceptor vector shown in Figure 1 allow easy manipulation of the flanking sequences of the 12 RSSs. Similar
5 vectors can be constructed to allow easy manipulation of the flanking sequences of the 23 bp RSSs.

EXAMPLE 7: Recombination And Expression Of A Recombination Substrate

[00227] In brief, HEK293 cells, containing an integrated LoxP sequence (Fukushige *et al.*, 1992, *PNAS USA*, 89:7905-7909; Baubonis *et al.*, 1993, *NAR*, 21(9):2025-2029; Thomson *et al.*, 2003, *Genesis*, 36:162-167) were maintained in DMEM media with 10% FBS. Integration into the LoxP site was shown to support high protein expression and also support V(D)J recombination of inserted substrates and provides an easy method to generate integrants with the required properties. Vectors comprising the recombination substrate were designed to include a LoxP site for targeted integration
10 which is in-frame with a codon-optimized hygromycin open reading frame. Bipartite vectors were also designed so that productive rearrangements will be in-frame with the selectable marker neomycin. The neomycin gene is cloned in-frame with a transmembrane domain, both of which are downstream of a furin cleavage site that allows for secretion of the encoded protein (see Figure 9 and SEQ ID NO:64, as an
15 example).

[00228] For example, for bipartite substrates, HEK293 cells containing the LoxP site were co-transfected with the bipartite substrate containing the hygromycin gene for selection of stable integrants and a vector expressing the CRE protein at a ratio of 10:1 substrate to CRE expressing vector. Specifically, a 10 cm dish of cells was transfected
25 using a polyethylenimine (PEI; 1mg/ml) to DNA ratio of 3:1. 21.6ug of substrate DNA was mixed with 2.4ug of CRE expression vector and placed in 1.5 ml OptiMEM™ media and mixed with an equal volume of OptiMEM™ containing the 72 ul of PEI. The transfection was carried out for 24 hours and the following day the transfection media was removed and replaced with fresh DMEM media. The following
30 day the transfected cells were split into ten 10 cm² dishes and selection was carried out for approximately 2 weeks. A pool of stable hygromycin resistant cells were selected.

The cell line was subsequently expanded in the un-recombined state to approximately 10 million cells and transfected with RAG-1, RAG-2 and TdT. 72 hours post-transfection the cells were placed in neomycin selection (1mg/ml).

5 [00229] Tripartite recombination substrates used vectors designed such that puromycin could be used for in-frame selection. Tripartite vectors also included a modified neomycin cassette that allows for maintenance of the unrecombined substrate during expansion.

EXAMPLE 8: Preparation of Immunoglobulins Grafted with an Anti-TPO Receptor Peptide

10 [00230] Full length human IgGs comprising a peptide targeted to the TPO receptor (shown below) were prepared using a recombination substrate comprising the cassette shown in Figure 11A, which includes 5' and 3' flanking sequences that include degenerative nucleotide combinations to generate diversity at the ends in addition to the diversity that is generated via V(D)J recombination. The flanking VH gene segment in
15 the recombination substrate was VH1-69 and the flanking JH segment was JH6.

[00231] Anti-TPO receptor peptide: IEGPTLRQWLAARA [SEQ ID NO:71]

[00232] Cells incorporating a tripartite V(D)J recombination substrate containing the recombination cassette shown in Figure 11A were cultured and induced to generate peptide grafted variants as described in the preceding Examples. V(D)J recombined
20 cells expressing the peptide-grafted antibodies were incubated with biotinylated TPO receptor (R&D Systems). The cells were subsequently stained with Streptavidin-PE to identify cells that had successfully bound the receptor and FACS sorted. cDNA from the FACS sorted cells were subsequently cloned into pcDNA to express the novel antibody as a full length IgG containing a transmembrane domain.

25 [00233] Amino acid and nucleotide sequences of the peptide and flanking sequences, for exemplary peptide-grafted immunoglobulins isolated by this approach that were shown to bind to the TPO receptor are shown in Tables 3 and 4 below. In each case, the VH and JH segments utilized were IGHV1-69*01 and IGHJ6, respectively.

	GGCACCTTCAGCAGCTACGCCATCAGCTGGGTCCGCCAG GCTCCTGGACAGGGACTGGAATGGATGGGCGGCATC CCCATCTTCGGCACCCGCCAACTACGCCAGAAATTCCAG GGCAGAGTGACCATCACCGCCGACGAGAGCACCAGCACC GCCTACATGGAAGTGAAGCTTTCGAAGCGAGGACACC GCTGTGTATTACTGTGCGGGGCTCTGCGGATCGAGGGC CCTACCTGAGACAGTGGCTGGCCGCTAGAGCTTCCCTAT ACTACATGGACGTCTGGGGCAAAGGGACCACGGTCACCG TGTCTCAG	
Anti-TpoR 4	CAGGTGCAGCTGGTGCAGTCTGGCGCCGAAGTGAAGAAA CCCAGCAGCAGCGTGAAGGTGTCCTGCAAGGCCAGCGGC GGCACCTTCAGCAGCTACGCCATCAGCTGGGTCCGCCAG GCTCCTGGACAGGGACTGGAATGGATGGGCGGCATC CCCATCTTCGGCACCCGCCAACTACGCCAGAAATTCCAG GGCAGAGTGACCATCACCGCCGACGAGAGCACCAGCACC GCCTACATGGAAGTGAAGCTTTCGAAGCGAGGACACC GCTGTGTATTACTGTGCGAGAGCAAAGAGTAGTCAGATC GAGGGCCCTACCCTGAGACAGTGGCTGGCCGCTAGAGCT GAGCTGAGGCTGCAACTACTACATGGACGTCTGGGGC AAAGGGACCACGGTCACCGTGTCTCAG	82
Anti-TpoR 5	CAGGTGCAGCTGGTGCAGTCTGGCGCCGAAGTGAAGAAA CCCAGCAGCAGCGTGAAGGTGTCCTGCAAGGCCAGCGGC GGCACCTTCAGCAGCTACGCCATCAGCTGGGTCCGCCAG GCTCCTGGACAGGGACTGGAATGGATGGGCGGCATC CCCATCTTCGGCACCCGCCAAATACGCCAGAAATTCCAG GGCAGAGTGACCATCACCGCCGACGAGAGCACCAGCACC GCCTACATGGAAGTGAAGCTTTCGAAGCGAGGACACC GCTGTGTATTACTGTGCGAGGAAGCCGCAGGCTATCGAG GGCCCTACCCTGAGACAGTGGCTGGCCGCTAGAGCTCTG GGAACTACTACATGGACGTCTGGGGCAAAGGGACCACG GTCACCGTGTCTCAG	83
Anti-TpoR 6	CAGGTGCAGCTGGTGCAGTCTGGCGCCGAAGTGAAGAAA CCCAGCAGCAGCGTGAAGGTGTCCTGCAAGGCCAGCGGC GGCACCTTCAGCAGCTACGCCATCAGCTGGGTCCGCCAG GCTCCTGGACAGGGACTGGAATGGATGGGCGGCATC CCCATCTTCGGCACCCGCCAACTACGCCAGAAATTCCAG GGCAGAGTGACCATCACCGCCGACGAGAGCACCAGCACC GCCTACATGGAAGTGAAGCTTTCGAAGCGAGGACACC GCTGTGTATTACTGTGCGAGAACGTTGCGTATCGAGGGC CCTACCTGAGACAGTGGCTGGCCGCTAGAGCTCCGGCG GCCTACTACTACTACTACATGGACGTCTGGGGCAAAGGG ACCACGGTCACCGTGTCTCAG	84
Anti-TpoR 7	CAGGTGCAGCTGGTGCAGTCTGGCGCCGAAGTGAAGAAA CCCAGCAGCAGCGTGAAGGTGTCCTGCAAGGCCAGCGGC GGCACCTTCAGCAGCTACGCCATCAGCTGGGTCCGCCAG GCTCCTGGACAGGGACTGGAATGGATGGGCGGCATC CCCATCTTCGGCACCCGCCAACTACGCCAGAAATTCCAG GGCAGAGTGACCATCACCGCCGACGAGAGCACCAGCACC GCCTACATGGAAGTGAAGCTTTCGAAGCGAGGACACC GCTGTGTATTACTGTGCGAGTTGTAGGTCTATCGAGGGC CTACCCTGAGACAGTGGCTGGCCGCTCGAGCTTGTCTGG	85

	ATCTGATCGGGTACTACTACATGGACGTCTGGGGCAAAG GGACCACGGTCACCGTGTCTCAG	
--	--------------------------------------------------------------------	--

EXAMPLE 9: Binding of Immunoglobulins Grafted with an Anti-TPO Receptor Peptide to the TPO Receptor

[00234] HEK-293 cells were transfected with 4 peptide grafted antibody clones
5 isolated as described in Example 8 or a negative control antibody. The expression
vector expresses the IgGs as full length IgG1 with a transmembrane domain which
links the antibody to the cell surface. The clones used were Anti-TpoR 1, 3, 4 and 5
from Table 3 (clones 1-4, respectively).

[00235] 24hrs post transfection, cells were trypsinized and incubated with 1ug/ml
10 biotinylated soluble TPOR (R&D Systems) or 1ug/ml irrelevant biotinylated protein
diluted in PBS+2% FBS. Following a 1hr incubation, cells were pelleted, staining
media was aspirated and the cells resuspended into 1ug/ml R-Phycoerythrin labelled
Streptavidin and 1ug/ml Gt anti Human IgG Fc Alexa 647 conjugated antibody
(Jackson Laboratories) + 7AAD diluted into PBS+ 2% FBS and then incubated for an
15 additional 1hr.

[00236] Following the incubation cells were pelleted, staining media was aspirated, the
cells resuspended into PBS+2% FBS and then analyzed by flow cytometry. The results
are shown in Figure 12A and show that all four of the peptide-grafted IgGs bound the
TPO receptor and that the binding activity of the grafted peptide was affected by both
20 the composition and the length of the flanking sequences.

EXAMPLE 10: Preparation of Immunoglobulins Grafted with Anti-GLP-1 Receptor Peptides

[00237] Full length human IgGs comprising a peptide targeted to the GLP-1 receptor
were prepared using a recombination substrate incorporating the cassette shown in
25 Figure 11B, which includes 5' and 3' flanking sequences that include degenerative
nucleotide combinations to generate diversity at the ends in addition to the diversity
that is generated via V(D)J recombination. The peptide encoding sequences used in the
cassette are shown in Figure 11C [SEQ ID NOs:66-70]. These sequences encoded

peptides derived from GLP-1 or exendin-4 as indicated. The flanking VH gene segments in the recombination substrate were VH1-2*01, VH1-3*01, VH1-8*01, VH1-18*01, VH1-46*01, VH1-24*01, VH1-45*01, VH1-58*01, VH1-69*01, VH1-f*01, VH2-5*01, VH2-26*01, VH2-70*01, VH3-7*01, VH3-11*01, VH3-21*01, VH3-23*01, VH3-30*01, VH3-33*01, VH3-48*01, VH3-53*01, VH3-9*01, VH3-13*01, 5 VH3-20*01, VH3-43*01, VH3-64*01, VH3-66*01, VH3-74*01, VH3-d*01, VH3-15*01, VH3-49*01, VH3-72*01, VH3-73*01, VH4-30-1*01, VH4-30-4*01, VH4-31*01, VH4-39*01, VH4-59*01, VH4-61*01, VH4-4*02, VH4-28*01, VH4-30-2*01, VH4-34*01, VH4-b*01, VH5-51*01, VH5-a*01, VH6-1*01, VH7-4-1*01 and the 10 flanking JH segments used were JH1, JH2, JH3, JH4, JH5, JH6.

[00238] Cells incorporating a tripartite V(D)J recombination substrate containing the cassette shown in Figure 11B were cultured and induced to generate peptide grafted variants. V(D)J recombined cells expressing the peptide-grafted antibodies were incubated with FLAG(DDK) tagged soluble GLP-1 receptor. The cells were 15 subsequently stained with biotinylated mouse anti-FLAG antibody (Sigma Aldrich) and Streptavidin-PE to identify cells that had successfully bound the receptor and FACS sorted. cDNA from the FACS sorted cells were subsequently cloned into pcDNA to express the novel antibody as a full length IgG that would be secreted into the supernatant.

20 **[00239]** Greater than 100 anti-GLP1R binding variants were generated using this approach. Amino acid and nucleotide sequences of the peptide and flanking sequences for exemplary peptide-grafted immunoglobulins comprising an exendin-4 peptide isolated by this approach that were shown to bind to the GLP-1 receptor are shown in Tables 5A and 6 below. The VH and JH segments utilized in each case are shown in 25 Table 5B.

Table 5A: Amino Acid Sequences of Anti-GLP-1 Receptor Exendin-4 Peptides and Flanking Sequences (Anti-GLPR 1-9) [SEQ ID NOs:72-80]

	CDR3 AA Sequence			SEQ ID NO
	5' FLANK	Exendin sequences incorporated	3' FLANK	
Anti GLP1R 1	ARRASSV	HGEGTFTSDLSKQMEEEEAVRLFIEWLK NGGPSSGAPPPS	DGSWFDP	72
Anti GLP1R 2	AST	HGEGTFTSDLSKQMEEEEAVRLFIEWLK NGGPSSGAPPPS	DWYFDL	73
Anti GLP1R 3	A	HGEGTFTSDLSKQMEEEEAVRLFIEWLK NGGPSSGAPPPS	SMLDAFDI	74
Anti GLP1R 4	AREL	HGEGTFTSDLSKQMEEEEAVRLFIEWLK NGGPSSGAPPPS	DDAFDI	75
Anti GLP1R 5	TSF	HGEGTFTSDLSKQMEEEEAVRLFIEWLK NGGPSSGAPPPS	QTLEYFQH	76
Anti GLP1R 6	ARDG	HGEGTFTSDLSKQMEEEEAVRLFIEWLK NGGPSSGAPPPS	WWPPDAFDI	77
Anti GLP1R 7	ATD	- GEGTFTSDLSKQMEEEEAVRLFIEWLKN GGPSSGAPPPS	MYDAFDI	78
Anti GLP1R 8	ARL	HGEGTFTSDLSKQMEEEEAVRLFIEWLK NGGPSSGAPPPS	LANNWFDP	79
Anti GLP1R 9	ARDH	----- LSKQMEEEEAVRLFIEWLKNGGPSSGAP PPS	YWYFDL	80

Table 5B: VH and JH Segments Utilized in Anti GLPR 1-9

	VH Segment	JH Segment
Anti GLP1R 1	IGHV1-69*01	IGHJ5
Anti GLP1R 2	IGHV1-69*01	IGHJ2
Anti GLP1R 3	IGHV3-64*01	IGHJ3
Anti GLP1R 4	IGHV3-64*01	IGHJ3

Anti GLP1R 5	IGHV3-73*01	IGHJ1
Anti GLP1R 6	IGHV1-69*01	IGHJ3
Anti GLP1R 7	IGHV1-f*01	IGHJ3
Anti GLP1R 8	IGHV1-69*01	IGHJ5
Anti GLP1R 9	IGHV3-20*01	IGHJ2

Table 6: Nucleotide Sequences of Anti-GLP-1 Receptor Exendin-4 Peptides and Flanking Sequences [SEQ ID NOs:86-94]

	Variable Gene Nucleotide Sequences (Heavy Chain)	SEQ ID NO
Anti GLP1R 1	CAGGTGCAGCTGGTGCAGTCTGGCGCCGAAGTGAAGAAACCCG GCAGCAGCGTGAAGGTGTCCTGCAAGGCCAGCGGCGGCACCTT CAGCAGCTACGCCATCAGCTGGGTCCGCCAGGCTCCTGGACAG GGACTGGAATGGATGGGCGGCATCATCCCCATCTTCGGCACCG CCAACTACGCCCAGAAATTCCAGGGCAGAGTGACCATCACCGC CGACGAGAGCACCAGCACCGCCTACATGGAAGTGCAGCAGCCTT CGAAGCGAGGACACCGCTGTGTATTACTGTGCGAGACGGGCCT CTAGTGTGCATGGCGAGGGCACCTTCACCTCCGACCTGTCCAAA CAAATGGAAGAAGAAGCCGTCCGGCTGTTTCATCGAATGGCTGA AAAATGGCGGCCCTTCTCTGGCGCCCCCTCCTTCTGATGGA AGCTGGTTCGACCCCTGGGGCCAGGGAACCCCTGGTCACCGTGT CTCAG	86
Anti GLP1R 2	CAGGTGCAGCTGGTGCAGTCTGGCGCCGAAGTGAAGAAACCCG GCAGCAGCGTGAAGGTGTCCTGCAAGGCCAGCGGCGGCACCTT CAGCAGCTACGCCATCAGCTGGGTCCGCCAGGCTCCTGGACAG GGACTGGAATGGATGGGCGGCATCATCCCCATCTTCGGCACCG CCAACTACGCCCAGAAATTCCAGGGCAGAGTGACCATCACCGC CGACGAGAGCACCAGCACCGCCTACATGGAAGTGCAGCAGCCTT CGAAGCGAGGACACCGCTGTGTATTACTGTGCGAGTACTCATG GCGAGGGCACCTTCACCTCCGACCTGTCCAAACAAATGGAAGA AGAAGCCGTCCGGCTGTTTCATCGAATGGCTGAAAAATGGCGGC CCTTCTCTGGCGCCCCCTCCTCCTTCTGACTGGTACTTTCGATCTC TGGGGCCCGTGGCACCCCTGGTCACTGTGTCCTCAG	87
Anti GLP1R 3	GAAGTGCAGCTGGTGGAAAGCGGCGGAGGCCTGGTGCAGCCTG GCGGCAGCCTGAGACTGTCTTGCGCCGCCAGCGGCTTCACCTTC AGCAGCTACGCCATGCACTGGGTCCGCCAGGCCCCCTGGCAAGG GACTGGAATACGTGTCCGCCATCAGCTCGAACGGCGGCAGCAC CTACTACGCCAACAGCGTGAAGGGCCGGTTCACCATCAGCCGG GACAACGCCAAGAACAGCCTGTACCTGCAGATGAACAGCCTGC GGGCCGAGGACACCGCCGTGTATTACTGTGCGCATGGCGAGGG CACCTTCACCTCCGACCTGTCCAAACAAATGGAAGAAGAAGCC GTCCGGCTGTTTCATCGAATGGCTGAAAAATGGCGGCCCTTCTC TGGCGCCCCCTCCTTCTAGTATGCTCGATGCTTTTGATATCTG GGGCCAAGGGACAATGGTCACCGTGTCTCCTCAG	88

	Variable Gene Nucleotide Sequences (Heavy Chain)	SEQ ID NO
Anti GLP1R 4	GAAGTGCAGCTGGTGGAAAGCGGCGGAGGCCTGGTGCAGCCTG GCGGCAGCCTGAGACTGTCTTGCGCCGCCAGCGGCTTCACCTTC AGCAGCTACGCCATGCACTGGGTCCGCCAGGCCCTGGCAAGG GACTGGAATACGTGTCCGCCATCAGCTCGAACGGCGGCAGCAC CTACTACGCCAACAGCGTGAAGGGCCGGTTCACCATCAGCCGG GACAACAGCAAGAACACCCTGTACCTGCAGATGGGCAGCCTGC GGGCCGAGGATATGGCCGTGTATTACTGTGCGAGAGAGCTGCA TGCGGAGGGCACCTTCACCTCCGACCTGTCCAAACAAATGGAA GAAGAAGCCGTCCGGCTGTTTCATCGAATGGCTGAAAAATGGCG GCCCTTCCTCTGGCGCCCCTCCTCCTTCTGATGATGCTTTTGATA TCTGGGGCCAAGGGACAATGGTCACCGTGTCTCAG	89
Anti GLP1R 5	GAAGTGCAGCTGGTGGAAAGCGGCGGAGGCCTGGTGCAGCCTG GCGGCAGCCTGAAACTGAGCTGCGGCCGCCAGCGGCTTCACCTTT AGCGGCAGCGCCATGCACTGGGTCCGCCAGGCCTCTGGCAAGG GACTGGAATGGGTTCGGACGGATTTCGAAGCAAGGCCAACAGCTA CGCCACCGCCTACGCCGCCTCCGTGAAGGGCCGGTTCACCATCA GCCGGGACGACAGCAAGAACACCGCCTACCTGCAGATGAACAG CCTGAAAACCGAGGACACCGCCGTGTATTACTGTACTAGTTTTTC ATGGCGAGGGCACCTTCACCTCCGACCTGTCCAAACAAATGGA AGAAGAAGCCGTCCGGCTGTTTCATCGAATGGCTGAAAAATGGC GGCCCTTCCTCTGGCGCCCCTCCTCCTTCTCAGACGCTGGAATA CTCCAGCACTGGGGCCAGGGCACCTGGTCACCGTGTCTCAG	90
Anti GLP1R 6	CAGGTGCAGCTGGTGCAGTCTGGCGCCGAAGTGAAGAAACCCG GCAGCAGCGTGAAGGTGTCTTGCAAGGCCAGCGGCGGCACCTT CAGCAGCTACGCCATCAGCTGGGTCCGCCAGGCTCCTGGACAG GGACTGGAATGGATGGGCGGCATCATCCCCATCTTCGGCACCG CCAATACTACGCCAGAAATTCCAGGGCAGAGTGACCATCACCGC CGACGAGAGCACCAGCACCGCCTACATGGAAGTGCAGCAGCCTT CGAAGCGAGGACACCGCTGTGTATTACTGTGCGAGAGATGGTC ATGGCGAGGGCACCTTCACCTCCGACCTGTCCAAACAAATGGA AGAAGAAGCCGTCCGGCTGTTTCATCGAATGGCTGAAAAATGGC GGCCCTTCCTCTGGCGCCCCTCCTCCTTCTTGGTGGCCACCCGAT GCTTTTGATATCTGGGGCCAAGGGACAATGGTCACCGTGTCTC AG	91
Anti GLP1R 7	GAAGTGCAGCTGGTGCAGTCTGGCGCCGAAGTGAAGAAACCTG GCGCCACCGTGAAGATCAGCTGCAAGGTGTCCGGCTACACCTTC ACCGACTACTACATGCACTGGGTGCAGCAGGCCCTGGCAAGG GACTGGAATGGATGGGCTGGTTCGATCCCGAGGACGGCGAGAC AATCTACGCCGAGAAGTTCCAGGGCAGAGTGACCATCACCGCC GACACCAGCACCGACACCGCCTACATGGAAGTGCAGCAGCCTGC GGAGCGAGGACACCGCTGTGTATTACTGTGCAACAGATGGCGA GGGCACCTTCACCTCCGACCTGTCCAAACAAATGGAAGAAGAA GCCGTCCGGCTGTTTCATCGAATGGCTGAAAAATGGCGGCCCTTC CTCTGGCGCCCCTCCTCCTTCTATGTATGATGCTTTTGATATCTG GGGCCAAGGGACAATGGTCACCGTGTCTCAG	92
Anti GLP1R 8	CAGGTGCAGCTGGTGCAGTCTGGCGCCGAAGTGAAGAAACCCG GCAGCAGCGTGAAGGTGTCTTGCAAGGCCAGCGGCGGCACCTT	93

	Variable Gene Nucleotide Sequences (Heavy Chain)	SEQ ID NO
	CAGCAGCTACGCCATCAGCTGGGTCCGCCAGGCTCCTGGACAG GGACTGGAATGGATGGGCGGCATCATCCCCATCTTCGGCACCG CCA ACTACGCCCAGAAATTCCAGGGCAGAGTGACCATCACCGC CGACGAGAGCACCAGCACCGCCTACATGGA ACTGAGCAGCCTT CGAAGCGAGGACACCGCTGTGTATTACTGTGCGCGGCTTCATGG CGAGGGCACCTTCACCTCCGACCTGTCCAAACAAATGGAAGAA GAAGCCGTCCGGCTGTTCATCGAATGGCTGAAAAATGGCGGCC CTTCCTCTGGCGCCCCTCCTCCTTCTTTGGCGAACAACTGGTTCCG ACCCCTGGGGCCAGGGAACCCTGGTCACCGTGTCTCAG	
Anti GLP1R 9	GAAGTGCAGCTGGTGGAAAGCGGAGGCGGAGTGGTTTCGACCTG GCGGAAGCCTGAGACTGTCTTGCGCCGCCAGCGGCTTCACCTTT GACGACTACGGCATGAGCTGGGTCCGCCAGGCCCTGGCAAGG GACTGGAATGGGTGTCCGGCATCAACTGGAACGGCGGCAGCAC CGGCTACGCCGACAGCGTGAAGGGCCGGTTCACCATCAGCCGG GACAACGCCAAGAACAGCCTGTACCTGCAGATGAACAGCCTGC GGGCCGAGGACACCGCCTTGTATCACTGTGCGAGAGATCACCT GTCCAAACAAATGGAAGAAGAAGCCGTCCGGCTGTTTCATCGAA TGGCTGAAAAATGGCGGCCCTCCTCTGGCGCCCCTCCTCCTTC TTACTGGTACTTCGATCTCTGGGGCCGTGGCACCCCTGGTCACTG TGTCTCAG	94

EXAMPLE 11: Binding of Immunoglobulins Grafted with an Anti-GLP-1 Receptor Peptide to Natively Expressed Full-Length GLP-1 Receptor

[00240] To generate soluble peptide-grafted antibody for analysis, HEK-293 cells were
 5 transfected with 4 peptide grafted antibody clones isolated as described in Example 10.
 All four antibodies were derived from a recombination substrate using Exendin-4 1-39
 (see Figure 11C). The clones used were Anti-GLP1R 1, 6, 7 and 9 from Table 5
 (clones 1-4, respectively). The negative control in this experiment was an irrelevant
 antibody. 24hrs post transfection supernatants from the transfected cells were
 10 harvested and spun down at 14000rpm for 5min to remove cell debris.

[00241] In order to generate target cells expressing the native GLP-1 receptor, HEK-
 293 cells were transfected with an expression construct which expresses full length
 GLP1R or mock transfected to serve as a negative control. 24hrs later both the mock
 transfected and GLP1R transfected cells were trypsinized and incubated with the 250ul
 15 supernatants isolated above. 1ul of Gt anti-Human IgG R-Phycoerythrin conjugated
 antibody (Jackson Laboratories, 1mg/ml Stock Solution) was added to the receptor

transfected cells/supernatant mixture. After a 1hr incubation cells were spun down, staining solution was aspirated, the cells resuspended into PBS+2% FBS+1ug/ml 7AAD and analyzed by flow cytometry. The results are shown in Figure 12B and show that that all four of the peptide-grafted IgGs bound the natively expressed full-length
5 GLP-1 receptor.

[00242] The disclosures of all patents, patent applications, publications and database entries referenced in this specification are hereby specifically incorporated by reference in their entirety to the same extent as if each such individual patent, patent application, publication and database entry were specifically and individually indicated to be
10 incorporated by reference.

[00243] Although the invention has been described with reference to certain specific embodiments, various modifications thereof will be apparent to those skilled in the art without departing from the spirit and scope of the invention. All such modifications as would be apparent to one skilled in the art are intended to be included within the scope
15 of the following claims.

THE EMBODIMENTS OF THE INVENTION IN WHICH AN EXCLUSIVE PROPERTY OR PRIVILEGE IS CLAIMED ARE DEFINED AS FOLLOWS:

1. A method of generating variants of a fusion protein comprising the steps of:
 - (i) providing a first nucleic acid sequence comprising a first coding sequence encoding a first portion of the fusion protein and further comprising a first recombination signal sequence (RSS);
 - (ii) providing a second nucleic acid sequence comprising a second coding sequence encoding a second portion of the fusion protein and further comprising a second RSS capable of functional recombination with the first RSS;
 - (iii) introducing the first and second nucleic acid sequence into a recombination-competent host cell, and
 - (iv) culturing the host cell *in vitro* under conditions allowing (a) recombination of the first and second RSS to generate a chimeric polynucleotide comprising the first and second coding sequences and (b) expression of the chimeric polynucleotide, thereby generating variants of the fusion protein.
2. The method according to claim 1, wherein the first and second nucleic acids are comprised by a single polynucleotide.
3. The method according to claim 2, wherein the first and second nucleic acid sequences further comprise a nucleotide sequence encoding a polypeptide linker and the first and second RSSs are positioned within or proximal to the nucleotide sequence encoding the polypeptide linker, whereby recombination in step (iv) introduces sequence diversity into the nucleotide sequence encoding the polypeptide linker.
4. The method according to claim 1 or 2, wherein the second nucleic acid sequence further comprises a third RSS and the method further comprises the steps of:
 - providing a third nucleic acid sequence comprising a third coding sequence encoding a third portion of the fusion protein and further comprising a fourth RSS capable of functional recombination with the third RSS, and

- introducing the third nucleic acid sequence into the recombination-competent host cell,

wherein culturing the host cell in step (iv) further allows for recombination of the third and fourth RSS and the chimeric polynucleotide comprises the first, second and third coding sequences.

5. The method according to claim 4, wherein the first, second and third nucleic acid sequences are comprised by a single polynucleotide.

6. The method according to any one of claims 1, 2 or 3, wherein the first and second portions of the fusion protein are derived from different proteins.

7. The method according to any one of claims 1, 2 or 3, wherein the first and second portions of the fusion protein are derived from different domains of the same protein.

8. The method according to claim 4 or 5, wherein the first, second and third portions of the fusion protein are derived from different proteins.

9. The method according to claim 4 or 5, wherein the first and third portions of the fusion protein are derived from different domains of the same protein.

10. The method according to any one of claims 4, 5 or 9, wherein the first and third portions of the fusion protein are derived from an antibody.

11. The method according to claim 10, wherein the first and third portions of the fusion protein are derived from the heavy chain of the antibody.

12. The method according to claim 10, wherein the first and third portions of the fusion protein are derived from the light chain of the antibody.

13. The method according to any one of claims 10, 11 or 12, wherein the first and fourth RSS are positioned within or proximal to a region encoding a CDR3 of the antibody.

14. The method according to any one of claims 1 to 13, wherein one or more of the nucleic acid sequences further comprises flanking sequences adjacent to the respective RSS.
15. The method according to any one of claims 1 to 14, wherein one or more of the nucleic acid sequences further comprises additional coding sequences that encode a membrane anchor domain peptide.
16. The method according to any one of claims 1 to 15 further comprising the step of screening the fusion protein variants for a desired function.
17. The method according to any one of claims 1 to 16, wherein the recombination-competent host cell is capable of expressing TdT.
18. The method according to any one of claims 1 to 17, wherein the polynucleotide is stably integrated into the genome of the host cell.
19. The method according to claim 18, wherein the polynucleotide is stably integrated into the genome of the host cell as a single copy.
20. A polynucleotide comprising a first nucleic acid sequence comprising a first coding sequence encoding a first portion of a fusion protein and further comprising a first recombination signal sequence (RSS) and a second nucleic acid sequence comprising a second coding sequence encoding a second portion of the fusion protein and further comprising a second RSS capable of functional recombination with the first RSS.
21. The polynucleotide according to claim 20, wherein the first and second nucleic acid sequences further comprise a nucleotide sequence encoding a polypeptide linker and the first and second RSSs are positioned within or proximal to the nucleotide sequence encoding the polypeptide linker.
22. The polynucleotide according to claim 20, wherein the second nucleic acid sequence further comprises a third RSS and the polynucleotide further comprises a third nucleic acid sequence comprising a third coding sequence encoding a third portion

of the fusion protein and further comprising a fourth RSS capable of functional recombination with the third RSS.

23. The polynucleotide according to claim 20 or 21, wherein the first and second portion of the fusion protein are derived from different proteins.

24. The polynucleotide according to claim 20 or 21, wherein the first and second portion of the fusion protein are derived from different domains of the same protein.

25. The polynucleotide according to claim 22, wherein the first, second and third portion of the fusion protein are derived from different proteins.

26. The polynucleotide according to claim 22, wherein the first and third portion of the fusion protein are derived from different domains of the same protein.

27. The polynucleotide according to claim 22 or 26, wherein the first and third portion of the fusion protein are derived from an antibody.

28. The polynucleotide according to claim 27, wherein the first and third portions of the fusion protein are derived from the heavy chain of the antibody.

29. The polynucleotide according to claim 27, wherein the first and third portions of the fusion protein are derived from the light chain of the antibody.

30. The polynucleotide according to any one of claims 27, 28 or 29, wherein the first and fourth RSS are positioned within or proximal to a region encoding a CDR3 of the antibody.

31. The polynucleotide according to any one of claims 20 to 30, wherein one or more of the nucleic acid sequences further comprises flanking sequences adjacent to the respective RSS.

32. The polynucleotide according to any one of claims 20 to 31, wherein the polynucleotide further comprises additional coding sequences that encode a membrane anchor domain peptide.

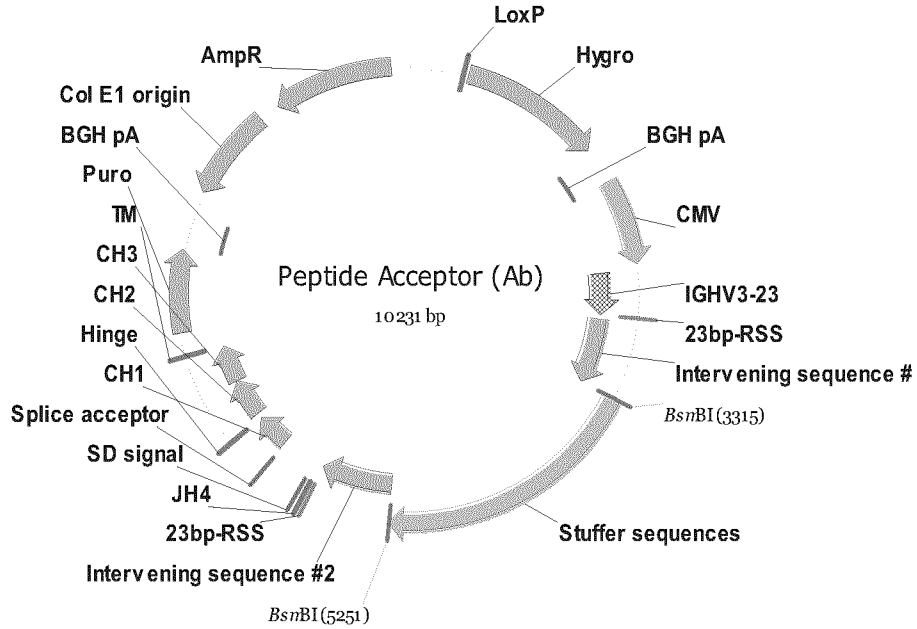
33. An isolated host cell comprising the polynucleotide of any one of claims 20 to 32.
34. The isolated host cell according to claim 33, wherein the host cell is recombination-competent.
35. A variant fusion protein produced by the method according to any one of claims 1 to 19.
36. A peptide-grafted immunoglobulin comprising an immunoglobulin scaffold and a heterologous polypeptide inserted into, or replacing, at least one CDR of the immunoglobulin scaffold, wherein the heterologous polypeptide comprises a peptide sequence capable of binding to a GPCR, an upstream flanking sequence comprising between about 1 and about 20 amino acids and a downstream flanking sequence comprising between about 1 and about 20 amino acids.
37. The peptide-grafted immunoglobulin according to claim 36, wherein the immunoglobulin scaffold is a full-length IgG.
38. The peptide-grafted immunoglobulin according to claim 37, wherein the immunoglobulin scaffold is a full-length human IgG.
39. The peptide-grafted immunoglobulin according to any one of claims 36, 37 or 38, wherein the heterologous polypeptide is inserted into, or replaces, at least one CDR3 of the immunoglobulin scaffold.
40. The peptide-grafted immunoglobulin according to claim 39, wherein the heterologous polypeptide replaces the D segment in the at least one CDR3.
41. The peptide-grafted immunoglobulin according to any one of claims 36 to 40, wherein the peptide sequence is derived from GLP-1 or exendin-4.
42. The peptide-grafted immunoglobulin according to any one of claims 36 to 40, wherein the peptide sequence comprises the sequence as set forth in SEQ ID NO:38 or 43, or a fragment thereof.

43. The peptide-grafted immunoglobulin according to any one of claims 36 to 40, wherein the heterologous polypeptide has a sequence as set forth in any one of SEQ ID NOs:73-80.

44. The peptide-grafted immunoglobulin according to any one of claims 36 to 40, wherein the peptide sequence is derived from TPO.

45. The peptide-grafted immunoglobulin according to any one of claims 36 to 40, wherein the peptide sequence comprises the sequence as set forth in SEQ ID NO:71, or a fragment thereof.

A.



B.

ctaaattgtaagcgftaatatfttgttaaaatfcgcgttaatfttgttaaatcagctcatttttaaccaataggccgaaatcggcaaaatccc
 ttataaatcaaaagaatagaccgagatagggttgagtggccgctacagggcgctcccattcgccattcaggctgcgcaactgttggga
 agggcggttcgggtcgggcctcttcgctattacgccagctggcgaaaaggggatgtgctgcaaggcgatfaagtgggtaacgccag
 ggtttccagtcacgacggtgtaaaacgacggccagtgaagcgcgacgtaatacgaactcactataggcgcaattggcggaggccgt
 caaggcctaggcgcgcctgaataactcgtatagcatacattatagcaattatcgaaaaagcctgaactcaccgcgacatccgtggag
 aaattcctcatgaaaaatcgactccgtgtccgatctcatgcagctgtccagggcgaggagagtagagcattctcattcgatgtggg
 cgggagaggctacgtgctgagagtgaactctgtgccgacggcttctacaaggaccgatacgtctaccggcattttgcttccgccgctc
 tgcctattccagaagtctggacattggggagtttagcagctcctcacttactgtattagccggcgagcccagggagtgacactccag
 gatctgctgaaactgaactgctgtgctcagcctgtcgtgaggcaatggatgctattgctgctgccgatctgagtcagactagc
 ggattcggcccattggaccccagggcattggccagtacacaacatggcgagacttcatctgtgctatcgccgatcctcacgtgtacca
 ttggcagactgtgatggacgatactgtgtctgtctgtggcacaggcactcgcgacgaactcatgctgtgggctgaggactgtcctgaagt
 gagacatctgggtccatgccgattttggctccaacaatgtgctcaccgataacgggagaaatcactgcccgtgatcgaactggagcaggca
 atgtttggcgattccagtagaagtggccaacatcttctttggcggccttggctggctgtatggaacagcagaccgggtactttgaac
 ggcgccaccctgagctggctgggagctctagactgagagcctacatgctccgaattggcctggatcagcttaccagtcactgggtgga
 tggcaatttcgacgatgctgcttgggacacaggggcgctgtgatgctattgtccgatccggcgctggaactgtggggagaacacagatc
 gctaggagatccgctgctgtctggaccgatggatgtgtggaagtgtggccgatagtggaaccggaggccttcaaccgacccccg
 ggcaaaaggagtaatgaccgttaaaaccgctgatcagcctcgaactgtgccttctagtggcagccatctgtttgcccctccccgtgc
 ctctctgacctggaagggtgcccactcccactgtccttcttaataaaatgaggaaattgcatcgcattgtctgagtaggtgcattctattct
 ggggggtgggtggggcaggacagcaagggggaggattgggaagacaatagcaggcatgctgggggatgcggtgggctctatgg
 ggatcccgttgacattgatttactagttattaatgtaataaattacggggctcattgattcatagcccatataggattccgcgttac
 ataacttacggtaaatggcccgcctggctgaccgcccacgacccccgccattgacgtcaataatgacgtatgttccatagtaacgc
 caatagggactttccattgacgtcaatgggtggagtatttacggtaaacgtcccacttggcagttacatcaagtgtatcatatgccaagtac

Fig. 1

gccccattgacgtcaatgacggtaaattggcccctggcattatgccagtacatgacctatgggacttctacttggcagtacatc
tacgtattagtcacgctattaccatgggtgatgcggtttggcagtacatcaatgggcgtggatagcggttgactcacgggatttcaa
gtctccacccattgacgtcaatgggagttggtttggcaccaaaatcaacgggactttcaaaatgctgtaacaactccgccccattgac
gcaaatgggcggtgagcgtgtacgggtgggaggtctatataagcagagctctctggtaactagagaacccactgctactgtctgacg
atctgatcaagagacaggataaggagccgccaccatggagttgggctgagctggcttttctgtggctattttaaagggtgtccagtgt
gaagtgcagctgctgaaagcggcgggagccctggcagcctggcggcagcctgagactgtcttgcgccgccagcggttcacctt
cagcagctacgcatgagctgggtccgccaggccccctggcaagggactcgaatgggtgtccgccatcagcggcagcggcggcag
cacctactacgccgacagcgtgaagggccggtcaccatcagccgggacaacagcaagaacacccctgtacctgcatgaacagc
ctgcccggccgaggacaccgccgtatattactgtgcgaaagacacagtggtagtactccactgtctgggtgtacaaaaacctccctgca
cgctcttaacctacaattctgtggcggccgcccgccaccatgattgaacaagatggattgacgcaggttctccggccgcttgg
gtggagaggctattcggctatgactgggcacaacagacaatcggctgtctgatgccccgtgtccggctgtcagcgcaggggccc
ccggttcttttgaagaccgacctgcccgggtccctgaatgaactgcaggacgaggcagcgcggctatctggctggccacgacg
ggcgttcttgcgcagctgtgctcagctgtgactgaagcgggaagggactggctgtattgggcgaagtccggggcaggatctc
ctgtcatctaccctgtctctccgagaaagtatccatcatggctgatgcaatgcggcggctgcatacgttatccggctacctgccca
ttcgaccacaagcgaacatcgcacgagcgcagcactcggatggaaagcggcttctgctgatcaggtgagtagcaggaggtgg
agagtacgcgtaaacactaaagcgtctctcaagatccttttaaccatcacatatacctgccgttactatttttagtgaatgagatatta
tgatatttctgaattgtgataaaaaggcaactttatgccatgcaacagaaactataaaaaatacagagaatgaaaagaacagataga
tttttagtcttttagcccgtagtctgcaaaccttttatgatttctatcaaacaaaaggaaaatagaccagttgcaatccaaacgaga
gtctaataagaatgaggtcgaagaaatcgcgcgggttggactgataaagcaggcaagacctaataatgtgtaaaggcgaagtgtgta
tactttggcgtcacccttacatatttttaggtctttttattgtgcgtaactaacttccatctcaaacaggagggtggaagaagcagac
cgtaaacacagtaataaaaaaggagacatgaacgatgaacatcaaaaagtgtgcaaaaagcaacagtaataaccttactaccgca
ctgctggcaggaggcgaactcaagcgttgcgaaagaaacgaacaaaagccatataaggaaacatacggcatttccatattaca
cgcatgatatgtgcaaatccctgaacagcaaaaaatgaaaaatcaagttctgagttcgattcgtccacaataaaaatctcttctt
tgcaaaaggcctggacgtttgggacagctggccattacaaaacgctgacggcactgtcgaactatcacggctaccacatcgtctttg
cattagccggagatcctaaaaatcgggatgacacatcgatttcatgttctatcaaaaagtcggcgaacttctattgacagctgga
aacgctggccgcgtcttaagacagcgcacaaatcgatgcaaatgattctatcctaaaagaccaaacacaagaatggtcagggtcagcc
acatttacatctgacggaataatccgtttattctactgatttctccggtaaacattacggcaaacaaactgacaactgcacaagttaa
cgtatcagcatcagacagctcttgaacatcaacgggtgtagaggattataatcaatctttgacggtgacggaaaaacgtatcaaatgt
acagcagttcatcagatgaagcaactacagctcaggcgcacaaccatacgtgagagatcctactacgtagaagataaaggccaca
atacttagtatttgaagcaaacactggaaactgaagatggctaccaaggcgaagaatctttatfaacaagcactatggcaaaagcac
atattctccgtcaagaaagtcaaaaactctgcaaacgataaaaaacgcacggctgagttagcaaacggcgtctcggatgattg
agctaaacgatgattacacactgaaaaagtgatgaaaccgctgattgcatctaacacagtaacagatgaaatgaacgcgcgaacgt
ctttaaataaacggcaaatggtatctgttactgactcccgggatcaaaaatgacgattgacggcattacgtctaacgatatttcatg
ctgggtatgttctaatcttaactggcccatacaagccgctgaacaaaactggccttgtgttaaaaatggatcttgatcctaacgatgta
ccttacttactcacacttctgtactcaagcgaaggaacaatgtcgtgattacaagctatatacaaacagaggattctacgcag
acaaacaatcaacgtttgcgctagcttctgctgaacatcaaggcaagaaaacatctgtgtcaaaagacagcatcttgaacaagg
caattaacagttaacaataaaaaacgcaaaagaaaatgccgattatggtgactctcagtaaatctgggtaagagacgtccggaggc
cagcccttctcatgttcagagaacatggftaactggftaagtatgctgctccacaggatgatctggacgaagagcatcaggggctcgc
gccagccgaactgttcgccaggctcaaggcgcgatcccgcagggcaggatctcgtcgtgacctatggcgtgcttgccttgcgga
atatcatgggtggaataatggcgttttctggattcatcactgtggccggctgggtgtggcggaccgtatcaggacatagcgttggct
accgtgatattgctgaagagcttggcggcgaatgggctgaccgcttctctgtctttacgggtatcgcgctcccgttccgcagcgc
cgcttctatcgcttcttgcagagttcttctgagtcgactgcaggagteccactgcacccccctccagcttctctgtccaggcaccag
gccaggtatctggggtgtcagccggcctgggtctggcctgaggccacaagcccgggggtctgtgtggctggggacagggacgcc
ggctgctctgctctgtcttggccatgtgaccttcaggtgctcctgcacgggcacaggttttgcacaccagacagtggagtacta
ccactgtgactactttgactactggggccagggaacctggctaccgtgctcctcaggttaagatggcttcttctgctccttctctggg
cccagcgtctctgtcctggagctgggagataatgtccgggggctccttggctgctgctggcggcaaggggtgggacagatcattgt
gctggggacaaaatgaccttgggacacggggctggctgccacggccggccgggacagtcggagatcaggttctagcgaacct

Fig. 1 (con.)

cgcgacagftaagaaccaggggctctgcgacctgggcccagctctgtcccacaccgggtcacatggcaccacctctctgcag
cctccaccaaggggccatcggtctccccctggcaccctctccaagagcacctctgggggcacagcggccctgggctgctgtgca
aggactacttccccgaaccgggtgacgggtgctgtggaactcaggcgcctgaccagcggcgtgcatacttccggctgctcagat
cctcaggacttactccctcagcagcgtggtgaccgtgacctccagcagcttgggcaccagacctacatctgcaacgtgaatcaaa
gcccagcaacccaaggtggacaagaagttagcccaaatctgtgacaaaactcacacatgccaccgtgccagcacctgaact
cctggggggaccgtcagcttctctcccccaaaaacccaaggacacctcatgatcttagaacccctgaggtcacatgctgtggtg
gtggacgtgagccacgaagacctgaggtcaagtcaactggtacgtggacggcgtggaggtgcataatgccaagacaaagccgc
gggaggagcagtacaacagcacgtaccgtgtggtcagcgtcctcaccgtctgaccaggactggctaatggcaaggagtacaag
tgcaagggtgccaacaaagccctccccagccccatcgagaaaaccatctccaaagccaaagggcagccccgagaaccacaggtg
acacctgccccatcccggtgagctgaccaagaaccaggtcagcctgacctgctgtgcaaggcttctatccagcgacatcg
ccgtggagtgaggagcaatgggcagccgggagaactacaagaccacgctccctgctggtgactccgacggctccttctctct
acagcaagctaccgtggacaagtctagatggcagcaggggaactcttctcatgctccgtgatgcatgaggctctgcacaaccacta
cacgcagaagagcctctccctgctcgggcaaaactggctctcattgtcctgggcggcgtggctggcctgctgctgtttattgggctgg
gcatctcttttgtgctccggtgctggcataggaggcgcgaaggaggtggcggatctggagggggaggatctggagggggctcagga
tcagggggaggatctggaggggatcaactgagtacaaccactgtgaggctcgtactagagatgatgtcctagagctgtccga
actctggctgctccttcgcccattacctgcccactgccataaccgtgatccgatcggcaccatgaacgagtcaccgaactccagga
gctgttctcactagatcgggctggatattggcaaaagtctgggtggcggatgacggagccgctgtcgtgctgtggtgactacactgagt
ctgtggaggctggcggcctgtttgctgaaattggacctcggatggtgactgtctggatctcactggctgccagcagcagatgga
gggactgctggcaccatagaccaaaaggaaacctgctggttctggcaactgtgggagtgtcaccgatcatcaggggcaaaaggact
gggatctgcccgtggtgctccctggcgtggaggccgctgaacgagctggcgtccccgctttctcgaacctctgccccgaaatctc
cctttctacgaacgactgggattcactgtcaccgccgatgtcgaagtgcctgaggggcctagaacatggtgatgaccggaaaccg
gagcttaaccgftaaaccgctgacgctcactgtgcttctagtggcaccatctgtgttggccctccccgctccttctga
ccctggaaggtgccactcccactgctcttcttaataaaatgaggaatgcatcgcattgtctgagtaggtgctattctattctgggggt
ggggtggggcaggacagcaaggggggaggtattgggaagacaatagcaggcatgctggggatgcccgtgggctctatggctcgagtt
aattaactggcctcatgggcttccgctcactgcccgttccagtcgggaaacctgtcgtgccagctgattaacatggtcatagtggt
tcttgcgtattgggctcctccgcttctcgtcactgactcgtcgcctcggctgtcgggtaaaacctggggtgctaatgagcaaa
aggccagcaaaaggccaggaaccgtaaaaaggccgctgtgctggcgttttccataggtccgccccctgacgagcatcaca
atcagcgtcaagtcagaggtggcgaaccgacaggactataaagataaccaggcgttccccctggaagctcctcgtgcgctc
ctgttccgacctgcccgttaccggatactgtccgcttctccttccgggaaagcgtggcgttctcactacgctgtaggtatctc
agttcgggtgtaggtcgttctcctcaagctgggctgtgtgcacgaacccccgttccagccaccgctgctccttaccggtaactatcg
tcttagtccaaccggtaagacacgactatcggcactggcagcagccactggtaacaggattagcagagcaggtatgtagggg
tgctacagagttctgaagtgggtggcctaactacggctacactagaagaacagatttggatctgcgctcgtgtaagccagttaccctc
ggaaaaagagttgtagctcttgatccggcaaacggtgtgctacaggcatcgtggtgacgctcgtctgttggtaggttctc
aaaaaaggatcgaagaagatcctttgatctttctacgggctgtgacgctcagtggaacgaaaactcacgttaagggatttggctatg
agattacaaaaaggatctcacctagatccttttaataaaatgaagtttaatacaatctaaagtataatgagtaaacctggtctgaca
gttaccatgcttaacagtgaggcacctatctcagcagctgtctatttcttaccatagttgctgactccccgctgctgtagataacta
cgatacgggagggttaccatctggccccagtgctgcaatgataccgcgagaaccacgctcaccggctccagattatcagcaataaa
ccagccagccggaaaggccgagcgcagaagtggctcctgcaactttaccgctccatccagctctattaattgttccgggaaagctaga
gtaagtagtccaggttaatagtttgcgaacggtgtgctacaggcatcgtggtgacgctcgtctgttggtaggttctc
agctccggttcccaacgatcaaggcgagtacatgatccccatgtgtgcaaaaaagcgggttagctcctcggctcctccgatcgtg
agaagtaagttggccgaggttatcactatggttatggcagcactgcataatctcttactgtcatgccatccgtaagatgctttctgtg
actggtgagtactcaaccaagctattctgagaatagtgatcgggcgaccgagttgctcttggccggctcaatacgggataataccgc
gccacatagcagaactttaaagtctcatattggaacggttctcggggcgaaaactcgaaggtatccgctgtttagatccag
ttcgatgaaccactcgtgcaccaactgatctcagcatctttactttaccagcgttctgggtgagcaaaaacaggaaggcaaaat
gcccgaaaaaagggaataaggccgacacggaaatgtgaatactatactctccttttcaatatttgaagcattatcagggtattgt
ctcatgagcgggatacatatttgaatgtatttagaaaaataacaaataggggttccgacattccccgaaaagtccac (10231
bp)

[SEQ ID NO:28]

Fig. 1 (con.)

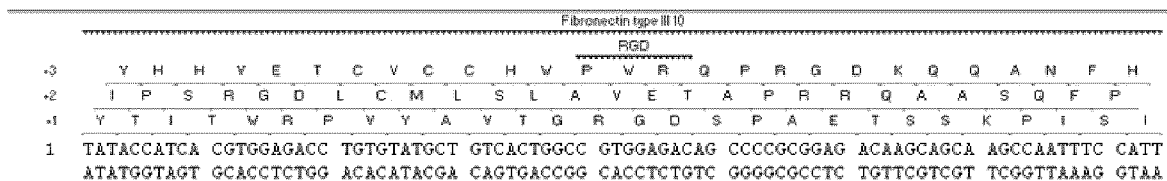
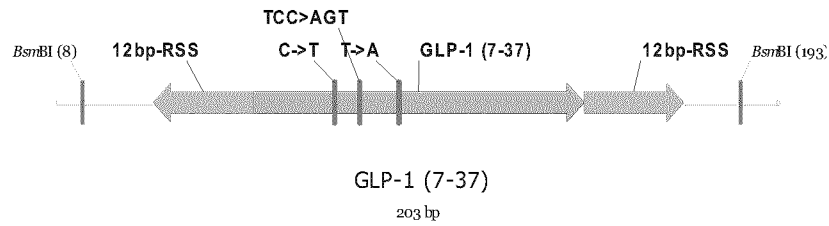


Fig. 2

A.



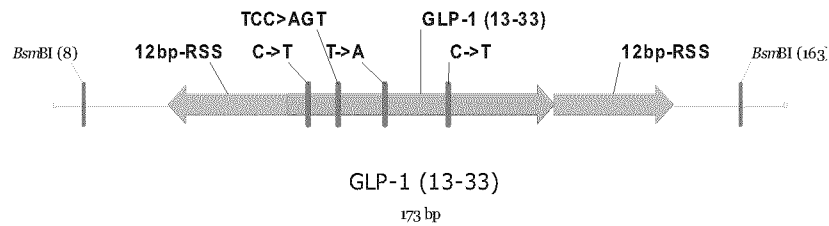
B.

cgtctctccaagtgcaaaggacaggagggttttgtaagggtgtatcactgtgcatgccgagggcacctttaccagtgacgtgagt
agctacctagaaggccaggccgccaagagttatcgctggctcgtgaaggccagaggccacagtatacagcccttaaca
 aaaccctactgcaacctggcggtaagagacg
 [SEQ ID NO:36]

C.

HAEGTFTSDVSSYLEGQAAKEFIAWLKGRG [SEQ ID NO:38]

D.



E.

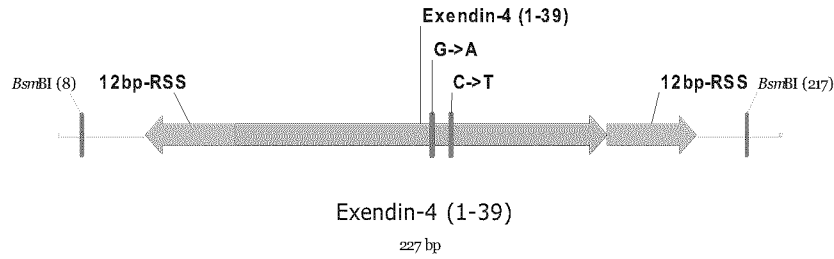
cgtctctccaagtgcaaaggacaggagggttttgtaagggtgtatcactgtg**accagtgacgtgagtagctacctagaaggcca**
ggccgctaaagagttatcgctggctcgtgcacagtatacagcccttaacaaaaaccctactgcaacctggcggtaagagacg
 [SEQ ID NO:37]

F.

TSDVSSYLEGQAAKEFIAWLV [SEQ ID NO:39]

Fig. 3

A.



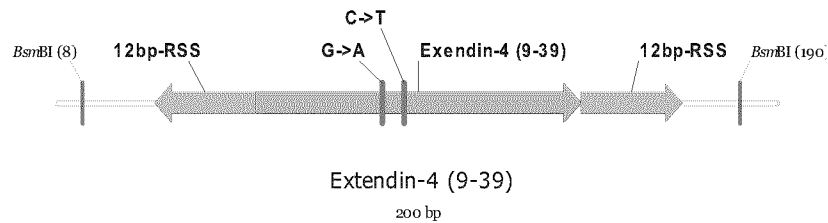
B.

cgtctctccaagtgcaaaggacaggagggttttgtaagggtgtatcactgtgcatggcgagggcacctcaccagcgacctgag
caagcagatggaagaggaagccgtgcggtattcattgagtggctgaagaatggcggccctagctctggcgccctctctctt
ctcacagtgatacagcccttaacaaaaaccctactgcaacctggcggtaagagacg
[SEQ ID NO:40]

C.

HGEGTFTSDLKQMEEEEAVRLFIEWLKNGGPSSGAPPPS [SEQ ID NO:43]

D.



E.

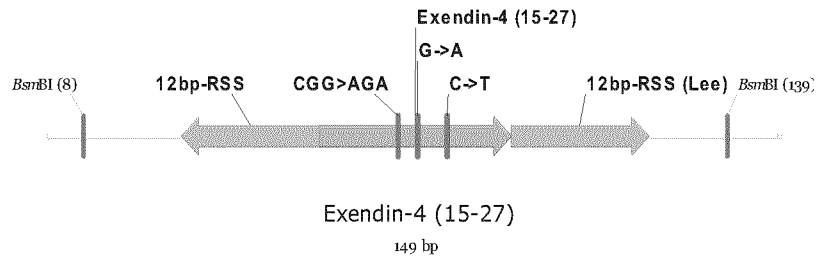
cgtctctccaagtgcaaaggacaggagggttttgtaagggtgtatcactgtgctgagcaagcagatggaagaggaagccgtgc
ggctattcattgagtggctgaagaatggcggccctagctctggcgccctctctcttctcacagtgatacagcccttaacaaaaac
ccctactgcaacctggcggtaagagacg
[SEQ ID NO:41]

F.

LSKQMEEEEAVRLFIEWLKNGGPSSGAPPPS [SEQ ID NO:44]

Fig. 4

G.



H.

cgtctccaagtgcaaaggacaggaggTTTTGTTAAGGGCTGTACTGTGGAAGAGGAAGCCGTGAGACTATTCATTGAGTG
gctgaagcacagtgatacagcccttaacaaaaaccctactgcaacctggcggtaagagacg
 [SEQ ID NO:42]

I.

EEEAVRLFIEWLK [SEQ ID NO:45]

Fig. 4 (con.)

A.

IL-8

A K E L R C Q C I K T Y S K P F H P K F I K E L R

TGCTA AAGAACTTAG ATGTCAGTGC ATAAAGACAT ACTCCAAACC TTCCACCCC AAATTTATCA AAGAACTGAG A
 ACGAT TTCTTGAATC TACAGTCACG TATTTCTGTA TGAGGTTTGG AAAGGTGGGG TTTAAATAGT TTCTTGACTC T

B.

AAAGAACTTAGATGTCAGTGCATAAAGACATACTCTAAACCTTTCCACCCTAAAT
 TTATCAAA
 [SEQ ID NO:12]

C.

CXCL1

Binding domain

Stop (TGA) Stop (TGA) Stop (TGA)

A T E L R C Q C L Q T L Q G I H P K N I Q S V

CGTGGC CACTGAACTGCGCTGC CAGTGCTTGCAGACCC TGCAGGGAATTCACCC CAAGAACATCCAAAAGT GTGAAACGTGAAGTCCC
 GCACCG GTGACTTGACGCGACG GTCACGAAACGTCTGGG ACGTCCCTTAAGTGGG GTTCTTGTAGGTTTCA CACTTGCACCTTCAGGG
 Stop (TGA)

D.

GCCACTGAACTGCGCTGCCAGTGCTTGCAGACCCTGCAGGGAATTCACCCCAAG
 AACATCCAAAGTGTG
 [SEQ ID NO:49]

E.

Stop (TGA)

Stop (TAG) Stop (TAA)

A R * F L P E I

C * V I S S G N L

L L G D F F R K S K

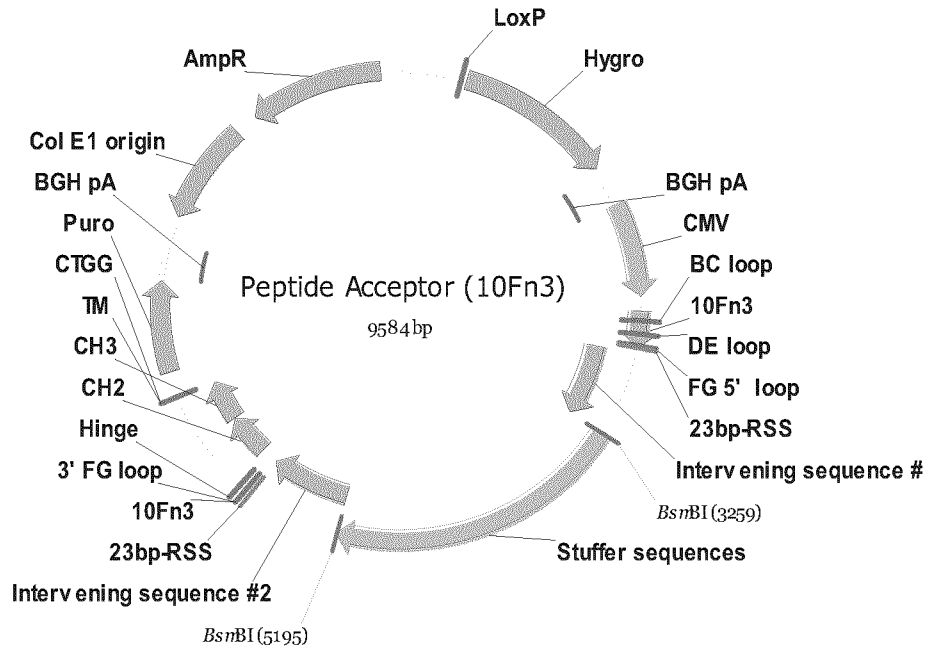
TGCCCTGCTA GGTGATTCT TCCGGAAATC TAAAGAGAAG
 ACGGGACGAT CCACTAAAGA AGGCCTTTAG ATTCTCTTC
 Stop (TAG) Stop (TAG)

Fig. 5

ErbB1	C P P L M L Y N P T T Y Q M D V N P E G K Y S F G A T C V 268
ErbB2	C P A L V T Y N T D T F E S M P N P E G R Y T F G A S C V 275
ErbB3	C P Q P L V Y N K L T F Q L E P N P H T K Y Q Y G G V C V 268
ErbB4	C P Q T F V Y N P T T F Q L E M N F N A K Y T Y G A F C V 265

FIG. 6

A.



B.

ctaaattgtaagcgttaatatatttgttaaaattcgcgttaaattttgttaaatcagctcatttttaaccaataggccgaaatcggcaaaatccc
 ttataaatcaaaagaatagaccgagataggggtgagtgccgctacagggcgctcccattcgccattcaggctgcgcaactgttggga
 agggcgcttcgggtgcgggectcttcgctattacgccagctggcgaaagggggatgtgctgcaaggcgattaagtgggtaacgccag
 ggtttcccagtcacgacgttgtaaacgacggccagtgagcgcgacgtaatacactcactatagggcgaattggcggaggccgt
 caaggcctagggcgcctgaataactcgtatagcatacattatagcaattatcgaaaagcctgaactaccgcgacatccgtggag
 aaattcctcatgaaaaatcagctccgtgtccgatctcatgagctgtccgagggcgaggagtagagcattctcattcgatgtggg
 cgggagaggctacgtgctgagagtgaactctgtgccagggcttctacaaggaccgatacgtctaccggcattttgcttccgccctc
 tgcctattccagaagtctggacattggggagtttagcgagtcctcacttactgtattagccggcgagcccaggagtgacactccag
 gatctgcctgaaactgaactgcctgctgtgctccagcctgtcgtgaggaatggatgctattgctgctgccgatctgagtcagactagc
 ggattcggcccattggaccccagggcattggccagtacacaacatggcgagacttcatctgtgctatcggcgtcctcacgtgtacca
 ttggcagactgtgatggacgatactgtgtctgtctgtggcacagggcactcgacgaactcatgctgtgggctgaggactgtcctgaagt
 gagacatctggccatgccgattttggctccaacaatgtgctcaccgataacggggagaatcactgccgtgacgactggagcaggca
 atgtttggcgattccagtagaagtggccaacatcttctttggcgccctggctgctgtatggaacagcagaccgggtactttgaac
 ggcgccaccctgagctggctgggagtcctagactgagagcctacatgctccgaattggcctggatcagctctaccagtcactggtgga
 tggcaattcagcagatgctgctgggacagggggcgtgtgatgtattgtccgatccggcgctggaactgtggggagaacacagatc
 gctagagatccgctgtgctggaccgatggatgtgtggaagtgtggccgatagtggaaccggaggccttcaacccgacccc
 ggcaaggagtaatgaccgtttaaaccgctgatcagcctcactgtgccttctagtggcagccatctgtgtttgccctccccgtgc
 ctcttgaccctggaaggtgccactcccactgtcttcttaataaaatgaggaaatgcatcgcattgtctgagtaggtgtcattctattct
 ggggggtgggtggggcaggacagcaagggggaggattgggaagacaatagcaggcatgctggggatgctgggtgggctctatgg
 ggatcccgcgttgacattgatttactagtttaataatagtaataatcaggggtcattagttcatagcccataatggagttccgcgttac
 ataacttacggtaaatggcccgcctggctgaccgccaacgacccccgccattgacgtcaataatgacgtatgttccatagtaacgc

FIG. 7

caatagggactttccattgacgtcaatgggtggagtatttacggtaaactgccacttggcagfacatcaagtgtatcatatgccaagtac
gccccctattgacgtcaatgacggtaaatggcccgcctggcattatgccagfacatgacctatgggactttcctacttggcagfacatc
tacgtattagtcacgctattaccatgggtgatcggttttggcagfacatcaatgggcgtggatagcggttgactcacgggatttccaa
gtctccacccattgacgtcaatgggagtttggcaccaaaatcaacgggactttccaaaatgctgtaaacactccgccccattgac
gcaaatgggcggttaggctgtacgggtggaggtctatataagcagagctctctggctaactagagaaccactgcttactgctcgacg
atctgatcaagagacaggataaggagccgccaccatggagtttgggctgagctggcttttctgtggctattttaaagggtgccagtgt
tccgatgtgccagggacctggaagtgggtggccgccacacctaccagcctgctgatctcttgggatgccctgccgtgaccgtcgg
tactacagaatcacctacggcgagacagggcggaacagccccgtgcaggagtttacagtgccggcagcaagagcaccgccacca
tctctggactgaagccggcgtggactacaccatcaccgtgtacggcgtgacaggcagaggcgatagccacagtggtagtactccac
tgtctgggtgtacaaaaacctccctgcacgcctcttaacctcacaattctgtggcgccgcgccaccatgattgaacaagatgga
ttgcacgcaggttctccggcgttgggtggagaggtatcggctatgactgggcacaacagacaatcggctgctctgatccgccg
tgttccggctgtcagcgacggggcgccccgttctttgtcaagaccgacctgcccgtgcccgaatgaactgcaggacgaggcagc
gcggtctatctggctggccacgacgggcttccctgcccagctgtgctcagctgtgcaactgaagcggggaagggactggctgctattg
ggcgaagtgccggggcaggatctctgtcatctcacctgtcctgcccagaaagtatccatcatggctgatgcaatggcgggctgc
atacgttgatccggctacctgccattcgaccaccaagcgaacatcgcatcagcgagcagctactcggatggaagccggctctgt
cgatcaggtgagtacaggaggtggaggtacgcgtaaaccttaagcgtctctcaagatccttttaaccatcacatatactgccgtt
cactattatttagtaaatgagatattatgataatttctgaattgtgattaaaaaggcaactttatgcccatgcaacagaataaaaaata
cagagaatgaaaagaacagatagatttttagtcttagccccgtagctgcaaatcctttatgattttctatcaacaaaaagggaata
tagaccagtgtcaatccaaacgagagcttaataagaatgaggtcgaaaagttaaatcgcgcgggtttgactgataaagcaggcaagac
ctaaaatgtgtaaaggcgaagtgtatacttggcgtcacccctacataatttaggtctttttattgtgcgtaactaactgccatctcaaa
caggagggtcgaagaagcagaccgtaacacagtaataaaaaaggagacatgaacgatgaacatcaaaaagtftgcaaaaacaa
gcaacagtattaaccttactaccgactgctggcaggagggcgaactcaagcgtttgcgaaagaacgaacaaaagccatataag
gaaacatacggcatttccatattacacgccatgatctgcaaatccctgaacgcaaaaaatgaaaaatacaagttcctgagttc
gattcgtccacaataaaaaatctcttctgcaaaaaggcctggacgcttgggacagctggccattacaaaacgctgacggcactgtcga
aactatcacggctaccacatcgtcttgcattagccggagatcctaaaaatgcccgatgacacatcgatttactatcaaaaagtcg
gcaaaacttctattgacagctggaaaaacgctggccgcgtctttaaagacagcgcaaaatcggatgcaaatgattctatcctaaaagac
caaacacaagaatggtcaggtcagccacattacatctgacggaaaaatccgtttattctactgatttctccggtaaacattacggca
aacaacactgacaactgcacaagftaacgtatcagcatcagacagctcttgaacatcaacgggtgtagaggattataaatcaatcttga
cgggtacggaaaaacgtatcaaaatgtacagcagttcatcgtgaaggcaactacagctcagggcacaaccatacgtgagatcc
tactacgtagaagataaaggccacaatacttagtatttgaagcaaacactggaactgaagatggctaccaaggcgaagaatctttatt
taacaaagcatactatggcaaaagcacatcattctcgtcaagaaagtcaaaaactctgcaaaagcataaaaaacgcacggctgagt
tagcaaacggcgtctcgggtatgattgagctaacgatgattacactgaaaaaagtgtatgaaccgctgattgcatctaacacagta
acagatgaaattgaacgcgcgaacgtctttaaagtgaacggcaaatggatctgttactgactcccgcggatcaaaaatgacgattga
cggcattacgtctaacgataattacatgcttgggtatgttcttaattcttaactggccatacaagccgtgaacaaaactggccttgtgta
aaaatggatctgatcctaacgatgtaaccttacttactcactcgcctgacccaagcgaaggaacaatgctggtgattacaagcta
tatgcaaacagaggattctacgcagacaaaacaatcaacgtttgcgctgacttctgctgaaacatcaaaaggcaagaaaacatctgtgt
caaagacagcatcctgaacaaggacaataacagftaacaataaaaaacgcaaaagaaaatgccgattatggtgactctcagtfaca
atctgggtaagagacgtccggaggccagccctctcatgttcagagaacatggttaactggttaagtcatgctgcccacaggatgatct
ggacgaagagcatcaggggctcgcgccagccgaactgttcgccaggctcaaggcgcgcatgcccgacggcgaggatctcgtcgt
gacctatggcgtgcttgccttccgaatatcatgggtgaaatggccgcttcttggattcatcactgtggccgggtgggtgtggcgg
accgctatcaggacatagcgttggctaccctgataattgctgaagagcttggcggcgaatgggctgaccgcttctcgtgctttacgggt
atcggcgtcccgatcgcagcgcategccttctatcgccttctgacgagttcttctgagtgactgcaggagteccactgcaccccc
tcccagttctctgtccaggcaccaggccaggtatctgggtgtgacggcggcctgggtctggcctgaggccacaagccgggggt
ctgtgtggctggggacagggacggcggctgctctgtgcttgggcatgtgaccattcagtgctcctgacgggcacaggtt
ttgtacaccagacagtgaggtactaccactgtgcctgcccagcagcaagccatcagcatcaactaccggaccggtaccgagccca
aatctgtgacaaaactcacacatgccaccgtgccagcactgaaactcctggggggaccgtcagcttctcttcccccaaaacc
aaggacacctcatgatcttagaaccttgaggtcacatgcgtggtggtggacgtgagccacgaagacctgaggtcaagttcaact

FIG. 7 (con.)

ggtacgtggacggcgtggaggtgcataatgccaagacaaagccgcgggaggagcagtacaacagcacgtaccgtgtggtcagcgt
 cctcaccgtcctgcaccaggactggctgaatggcaaggagtacaagtgaagggtccaacaaagccctcccagccccatcgaga
 aaaccatcctcaaaagccaaagggcagccccgagaaccacaggtgtacacctgccccatcccgggatgagctgaccaagaacca
 ggtcagcctgacctgctggtcaaaagcttctatcccagcgcacatcgccgtggagtgaggagcaatgggcagccggagaacaact
 acaagaccacgcctcccgtgctggactccgacggctccttctctacagcaagctcaccgtggacaagtctagatggcagcagg
 ggaacgtcttctcatgctccgtgatgcatgaggctctgcacaaccactacacgcagaagagcctcctcctgtctccgggcaaaactggct
 ctattgtcctggcggcgtggctggcctgctgtgtttattgggctgggcatcttctttgtgtccggtgctggcatagggaggcccaag
 gaggtggcggatctggagggggaggatctggagggggctcaggatcagggggaggatctggagggcgatcaactgagtacaaac
 ccactgtgaggctcgtactagatgatgtgcctagagctgtccgaactctggctgtgccttcgccgattaccctgccactgcata
 ccgtcgatcccgatccacattgaacgagtcaccgaactccaggagctgtttctactagagtcgggctggatattggcaaagtctgg
 gtggccgatgacggagccgctgtcgtgtgtggactacacctgagctgtggaggctggcggcgtgttctgaaattggacctcggga
 tggctgaactgtctggatctgactggctgccagcagcagatggagggactgtggcaccatagaccaaaaggaacctgcctgg
 tctggcaactgtgggagtgtcaccgatcatcagggcaaaaggactgggatctgccgtgggtgctcctggcgtggaggccgctgaac
 gagctggcgtccccgtttctgaaactctgccccgaaatctcctttctacgaacgactgggattactgtcaccgccgatgtcg
 aagtgcctgaggggctagaacatggtgtatgaccggaaacccggagcttaaccgttaaacccgctgatcagcctcactgtgcct
 tctagttgccagccatctgttgttccccctccccgtccttctgacctggaaggtgccactcccactgtccttcttaataaaatgag
 gaaattgcatcgattgtctgagtaggtgtcattctattctggggggtgggggtggggcaggacagcaagggggaggattgggaagac
 aatagcaggcatgctgggatgctgggtggctctatggctcagtttaactggcctcatggcctccgctcactgccgcttcca
 gtcgggaaacctgtcgtgccagctgcattaacatggtcatagctgttctcctgctattgggcgctcctccgctcctcgtcactgactcg
 ctgctcgtcggctgtcgggtaaaagcctggggtgcctaataagcaaaaaggccagcaaaaaggccaggaaccgtaaaaaggccgctgt
 gctggcgttttccataggtcggccccctgacgagcatcaaaaaatcgacgctcaagtcagaggtggcgaaccggacaggact
 ataaagataaccaggcgttccccctggaagctccctcgtgcgctcctgctccgacctgccgttaccggatacctgtccgctttctc
 ccttcgggaaagcgtggcgtttctcatagctcacgctgtaggtatctcagttcgggtgtaggtcttgcctcaagctgggctgtgtgac
 gaacccccgttcagcccagccgtgcgccttaccggtaactatcgtctttagtccaaccggtaagacacgacttatgccactggc
 agcagccactgtaaacaggattagcagagcaggtatgtaggcgggtctacagagtcttgaagtgggtgcctaactacggctacact
 agaagaacagtatgtgtatctgcgctctgctgaagccagttaccttggaaaaagagttgtagctcttgatccggcaacaaccac
 cgctggtagcgggtgggtttttgttgaagcagcagattacgcgcagaaaaaaggatctcaagaagatcctttgatctttctacggggt
 ctgacgctcagtggaacgaaaactcacgttaagggttttggctcatgagattatcaaaaaggatctcacctagatcctttaaattaaaa
 tgaagtttaaatcaatctaaagtatatagtaaaactggctgacagttaccaatgcttaacagtgaggcacctatctcagcgtctgt
 ctatttcttcatccatagttgctgactccccgtgtagataactacgatacgggagggttaccatctggccccagtgctgcaatga
 taccgcgagaaccacgctcaccggctccagattatcagcaataaacagccagccggaaggccgagcgcagaagtgtgctctgc
 aactttatccgctccatccagctctattaattgttggcgggaagctagagtaagtgttcgacagttaatagtttgcgcaacgttggccat
 tctacaggcatcgtggtgtcacgctcgtcttggatggctcattcagctccggttccaacgatcaaggcgagttacatgatcccc
 atgtgtgcaaaaaagcggtagctccttggctcctccgatcgtgtcagaagtaagtggccgagtgatcactcatggtatggcag
 cactgcataaattcttactgcatgccatccgtaagatccttctgtgactggtgagtaactcaaccaagtcattctgagaatagtgtatgc
 ggcgaccgagttgctcttcccggcgtcaatacgggataataccgcgccacatagcagaactttaaagtgtcatcattggaaaacgt
 tcttggggcgaaaactcgaaggatcttaccgctgtgagatccagttcgatgtaaccactcgtgcaaccaactgatcttcagcatcttt
 tactttcaccagcgttctgggtgagcaaaaacaggaaggcaaaaatgccgcaaaaaagggaataagggcgacaccgaaatgttgaat
 actcactcttctttcaatattattgaagcattatcagggtattgtctcatgagcggatacatattgaaatgtattgaaaaataaca
 aataggggtccgcgcacattccccgaaaagtgccac (9584 bp)

[SEQ ID NO:63]

FIG. 7 (con.)

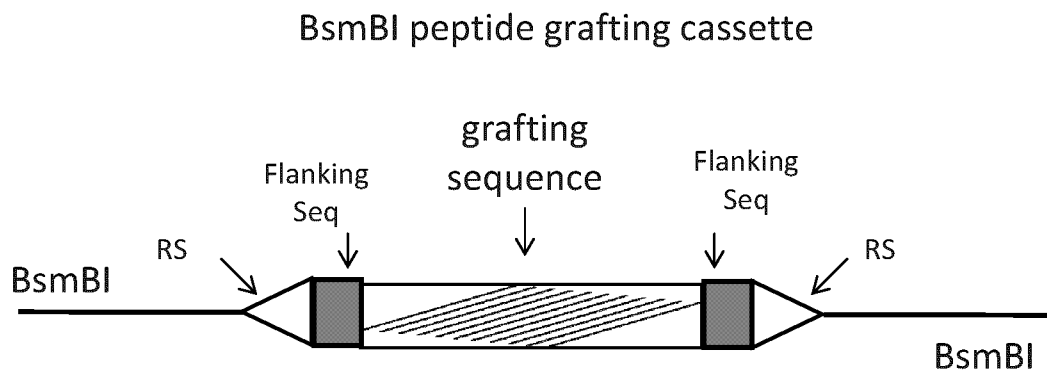
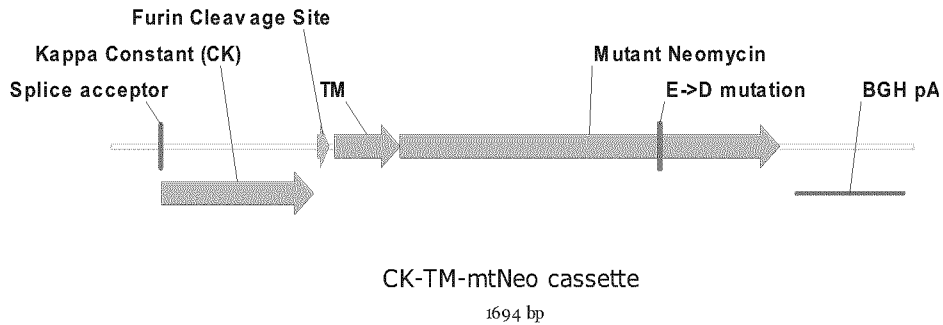


FIG. 8

A.



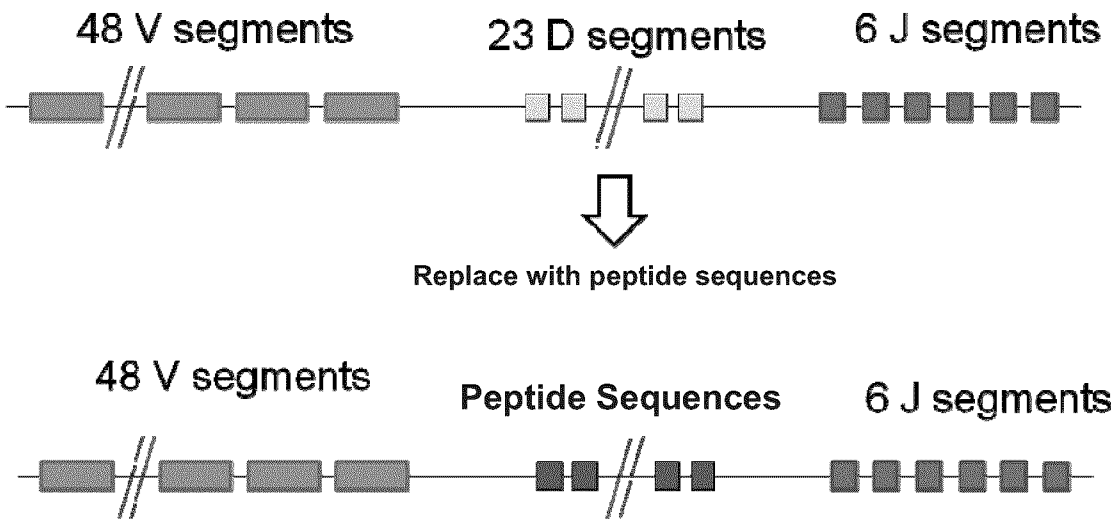
B.

aattcttctgtctgtccctaacaatgcctgtgattatccgaaacaacacaccaagggcagaactttgtacttaaacaccatcctgtttgc
 ttcttctcaggaactgtggctgaccatctgtcttcatctcccgccatctgatgagcagttgaaaagcgggaacagccagcgttgggtg
 cctgctgaataacttctatcccagagaggccaaagtacagtggaaagtgataacgccctccaatcgggtaactcccaggagagtgc
 acagagcaggacagcaaggacagcacctacagcctcagcagcaccctgacgctgagcaaacgagactacgagaaacacaaagtc
 tacgctgcgaagtcacccatcaggcctgagcagccccgtcacaagagctcaacaggggagagtgtggcggcgagc**gtccc**
ggcaccgcccagccctggggcggcgagcagcgtcccgtcaaatattgcaaaaattatcgcggccccctcatcttcttcttcttct
 ccgttgtgattggaagtattatctattctctgagaaagaggcagccagatgggcccgtgggaccgcttacgcttctggaagcgtattg
 aacaagatggattgcacgcaggttctccggccgcttgggtggagaggctattcggctatgactgggcacaacagacaatcggctgct
 ctgatcccgctgttccggctgtcagcgcaggggcccggcttctttgtcaagaccgacctgtccgggtgcccgaatgaactgcag
 gacgaggcagcgcggctatctggctggccacgacgggcttcttgcgcagctgtgctcgcgcttctactgaagcgggaaggga
 ctggctgctattggcggaagtccgggagcagatctctgtcatctcacctgctcctgcccagaaagtatccatcatggctgatgcaat
 gcggcggtgcatacgttgcctgacccgctacccattcaccaccaagcgaacatcgcacgagcagcagcactcggatgg
 aagccggtcttctgatcaggatgatctggacgaagagcagcaggggctcgcgccagccgaactgttccagggctcaaggcgcgc
 atcccagcggcgatgatctcgtgtgacctatggcagcctgcttccgaatcatgggtggaatggccgcttcttctggattcgc
 gactgtggccggctgggtgtggcggaccgctatcaggacatagcgttggctaccctgtatattgctgaagagcttggcggcgaatgg
 gctgaccgcttctcgtcttaccggtatcggcgtcccgattcgcagcgcacgccttctatgccttcttgcagagttcttctgatcaga
 gggccccgtttaaaccgctgatcagcctcactgtgccttctagtggcagccatctgtgtttgccccctccccctgcttcttgcacct
 ggaaggtgccactcccactgtccttcttaataaaatgaggaaattgcacgcattgtctgagtaggtgcattctattctggggggtggg
 gtggggcaggacagcaagggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatggcttctgagggc
 gctgcagttatg

[SEQ ID NO:64]

FIG. 9

A.



B.

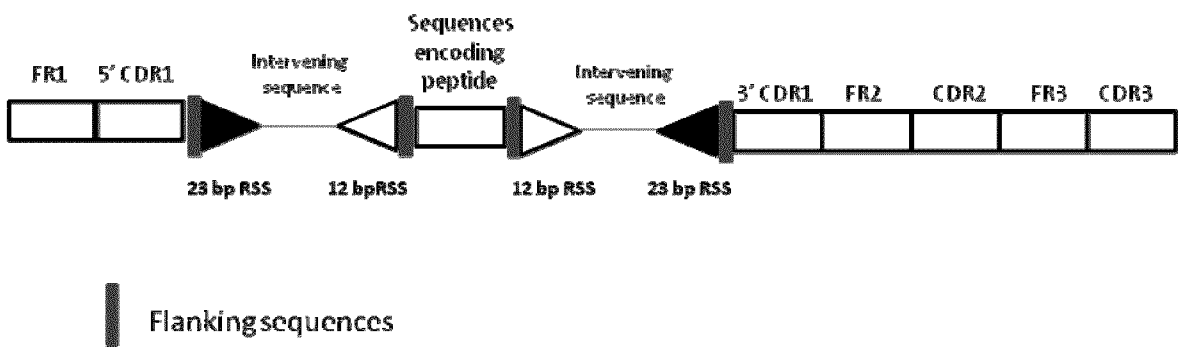


FIG. 10

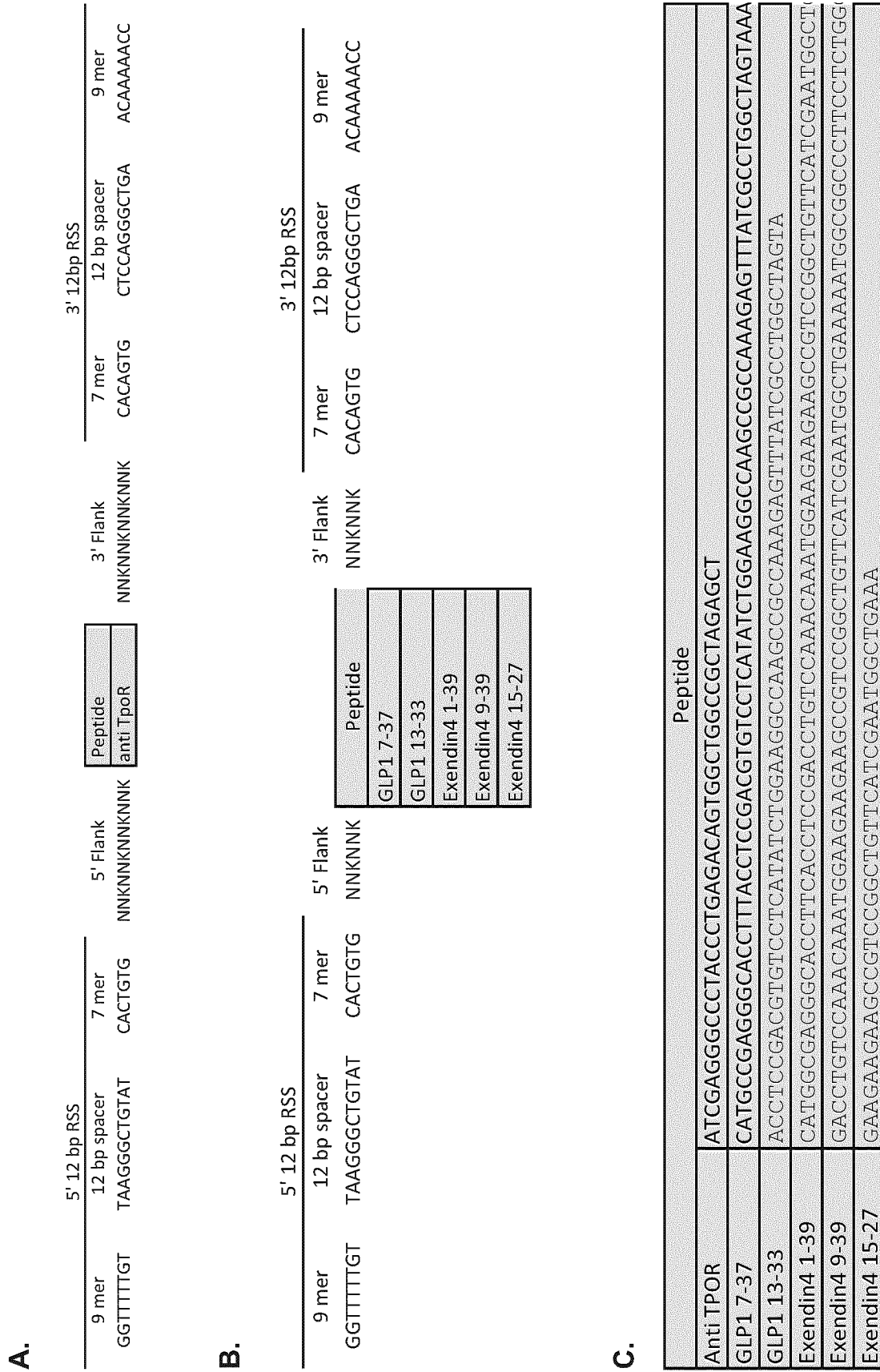
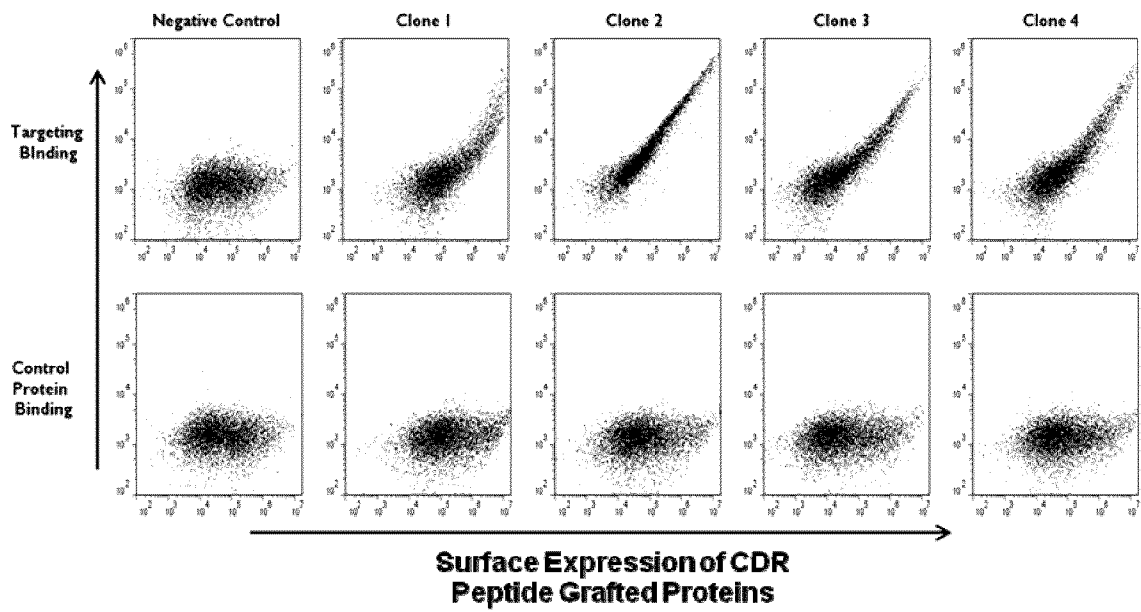


FIG. 11

A.



B.

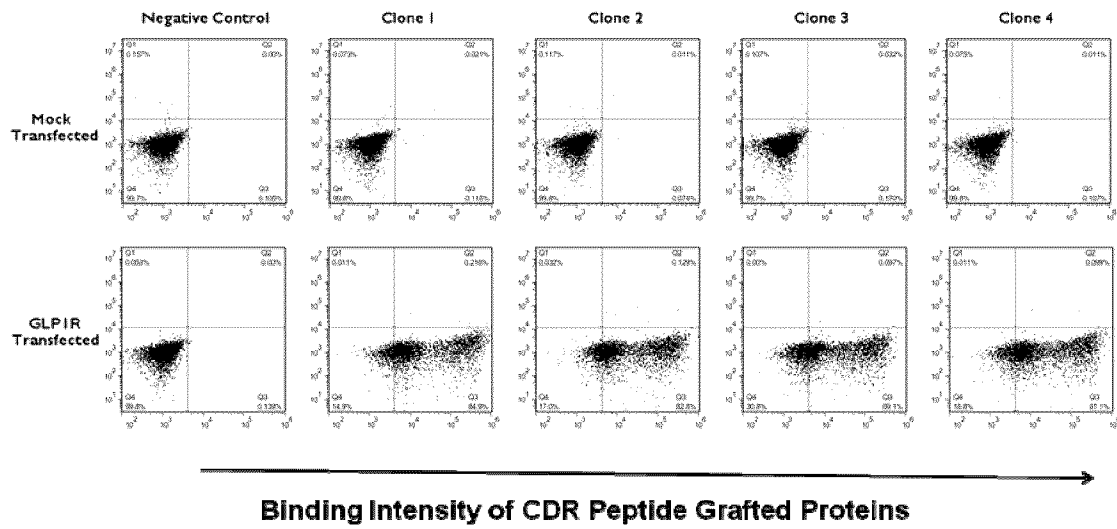


FIG. 12

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CA2013/050204

<p>A. CLASSIFICATION OF SUBJECT MATTER IPC: <i>C12N 15/63</i> (2006.01) , <i>C07K 14/575</i> (2006.01) , <i>C07K 14/605</i> (2006.01) , <i>C07K 16/00</i> (2006.01) , <i>C07K 19/00</i> (2006.01) , <i>C12N 15/09</i> (2006.01) (more IPCs on the last page) According to International Patent Classification (IPC) or to both national classification and IPC</p>														
<p>B. FIELDS SEARCHED</p> <p>Minimum documentation searched (classification system followed by classification symbols) IPC: <i>C12N 15/63</i> (2006.01), <i>C07K 14/575</i> (2006.01), <i>C07K 14/605</i> (2006.01), <i>C07K 16/00</i> (2006.01), <i>C07K 19/00</i> (2006.01) , <i>C12N 15/09</i> (2006.01) (more IPCs on the last page)</p> <p>Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched</p>														
<p>Electronic database(s) consulted during the international search (name of database(s) and, where practicable, search terms used) TotalPatent, Scopus, Canadian Patent Database, Medline (keywords: recombination signal sequence, fusion protein, transgenic, IgG, immunoglobulin, agonist, antagonist, exendin-4, GLP-1, glucagon-like peptide-1, thrombopoietin, somatic recombination, directed</p>														
<p>C. DOCUMENTS CONSIDERED TO BE RELEVANT</p> <table border="1" style="width:100%; border-collapse: collapse;"> <thead> <tr> <th style="width:10%;">Category*</th> <th style="width:70%;">Citation of document, with indication, where appropriate, of the relevant passages</th> <th style="width:20%;">Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td align="center">X</td> <td>ENGLER et al., "Influence of CpG Methylation and Target Spacing on V(D)J Recombination in a Transgenic Substrate" Molecular and Cellular Biology, January 1993 (01-1993), volume 13, pages 571-577. ISSN 0270-7306 *whole document*</td> <td align="center">1-35</td> </tr> <tr> <td align="center">X</td> <td>WO 2009/129247 A2 (GALLO et al.) 22 October 2009 (22-10-2009) *whole document*</td> <td align="center">1-35</td> </tr> <tr> <td align="center">X</td> <td>WO 2005/060642 A2 (BOWDISH et al.) 7 July 2005 (07-07-2005) *whole document*</td> <td align="center">36-45</td> </tr> </tbody> </table>			Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	X	ENGLER et al., "Influence of CpG Methylation and Target Spacing on V(D)J Recombination in a Transgenic Substrate" Molecular and Cellular Biology, January 1993 (01-1993), volume 13, pages 571-577. ISSN 0270-7306 *whole document*	1-35	X	WO 2009/129247 A2 (GALLO et al.) 22 October 2009 (22-10-2009) *whole document*	1-35	X	WO 2005/060642 A2 (BOWDISH et al.) 7 July 2005 (07-07-2005) *whole document*	36-45
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.												
X	ENGLER et al., "Influence of CpG Methylation and Target Spacing on V(D)J Recombination in a Transgenic Substrate" Molecular and Cellular Biology, January 1993 (01-1993), volume 13, pages 571-577. ISSN 0270-7306 *whole document*	1-35												
X	WO 2009/129247 A2 (GALLO et al.) 22 October 2009 (22-10-2009) *whole document*	1-35												
X	WO 2005/060642 A2 (BOWDISH et al.) 7 July 2005 (07-07-2005) *whole document*	36-45												
<p><input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.</p>														
<table border="0" style="width:100%;"> <tr> <td style="width:50%; vertical-align: top;"> <p>* Special categories of cited documents :</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier application or patent but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> </td> <td style="width:50%; vertical-align: top;"> <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance, the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>"Y" document of particular relevance, the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>"&" document member of the same patent family</p> </td> </tr> </table>			<p>* Special categories of cited documents :</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier application or patent but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p>	<p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance, the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>"Y" document of particular relevance, the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>"&" document member of the same patent family</p>										
<p>* Special categories of cited documents :</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier application or patent but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p>	<p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance, the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>"Y" document of particular relevance, the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>"&" document member of the same patent family</p>													
<p>Date of the actual completion of the international search 24 May 2013 (24-05-2013)</p>		<p>Date of mailing of the international search report 14 June 2013 (14-06-2013)</p>												
<p>Name and mailing address of the ISA/CA Canadian Intellectual Property Office Place du Portage I, C114 - 1st Floor, Box PCT 50 Victoria Street Gatineau, Quebec K1A 0C9 Facsimile No.: 001-819-953-2476</p>		<p>Authorized officer Jeremy McLean (819) 994-9302</p>												

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of the first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons :

1. Claim Nos. :
because they relate to subject matter not required to be searched by this Authority, namely :

2. Claim Nos. :
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically :

3. Claim Nos. :
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows :

continued in extra sheet

1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claim Nos. :
4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim Nos. :

- Remark on Protest** The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- No protest accompanied the payment of additional search fees.

continuation of Box No. III:

Group A: Claims 1-35 featuring a method for generating variants of a fusion protein comprising the use of recombination signal sequences as well as polynucleotides comprising portions of said fusion protein and said sequences, host cells comprising said polynucleotides and fusion proteins produced by said method; and,

Group B: Claims 36-45 featuring a peptide-grafted immunoglobulin comprising an immunoglobulin scaffold wherein a peptide capable of binding to a G-protein coupled receptor as well as flanking sequences are inserted into or replace a complementarity determining region of the immunoglobulin scaffold.

Though the method, polynucleotides and host cells of Group A could be used to produce proteins featured in Group B, Group B encompasses all proteins that possess the defined elements regardless of the method by which the proteins are made. Therefore, the method of production of the proteins cannot be the special technical feature that renders all claims part of the same inventive concept.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CA2013/050204

Patent Document Cited in Search Report	Publication Date	Patent Family Member(s)	Publication Date
WO2009129247A2	22 October 2009 (22-10-2009)	CA2721231A1 EP2271758A2 EP2271758A4 US2009258392A1 US8012714B2 US2012258495A1	22 October 2009 (22-10-2009) 12 January 2011 (12-01-2011) 23 May 2012 (23-05-2012) 15 October 2009 (15-10-2009) 06 September 2011 (06-09-2011) 11 October 2012 (11-10-2012)
WO2005060642A2	07 July 2005 (07-07-2005)	AT320450T AT414720T AT544785T AU3400102A AU2002234001B2 AU2003295623A1 AU2003295623B2 CA2436671A1 CA2507494A1 DE60118023D1 DE60118023T2 DE60136656D1 DK1642910T3 EP1370589A2 EP1370589B1 EP1498429A2 EP1498429A3 EP1567541A2 EP1567541A4 EP1589034A2 EP1589034A3 EP1589034B1 EP1642910A1 EP1642910B1 ES2258558T3 ES2316919T3 ES2380367T3 JP2006518585A PT1642910E US2003232972A1 US7396917B2 US2003049683A1 US7482435B2 US2004253242A1 US2009130716A1 US2010041012A1 WO0246238A2 WO0246238A3 WO2004050017A2 WO2004050017A3 WO2004108078A2 WO2004108078A3	15 April 2006 (15-04-2006) 15 December 2008 (15-12-2008) 15 February 2012 (15-02-2012) 18 June 2002 (18-06-2002) 01 May 2008 (01-05-2008) 23 June 2004 (23-06-2004) 05 June 2008 (05-06-2008) 13 June 2002 (13-06-2002) 17 June 2004 (17-06-2004) 11 May 2006 (11-05-2006) 28 December 2006 02 January 2009 07 May 2012 (07-05-2012) 17 December 2003 (17-12-2003) 15 March 2006 (15-03-2006) 19 January 2005 (19-01-2005) 03 August 2005 (03-08-2005) 31 August 2005 (31-08-2005) 02 August 2006 (02-08-2006) 26 October 2005 (26-10-2005) 02 November 2005 (02-11-2005) 19 November 2008 (19-11-2008) 05 April 2006 (05-04-2006) 08 February 2012 (08-02-2012) 01 September 2006 (01-09-2006) 16 April 2009 (16-04-2009) 11 May 2012 (11-05-2012) 17 August 2006 (17-08-2006) 22 March 2012 (22-03-2012) 18 December 2003 (18-12-2003) 08 July 2008 (08-07-2008) 13 March 2003 (13-03-2003) 27 January 2009 (27-01-2009) 16 December 2004 (16-12-2004) 21 May 2009 (21-05-2009) 18 February 2010 (18-02-2010) 13 June 2002 (13-06-2002) 10 July 2003 (10-07-2003) 17 June 2004 (17-06-2004) 13 January 2005 (13-01-2005) 16 December 2004 (16-12-2004) 06 April 2006 (06-04-2006)
(28-12-2006)			
(02-01-2009)			

C12N 15/11 (2006.01) , *C12N 15/13* (2006.01) , *C12N 15/62* (2006.01) , *C12N 15/85* (2006.01)