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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 December 2005 (15.12.2005)

PCT

(10) International Publication Number
WO 2005/118874 A1

(51) International Patent Classification[7]: C12Q 1/68, C12N 1/00, 15/09, 15/63, 15/70, 15/74, 5/00, 5/02, C07H 21/02, 21/04

(21) International Application Number: PCT/US2005/019592

(22) International Filing Date: 6 June 2005 (06.06.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/576,819 4 June 2004 (04.06.2004) US

(71) Applicant *(for all designated States except US)*: **WYETH** [US/US]; Five Giralda Farms, Madison, NJ 07940 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **SMITH, Larry, Russell** [US/US]; 11410 Turtleback Lane, San Diego, CA 92127 (US). **SHAHABI, Vafa** [US/US]; 200 Jug Hollow Road, Valley Forge, PA 19481 (US). **SIDHU, Maninder, K.**; 401 N. Middletown Road, Pearl River, NY 10965 (US).

(74) Agents: **SCHULMAN, Robert, M.** et al.; Hunton & Williams, LLP, 1900 K Street, N.W., Suite 1200, Washington, DC 20006-1109 (US).

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

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 2005/118874 A1

(54) **Title:** ENHANCING PROTEIN EXPRESSION

(57) **Abstract:** Modified polynucleotide compositions providing enhanced gene expression and methods for preparing said compositions are disclosed. Methods of using the compositions, such as in screening assays, diagnostic tools, kits, etc. and for prevention and/or treatment of diseases and disorders are also disclosed.

ENHANCING PROTEIN EXPRESSION

FIELD OF THE INVENTION

[0001] The present invention relates to polynucleotide compositions that provide enhanced efficiency in the expression of proteins or polypeptides by genes in mammalian cells (i.e., resulting in an increase in the levels of the proteins or polypeptides encoded by the genes), such as viral, bacterial and mammalian genes, as well as methods for preparing said compositions. In particular, the invention provides polynucleotide sequences that provide enhanced gene expression over the corresponding wild-type polynucleotides. Also provided are methods of using the polynucleotide compositions in prevention and treatment of diseases and disorders (e.g., immuno-therapeutic, immuno-prophylactic and genetic therapy uses and the like), such as in DNA and RNA vaccines (e.g., DNA vaccines for preventing/treating HIV/AIDS) as well as in biological assays, diagnostics and the like.

BACKGROUND OF THE INVENTION

[0002] The level of protein expressed by a gene is crucial to *in vivo* responses/effects involving the protein, as well as *in vitro* assays involving the protein. Under some circumstances and for reasons not fully characterized, however, *in vitro* and/or *in vivo* benefits of the protein product of a gene are compromised because the gene is not adequately expressed in cells. Poor protein expression is encountered in a number of different contexts. For example, poor expression of proteins by eukaryotic genes in prokaryotic cells has been previously reported (see Seed et al., U.S. Patent Nos. 5,786,464 and 5,795,737). The poor expression of proteins by viral genes in mammalian cells has also been described (see Schwartz et al., *J. Virol.* 66(12):7176-7182 (1992), Schneider et al., *J Virol.*, 71(7):4892-4903 (1997) and Pavlakis et al., U.S. Patent No. 6,414,132 B1). However, the poor expression of certain viral, bacterial and mammalian genes, in mammalian cells remains a significant problem from the standpoint of both *in vivo* uses of the protein products and *in vitro* uses in assays and the like.

[0003] There are a number of factors that influence the levels of gene expression of proteins in mammalian cells and that account for, or at least contribute to, the poor expression

observed for certain genes in these cells. In some instances, translational mechanisms are responsible for the poor expression. For example, it has been recognized that in certain wild-type genes, the naturally occurring nucleic acid sequences of the genes are rich in adenine (A) and/or uracil (U) (if the polynucleotide is RNA) or adenine (A) and/or thymine (T) (if the polynucleotide is DNA) and biased toward "disfavored codons". The term "disfavored codons," as used herein, refers to codons that contain A, U, or T in the third ("wobble") position of the codon nucleotide triplet. It has been suggested in the art (see Haas et al., Current Biol. 6:315-324, 1996) that certain wild-type genes are not handled efficiently by the translational machinery of mammalian cells.

[0004] Also, in addition to translational mechanisms accounting for poorly-expressed genes, there have been various AU rich RNA instability sequences discovered in several messenger RNAs (mRNAs) which do not directly impact the translatability of a given mRNA, but limit protein expression by increasing mRNA turnover. Further, several specific "inhibitory" sequences contained within the HIV-1 gag ORF have been described (see Pavlakis, U.S. Patent No. 6,414,132 B1) which limit the expression levels of gag by inhibiting nuclear export of these transcripts.

[0005] IL-15 exemplifies the problem inherent in poor gene expression. IL-15 is a pluripotent cytokine that is secreted by antigen presenting cells such as monocytes/macrophages and dendritic cells, but also a variety of nonlymphoid tissues. IL-15, in addition to being a growth and survival factor for memory CD8+ T cells, is also a potent activator of effector-memory CD8+ T cells, both in healthy and HIV-infected individuals. Because IL-15 is a prototypic Th1 cytokine, and by virtue of its activity as a stimulator of T cells, NK cells, LAK (lymphokine-activated killer) and TILs (tumor infiltrating lymphocytes), IL-15 is a potential candidate for use as a molecular adjuvant along with HIV DNA vaccines to enhance cellular immune responses. However, one major limiting factor for its use as a genetic adjuvant, remains its poor expression due to its complex regulation at the levels of mRNA transcription and translation and, protein translocation and secretion.

[0006] Further, DNA vaccines, which are being studied for many diseases, including HIV, influenza, tuberculosis and malaria, usually work by injecting specially reproduced genetic material of the organism directly into the body. This genetic material encodes information that gets the individual's own cells to make the vaccine. DNA vaccines have shown some impressive results in animals. Studies by Merck & Co. demonstrated that a DNA vaccine can prevent influenza in animals.

[0007] In the area of HIV disease, DNA vaccines have generally not been able to stimulate strong immune responses in people. It has been suggested that DNA vaccines are less effective in humans than in smaller animals as a result of the problem of scaling up doses, where it is not practical to give large enough amounts of these vaccines to match the doses given to mice or monkeys. Interest in DNA vaccines either for prevention or treatment is therefore likely to depend on finding new and more efficient ways to present them to the immune system. An approach that improves the expression of a protein, such as IL-15 for use as an adjuvant in a DNA vaccine against HIV/AIDS, for example, is thus highly desireable.

[0008] Various techniques have been proposed for optimizing expression of genes, particularly for poorly expressed genes. For example, one approach involved selectively replacing wild-type codons encompassing inhibitory sequences with other codons to eliminate the inhibitory effect. However, the sequence motifs that define either instability or inhibitory sequences are not readily apparent and therefore not easily identified. Several genes (e.g. E7 and En among others) which appear to also contain inhibitory sequences have not yet been mapped to identify the location of inhibitory sequences and there are no straightforward prescriptions from the gag work to predict how to eliminate inhibitory sequences from these genes.

[0009] Further, a complete "codon optimized" version of gp120 envelope has been described (see Haas et al., *Current Biology*, 6:315-324, 1996; Andre et al., *J. Virology*, 72:1497-1503) in which all "non-preferred" wild-type codons from env were replaced with "preferred" codons and found to enhance expression levels.

[0010] Previously available approaches, as described above, impose stringent requirements in their application. In particular, these approaches require the use of "preferred codons," or alternatively, identification of specific "inhibitory sequences." For example, the technology described by Seed requires incorporation of "preferred codons" and purportedly depends on invoking the translational enhancement as the mechanism of increased protein levels. "Preferred codons," as defined by Seed, are GCC for Ala, CGC for Arg, AAC for Asn, GAC for Asp, TGC for Cys, CAG for Gln, GGC for Gly, CAC for His, ATC for Ile, CTG for Leu, AAG for Lys, CCC for Pro, TTC for Phe, AGC for Ser, ACC for Thr, TAC for Tyr, and GTG for Val. According to Seed, "less preferred codons" are GGG for Gly, ATT for Ile, CTC for Leu, TCC for Ser, and GTC for Val. Seed also teaches that all codons which do not fit the description of preferred codons or less preferred condons are "non-preferred codons."

Accordingly, Seed's approach demands the use of the one specific codon prescribed in each instance and the replacement of every codon or nearly every codon in a sequence.

[0011] Likewise, the technology described by Pavlakis requires identification of inhibitory/instability sequences and the alteration of those specifically identified inhibitory/instability sequences. According to Pavlakis, an inhibitory/instability sequence of a transcript is a regulatory sequence that resides within an mRNA transcript and is either (1) responsible for rapid turnover of that mRNA and can destabilize a second indicator/reporter mRNA when fused to that indicator/reporter mRNA, or is (2) responsible for underutilization of a mRNA and can cause decreased protein production from a second indicator/reporter mRNA when fused to that second indicator/reporter mRNA or (3) both of the above. The procedures to locate and mutate the inhibitory/instability sequences are described in detail by Pavlakis. Accordingly, this approach is experimental result-dependent in that it requires preliminary experimentation to identify specific regions of sequence for targeted mutation.

[0012] Polynucleotide compositions that provide enhanced gene expression while obviating any requirement to alter each codon to a "preferred codon" or identify "inhibitory sequences" provide certain benefits. These benefits include not only improved efficiency, cost-effectiveness, consistency and accuracy in improving the expression of certain genes, but also the ability to achieve a far greater scope of applicability (i.e., the ability to attain such improved gene expression possible for genes for which it was previously not possible (or at least highly inefficient) using previously available technology). It would be desirable to have an approach to attain enhanced gene expression that avoids the stringent requirements of previous approaches. Accordingly, it would be desirable to have an approach to attain enhanced gene expression without having to alter all the codons of the gene to preferred codons or identify inhibitory sequences of the gene and then altering those sequences. Moreover, it would be desirable to have an approach that does not target, define, nor rely upon a specific transcriptional or translational mechanism for improved gene expression.

SUMMARY OF THE INVENTION

[0013] The present invention provides enhanced gene expression in mammalian cells. In particular, the present invention provides modified polynucleotides with significantly improved expression over their wild-type counterparts. The present invention also provides compositions for preventing and treating conditions , as well as compositions for use in assays, vectors, diagnostic tools and the like.

[0014] According to an embodiment, the present invention provides a method of preventing or treating a disease in a mammal comprising: administering to the mammal an effective amount of one or more compositions of the invention.

[0015] According to a further embodiment, the present invention provides a method for enhancing expression of a gene comprising: expressing *in vivo* or *in vitro* a modified polynucleotide of the invention.

[0016] According to another embodiment, the present invention provides a method for preparing a polynucleotide that provides enhanced expression of a gene comprising: assembling oligonucleotides comprising surrogate codons to form a modified polynucleotide comprising a predetermined nucleic acid sequence wherein the nucleotides cytosine (C) or guanine (G) occupy the wobble position of each of said surrogate codons in place of the corresponding nucleotides adenine (A), uracil (U) or thymine (T) of a naturally-occurring polynucleotide that expresses the same protein or polypeptide as said modified polynucleotide.

[0017] According to yet another embodiment, the present invention provides a method for preparing a polynucleotide that provides enhanced expression of a gene comprising: (1) determining for said gene a modified nucleic acid sequence comprising surrogate codons in which the nucleotides cytosine (C) or guanine (G) occupy the wobble position in place of the corresponding nucleotides adenine (A) or uracil (U) or thymine (T) of a naturally-occurring polynucleotide that expresses the same protein or polypeptide as said modified polynucleotide; (2) selecting oligonucleotides having nucleotide sequences corresponding to portions of said determined recombinant nucleic acid sequence; and (3) assembling the oligonucleotides to form a recombinant polynucleotide comprising the determined recombinant nucleic acid sequence.

[0018] According to a still further embodiment, the present invention provides a method for enhancing expression of a gene comprising: altering a wild-type polynucleotide so that a naturally-occurring codon having adenine (A), uracil (U) or thymine (T) in the wobble position is replaced by a surrogate codon having cytosine (C) or guanine (G) in the wobble position, said surrogate codon encoding the same amino acid as the naturally-occurring codon.

[0019] According to another embodiment, the present invention provides a modified polynucleotide comprising a nucleic acid sequence comprising surrogate codons in which the nucleotides cytosine (C)or guanine (G) occupy the wobble position in place of the

corresponding nucleotides adenine (A) or uracil (U), in RNA, or adenine (A) or thymine (T), in DNA, of a naturally-occurring polynucleotide that expresses the same protein or polypeptide as said modified polynucleotide.

[0020] According to a further embodiment, the present invention provides a modified polynucleotide comprising a nucleic acid sequence in which each codon encoding alanine is GCG, each codon encoding arginine is CGG or AGG, each codon encoding leucine is CTC, each codon encoding proline is CCT or CCG, each codon encoding glutamic acid is GAG, each codon encoding glycine is GGG, each codon encoding isoleucine is ATT, each codon encoding serine is TCC, each codon encoding threonine is ACG, and each codon encoding valine is GTC.

[0021] According to still another embodiment, the present invention provides a modified polynucleotide comprising a nucleic acid sequence having the general formula: $-(X)_i - (Y)_j - (X)_i -$, wherein X represents non-surrogate codons having the nucleic acid sequence of any of the corresponding wild-type codons in the naturally-occurring polynucleotide that encode the same protein or polypeptide as said recombinant polynucleotide, said wild-type codons having cytosine (C) or guanine (G) in the wobble position, wherein Y represents surrogate codons having a nucleic acid sequence that is different from the corresponding wild-type codons in the naturally-occurring polynucleotide that encode the same protein or polypeptide as said recombinant polynucleotide, said wild-type codons having adenine (A) or uracil (U) or thymine (T) in the wobble position, said surrogate codons having cytosine (C), guanine (G) or thymine (T) in the wobble position and encoding the same amino acid as the corresponding wild-type codons in the naturally-occurring polypeptide that encodes the same protein or polypeptide as said modified polynucleotide, wherein i is any positive integer of at least 0; and wherein j is any positive integer of at least 1.

[0022] According to a still further embodiment, the present invention provides a modified polynucleotide comprising: (a) the nucleic acid sequence of any of SEQ ID NOS: 1, 3 or 5; (b) an immunogenic encoding portion of (a); or (c) a nucleic acid sequence that hybridizes under stringent conditions to any of (a) or (b).

[0023] According to another embodiment, the present invention provides a composition comprising: a modified polynucleotide comprising a nucleic acid sequence in which the nucleotides cytosine (C) or guanine (G) occupy the wobble position of surrogate codons in place of the corresponding nucleotides adenine (A), thymine (T) or uracil (U) in the nucleic

acid sequence of a naturally-occurring polynucleotide that expresses the same protein or polypeptide as said recombinant polynucleotide; and a pharmaceutically acceptable buffer, diluent, adjuvant, carrier and/or vector.

[0024] According to yet another embodiment, the present invention provides a composition comprising a modified polynucleotide comprising a nucleic acid sequence in which each codon encoding alanine is GCG, each codon encoding arginine is CGG or AGG, each codon encoding leucine is CTC, each codon encoding proline is CCT or CCG, each codon encoding glutamic acid is GAG, each codon encoding glycine is GGG, each codon encoding isoleucine is ATT, each codon encoding serine is TCC, each codon encoding threonine is ACG, and each codon encoding valine is GTC; and a pharmaceutically acceptable buffer, diluent, adjuvant, carrier and/or vector.

[0025] According to a further embodiment, the present invention provides a composition comprising a pharmaceutically acceptable buffer, diluent, adjuvant, carrier and/or vector; and a modified polynucleotide comprising a nucleic acid sequence having the general formula: -$(X)_i$ - $(Y)_j$ -$(X)_i$-; wherein X represents non-surrogate codons having the nucleic acid sequence of any of the corresponding wild-type codons in the naturally-occurring polynucleotide that encode the same protein or polypeptide as said modified polynucleotide, said wild-type codons having cytosine (C) or guanine (G) in the wobble position; wherein Y represents surrogate codons having a nucleic acid sequence that is different from the corresponding wild-type codons in the naturally-occurring polynucleotide that encode the same protein or polypeptide as said modified polynucleotide, said wild-type codons having adenine (A), uracil (U) or thymine (T) in the wobble position, said surrogate codons having cytosine (C) or guanine (G) in the wobble position and encoding the same amino acid as the corresponding wild-type codons in the naturally-occurring polynucleotide that encodes the same protein or polypeptide as said modified polynucleotide; wherein i is any positive integer of at least 0; and wherein j is any positive integer of at least 1.

[0026] According to another embodiment, the present invention provides a composition comprising: (a) the nucleic acid sequence of any of SEQ ID NOS: 1, 3 or 5; (b) an immunogenic encoding portion of (a); or (c) a nucleic acid sequence that hybridizes under stringent conditions to any of (a) or (b).

[0027] According to a still further embodiment, the present invention provides a composition comprising a polynucleotide comprising the nucleic acid sequence of any of SEQ ID NOS: 1, 3 or 5; and a vector.

[0028] According to another embodiment, the present invention provides a composition comprising: a recombinantly expressed protein or polypeptide encoded by a modified polynucleotide comprising any of: (a) the nucleic acid sequence of any of SEQ ID NOS: 1, 3 or 5; (b) an immunogenic encoding portion of (a); or (c) a nucleic acid sequence that hybridizes under stringent conditions to any of (a) or (b).

[0029] According to yet another embodiment, the present invention provides a composition comprising a recombinantly expressed protein or polypeptide encoded by a modified polynucleotide comprising a nucleic acid sequence comprising surrogate codons in which the nucleotides cytosine (C) or guanine (G) occupy the wobble position in place of the corresponding nucleotides adenine (A), uracil (U) or thymine (T) of a naturally-occurring polynucleotide that expresses the same protein or polypeptide as said recombinant polynucleotide.

[0030] According to a further embodiment, the present invention provides a composition comprising an antibody that immunospecifically binds to a recombinantly expressed protein of the invention.

[0031] According to an even further embodiment, the present invention provides a composition prepared by a process comprising inserting into a vector a modified nucleic acid sequence comprising surrogate codons in which the nucleotides cytosine (C)or guanine (G) occupy the wobble position in place of the corresponding nucleotides adenine (A), uracil (U) or thymine (T) of a naturally-occurring polynucleotide that expresses the same protein or polypeptide as said modified polynucleotide.

[0032] According to a still further embodiment, the present invention provides a composition prepared by a process comprising: inserting into a vector a modified nucleic acid sequence in which each codon encoding alanine is GCG, each codon encoding arginine is CGG or AGG, each codon encoding leucine is CTC, each codon encoding proline is CCT or CCG, each codon encoding glutamic acid is GAG, each codon encoding glycine is GGG, each codon encoding isoleucine is ATT, each codon encoding serine is TCC, each codon encoding threonine is ACG, and each codon encoding valine is GTC.

[0033] According to another embodiment, the present invention provides a composition prepared by a process comprising: inserting into a vector a polynucleotide comprising a modified nucleic acid sequence having the general formula: $-(X)_i - (Y)_j - (X)_i -$; wherein X represents non-surrogate codons having the nucleic acid sequence of any of the corresponding wild-type codons in the naturally-occurring polynucleotide that encode the same protein or polypeptide as said modified polynucleotide, said wild-type codons having cytosine (C) or guanine (G) in the wobble position; wherein Y represents surrogate codons having a nucleic acid sequence that is different from the corresponding wild-type codons in the naturally-occurring polynucleotide that encode the same protein or polypeptide as said modified polynucleotide, said wild-type codons having adenine (A) or uracil (U) in the wobble position, said surrogate codons having cytosine (C), guanine (G) or thymine (T) in the wobble position and encoding the same amino acid as the corresponding wild-type codons in the naturally-occurring polypeptide that encodes the same protein or polypeptide as said modified polynucleotide; wherein i is any positive integer of at least 0; and wherein j is any positive integer of at least 1.

[0034] According to yet another embodiment, the present invention provides a composition prepared by a process comprising: inserting into a vector any of: (a) the nucleic acid sequence of any of SEQ ID NOS: 1, 3 or 5; (b) an immunogenic encoding portion of (a); or (c) a nucleic acid sequence that hybridizes under stringent conditions to any of (a) or (b).

[0035] According to a further embodiment, the present invention provides for the use of a composition in the preparation of a medicament for inducing an immune response in a mammal.

[0036] According to another embodiment, the present invention provides for the use of a composition in the preparation of a medicament for treating a condition in a mammal.

[0037] According to a still further embodiment, the present invention provides a transformed, transfected, lipofected or infected cell line comprising: a recombinant cell that expresses any of: (a) the nucleic acid sequence of any of SEQ ID NOS: 1, 3 or 5; (b) an immunogenic encoding portion of (a); or (c) a nucleic acid sequence that hybridizes under stringent conditions to any of (a) or (b).

[0038] According to another embodiment, the present invention provides a modified polynucleotide comprising: (a) the nucleic acid sequence of any of SEQ ID NOS: 12-16; (b)

an immunogenic encoding portion of (a); or (c) a nucleic acid sequence that hybridizes under stringent conditions to any of (a) or (b).

[0039] According to yet another embodiment, the present invention provides a composition that comprises a modified polynucleotide comprising: (a) a non-native leader sequence; and (b) a nucleic acid sequence comprising cytosine (C) or guanine (G) at the wobble position of at least one codon that encodes any of the amino acids alanine, arginine, leucine, proline, glutamic acid, glycine, isoleucine, serine, threonine, or valine where adenine (A), uracil (U) or thymine (T) occupy the wobble position of the corresponding codon of the naturally-occuring nucleic acid sequence.

[0040] According to a further embodiment, the present invention provides a composition that comprises a recombinant polynucleotide comprising: (a) an IgE leader sequence; and (b) a nucleic acid sequence comprising cytosine (C) or guanine (G) at the wobble position of at least one codon that encodes any of the amino acids alanine, arginine, leucine, proline, glutamic acid, glycine, isoleucine, serine, threonine, or valine where adenine (A), uracil (U) or thymine (T) occupy the wobble position of the corresponding codon of the naturally-occuring nucleic acid sequence.

[0041] According to a still further embodiment, the present invention provides a composition comprising: a polynucleotide comprising (a) a nucleic acid sequence having at least about 70% sequence identity to the nucleic acid sequence of SEQ ID NO:14; or (b) a nucleic acid sequence that hybridizes to SEQ ID NO:14 under stringent conditions.

BRIEF DESCRIPTION OF THE DRAWINGS

[0042] Figure 1 is a graph comparing the expression of protein from the recombinant HIV-1 6106 env gp160 gene prepared in accordance with an embodiment of the present invention relative to the expression of protein from the wild-type gp160 gene and gp160 gene having modified inhibitory sequences.

[0043] Figure 2 is a plasmid map of the plasmid construct of SEQ ID NO:7.

[0044] Figure 3 is a plasmid map of the plasmid construct of SEQ ID NO:8.

[0045] Figure 4 is a plasmid map of the plasmid construct of SEQ ID NO:9.

[0046] Figure 5 is a plasmid map of the plasmid construct of SEQ ID NO:10.

[0047] Figure 6 is a graph comparing expression of protein from IL-15 modified polypeptide (LP) with an IgE leader sequence in accordance with an embodiment of the present invention relative to the expression of protein from alternative IL-15 constructs in (a) RD cells; (b) COS7 cells, and (c) Hela cells.

[0048] Figure 7 is a graph comparing expression of protein from IL-15 modified polypeptide (LP) with an IgE leader sequence in accordance with an embodiment of the present invention relative to the expression of protein from alternative IL-15 constructs in (a) RD cells, and (b) 293 cells.

[0049] Figure 8 is a table comparing expression (fold increase) of protein from IL-15 modified polypeptide (LP) with an IgE leader sequence in accordance with an embodiment of the present invention relative to the expression of protein from alternative IL-15 constructs in RD cells, COS7 cells, Hela cells, and 293 cells.

[0050] Figure 9 is a graph comparing expression of protein from IL-15 modified polypeptide (LP) with an IgE leader sequence in accordance with an embodiment of the present invention relative to the expression of protein from alternative IL-15 constructs in a CTLL2 mouse cell proliferation assay.

[0051] Figure 10 is a graph comparing *in vivo* expression of protein from IL-15 modified polypeptide (LP) with an IgE leader sequence in accordance with an embodiment of the present invention relative to the expression of protein from alternative IL-15 over time.

[0052] Figure 11 is a plasmid map for the O-IL-15-IgE leader plasmid construct according to an embodiment of the present invention.

[0053] Figure 12 is a plasmid map for the LP-IL-15-IgE leader plasmid construct according to an embodiment of the present invention.

[0054] Figure 13 is a plasmid map for the BH-15-IgE leader plasmid construct according to an embodiment of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

[0055] An appropriate level of a protein in mammalian cells is essential *in vivo* for enhanced immunological and/or therapeutic responses, e.g., the use of the gene and its protein

product as an immunogen, DNA vaccine, co-immunogen, adjuvant, carrier protein or vector, therapeutic agent, diagnostic agent, therapeutic, immuno-prophylactic, immuno-therapeutic, etc., as well as for *in vitro* recombinant protein expression purposes, e.g., the use of the gene and its protein product in assays, tests, diagnostics, research tools, etc. The efficiency of a gene in expressing its protein product is a controlling factor in the attainment of appropriate levels of the protein in cells. Certain wild-type genes fail to provide appropriate protein levels in mammalian cells. The present invention is directed to improving the expression efficiency of such genes.

[0056] An effective IL-15 plasmid for DNA vaccination that secretes enhanced levels of IL-15 was unexpectedly identified. In particular, it was found that 1) the replacement of native signal peptide with the Human IgE leader sequence; 2) non preferred codons are replaced with either optimized or less preferred codons while preserving the native amino acid sequence; 3) the nucleotide sequence was modified to reduce the secondary mRNA structure for improved translation.

Modified Polynucleotides

[0057] As described herein, the inventors have devised modifiedpolynucleotides that provide unexpectedly improved gene expression in mammalian cells both *in vitro* and *in vivo* for various poorly-expressed genes.

[0058] These polynucleotides represent a new version of a wild-type gene. In particular, the inventors discovered that enhanced expression was unexpectedly provided by a new version of a gene in the form of a synthesized polynucleotide which comprises "surrogate codons" in the open reading frame (ORF) of the gene sequence, wherein the "surrogate codons" still encode identical amino acid residues (although biologically equivalent amino acid sequences/proteins, substantially identical amino acid sequences/proteins, etc. are also contemplated by the present invention, as described in further detail below).

[0059] A "surrogate codon", as used herein, refers to a codon for an ORF, other than the naturally occurring (i.e., wild-type) codon when that wild-type codon has an A, T (in the case of DNA) or U (in the case of RNA) in the wobble position, but encoding the same amino acid as that corresponding naturally occurring codon (i.e., the codon at the same position in the wild-type ORF). As used herein, the terms, "naturally-occurring and "wild-type" are used

interchangeably herein. In certain embodiments, the surrogate codon has C or G in its wobble position. In another embodiment, the surrogate codon is not a "preferred codon" as defined by Seed et al. The surrogate codons of the present invention are used in modified polynucleotides in place of corresponding disfavored codons, e.g., the naturally-occurring codon with A or T (if DNA) or U (if RNA) in the wobble position, of the wild-type form of the gene, for certain of the amino acids as described below. As used herein, the "wobble" position of a codon is the third nucleotide position of a codon triplet, as read in the 5' to 3' direction.

[0060] The invention disclosed herein utilizes a general approach directed to modified forms of a gene (i.e., recombinant polynucleotides). According to this general approach, modified polynucleotides are formed. These polynucleotides comprise a nucleic acid sequence comprising surrogate codons in place of at least some of the codons of the corresponding wild-type polynucleotide for the gene. For example, in accordance with embodiments of the invention, a modified polynucleotide comprises a nucleic acid sequence comprising surrogate codons in which the nucleotides cytosine (C) or guanine (G) occupy the wobble position in place of the corresponding nucleotides adenine (A) or uracil (U) or thymine (T) of a naturally-occurring polynucleotide that expresses substantially the same protein or polypeptide as said modified polynucleotide (or a functionally equivalent protein or polypeptide, as would be known to a person of skill in the art). The modified polynucleotide of the invention need not be an exact replica of the wild-type ORF wherein every codon having A or U in the wobble position is substituted with a surrogate codon. Merely a sufficient number of surrogate codons in place of naturally occurring codons to achieve enhanced gene expression is necessary.

[0061] A minimally sufficient number of surrogate codons or any number greater than that amount is contemplated by the invention. A suitable number of surrogate codons for a polynucleotide in accordance with the present invention is readily determined by one of skill by routine testing. It is not necessary that a predetermination of a specific number of surrogate codons be made. However, a predetermined number of replacements may be used in the interest of efficiency. For example, in constructing a polynucleotide of the invention, one may predetermine that a specified percentage of the codons of the ORF may be re-engineered, for example, about 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 96%, 97%, 98%, 99% or 100% of the codons, without limitation, may be the subject of re-engineering. Normally, at least 10% of the codons are the subject of re-engineering (e.g., 10% of the ORF is the new version of the gene while the remaining 90% is

the same as or functionally the same as the wild-type ORF). In certain embodiments, at least about 50% of the codons are the subject of re-engineering. In other embodiments, at least about 90% of the codons are the subject of re-engineering with surrogate codons.

[0062] The surrogate codons of the present invention are the non-naturally-occurring codons (of a gene) that encode for the following amino acids: alanine (Ala), asparagine or aspartate (Asx), cysteine (Cys), aspartate (Asp), glutamate (Glu), phenylalanine (Phe), glycine (Gly), histidine (His), isoleucine (Ile), lysine (Lys), leucine (Leu), methionine (Met), asparagine (Asn), proline (Pro), glutamine (Gln), arginine (Arg), serine (Ser), threonine (Thr), tyrosine (Tyr), or glutamine or glutamate (Glx). In a particular embodiment, the surrogate codons of the invention are the non-naturally-occurring codons (of a gene) with C or G in the wobble position that encode for any of alanine (Ala), asparagine or aspartate (Asx), cysteine (Cys), aspartate (Asp), glutamate (Glu), phenylalanine (Phe), glycine (Gly), histidine (His), lysine (Lys), leucine (Leu), methionine (Met), asparagine (Asn), proline (Pro), glutamine (Gln), arginine (Arg), serine (Ser), threonine (Thr), tyrosine (Tyr), or glutamine or glutamate (Glx), without limitation. A recombinant polynucleotide of the invention need not include surrogate codons for each amino acid encoded. Select surrogate codons that encode any number of amino acids may be predetermined for inclusion in the recombinant version of the gene provided that the objective of improving expression of the gene is achieved. A person of skill in the art would be able to determine through routine testing a minimally effective number. In one particular embodiment, each of the codons for alanine (Ala), asparagine or aspartate (Asx), cysteine (Cys), aspartate (Asp), glutamate (Glu), phenylalanine (Phe), glycine (Gly), histidine (His), isoleucine (Ile), lysine (Lys), leucine (Leu), methionine (Met), asparagine (Asn), proline (Pro), glutamine (Gln), arginine (Arg), serine (Ser), threonine (Thr), tyrosine (Tyr), or glutamine or glutamate (Glx) is replaced with a surrogate codon to form the recombinant version of the gene in accordance with an embodiment of the invention.

[0063] Accordingly, in the present invention, it is unnecessary to replace each codon that has A, T or U in the wobble position for every amino acid, substitute in specifically determined “preferred codons” or remove inhibitory sequences.

[0064] In certain embodiments, the surrogate codons used in the modified polynucleotides of the present invention are those that encode alanine, arginine, leucine, proline, glutamic acid, glycine, isoleucine, serine, threonine and valine. In other embodiments, the surrogate codons used in the polynucleotides of the invention are those that encode alanine, arginine, leucine, proline, glycine, isoleucine, serine, threonine and valine. In one particular

embodiment, the surrogate codons used in the modified polynucleotides of the invention are those that encode alanine, arginine, leucine, proline, glycine, serine, threonine and valine.

[0065] In accordance with an embodiment of the invention, the surrogate codons are a randomized selection of at least about 10% of the codons in said modified polynucleotide that encode for any of the amino acids alanine, arginine, leucine, proline, glycine, isoleucine, serine, threonine and valine. In accordance with another embodiment, the surrogate codons are a randomized selection of at least about 50% of the codons in said polynucleotide that encode for any of the amino acids alanine, arginine, leucine, proline, glycine, isoleucine, serine, threonine and valine. In a further embodiment, the surrogate codons are a randomized selection of at least about 90% of the codons in said polynucleotide that encode for any of the amino acids alanine, arginine, leucine, proline, glycine, isoleucine, serine, threonine and valine. In yet another embodiment, the surrogate codons are each of the codons in said polynucleotide (i.e., 100%) that encode for the amino acids alanine, arginine, leucine, proline, glycine, isoleucine, serine, threonine and valine.

[0066] The present invention contemplates embodiments directed to any gene that is poorly expressed or any gene for which improved levels of protein expression is desirable for *in vivo* and/or *in vitro* uses. For example, a subject gene may be a viral, bacterial, protist, fungal, plant or animal gene, without limitation. Any such gene that is poorly expressed in mammalian cells is contemplated by the present invention.

[0067] In the case of viral genes, without limitation, the viral gene may be associated with a DNA (double stranded or single stranded) or RNA (double stranded or single stranded) virus , without limitation. Viral genes of viruses from any viral family are contemplated by the present invention, including, for example, Adenoviridae, Arenaviridae, Arterivirus, Astroviridae, Baculoviridae, Badnavirus, Barnaviridae, Brinaviridae, Bromoviridae, Bunyaviridae, Caliciviridae, Capillovirus, Carlavirus, Caulimovirus, Circoviridae, Closteroviridae, Comoviridae, Coronaviridae, Corticoviridae, Cystoviridae, Deltavirus, Dianthovirus, Enamovirus, Filoviridae, Flaviviridae, Furovirus, Fuselloviridae, Geminiviridae, Hepadnaviridae, Herpesviridae, Hordeivirus, Hypoviridae, Idaeovirus, Inoviridae, Iridoviridae, Leviviridae, Lipothrixviridae, Luteovirus, Machlomovirus, Marafivirus, Microviridae, Myoviridae, Necrovirus, Nodaviridae, Orthomyxoviridae, Papovaviridae, Paramyxoviridae, Partitiviridae, Parvaviridae, Phycodnaviridae, Picornaviridae, Plasmaviridae, Podoviridae, Polydnaviridae, Potexvirus, Potyviridae, Poxviridae, , Reoviridae, Retroviridae, Rhabdoviridae, Rhizidiovirus, , Sequiviridae, Siphoviridae, Sobemovirus, Tectiviridae,

Tenuivirus, Tetraviridae, Tobamovirus, Tobravirus, Togavridae, Tombusviridae, Totiviridae, Trichovirus, Tymovirus, Umbravirus, Viroids, Mononegavirales, Tailed Phages, and as yet unclassified viruses, without limitation.

[0068] In one embodiment of the invention, a viral gene is associated with lentiviruses, retroviruses, herpes viruses, adenoviruses, adeno-associated viruses, vaccinia virus, or baculovirus, without limitation. In certain embodiments, viral genes include, for example, those of Human immunodeficiency virus, Simian immunodeficiency virus, Respiratory syncytial virus, Parainfluenza virus types 1-3, Influenza virus, Herpes simplex virus, Human cytomegalovirus, Hepatitis A virus, Hepatitis B virus, Hepatitis C virus, Human papillomavirus, poliovirus, rotavirus, caliciviruses, Measles virus, Mumps virus, Rubella virus, adenovirus, rabies virus, vesicular stomatitis virus, canine distemper virus, rinderpest virus, Human metapneumovirus, avian pneumovirus (formerly turkey rhinotracheitis virus), Hendra virus, Nipah virus, coronavirus, parvovirus, infectious rhinotracheitis viruses, feline leukemia virus, feline infectious peritonitis virus, avian infectious bursal disease virus, Newcastle disease virus, Marek's disease virus, porcine respiratory and reproductive syndrome virus, equine arteritis virus and various Encephalitis viruses, without limitation.

[0069] Specific viral genes contemplated by the present invention include, for example, any of the genes of HIV or any of the genotypes of HPV, including high-risk and low-risk genotypes. For example, genes of HIV contemplated by the invention include gag, pol, env, tat, rev, vif, nef, vpr, vpu and vpx, without limitation. Genes of HPV contemplated by the invention include, for example, E1, E2, L1, L2, E6 and E7 without limitation. The genotypes of HPV contemplated by the present invention include, for example, high-risk genotypes, such as HPV 16, 18, 31, 33, 45, 52, 56 or 58 and low-risk genotypes, such as 6 and 11, without limitation. According to an embodiment, the gene is the human papillomavirus 16 (HPV16) E7 gene (E7), or human immuno-deficiency virus (HIV-1) gag gene (gag) or gp160 envelope gene (env). Compositions, fusion constructs or any other multi-gene structures containing any combination of the foregoing are also contemplated by the present invention.

[0070] Specific bacterial genes include the genes of any bacterial species, including for example, without limitation, *Haemophilus influenzae* (both typable and nontypable), *Haemophilus somnus*, *Moraxella catarrhalis*, *Streptococcus pneumoniae*, *Streptococcus pyogenes*, *Streptococcus agalactiae*, *Streptococcus faecalis*, *Helicobacter pylori*, *Neisseria meningitidis*, *Neisseria gonorrhoeae*, *Chlamydia trachomatis*, *Chlamydia pneumoniae*, *Chlamydia* psittaci, *Bordetella pertussis*, *Alloiococcus otiditis*, *Salmonella typhi*, *Salmonella*

typhimurium, *Salmonella choleraesuis*, *Escherichia coli*, *Shigella*, *Vibrio cholerae*, *Corynebacterium diphtheriae*, *Mycobacterium tuberculosis*, *Mycobacterium avium-Mycobacterium intracellulare* complex, *Proteus mirabilis*, *Proteus vulgaris*, *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Clostridium tetani*, *Leptospira interrogans*, *Borrelia burgdorferi*, *Pasteurella haemolytica*, *Pasteurella multocida*, *Actinobacillus pleuropneumoniae* and *Mycoplasma gallisepticum*.

[0071] Further, the present invention is applicable to any gene which is a suitable subject for improved efficiency in the manner of the present invention, i.e., engineering a recombinant polynucleotide for the gene with surrogate codons in place of naturally occurring codons with A or U in the wobble position. Thus, although the term "poorly-expressed" genes is used throughout, the present invention is by no means intended to be limited to genes that meet some threshold requirement of poor expression. Instead, modified polynucleotides directed to poorly-expressed genes are merely exemplary to illustrate the dramatic improvement in protein levels in the circumstances where such improvement is most pertinent. Therefore, the present invention contemplates applicability to genes that may not be considered to be poorly-expressed by persons skilled in the art, as well as to those that are generally considered or proven to be poorly-expressed, without limitation.

[0072] Upon selection of a desired target gene of a desired species (*e.g.*, the E1 gene of HPV 16), a person of skill in the art, based upon the guidance provided herein, would be able to formulate the sequence of a desired recombinant in accordance with an embodiment of the present invention. The sequencing is performed for example, by hand or is computer-assisted. A person of skill in the art may make a replacement at each disfavored wobble position, or at some percentage of the disfavored wobble positions. For example, the first 50% of disfavored wobble positions or the second 50% of disfavored wobble positions. The modified sequence is tested by routine methods to determine whether the percentage change provides a desired level of expression. The examples herein provide guidance as to such testing, however, it is well within the abilities of a person of skill in the art to conduct such routine testing in a variety of ways. In certain embodiments, replacement is made at each disfavored wobble position, thus eliminating the need to select certain portions of the gene and certain percentages of wobble positions for replacement. Once the sequence of the polynucleotide is determined, it is well within the ability of a person of skill in the art to prepare the modified polynucleotide using well known techniques and methods, as further described in the examples below.

[0073] Several poorly-expressed viral genes illustrate the benefits of the present invention. For example, the following wild-type viral genes demonstrate poor expression in mammalian cells: human papillomavirus 16 (HPV16) E7, human immuno-deficiency virus type-1 (HIV-1) gag and gp160 (envelope) (hereafter denoted E7, gag, and env, respectively). In each of these wild-type genes, the naturally occurring nucleic acid sequences of the genes are AU rich and biased toward "disfavored codons" (containing an A or U in the 3d or "wobble" position of the codon nucleotide triplet). As noted above, mammalian genes that express proteins at high levels have a G/C preference in the wobble position. Thus, these wild-type genes with A or U in the wobble position may not be handled efficiently by the mammalian translational machinery.

[0074] Further, as discussed above, separately from the translational mechanisms accounting for poorly-expressed genes, there have been various AU rich RNA instability sequences discovered in several messenger RNAs (mRNAs) which do not directly impact the translatability of a given mRNA but limit protein expression by increasing mRNA turnover. In addition, several specific "inhibitory" sequences contained within the HIV-1 gag ORF have been described (see Pavlakis) which limit the expression levels of gag by inhibiting nuclear export of these transcripts. Codons encompassing these inhibitory sequences are difficult to selectively replace to eliminate the inhibitory effect because the sequence motifs that define either instability or inhibitory sequences are not easily identified. Moreover, several genes (e.g. E7 and En among others) which appear to also contain inhibitory sequences have not yet been mapped to identify the location of inhibitory sequences and there are no straightforward prescriptions from the gag work to predict how to eliminate inhibitory sequences from these genes.

[0075] According to an embodiment of the present invention, codons throughout a gene sequence are replaced (e.g., surrogate codons replace wild-type codons in a modified construct) without the need to identify and then mutate inhibitory sequences (as performed for gag) and without altering every codon by use of preferred codons (as performed for env). When a naturally occurring disfavored codon (e.g., with A or U in the wobble position) is replaced with (i.e., its position in the modified form is occupied by) a "surrogate codon" encoding the same amino said, there is an opportunity to eradicate inhibitory sequence(s), instability sequence(s), and/or provide codons that are more efficiently translated than their naturally occurring counterparts.

[0076] It was surprisingly discovered that alteration of all possible codons and utilization of "preferred" codons was not necessary to achieve improved protein levels expressed by the genes cited above. Thus, it is possible to exploit the degeneracy of the genetic code to develop recombinant polynucleotides with improved protein expression of a gene relative to the wild-type polynucleotide of the gene (or other recombinant polynucleotides for the gene). Thus, it is unnecessary to construct a complete "codon optimized" version of gp120 envelope as previously described (see Haas et al., Andre et al.) in which non-preferred wild-type codons from env were replaced with "preferred" codons to enhance protein levels expressed by the gene.

[0077] Table I below lists non-limiting examples of surrogate codons of the present invention. In particular, Table I shows the surrogate codons for ten of the twenty L-amino acids that have been utilized as replacements for existing disfavored codons, according to an implementation of the present invention. In accordance with this embodiment of the invention, codons encoding the remaining ten amino acids were not replaced by surrogate codons in the modified form of the gene.

TABLE I-SURROGATE CODONS

Codon	Amino acid encoded	Codon	Amino acid encoded
GCG	Alanine	GAG	Glutamic Acid
CGG or AGG	Arginine	GGG	Glycine
CTC	Leucine	ATT	Isoleucine
CCT or CCG	Proline	TCC	Serine
ACG	Threonine	GTC	Valine

[0078] In accordance with an embodiment of the present invention, recombinant polynucleotides were prepared in which disfavored codons (A or U at the wobble position) were replaced by the surrogate codons listed in Table I above for the amino acid encoded by the disfavored codon, and the corresponding new (i.e. modified) nucleic acid sequence was created by joining oligonucleotides encoding the new sequence and assembling the fragments to create the modified polynucleotide comprising the new sequence.

[0079] The recombinant ORF was cloned into a plasmid DNA expression vector that allowed *in vitro* expression-studies for comparing the levels of protein expression of the modified polynucleotide and the wild-type polynucleotide. Transient transfection assays (data not shown) performed with several cell lines revealed increases in protein expression levels for three gene products (i.e., E7, gag, and env) when their gene sequence was modified as described above. The increased protein expression (as measured by Western blot, ELISA and the like) demonstrated by the altered codon constructs compared to wild-type (naturally occurring) construct for three different genes indicated that this method is applicable to a variety of poorly expressed proteins.

[0080] In recognition that several codon choices are possible for some of the twenty amino acids, for example, the amino acids alanine, arginine, glycine, glutamic acid, isoleucine, leucine, proline, serine, threonine, and valine, an embodiment of the present invention is directed to the codons encoding those amino acids. Thus, in accordance with an embodiment of the invention, a modified polynucleotide has a nucleic acid sequence, which differs from that of the wild-type sequence, in which each codon, that corresponds to a naturally-occurring codon having A, U or T in the wobble position, encoding alanine is GCG, each codon encoding arginine is CGG or AGG, each codon encoding leucine is CTC, each codon encoding proline is CCT or CCG, each codon encoding glutamic acid is GAG, each codon encoding glycine is GGG, each codon encoding isoleucine is ATT, each codon encoding serine is TCC, each codon encoding threonine is ACG, and each codon encoding valine is GTC.

[0081] In certain other embodiments, codons for amino acids other than the ten listed above also serve as surrogate codons. In other words, replacement of the naturally-occurring codons, with A, U or T in the wobble position, encoding other amino acids is contemplated. It is also contemplated that certain embodiments of the invention provide surrogate codons for only some of the ten amino acids listed in Table I. Upon grasping the concept of the invention as fully described herein, a person skilled in the art would routinely be able to determine a minimally or optimally desired number of codons through routine methods, based upon the guidance provided herein. In certain embodiments, the polynucleotides of the present invention comprise surrogate codons for just the nine amino acids, alanine, arginine, glycine, isoleucine, leucine, proline, serine, threonine, and valine in place of each of the corresponding codons having A or U in the wobble position. It should be noted, however, that any changes to those changed codons and/or the other codons that permit the

protein to retain its functionality are contemplated by the present invention. Examples of such changes are provided below.

[0082] The modified polynucleotides of the invention are prepared in any suitable manner as would be known to persons skilled in the art. For example, the present invention contemplates the use of chemical synthesis, nucleotide substitution, codon substitution, DNA libraries, mutagenesis, isolation and purification from native entity, etc. and any combinations thereof, without limitation.

[0083] In one embodiment, a full length polynucleotide sequence is determined by selecting surrogate codons for the disfavored codons. This may be done by hand, computer-assisted or any other method. Once the desired sequence is determined, then oligonucleotides comprising fragments of the determined sequence are obtained or prepared. Such oligonucleotides are readily obtained from commercial vendors, such as Invitrogen™ (Carlsbad, CA). The fragments are selected such that they can form a staggered, overlapping arrangement. The modified polynucleotides are synthesized by joining oligonucleotides that comprise fragments of the recombinant nucleic acid sequence. The fragments are hybridized and subsequently filled in by a DNA polymerase (such as Pfx Turbo, Invitrogen). This staggered, overlapping arrangement of the fragments is then ligated, for example, using a heat stable ligase (Ampligase).

[0084] Specific protocols for preparing the polynucleotides of the present invention are provided in the Examples below. These specific protocols are merely illustrative. A person skilled in the art would readily be able to employ a variety of suitable techniques to accomplish the objectives of the present invention, upon grasping the inventive concepts disclosed herein. All such suitable techniques for preparing recombinant polynucleotides are contemplated by the present invention.

According to an embodiment of the invention, the leader sequence of the polynucleotide is altered or substituted with a non-native leaders sequence. For example, a non-native leader sequence is added to a modified polynucleotide of the presents invention and replaces the native leader sequence of the polynucleotide. Thus, the present invention contemplates a modified polynucleotide comprising a non-native leader sequence. The non-native leader sequence may be any suitable sequence or combination thereof that provides enhanced expression. It has been suprisingly found that the combination of modifying the polynucleotides using surrogate codons as described herein with the use of a non-native

leader sequence provides synergistically improved expression, as described in Example 5 below. The non-native leader sequence may be human non-native leader sequence. The non-native leader sequence may be an immunoglobulin leader sequence.

[0085] According to an embodiment, the non-native leader sequence is (a) an IgE leader sequence or (b) a leader sequence that hybridizes to an IgE leader sequence under stringent conditions. According to another embodiment, the non-native leader sequence is: (a) a leader sequence having SEQ ID NO:11; or (b) a leader sequence that hybridizes to SEQ ID NO:11 under stringent conditions. The non-native leader sequence has at least 70%, 80%, 90%, 95%, 97%, 98% or 99% sequence identity to the nucleic acid sequence of SEQ ID NO:11 according to other embodiments of the present invention. According to another embodiment, the non-native leader sequence has the nucleic acid sequence of SEQ ID NO:11. A person skilled in the art would readily be able to construct or alter a polynucleotide to include a non-native leader sequence in the manner of the present invention, based upon the guidance provided herein.

[0086] The polynucleotides are prepared in various forms (e.g., single-stranded, double-stranded, vectors, probes, primers) as desired. The term "polynucleotide" includes any strand of DNA and RNA, single stranded and double stranded, and also their analogs, such as those containing modified backbones. The term "modified polynucleotide" as used herein, describes any strand of DNA or RNA, including single or double stranded, that are recombinantly prepared or that have been altered from their naturally-occurring state (through insertion, deletion, substitution, etc.) with surrogate codons or as otherwise consistent with the embodiments of the present invention as described herein. The DNA may be of any type, such as cDNA, genomic DNA, synthesized DNA, isolated DNA or a hybrid thereof. The RNA may be also be of any type RNA molecule such as mRNA. The constructs of the present invention contemplate any regulator elements necessary or desireable for expression of the sequence, such as a promoter, an initiation codon, a stop codon, and a polyadenylation signal, for example, without limitation. Any suitable enhancer is also contemplated by the present invention. Non-limiting exemplary enhancers include human Actin, human Myosin, human Hemolobin, human muscle creatine, and viral enhancers such as those from CMV, RSV and EBV.

[0087] Several specific recombinant polynucleotides, including specific nucleic acid sequences, for various viral genes are provided herein. These are merely exemplary and the invention is not intended to be limited thereto. Rather, the inventive concept is broadly

applicable as described herein. Moreover, the present invention contemplates modified polynucleotides which are variations on any of the recombinant polynucleotides described herein, such as, for example, the specifically disclosed sequences, without limitation. For example, these would include variations wherein the variant nucleic sequence encodes a different amino acid sequence than the specifically disclosed sequence, however, the functionality of the different amino acid sequence is the same as that encoded by the sequence described herein.

[0088] According to an embodiment the modified polynucleotide expresses a viral polypeptide. The present invention contemplates modified polynucleotides from any agent or organism, such as pathogenic organisms, for example, HIV, HSV, HCV, WNV or HBV. For example, according to an embodiment immunogenic compositions are prepared from the pathogenic organisms for the purpose of immunizing an individual against the pathogen. For example, the modified polynucleotide may express the viral polypeptides HPV16, HIV-1 or gp160 or any combinations thereof, without limitation. According to an embodiment, a modified polynucleotide may comprise the ORF for HPV16 E7 gene. According to another embodiment, a modified polynucleotide comprises the ORF for the HIV-1 gag gene. According to another embodiment, a modified polynucleotide comprises the ORF for the gp160 envelope gene.

[0089] According to an embodiment, the modified polynucleotide encodes for a cytokine, growth factor, lymphokine, such as alpha-interferon, gamma-interferon, GM-CSF, platelet derived growth factor, TNF, EGF, ILA, IL-2, IL-4, IL-6, IL-10, IL-12, IL-15 as well as fibroblast growth factor, surface active agents such as immune-stimulating complexes (ISCOMS), Freunds incomplete adjuvant, LPS analog inlcuding monophsphoryl Lipid A (WL), muramyl peptides, quinone analogs and vesicles such as squalene and squalene, and hyaluronic acid. Any cytokine is contemplated by the present invention. According to another embodiment, the cytokine is an interleukin. According to another embodiment, polynucleotide encodes for IL-15 or a peptide or polypeptide having the activity of IL-15. According to another embodiment, the modified polynucleotide encodes for IL-15. According to another embodiment, the modified polynucleotide comprises the nucleic acid sequence of any of SEQ ID NOS: 12-16. According to another embodiment, the modified polynucleotide comprises the nucleic acid sequence of SEQ ID NO:14. The nucleotide and amino acid sequences of IL-15 are well known and set forth in Campbell, et al. (1987) Proc. Natl. Acad. Sci. USA 84:6629-6633, Tanabe, et al. (1987) J.Biol. Chem. 262:16580-16584, Campbell, et al. (1988) Eur. J. Biochem. 174:345-352, Azuma, et al. (1986) Nucl. Acids Res. 14:9149-

9158, Yokota, et al. (1986) Proc. Natl. Acad. Sci. USA 84:7388-7392, and accession code Swissprot PO5113, which are each incorporated herein by reference in their entirety.

[0090] For example, according to an embodiment of the present invention, the modified polynucleotides comprise a nucleic acid sequence that is identical to any of the reference sequences of odd numbered SEQ ID NOS:1-5 or any of SEQ ID NOS:12-16 (which are sequences modified in accordance with the invention), that is 100% identical, or it may include a number of nucleotide alterations (e.g. at least 99%, 98%, 97%, 96%, 95%, 94%, 90%, 85%, 80%, 70%, or 60% identical, etc.) as compared to the reference sequence. Such alterations are selected from the group consisting of at least one nucleotide deletion, substitution, including transition and transversion, or insertion, and wherein said alterations occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among the nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence. The number of nucleotide alterations is determined by multiplying the total number of nucleotides in any of odd numbered SEQ ID NOS:1-5 or any of SEQ ID NOS:12-16 by the numerical percent of the respective percent identity (divided by 100) and subtracting that product from said total number of nucleotides in said sequence.

[0091] Certain embodiments of the invention relate to polynucleotides and sequence modifications thereof. In one embodiment, a polynucleotide of the invention is a polynucleotide comprising a nucleotide sequence having functional equivalency and at least about 95% identity to a nucleotide sequence chosen from one of the odd numbered SEQ ID NO:1-5 or any of SEQ ID NOS:12-16 , a degenerate variant thereof, or a fragment thereof. As defined herein, a "degenerate variant" is defined as a polynucleotide that differs from the nucleotide sequence shown in the odd numbered SEQ ID NOS:1-5 or any of SEQ ID NOS:12-16 (and fragments thereof) due to degeneracy of the genetic code, but still encodes the same protein (*e.g.*, the even numbered SEQ ID NOS: 2-6) as that encoded by the nucleotide sequence shown in the odd numbered SEQ ID NOS: 1-5 or any of SEQ ID NOS:12-16 .

[0092] In other embodiments, the polynucleotide is a complement to a nucleotide sequence chosen from one of the odd numbered SEQ ID NOS: 1-5 or any of SEQ ID NOS:12-16 , a degenerate variant thereof, or a fragment thereof. In yet other embodiments, the polynucleotide is selected from the group consisting of DNA, chromosomal DNA, cDNA and RNA and may further comprises heterologous nucleotides. In another embodiment, an

isolated polynucleotide hybridizes to a nucleotide sequence chosen from one of odd numbered SEQ ID NOS: 1-5 or any of SEQ ID NOS:12-16 , a complement thereof, a degenerate variant thereof, or a fragment thereof, under high stringency hybridization conditions. In yet other embodiments, the polynucleotide hybridizes under intermediate stringency hybridization conditions.

[0093] It will be appreciated that polynucleotides of the present invention are obtained from natural sources (and then altered) or are synthetic or semi-synthetic or some combination thereof. Furthermore, the nucleotide sequence is related by mutation, including single or multiple base substitutions, deletions, insertions and inversions, to a naturally occurring sequence, provided always that the nucleic acid molecule comprising such a sequence is capable of being expressed as a functionally equivalent polypeptide as described above. A nucleic acid molecule of the invention is RNA, DNA, single stranded or double stranded, linear or covalently closed circular form. In certain embodiments, the nucleotide sequence has expression control sequences positioned adjacent to it, such control sequences usually being derived from a heterologous source. In other embodiments, the recombinant expression of a nucleic acid sequence of the invention include a stop codon sequence, such as TAA, at the end of the nucleic acid sequence.

[0094] According to an embodiment, the invention also includes polynucleotides capable of hybridizing under reduced stringency conditions. According to another embodiment the invention includes polynucleotides capable of hybridizing under stringent conditions, and under another embodiment the present invention includes polynucleotides capable of hybridizing under highly stringent conditions, to the polynucleotides described above. Examples of stringency conditions are shown in the Stringency Conditions Table below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

TABLE II-HYBRIDIZATION STRINGENCY CONDITIONS

Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp)I	Hybridization Temperature and BufferH	Wash Temperature and BufferH
A	DNA:DNA	> 50	65 C; 1xSSC -or- 42 C; 1xSSC, 50%	65 C; 0.3xSSC

Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp)I	Hybridization Temperature and BufferH	Wash Temperature and BufferH
			formamide	
B	DNA:DNA	< 50	TB; 1xSSC	TB; 1xSSC
C	DNA:RNA	> 50	67 C; 1xSSC -or- 45 C; 1xSSC, 50% formamide	67 C; 0.3xSSC
D	DNA:RNA	< 50	TD; 1xSSC	TD; 1xSSC
E	RNA:RNA	> 50	70 C; 1xSSC -or- 50 C; 1xSSC, 50% formamide	70 C; 0.3xSSC
F	RNA:RNA	< 50	TF; 1xSSC	Tf; 1xSSC
G	DNA:DNA	> 50	65 C; 4xSSC -or- 42 C; 4xSSC, 50% formamide	65 C; 1xSSC
H	DNA:DNA	< 50	TH; 4xSSC	TH; 4xSSC
I	DNA:RNA	> 50	67 C; 4xSSC -or- 45 C; 4xSSC, 50% formamide	67 C; 1xSSC
J	DNA:RNA	< 50	TJ; 4Xssc	TJ; 4xSSC
K	RNA:RNA	> 50	70 C; 4xSSC -or- 50 C; 4xSSC, 50% formamide	67 C; 1xSSC
L	RNA:RNA	< 50	TL; 2Xssc	TL; 2xSSC
M	DNA:DNA	> 50	50 C; 4xSSC -or- 40 C; 6xSSC, 50% formamide	50 C; 2xSSC
N	DNA:DNA	< 50	TN; 6xSSC	TN; 6xSSC
O	DNA:RNA	> 50	55 C; 4xSSC -or-	55 C; 2xSSC

Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp)I	Hybridization Temperature and BufferH	Wash Temperature and BufferH
			42 C; 6xSSC, 50% formamide	
P	DNA:RNA	< 50	TP; 6xSSC	TP; 6xSSC
Q	RNA:RNA	> 50	60 C; 4xSSC -or- 45 C; 6xSSC, 50% formamide	60 C; 2xSSC
R	RNA:RNA	< 50	TR; 4xSSC	TR; 4xSSC

[0095] bpI: The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarities.

[0096] bufferH: SSPE (1xSSPE is 0.15M NaCl, 10mM NaH2PO4, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

[0097] TB through TR: The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be about 5-10 C less than the melting temperature (Tm) of the hybrid, where Tm is determined according to the following equations. For hybrids less than 18 base pairs in length, Tm(C) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, Tm(C) = 81.5 + 16.6(log10[Na+]) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na+] is the concentration of sodium ions in the hybridization buffer ([Na+] for 1xSSC = 0.165 M).

[0098] Additional examples of stringency conditions for polynucleotide hybridization are provided in Sambrook, J., E.F. Fritsch, and T. Maniatis, 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters

9 and 11, and Current Protocols in Molecular Biology, 1995, F.M. Ausubel et al., eds., John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

[0099] In certain embodiments, modifications and changes are made in the structure of a polynucleotide of the present invention while retaining functional equivalency (such as immunogenicity, therapeutic benefit, binding affinity, etc) of the protein product encoded by the modified polypeptide. Such modifications and changes are fully contemplated by the present invention. For example, without limitation, certain amino acids can be substituted for other amino acids, including nonconserved and conserved substitution, in an amino sequence without appreciable loss of functionality/utility (e.g., immunogenicity, therapeutic benefit, etc.) and thus in the polynucleotide the corresponding codon encoding those amino acids can be changed accordingly, as would be understood by a person skilled in the art.

[0100] In fact, as it is the interactive capacity and nature of a polypeptide that defines that polypeptide's biological functional activity, a number of amino acid sequence substitutions are made in a polypeptide sequence, and thus its underlying nucleic acid coding sequence, and nevertheless obtain a polypeptide with like properties. The present invention contemplates any changes to the structure of the nucleic acid sequences encoding the subject polypeptides or proteins, wherein the polypeptide or protein retains its functionality or a biologically equivalent functionality. A person of ordinary skill in the art would be readily able to routinely modify the disclosed polypeptides and polynucleotides accordingly, based upon the guidance provided herein, while remaining consistent with the inventive concept and the purposes of the present invention (e.g., the use of the surrogate codons to enhance expression).

[0101] In making such changes, any techniques known to persons of skill in the art are utilized. For example, without intending to be limited thereto, the hydropathic index of amino acids can be considered, as described below with regard to the recombinant proteins and polypeptides of the present invention. The importance of the hydropathic amino acid index in conferring interactive biologic function on polypeptides is generally understood in the art. Kyte et al. 1982. J. Mol. Bio. 157:105-132.

[0102] According to further implementations of the invention, the polynucleotides comprise a polynucleotide library, such as a cDNA library. The preparation of such a library of polynucleotides is well known to persons of skill in the art. A person skilled in the art could readily prepare such a library in accordance with an embodiment of the present invention,

using well known techniques and based upon the guidance provided herein. As described in further detail below, the polynucleotides of the invention are used in any suitable context, such as in vectors, immunogenic compositions, therapeutic compositions, recombinant cells and cell lines, assays, kits, tools, etc., as would be well understood by persons skilled in the art.

Proteins and Polypeptides

[0103] The present invention also provides recombinant proteins or polypeptides encoded by the modified polynucleotides of the invention described herein. For example, in certain embodiments, a recombinant polypeptide or protein of the invention is a recombinant that is identical to the reference sequence of even numbered SEQ ID NOS: 2-6 or amino acid sequences encoded by any of odd numbered SEQ ID NOS:1-5 or any of SEQ ID NOS: 12-16 (which are sequences modified in accordance with the invention), that is, 100% identical, or it may include a number of amino acid alterations as compared to the reference sequence such that the percent identity is less than 100%. Such alterations include at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion. The alterations occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between those terminal positions, interspersed either individually among the amino acids in the reference amino acid sequence or in one or more contiguous groups within the reference amino acid sequence.

[0104] Thus, the invention also provides proteins having sequence identity to an amino acid sequence of the invention, (e.g. even numbered SEQ ID NOS: 2-6 or proteins encoded by any of odd numbered SEQ ID NOS:1-5 or any of SEQ ID NOS:12-16). Depending on the particular sequence, the degree of sequence identity is greater than 60% (e.g., 60%, 70%, 80%, 85%, 90%, 94%, 95%, 97%, 98%, 99%, 99.9% or more). These homologous proteins include mutants and allelic variants.

[0105] In certain embodiments of the invention, the proteins or polypeptides (e.g., immunological portions and biological equivalents) generate antibodies. Specifically, the antibodies to the polypeptides protect from a challenge, such as intranasal. In further preferred embodiments, the polypeptides exhibit such protection for homologous strains and at least one heterologous strain. The polypeptide may be selected from even numbered SEQ ID NOS: 2-6 or amino acid sequences encoded by any of odd numbered SEQ ID NOS:1-5 or any of SEQ ID NOS: 12-16 , or the polypeptide may be any immunological fragment or biological equivalent of the listed polypeptides. According to an embodiment,

the polypeptide is selected from any of the even numbered SEQ ID NOS: 2-6 or amino acid sequences encoded by any of odd numbered SEQ ID NOS:1-5 or any of SEQ ID NOS: 12-16.

[0106] In certain embodiments, the invention relates to allelic or other variants of the polypeptides, which are biological equivalents. Suitable biological equivalents exhibit the ability to (1) elicit antibodies; (2) react with the surface of homologous strains and/or heterologous strains; (3) confer protection against a live challenge; and/or (4) prevent colonization.

[0107] Suitable biological equivalents have at least about 60% to about 100% similarity to one of the polypeptides specified herein (i.e., the even numbered SEQ ID NOS: 2-6 or amino acid sequences encoded by any of odd numbered SEQ ID NOS:1-5 or any of SEQ ID NOS: 12-16), provided the equivalent is capable of eliciting substantially the same immunogenic properties as one of the proteins of this invention.

[0108] Alternatively, the biological equivalents have substantially the same immunogenic properties of one of the proteins in the even numbered SEQ ID NOS: 2-6 or amino acid sequences encoded by any of odd numbered SEQ ID NOS:1-5 or any of SEQ ID NOS: 12-16. According to certain embodiments of the present invention, the biological equivalents have the same immunogenic properties as the even numbered SEQ ID NOS 2-6 or amino acid sequences encoded by any of odd numbered SEQ ID NOS:1-5 or any of SEQ ID NOS: 12-16 .

[0109] The biological equivalents are obtained by generating variants and modifications to the proteins of this invention. These variants and modifications to the proteins are obtained by altering the amino acid sequences by insertion, deletion or substitution of one or more amino acids. The amino acid sequence is modified, for example by substitution in order to create a polypeptide having substantially the same or improved qualities. In a particular embodiment, a means of introducing alterations comprises making predetermined mutations of the nucleic acid sequence of the polypeptide by site-directed mutagenesis.

[0110] Modifications and changes can be made in the structure of a polypeptide of the present invention while retaining functional equivalency (such as immunogenicity, therapeutic benefit, binding affinity, etc). Such modifications and changes are fully contemplated by the present invention. For example, without limitation, certain amino acids can be substituted for other amino acids, including nonconserved and conserved substitution, in a sequence without appreciable loss of functionality/utility (e.g.,

immunogenicity, therapeutic benefit, etc.). The present invention contemplates any changes to the structure of the polypeptides herein, as well as the nucleic acid sequences encoding said polypeptides, wherein the polypeptide retains its functionality or a biologically equivalent functionality.

[0111] In making such changes, any techniques known to persons of skill in the art may be utilized. For example, without intending to be limited thereto, the hydropathic index, hydrophilicity, and the like, of amino acids are considered (Kyte *et al.* 1982. *J. Mol. Bio.* 157:105-132, U.S. Patent 4,554,101).

[0112] Biological equivalents of a polypeptide are also prepared using site-specific mutagenesis. Site-specific mutagenesis is a technique useful in the preparation of second generation polypeptides, or biologically functional equivalent polypeptides or peptides, derived from the sequences thereof, through specific mutagenesis of the underlying DNA. Such changes are desirable where amino acid substitutions are desirable. The technique further provides a ready ability to prepare and test sequence variants, for example, incorporating one or more of the foregoing considerations, by introducing one or more nucleotide sequence changes into the DNA. Site-specific mutagenesis allows the production of mutants through the use of specific oligonucleotide sequences which encode the DNA sequence of the desired mutation, as well as a sufficient number of adjacent nucleotides, to provide a primer sequence of sufficient size and sequence complexity to form a stable duplex on both sides of the deletion junction being traversed. Typically, a primer of about 17 to 25 nucleotides in length is used, with about 5 to 10 residues on both sides of the junction of the sequence being altered.

[0113] In general, the technique of site-specific mutagenesis is well known in the art. As will be appreciated, the technique typically employs a phage vector which can exist in both a single stranded and double stranded form. Typically, site-directed mutagenesis in accordance herewith is performed by first obtaining a single-stranded vector which includes within its sequence a DNA sequence which encodes all or a portion of the polypeptide sequence selected. An oligonucleotide primer bearing the desired mutated sequence is prepared (*e.g.*, synthetically). This primer is then annealed to the single-stranded vector, and extended by the use of enzymes such as *E. coli* polymerase I Klenow fragment, in order to complete the synthesis of the mutation-bearing strand. Thus, a heteroduplex is formed wherein one strand encodes the original non-mutated sequence and the second strand bears the desired mutation. This heteroduplex vector is then used to transform appropriate cells such as *E. coli* cells and clones are selected which include recombinant vectors bearing

the mutation. Commercially available kits come with all the reagents necessary, except the oligonucleotide primers.

[0114] The polypeptides of the invention include any protein or polypeptide comprising substantial sequence similarity and/or biological equivalence to a protein having an amino acid sequence of any of the proteins of the embodiments of the invention such as any of even numbered SEQ ID NOS 2-6 or proteins encoded by any of odd numbered SEQ ID NOS:1-5 and 12-16. In addition, the polypeptides of the invention are not limited to a particular source. Also, the polypeptides can be prepared recombinantly using any such technique in accordance with the purpose of the invention as described herein, as is well within the skill in the art, based upon the guidance provided herein, or in any other synthetic manner, as known in the art.

[0115] In certain embodiments, a polypeptide is cleaved into fragments for use in further structural or functional analysis, or in the generation of reagents such as related polypeptides and specific antibodies. This is accomplished by treating purified or unpurified polypeptides with a proteolytic enzyme (*i.e.*, a proteinase) including, but not limited to, serine proteinases (*e.g.*, chymotrypsin, trypsin, plasmin, elastase, thrombin, substilin) metal proteinases (*e.g.*, carboxypeptidase A, carboxypeptidase B, leucine aminopeptidase, thermolysin, collagenase), thiol proteinases (*e.g.*, papain, bromelain, Streptococcal proteinase, clostripain) and/or acid proteinases (*e.g.*, pepsin, gastricsin, trypsinogen). Polypeptide fragments are also generated using chemical means such as treatment of the polypeptide with cyanogen bromide (CNBr), 2-nitro-5-thiocyanobenzoic acid, isobenzoic acid, BNPA-skatole, hydroxylamine or a dilute acid solution. In other embodiments, the polypeptide fragments of the invention are recombinantly expressed or prepared *via* peptide synthesis methods known in the art (Barany *et al.*, 1997; U.S. Patent 5,258,454).

[0116] "Variant" as the term is used herein, is a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide respectively, but retains essential properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical (*i.e.*, biologically equivalent). A variant and

reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques or by direct synthesis.

[0117] "Identity," as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. "Identity" and "similarity" can be readily calculated by known methods, including but not limited to those described in Computational Molecular Biology, Lesk, A. M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D. W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part I, Griffin, A. M., and Griffin, H. G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; and Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., SIAM J. *Applied Math.*, 48:1073 (1988). Preferred methods to determine identity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs. Preferred computer program methods to determine identity and similarity between two sequences include, but are not limited to, the GCG program package (Devereux, J., *et al* 1984), BLASTP, BLASTN, and FASTA (Altschul, S. F., *et al.*, 1990). The BLASTX program is publicly available from NCBI and other sources (BLAST Manual, Altschul, S., *et al.*, NCBI NLM NIH Bethesda, Md. 20894; Altschul, S., *et al.*, 1990). The well known Smith Waterman algorithm may also be used to determine identity.

[0118] In certain embodiments, a polypeptide of the invention (e.g. any of the even numbered SEQ ID NOS:2-6) comprises modifications such as a mature processed form of a protein, lipidation, glycosylation, de-O-acylation, phosphorylation and the like.

[0119] In one particular embodiment, the polypeptides and nucleic acids encoding such polypeptides are used in immunogenic compositions for preventing or ameliorating infection.

[0120] The proteins of the invention, including the amino acid sequences of even numbered SEQ ID NOS: 2-6, their fragments, and analogs thereof, or cells expressing them, are also

used as immunogens to produce antibodies immunospecific for the polypeptides of the invention.

Antigens

[0121] In certain embodiments, an immunogenic composition, including proteins, polynucleotides and equivalents of the present invention, is administered as a sole active immunogen or alternatively, the composition includes other active immunogens and/or therapeutics, including other immunogenic polynucleotides, polypeptides, or immunologically-active proteins of one or more other microbial pathogens (e.g. virus, prion, bacterium, or fungus, without limitation) or capsular polysaccharide. The compositions may comprise one or more desired proteins, fragments or pharmaceutical compounds as desired for a chosen indication. In the same manner, the compositions of this invention which employ one or more nucleic acids in the composition may also include nucleic acids which encode the same diverse group of proteins, as noted above. In certain embodiments, a modified polynucleotide of the invention comprises a plasmid or a viral vector.

[0122] Any antigen, multi-antigen or multi-valent immunogenic composition is contemplated by the present invention. For example, the compositions of the present invention comprise a single protein, combinations of two or more proteins, one or more polysaccharides, a combination of one or more proteins, and one or more polysaccharides or any combination thereof. Persons of skill in the art would be readily able to formulate such immunogenic or therapeutic compositions.

[0123] The present invention also contemplates multi-immunization (e.g., a prime/boost regimen) or therapeutic regimens wherein any composition useful against a pathogen may be combined therein or therewith the compositions of the present invention. For example, without limitation, a mammalian subject is administered an immunogenic composition of the present invention and another composition, as part of a multi-drug regimen. Persons of skill in the art would be readily able to select compositions for use in conjunction with the immunogenic and/or therapeutic compositions of the present invention for the purposes of developing and implementing multi-drug regimens.

[0124] Specific embodiments of this invention relate to the use of one or more polypeptides of this invention, or nucleic acids encoding such, in a composition or as part of a treatment regimen for the prevention or amelioration of infection. One can combine the polypeptides or polynucleotides with any immunogenic composition for use against infection. One can

also combine the polypeptides or polynucleotides with any other protein or polysaccharide-based immunogenic composition.

[0125] In certain embodiments, the polypeptides, fragments and equivalents are used as part of a conjugate immunogenic composition; wherein one or more proteins or polypeptides are conjugated to a carrier protein in order to generate a composition that has immunogenic properties against several serotypes and/or against several diseases. Alternatively, one of the polypeptides is used as a carrier protein for other immunogenic polypeptides.

[0126] The present invention also relates to a method of inducing immune responses in a mammal comprising the step of providing to said mammal an immunogenic composition of this invention. The immunogenic composition is a composition which is antigenic in the treated mammal such that an immunologically effective amount of the polypeptide(s) contained in such composition brings about the desired immune response against infection. Certain embodiments relate to a method for the treatment, including amelioration, or prevention of infection in a human comprising administering to a human an immunologically effective amount of the composition.

[0127] The phrase "immunologically effective amount," as used herein, refers to the administration of that amount to a mammalian host (*e.g.*, a human), either in a single dose or as part of a series of doses, sufficient to at least cause the immune system of the individual treated to generate a response that reduces the clinical impact of the bacterial or viral infection. This may range from a minimal decrease in bacterial or viral burden to prevention of the infection. Ideally, the treated individual will not exhibit the more serious clinical manifestations of the bacterial or viral infection. The dosage amount varies depending upon specific conditions of the individual. This amount is determined in routine trials or otherwise by means known to those skilled in the art.

[0128] The phrase "therapeutically effective amount", as used herein, refers to the administration of that amount to a mammalian host (*e.g.,* a human), either in a single dose or as part of a series of doses, sufficient to at least generate a response that reduces the impact of the pathogen on the host. The dosage amount can vary depending on the specific conditions of the host. The amount is determined through routine testing or otherwise as known to persons skilled in the art.

[0129] Another specific aspect of the present invention relates to using as the composition a vector or plasmid which expresses a protein of this invention, or an immunogenic or therapeutic portion thereof. Accordingly, a further aspect of the invention provides a method

of inducing a desired response, e.g., immunogenic, in a mammal, which comprises providing to a mammal a vector or plasmid expressing at least one isolated polypeptide. The protein of the present invention is delivered to the mammal using a live, or live attenuted vectors. In certain embodiments, the virus is attenuated and comprises a modified polynucleotide encoding a bacterial protein, viral protein and the like, containing the genetic material necessary for the expression of the polypeptide or immunogenic portion as a foreign polypeptide.

Viral and Non-Viral Vectors

[0130] The present invention also provides vectors comprising the polynucleotides of the present invention. According to various embodiments of the invention, vectors are used to transport recombinants of the invention to site of expression (e.g., transcription, translation/protein synthesis). Thus, the vectors are used *in vivo* or *in vitro* depending upon the desired objective. Any suitable vectors for accomplishing the objectives consistent with the inventive concept are contemplated by the present invention.

[0131] Viral vectors such as lentiviruses, retroviruses, herpes viruses, adenoviruses, adeno-associated viruses, vaccinia virus, baculovirus, and other recombinant viruses with desirable cellular tropism, are particularly useful for cellular assays *in vitro* and *in vivo*. Thus, a nucleic acid encoding a protein or immunogenic fragment thereof can be introduced *in vivo*, *ex vivo*, or *in vitro* using a viral vector or through direct introduction of DNA. Expression in targeted tissues can be effected by targeting the transgenic vector to specific cells, such as with a viral vector or a receptor ligand, or by using a tissue-specific promoter, or both. Targeted gene delivery is described in PCT Publication No. WO 95/28494, which is incorporated herein by reference in its entirety.

[0132] Viral vectors commonly used for *in vivo* or *ex vivo* targeting and therapy procedures include DNA vectors and RNA vectors. Methods for constructing and using viral vectors are known in the art (*e.g.*, Miller and Rosman, *BioTechniques*, 1992, 7:980-990). In certain embodiments, the viral vectors are replication-defective, that is, they are unable to replicate autonomously in the target cell. In other embodiments, the viral vector is a live attenuated virus. In one particular embodiment, the replication defective virus is a minimal virus, *i.e.*, it retains only the sequences of its genome which are necessary for encapsulating the genome to produce viral particles.

[0133] Various companies produce viral vectors commercially, including, but not limited to, Avigen, Inc. (Alameda, CA; AAV vectors), Cell Genesys (Foster City, CA; retroviral,

adenoviral, AAV vectors, and lentiviral vectors), Clontech (retroviral and baculoviral vectors), Genovo, Inc. (Sharon Hill, PA; adenoviral and AAV vectors), Genvec (adenoviral vectors), IntroGene (Leiden, Netherlands; adenoviral vectors), Molecular Medicine (retroviral, adenoviral, AAV, and herpes viral vectors), Norgen (adenoviral vectors), Oxford BioMedica (Oxford, United Kingdom; lentiviral vectors), and Transgene (Strasbourg, France; adenoviral, vaccinia, retroviral, and lentiviral vectors), incorporated by reference herein in its entirety.

[0134] **Adenovirus vectors.** Adenoviruses are eukaryotic DNA viruses that can be modified to efficiently deliver a nucleic acid of this invention to a variety of cell types. Various serotypes of adenovirus exist. In one particular embodiment, an adenovirus (Ad) is a type 2, type 4, type 5, or type 7 human adenoviruses (Ad 2, Ad 4, Ad 5 or Ad 7) or adenoviruses of animal origin (*see* PCT Publication No. WO 94/26914). Those adenoviruses of animal origin which can be used within the scope of the present invention include adenoviruses of canine, bovine, murine (*e.g.,* Mav1, Beard *et al.*, Virology, 1990, 75-81) bovine, porcine, avian, and simian (*e.g.,* SAV) origin. In one embodiment, the adenovirus of animal origin is a canine adenovirus, such as a CAV2 adenovirus (*e.g.*, Manhattan or A26/61 strain, ATCC VR-800). Various replication defective adenovirus and minimum adenovirus vectors have been described (PCT Publication Nos. WO 94/26914, WO 95/02697, WO 94/28938, WO 94/28152, WO 94/12649, WO 95/02697, WO 96/22378). The replication defective recombinant adenoviruses according to the invention can be prepared by any technique known to the person skilled in the art (Levrero *et al.*, *Gene*, 1991, 101:195; European Publication No. EP 185 573; Graham, EMBO J., 1984, 3:2917; Graham *et al.*, *J. Gen. Virol.*, 1977, 36:59). Recombinant adenoviruses are recovered and purified using standard molecular biological techniques, which are well known to persons of ordinary skill in the art.

[0135] **Adeno-associated viruses.** The adeno-associated viruses (AAV) are DNA viruses of relatively small size that can integrate, in a stable and site-specific manner, into the genome of the cells which they infect. They are able to infect a wide spectrum of cells without inducing any effects on cellular growth, morphology or differentiation, and they do not appear to be involved in human pathologies. The AAV genome has been cloned, sequenced and characterized. The use of vectors derived from the AAVs for transferring genes *in vitro* and *in vivo* has been described (*see*, PCT Publication Nos. WO 91/18088 and WO 93/09239; U.S. Patent Nos. 4,797,368 and 5,139,941; European Publication No. EP 488 528). The replication defective recombinant AAVs according to the invention can be prepared by cotransfecting a plasmid containing the nucleic acid sequence of interest flanked by two AAV inverted terminal repeat (ITR) regions, and a plasmid carrying the AAV encapsidation genes

(rep and cap genes), into a cell line which is infected with a human helper virus (for example an adenovirus). The AAV recombinants which are produced are then purified by standard techniques.

[0136] ***Retrovirus vectors.*** In another implementation of the present invention, the nucleic acid can be introduced in a retroviral vector, *e.g.*, as described in U.S. Patent No. 5,399,346; Mann *et al.*, *Cell*, 1983, 33:153; U.S. Patent Nos. 4,650,764 and 4,980,289; Markowitz *et al.*, *J. Virol.*, 1988, 62:1120; U.S. Patent No. 5,124,263; European Publication Nos. EP 453 242 and EP178 220; Bernstein *et al., Genet. Eng.*,1985, 7:235; McCormick, *BioTechnology*, 1985, 3:689; PCT Publication No. WO 95/07358; and Kuo *et al.*, *Blood*, 1993, 82:845, each of which is incorporated by reference in its entirety. The retroviruses are integrating viruses that infect dividing cells. The retrovirus genome includes two LTRs, an encapsidation sequence and three coding regions (gag, pol and env). In recombinant retroviral vectors, the *gag*, *pol* and *env* genes are generally deleted, in whole or in part, and replaced with a heterologous nucleic acid sequence of interest. These vectors can be constructed from different types of retrovirus, such as, HIV, MoMuLV ("murine Moloney leukaemia virus"), MSV ("murine Moloney sarcoma virus"), HaSV ("Harvey sarcoma virus"); SNV ("spleen necrosis virus"); RSV ("Rous sarcoma virus") and Friend virus. Suitable packaging cell lines have been described in the prior art, in particular the cell line PA317 (U.S. Patent No. 4,861,719); the PsiCRIP cell line (PCT Publication No. WO 90/02806) and the GP+envAm-12 cell line (PCT Publication No. WO 89/07150). In addition, the recombinant retroviral vectors can contain modifications within the LTRs for suppressing transcriptional activity as well as extensive encapsidation sequences which may include a part of the gag gene (Bender *et al.*, *J. Virol.*, 1987, 61:1639). Recombinant retroviral vectors are purified by standard techniques known to those having ordinary skill in the art.

[0137] Retroviral vectors can be constructed to function as infectious particles or to undergo a single round of transfection. In the former case, the virus is modified to retain all of its genes except for those responsible for oncogenic transformation properties, and to express the heterologous gene. Non-infectious viral vectors are manipulated to destroy the viral packaging signal, but retain the structural genes required to package the co-introduced virus engineered to contain the heterologous gene and the packaging signals. Thus, the viral particles that are produced are not capable of producing additional virus.

[0138] Retrovirus vectors can also be introduced by DNA viruses, which permits one cycle of retroviral replication and amplifies transfection efficiency (*see* PCT Publication Nos. WO 95/22617, WO 95/26411, WO 96/39036 and WO 97/19182).

[0139] **Lentivirus vectors.** In another implementation of the present invention, lentiviral vectors are used as agents for the direct delivery and sustained expression of a transgene in several tissue types, including brain, retina, muscle, liver and blood. The vectors efficiently transduce dividing and nondividing cells in these tissues, and effect long-term expression of the gene of interest. For a review, *see*, Naldini, Curr. Opin. Biotechnol., 1998, 9:457-63; *see also* Zufferey, *et al.*, *J. Virol.*, 1998, 72:9873-80). Lentiviral packaging cell lines are available and known generally in the art. They facilitate the production of high-titer lentivirus vectors for gene therapy. An example is a tetracycline-inducible VSV-G pseudotyped lentivirus packaging cell line that can generate virus particles at titers greater than 106 IU/mL for at least 3 to 4 days (Kafri, *et al.*, *J. Virol.*, 1999, 73: 576-584). The vector produced by the inducible cell line can be concentrated as needed for efficiently transducing non-dividing cells *in vitro* and *in vivo*.

[0140] In another implementation of the present invention, a modified polynucleotide of the invention is delivered via Mononegavirales. Viruses of the Order Mononegavirales are non-segmented, negative dtranded RNA viruses (*e.g.*, described in U.S. Patent 6,033,886, incorporated herein by reference)

[0141] In one particular embodiment, a modified polynucleotide of the invention is delivered via Vesicular Stomatitis Virus (VSV). Genetically modified VSV strains, attenuating VSV mutations and VSV rescue methods are well known in the art, e.g. see U.S. Patent No.. 6,033,886; 6,168,943; 6,596,529.

[0142] **Non-viral vectors.** In another implementation of the present invention, the vector can be introduced *in vivo* by lipofection, as "naked" DNA, or with other transfection facilitating agents (peptides, polymers, etc.). Synthetic cationic lipids are used to prepare liposomes for *in vivo* transfection of a gene encoding a marker (Felgner, et. al., Proc. Natl. Acad. Sci. U.S.A., 1987, 84:7413-7417; Felgner and Ringold, Science, 1989, 337:387-388; see Mackey, *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, 1988, 85:8027-8031; Ulmer *et al.*, *Science*, 1993, 259:1745-1748). Useful lipid compounds and compositions for transfer of nucleic acids are described in PCT Patent Publication Nos. WO 95/18863 and WO 96/17823, and in U.S. Patent No. 5,459,127. Lipids may be chemically coupled to other molecules for the

purpose of targeting (see Mackey, *et. al., supra*). Targeted peptides, *e.g.*, hormones or neurotransmitters, and proteins such as antibodies, or non-peptide molecules could be coupled to liposomes chemically.

[0143] Other molecules are also useful for facilitating transfection of a nucleic acid *in vivo*, such as a cationic oligopeptide (*e.g.*, PCT Patent Publication No. WO 95/21931), peptides derived from DNA binding proteins (*e.g.*, PCT Patent Publication No. WO 96/25508), or a cationic polymer (*e.g.*, PCT Patent Publication No. WO 95/21931).

[0144] In certain embodiments, a polynucleotide modified for optimal expression in a mammalian host (*i.e.*, comprising surrogate codons) is administered directly to the host as an immunogenic composition. The polynucleotide is introduced directly into the host either as "naked" DNA (U.S. Patent 5,580,859) or formulated in compositions with agents which facilitate immunization, such as bupivicaine and other local anesthetics (U.S. Patent 5,593,972) and cationic polyamines (U.S. Patent 6,127,170).

[0145] In this polynucleotide immunization procedure, the polypeptides of the invention are expressed on a transient basis *in vivo*; no genetic material is inserted or integrated into the chromosomes of the host. This procedure is to be distinguished from gene therapy, where the goal is to insert or integrate the genetic material of interest into the chromosome. An assay is used to confirm that the polynucleotides administered by immunization do not give rise to a transformed phenotype in the host (U.S. Patent 6,168,918).

[0146] It is also possible to introduce the vector *in vivo* as a naked DNA plasmid. Naked DNA vectors for vaccine purposes or gene therapy can be introduced into the desired host cells by methods known in the art, *e.g.*, electroporation, microinjection, cell fusion, DEAE dextran, calcium phosphate precipitation, use of a gene gun, or use of a DNA vector transporter (*e.g.*, Wu *et al.*, *J. Biol. Chem.*, 1992, 267:963-967; Wu and Wu, *J. Biol. Chem.*, 1988, 263:14621-14624; Canadian Patent Application No. 2,012,311; Williams *et al.*, Proc. Natl. Acad. Sci. USA, 1991, 88:2726-2730). Receptor-mediated DNA delivery approaches can also be used (Curiel *et al.*, *Hum. Gene Ther.*, 1992, 3:147-154; Wu and Wu, *J. Biol. Chem.*, 1987, 262:4429-4432). U.S. Patent Nos. 5,580,859 and 5,589,466 disclose delivery of exogenous DNA sequences, free of transfection facilitating agents, in a mammal. More recently, a relatively low voltage, high efficiency *in vivo* DNA transfer technique, termed electrotransfer, has been described (Mir *et al.*, *C.P. Acad. Sci.*, 1988, 321:893; PCT Publication Nos. WO 99/01157; WO 99/01158; WO 99/01175). Accordingly, additional embodiments of the present invention relates to a method of inducing an immune response

in a human comprising administering to said human an amount of a DNA molecule encoding a polypeptide of this invention, optionally with a transfection-facilitating agent, where said polypeptide, when expressed, retains the desired functionality and, when incorporated into an immunogenic composition and administered to a human, provides protection without inducing enhanced disease upon subsequent infection of the human with a pathogen. Transfection-facilitating agents are known in the art and include bupivicaine, and other local anesthetics (for examples see U.S. Patent No. 5,739,118) and cationic polyamines (as published in International Patent Application WO 96/10038), which are hereby incorporated by reference.

[0147] According to an embodiment of the present invention, the IL-15 constructs as described herein are administered in a plasmid. According to an embodiment, the plasmid of the present invention comprises SEQ ID NOS: 18, 19, 20 or combinations thereof. The preparation of plasmids is well known in the art. A person of ordinary skill in the art could readily prepare a plasmid having the modified polynucleotide, such as the IL-15 constructs, for example, in accordance with the present invention, based upon the guidance provided herein. For example, the preparation of plasmids is described in U.S. Patent No. 5,593,972, which is incorporated by reference in its entirety.

Adjuvants

[0148] According to an embodiment of the present invention, the polynucleotides of the present invention may be used as adjuvants, for example, as adjuvants for vaccines, such as DNA and/or RNA vaccines. Techniques for the preparation of adjuvants, DNA vaccines and RNA vaccines are well known in the art. A person of skill in the art would readily be able to prepare an adjuvant, DNA vaccine and/or RNA vaccine and the like, using the embodiments of the present invention, based upon the guidance provided herein.

[0149] The present invention contemplates that the modified polynucleotides of the present invention may be used alone or in combination with other compounds or compositions for any desired effect. For example, the modified polynucleotides of the present invention may be administered in combination with a DNA and/or RNA vaccine or as part of the DNA and/or RNA vaccine (e.g., as part of a plasmid containing the DNA and/or RNA vaccine). The modified polyncleotides of the present invention may be administered separately but contemporaneously with the administration of the DNA and/or RNA vaccine, include

administering during, before or after. Further, the polynucleotides of the present invention may be administered alone.

[0150] Exemplary DNA vaccines with which the present invention may be combined in any manner include, without limitation, nucleotides coding for the Plasmodium (malarial agent) proteins such as P. falciparum, P. vivax, P. malariae, and P. ovale CSP; SSP2(TRAP); Pfs16 (Sheba); LSA-1; LSA-2; LSA-3; STARP; MSA-1 (MSP-1, PMMSA, PSA, p185, p190); MSA-2 (MSP-2, Gymmsa, gp56, 38-45 kDa antigen); RESA (Pf155); EBA-175; AMA-1 (Pf83); SERA (p113, p126, SERP, Pf140); RAP-1; RAP-2; RhopH3; PfHRP-II; Pf55; Pf35; GBP (96-R); ABRA (p101); Exp-1 (CRA, Ag5.1); Aldolase; Duffy binding protein of P. vivax; Reticulocyte binding proteins; HSP70-1 (p75); Pfg25; Pfg28; Pfg48/45; and Pfg230. DNA and RNA vaccines also may comprise nucleotides coding for proteins associated with the GP or NP genes from the ebola virus; and the HPV6a L2, HPV6a E1, HPV6a E2, HPV6a E4, HPV6a E5, HPV6a E6, and HPV6a E7 proteins from the human Papillomavirus 6a (HPV6a). According to an embodiment, the DNA and RNA vaccines code for HIV proteins, including, but not limited to, the glycoproteins gp41, gp120, gp 140, and gp160; and proteins encoded by the gag (the proteins p55, p39, p24, p17 and p15), env, rev, tat, nef, vpr, vpx, prot, and pol (the proteins p66/p51 and p31-34) genes found in HIV.

[0151] According to an embodiment of the present invention, the IL-15 constructs of the present invention (e.g., SEQ ID NOS:12-16) is used in combination with DNA and/or RNA vaccine.e.g, a DNA vaccine against HIV/AIDS. According to an embodiment, SEQ ID NO:14 is used (e.g., administered contemporaneously and/or combined in a plasmid or other vector or compostion) in combination with a DNA vaccine against HIV/AIDS.

Compositions

[0152] One aspect of the present invention provides compositions, such as immunogenic compositions and therapeutic compositions, etc., which comprise a modified polynucleotide of the present invention, a protein or polypeptide encoded by said recombinant polynucleotide, an antibody to said protein or polypeptide, or the like, including any combinations thereof. For example, compositions that have the ability to confer protection against a live challenge and/or prevent colonization are contemplated by the present invention.

[0153] The formulation of such compositions is well known to persons skilled in this field. Compositions of the invention, according to an embodiment, include a pharmaceutically

acceptable carrier. Suitable pharmaceutically acceptable carriers and/or diluents include any and all conventional solvents, dispersion media, fillers, solid carriers, aqueous solutions, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like. Suitable pharmaceutically acceptable carriers include, for example, one or more of water, saline, phosphate buffered saline, dextrose, glycerol, ethanol and the like, as well as combinations thereof. Pharmaceutically acceptable carriers may further comprise minor amounts of auxiliary substances such as wetting or emulsifying agents, preservatives or buffers, which enhance the shelf life or effectiveness of the antibody. The preparation and use of pharmaceutically acceptable carriers is well known in the art. Except insofar as any conventional media or agent is incompatible with the active ingredient, use thereof in the compositions of the present invention is contemplated.

[0154] An immunogenic composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral (*e.g.*, intravenous, intradermal, subcutaneous, intramuscular, intraperitoneal), mucosal (*e.g.*, oral, rectal, intranasal, buccal, vaginal, respiratory) and transdermal (topical). Other modes of administration employ oral formulations, pulmonary formulations, suppositories, and transdermal applications, for example, without limitation. Oral formulations, for example, include such normally employed excipients as, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, and the like, without limitation.

[0155] The present invention contemplates the use of embodiments of the invention as adjuvants or co-adjuvants, for example, as adjuvants to DNA or RNA vaccines/immunogenic composition. The immunogenic compositions of the invention can include one or more adjuvants, or be administered along with one or more adjuvants, including, but not limited to aluminum salts (alum) such as aluminum phosphate and aluminum hydroxide, *Mycobacterium tuberculosis*, *Bordetella pertussis*, bacterial lipopolysaccharides, aminoalkyl glucosamine phosphate compounds (AGP), or derivatives or analogs thereof, which are available from Corixa (Hamilton, MT), and which are described in U.S. Patent Number 6,113,918; one such AGP is 2-[(R)-3-Tetradecanoyloxytetradecanoylamino]ethyl 2-Deoxy-4-O-phosphono-3-O-[(R)-3-tetradecanoyoxytetradecanoyl]-2-[(R)-3-tetradecanoyoxytetradecanoylamino]-b-D-glucopyranoside, which is also known as 529 (formerly known as RC529), which is formulated as an aqueous form or as a stable emulsion, MPL™ (3-O-deacylated monophosphoryl lipid A) (Corixa) described in U.S. Patent Number 4,912,094, synthetic polynucleotides such as oligonucleotides containing a CpG

motif (U.S. Patent Number 6,207,646), polypeptides, saponins such as Quil A or STIMULON™ QS-21 (Antigenics, Framingham, Massachusetts), described in U.S. Patent Number 5,057,540, a pertussis toxin (PT), an E. coli heat-labile toxin (LT), particularly LT-K63, LT-R72, CT-S109, PT-K9/G129; *see, e.g.*, International Patent Publication Nos. WO 93/13302 and WO 92/19265, cholera toxin (either in a wild-type or mutant form, *e.g.*, wherein the glutamic acid at amino acid position 29 is replaced by another amino acid, such as a histidine, in accordance with published International Patent Application number WO 00/18434).

[0156] Various cytokines and lymphokines are suitable for use as adjuvants. One such adjuvant is granulocyte-macrophage colony stimulating factor (GM-CSF), which has a nucleotide sequence as described in U.S. Patent Number 5,078,996. A plasmid containing GM-CSF cDNA has been transformed into *E. coli* and has been deposited with the American Type Culture Collection (ATCC), 1081 University Boulevard, Manassas, VA 20110-2209, under Accession Number 39900. The cytokine Interleukin-12 (IL-12) is another adjuvant which is described in U.S. Patent Number 5,723,127. Other cytokines or lymphokines have been shown to have immune modulating activity, including, but not limited to, the interleukins 1-α, 1-β, 2, 4, 5,6, 7, 8, 10, 13, 14, 15, 16, 17 and 18, the interferons-α, β and γ, granulocyte colony stimulating factor, and the tumor necrosis factors α and β, and are suitable for use as adjuvants.

[0157] In certain embodiments, the proteins of this invention are used in a composition for oral administration which includes a mucosal adjuvant and used for the treatment or prevention of infection in a mammalian host (e.g., a human). The mucosal adjuvant can be a wild-type cholera toxin or; a derivative of a cholera holotoxin, wherein the A subunit is mutagenized or chemically modified. For a specific cholera toxin which may be particularly useful in preparing immunogenic compositions of this invention, see the mutant cholera holotoxin E29H, as disclosed in Published International Application WO 00/18434, which is hereby incorporated herein by reference in its entirety. These may be added to, or conjugated with, the polypeptides of this invention. The same techniques are applied to other molecules with mucosal adjuvant or delivery properties such as Escherichia coli heat labile toxin (LT). Other compounds with mucosal adjuvant or delivery activity may be used such as bile; polycations such as DEAE-dextran and polyornithine; detergents such as sodium dodecyl benzene sulphate; lipid-conjugated materials; antibiotics such as streptomycin; vitamin A; and other compounds that alter the structural or functional integrity

of mucosal surfaces. Other mucosally active compounds include derivatives of microbial structures such as MDP; acridine and cimetidine. STIMULON™ QS-21, MPL, and IL-12, as described above, may also be used.

[0158] The compositions of this invention may be delivered in the form of ISCOMS (immune stimulating complexes), ISCOMS containing CTB, liposomes or encapsulated in compounds such as acrylates or poly(DL-lactide-co-glycoside) to form microspheres of a size suited to adsorption. The proteins of this invention may also be incorporated into oily emulsions.

[0159] Recombinant cells, recombinant cell lines, assays and kits that provide or use same and the like are also contemplated by the present invention. A person skilled in the art would readily understand how to prepare and use such embodiments of the present invention, based upon the guidance provided herein.

[0160] The present invention also relates to an antibody, which may either be a monoclonal or polyclonal antibody, specific for polypeptides as described above. Such antibodies may be produced by methods which are well known to those skilled in the art.

[0161] According to a further implementation of the present invention, a method is provided for diagnosing a condition in a mammal comprising: detecting the presence of immune complexes in the mammal or a tissue sample from said mammal, said mammal or tissue sample being contacted with an antibody composition comprising antibodies that immunospecifically bind with at least one polypeptide comprising the amino acid sequence of any of the even numbered SEQ ID NOS: 2-6; wherein the mammal or tissue sample is contacted with the antibody composition under conditions suitable for the formation of the immune complexes.

[0162] The description of the specific embodiments will so fully reveal the general nature of the invention that others can, by applying knowledge within the skill of the art, readily modify and/or adapt for various applications such specific embodiments, without undue experimentation, without departing from the general concept of the present invention. Therefore, such adaptations and modifications are intended to be within the meaning and range of equivalents of the disclosed embodiments, based on the teaching and guidance presented herein. It is to be understood that the phraseology or terminology herein is for the purpose of description and not of limitation, such that the terminology or phraseology of the present specification is to be interpreted by the skilled artisan in light of the teachings and guidance presented herein, in combination with the knowledge of one of ordinary skill in the

art. A person skilled in the art would know, or be able to ascertain, using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein, based upon the guidance provided herein.

[0163] The following examples are included to demonstrate particular embodiments of the invention. However, those of skill in the art should, in view of the present disclosure, appreciate that many changes can be made in the specific embodiments which are disclosed and still obtain a like or similar result without departing from the spirit and scope of the invention. The following examples are offered by way of illustration and are not intended to limit the invention in any way.

EXAMPLES

Example 1-Enhancement of HPV16 E7 expression

[0164] a. One example of a "modified" polynucleotide sequence demonstrating "enhanced" levels of protein expression is shown below in SEQ ID NO:1. The modified polynucleotide's sequence incorporates surrogate codons encoding the 98 amino acid human papillomavirus (HPV)16 E7 protein sequence (e.g., see HPV16 Accession No. K02718 in NCBI database).

[0165] The enhanced sequence of the polynucleotide in accordance with an embodiment of the invention is determined by selecting suitable surrogate codons. Surrogate codons were selected in order to alter the A and T (or A and U in the case of RNA) content of the naturally-occurring (wild-type) gene. The surrogate codons are those that encode the amino acids alanine, arginine, glutamic acid, glycine, isoleucine, leucine, proline, serine, threonine, and valine. Accordingly, the modified nucleic acid sequence had surrogate codons for each of these amino acids throughout the sequence. For the remaining 11 amino acids, no alterations were made, thereby leaving the corresponding naturally-occurring codons in place.

[0166] The modified sequence may be determined manually or by computer-assisted methods. As such, the information technology, including hardware, software, algorithms, arrays, datebases and the like, directed to the determination of the modified sequences of the present invention are contemplated herein.

SEQ ID NO:1 (polynucleotide) and SEQ ID NO:2 (protein)

```
1   ATGCATGGGGATACGCCTACGCTCCATGAATATATGCTCGATCTCCAACCTGA
1    M  H  G  D  T  P  T  L  H  E  Y  M  L  D  L  Q  P  E
54  GACGACGGATCTCTACTGTTATGAGCAACTCAATGACAGCTCCGAGGAGGAGG
18    T  T  D  L  Y  C  Y  E  Q  L  N  D  S  S  E  E  E
107 ATGAAATTGATGGGCCTGCGGGGCAAGCGGAACCTGACCGGGCCCATTACAAT
36  D  E  I  D  G  P  A  G  Q  A  E  P  D  R  A  H  Y  N
160 ATTGTCACCTTTTGTTGCAAGTGTGACTCCACGCTCCGGCTCTGCGTCCAAAG
54   I  V  T  F  C  C  K  C  D  S  T  L  R  L  C  V  Q  S
213 CACGCACGTCGACATTCGGACGCTCGAAGACCTGCTCATGGGCACGCTCGGGA
71    T  H  V  D  I  R  T  L  E  D  L  L  M  G  T  L  G
266 TTGTGTGCCCCATCTGTTCCCAGAAACCTTAATAG
89  I  V  C  P  I  C  S  Q  K  P
```

[0167] Referring to SEQ ID NO:1 above, the recombinant nucleotide sequence of HPV16 E7 (Accession No. K02718) incorporates surrogate codons but retains the capacity to encode the wild type E7 protein.

[0168] b. The nucleic acid sequence of SEQ ID NO:1 was assembled from oligonucleotides that were 100 nucleotides in length and corresponding in polarity to the positive (sense) strand sequence shown above. A person of skill in the art would readily be able to select suitable oligonucleotides depending upon the desired sequence in accordance with the present invention. Suitable oligonucleotides are available from a variety of commercial vendors, such as Invitrogen™ (Carlsbad, CA).

[0169] "Bridge" oligos 50 nucleotides in length and antisense in polarity were designed to straddle the joints at the ends of each sense 100-mer oligo. This strategy facilitated the hybridization of 25 nucleotides at the ends of each 100-mer targeted for ligation. A heat stable ligase (Ampligase, Epicentre, WI) was used at 68°C to ligate the 100-mer sense oligos together. The entire open reading frame (for HPV16 E7, approximately 300 nucleotides) was then PCR amplified using oligos corresponding to the 5' and 3' boundaries of the ORF. The fidelity of the intended final ORF was verified by sequencing reactions.

[0170] This HPV16E7 gene containing surrogate codons was tested for expression levels by Western blot (data not shown). Rhabdomyosarcoma (RD) cells (American Type Culture Collection, Manassas, VA ATCC# CCL136) were transfected with the indicated plasmid DNA expression vectors. All HPV16 E7 genes were cloned into pcDNA3.1 (Invitrogen, Carlsbad, CA). While a variety of different transfecting agents could be utilized, the experiments listed herein were performed using Lipofectamine (Invitrogen) according to

manufacturer's instructions. Total cell lysates were harvested 48 hours after transfection in SDS-sample buffer containing 1% SDS and 2-mercaptoethanol. Equivalent amounts of each transfectant lysate were loaded and electrophoresed on 4-20% tris-glycine gradient SDS-polyacrylamide gels. HPV16 E7 protein was detected by an E7-specific monoclonal antibody (Zymed Laboratories, San Francisco, CA).

[0171] The expression levels of the surrogate codon modified HPV16 E7 gene (SEQ ID NO:1) were markedly enhanced compared to the expression levels of the wild type HPV16 E7 gene. The expression levels of the surrogate codon modified HPV16E7 was comparable to the expression level of the "preferred" codon modified HPV16E7 (data not shown).

Example 2-Enhancement of HIV-1 gag p37 expression

[0172] A second example demonstrating the unexpected results of using "surrogate" codons in lieu of wild-type codons in a nucleic acid sequence was found for the HIV-1 gag gene, specifically the p37 component of the full-length p55 protein.

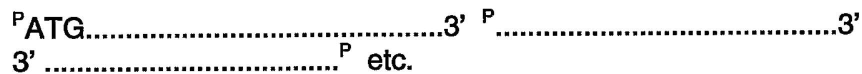
[0173] a. The amino acid sequence of the HXB2 strain of HIV-1 (NCBI Accession No. K03455) was selected as a representative HIV-1 gag gene.

SEQ ID NO:3 (polynucleotide) and SEQ ID NO:4 (protein)

```
1    ATGGGGGCGCGGGCGTCCGTCCTCTCCGGGGGGGAGCTCGATCGGTGGGAGAAA
1     M  G  A  R  A  S  V  L  S  G  G  E  L  D  R  W  E  K
55   ATTCGGCTCCGGCCGGGGGGGAAGAAAAAATATAAACTCAAACATATTGTCTGG
19    I  R  L  R  P  G  G  K  K  K  Y  K  L  K  H  I  V  W
109  GCGTCCCGGGAGCTCGAGCGGTTCGCGGTCAATCCGGGGCTGCTCGAGACGTCC
37    A  S  R  E  L  E  R  F  A  V  N  P  G  L  L  E  T  S
163  GAGGGCTGTCGGCAAATTCTCGGGCAGCTCCAACCGTCCCTCCAGACGGGGTCC
55    E  G  C  R  Q  I  L  G  Q  L  Q  P  S  L  Q  T  G  S
217  GAGGAGCTCCGGTCCCTCTATAATACGGTCGCGACGCTCTATTGTGTCCATCAA
73    E  E  L  R  S  L  Y  N  T  V  A  T  L  Y  C  V  H  Q
271  CGGATTGAGATTAAAGACACGAAGGAGGCGCTCGACAAGATTGAGGAGGAGCAA
91    R  I  E  I  K  D  T  K  E  A  L  D  K  I  E  E  E  Q
325  AACAAATCCAAGAAAAAAGCGCAGCAAGCGGCGGCGGACACGGGGCACTCCAAT
109   N  K  S  K  K  K  A  Q  Q  A  A  A  D  T  G  H  S  N
379  CAGGTCTCCCAAAATTACCCGATTGTCCAGAACATTCAGGGGCAAATGGTCCAT
127   Q  V  S  Q  N  Y  P  I  V  Q  N  I  Q  G  Q  M  V  H
433  CAGGCGATTTCCCCGCGGACGCTCAATGCGTGGGTCAAAGTCGTCGAGGAGAAG
145   Q  A  I  S  P  R  T  L  N  A  W  V  K  V  V  E  E  K
487  GCGTTCTCCCCGGAGGTCATTCCGATGTTTTCAGCGCTCTCCGAGGGGGCGACG
163   A  F  S  P  E  V  I  P  M  F  S  A  L  S  E  G  A  T
541  CCGCAAGATCTCAACACGATGCTCAACACGGTCGGGGGGCATCAAGCGGCGATG
181   P  Q  D  L  N  T  M  L  N  T  V  G  G  H  Q  A  A  M
595  CAAATGCTCAAAGAGACGATTAATGAGGAGGCGGCGGAGTGGGATCGGGTCCAT
199   Q  M  L  K  E  T  I  N  E  E  A  A  E  W  D  R  V  H
649  CCGGTCCATGCGGGGCCGATTGCGCCGGGGCAGATGCGGGAGCCGCGGGGGTCC
217   P  V  H  A  G  P  I  A  P  G  Q  M  R  E  P  R  G  S
703  GACATTGCGGGGACGACGTCCACGCTCCAGGAGCAAATTGGGTGGATGACGAAT
235   D  I  A  G  T  T  S  T  L  Q  E  Q  I  G  W  M  T  N
757  AATCCGCCGATTCCGGTCGGGGAGATTTATAAACGGTGGATTATTCTCGGGCTC
253   N  P  P  I  P  V  G  E  I  Y  K  R  W  I  I  L  G  L
811  AATAAAATTGTCCGGATGTATTCCCCGACGTCCATTCTCGACATTCGGCAAGGG
271   N  K  I  V  R  M  Y  S  P  T  S  I  L  D  I  R  Q  G
865  CCCAAGGAGCCGTTTCGGGACTATGTAGACCGGTTCTATAAAACGCTCCGGGCG
289   P  K  E  P  F  R  D  Y  V  D  R  F  Y  K  T  L  R  A
919  GAGCAAGCGTCCCAGGAGGTCAAAAATTGGATGACGGAGACGCTCCTCGTCCAA
307   E  Q  A  S  Q  E  V  K  N  W  M  T  E  T  L  L  V  Q
973  AATGCGAACCCGGATTGTAAGACGATTCTCAAAGCGCTCGGGCCGGCGGCTACG
325   N  A  N  P  D  C  K  T  I  L  K  A  L  G  P  A  A  T
1027 CTCGAGGAGATGATGACGGCGTGTCAGGGGGTCGGGGGGCCGGGGCATAAGGCG
343   L  E  E  M  M  T  A  C  Q  G  V  G  G  P  G  H  K  A
1081 CGGGTCCTCTAA
361   R  V  L
```

[0174] Referring to SEQ ID NO:3, an altered nucleotide sequence of the HXB2 strain of HIV-1 gag gene (Accession No. K03455) incorporating surrogate codons but retaining the capacity to encode the 363 amino acid wild type p37 component of the gag protein, was constructed.

[0175] The HIV-1 gag p37 gene incorporating surrogate codons was assembled by a different method than that used for the HPV16 E7 (Example 1). This gene was assembled using a series of 100-mer sense and antisense oligos containing overlapping 25 nucleotides of sequence as illustrated below.



[0176] Each 100 mer was phosphorylated (P) on the 5' end to facilitate downstream ligation. For reference, the 5' end of the gag gene, containing the initiation codon ATG, is depicted (sense oligo); an antisense oligo beneath it was designed to contain complementary sequence of 25 nucleotides to facilitate hybridization and subsequent fill in by a DNA polymerase (Pfx Turbo, Invitrogen). This staggered, overlapping arrangement was performed to assemble the entire ~1.1 kb gag gene encoding p37. The double stranded but "nicked" assembled gene was then ligated using a heat stable ligase (Ampligase).

[0177] PCR oligos representing the 5' and 3' most regions of the p37 ORF were then used to amplify the entire gene, which was subsequently cloned into the vector and sequenced to confirm the fidelity in assembly of the predicted sequence.

[0178] The expression levels of a plasmid DNA construct containing the altered/"surrogate" gag p37 gene shown above were tested by transfection in Cos7 cells (ATCC CRL 1651). The levels of gag present in the supernatant 48 hours post infection was quantified with an ELISA assay using a commercially available kit (Coulter p24 kit, Beckman Coulter catalog # PN6604535). The plasmid construct set forth in SEQ ID NO:7 was used for transfection of the wild-type gag p37. The plasmid construct set forth in SEQ ID NO:8 was used for transfection of the recombinant gag gene (modified in accordance with an embodiment of the present invention).

SEQ ID NO:7

aaatgggggc gctgaggtct gcctcgtgaa gaaggtgttg ctgactcata ccaggcctga 60
atcgccccat catccagcca gaaagtgagg gagccacggt tgatgagagc tttgttgtag 120
gtggaccagt tggtgatttt gaacttttgc tttgccacgg aacggtctgc gttgtcggga 180
agatgcgtga tctgatcctt caactcagca aaagttcgat ttattcaaca aagccgccgt 240
cccgtcaagt cagcgtaatg ctctgccagt gttacaacca attaaccaat tctgcgttca 300
aaatggtatg cgttttgaca catccactat atatccgtgt cgttctgtcc actcctgaat 360
cccattccag aaattctcta gcgattccag aagtttctca gagtcggaaa gttgaccaga 420
cattacgaac tggcacagat ggtcataacc tgaaggaaga tctgattgct taactgcttc 480
agttaagacc gacgcgctcg tcgtataaca gatgcgatga tgcagaccaa tcaacatggc 540
acctgccatt gctacctgta cagtcaagga tggtagaaat gttgtcggtc cttgcacacg 600
aatattacgc catttgcctg catattcaaa cagctcttct acgataaggg cacaaatcgc 660
atcgtggaac gtttgggctt ctaccgattt agcagtttga tacactttct ctaagtatcc 720
acctgaatca taaatcggca aaatagagaa aaattgacca tgtgtaagcg gccaatctga 780
ttccacctga gatgcataat ctagtagaat ctcttcgcta tcaaaattca cttccacctt 840
ccactcaccg gttgtccatt catggctgaa ctctgcttcc tctgttgaca tgacacacat 900
catctcaata tccgaatacg gaccatcagt ctgacgacca agagagccat aaacaccaat 960
agccttaaca tcatccccat atttatccaa tattcgttcc ttaatttcat gaacaatctt 1020
cattctttct tctctagtca ttattattgg tccgttcata acaccccttg tattactgtt 1080
tatgtaagca gacagtttta ttgttcatga tgatatattt ttatcttgtg caatgtaaca 1140
tcagagattt tgagacacaa cgtggctttc cccggcccat gaccaaaatc ccttaacgtg 1200
agttttcgtt ccactgagcg tcagaccccg tagaaaagat caaaggatct tcttgagatc 1260
cttttttct gcgcgtaatc tgctgcttgc aaacaaaaaa accaccgcta ccagcggtgg 1320
tttgtttgcc ggatcaagag ctaccaactc tttttccgaa ggtaactggc ttcagcagag 1380
cgcagatacc aaatactgtc cttctagtgt agccgtagtt aggccaccac ttcaagaact 1440
ctgtagcacc gcctacatac ctcgctctgc taatcctgtt accagtggct gctgccagtg 1500
gcgataagtc gtgtcttacc gggttggact caagacgata gttaccggat aaggcgcagc 1560
ggtcgggctg aacggggggt tcgtgcacac agcccagctt ggagcgaacg acctacaccg 1620
aactgagata cctacagcgt gagctatgag aaagcgccac gcttcccgaa gggagaaagg 1680
cggacaggta tccggtaagc ggcagggtcg gaacaggaga gcgcacgagg gagcttccag 1740
ggggaaacgc ctggtatctt tatagtcctg tcgggtttcg ccacctctga cttgagcgtc 1800
gatttttgtg atgctcgtca ggggggcgga gcctatggaa aaacgccagc aacgcggcct 1860
ttttacggtt cctggccttt tgctggcctt ttgctcacat gttctttcct gcgttatccc 1920
ctgattctgt ggataaccgt attaccgcct ttgagtgagc tgataccgct cgccgcagcc 1980
gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga agagcgcctg atgcggtatt 2040
ttctccttac gcatctgtgc ggtatttcac accgcatatg gtgcactctc agtacaatct 2100
gctctgatgc cgcatagtta agccagtatc tgctccctgc ttgtgtgttg gaggtcgctg 2160
agtagtgcgc gagcaaaatt taagctacaa caaggcaagg cttgaccgac aattgcatga 2220
agaatctgct tagggttagg cgttttgcgc tgcttcgcga tgtacgggcc agatatagcc 2280
gcggcatcga tgatatccat tgcatacgtt gtatctatat cataatatgt acatttatat 2340
tggctcatgt ccaatatgac cgccatgttg acattgatta ttgactagtt attaatagta 2400
atcaattacg gggtcattag ttcatagccc atatatggag ttccgcgtta cataacttac 2460
ggtaaatggc ccgcctggct gaccgcccaa cgacccccgc ccattgacgt caataatgac 2520
gtatgttccc atagtaacgc caatagggac tttccattga cgtcaatggg tggagtattt 2580
acggtaaact gcccacttgg cagtacatca agtgtatcat atgccaagtc cgccccctat 2640
tgacgtcaat gacggtaaat ggcccgcctg gcattatgcc cagtacatga ccttacggga 2700

ctttcctact tggcagtaca tctacgtatt agtcatcgct attaccatgg tgatgcggtt 2760
ttggcagtac atcaatgggc gtggatagcg gtttgactca cggggatttc caagtctcca 2820
ccccattgac gtcaatggga gtttgttttg gcaccaaaat caacgggact ttccaaaatg 2880
tcgtaacaac tccgccccat tgacgcaaat gggcggtagg cgtgtacggt gggaggtcta 2940
tataagcaga gctcgtttag tgaaccgtca gatcgcctgg agacgccatc cacgctgttt 3000
tgacctccat agaagacacc gggaccgatc cagcctccgc gggcgcgcgt cgacagagag 3060
atgggtgcga gagcgtcagt attaagcggg ggagaattag atcgatggga aaaaattcgg 3120
ttaaggccag ggggaaagaa aaaatataaa ttaaaacata tagtatgggc aagcagggag 3180
ctagaacgat tcgcagttaa tcctggcctg ttagaaacat cagaaggctg tagacaaata 3240
ctgggacagc tacaaccatc ccttcagaca ggatcagaag aacttagatc attatataat 3300
acagtagcaa ccctctattg tgtgcatcaa aggatagaga taaaagacac caaggaagct 3360
ttagacaaga tagaggaaga gcaaaacaaa agtaagaaaa aagcacagca agcagcagct 3420
gacacaggac acagcaatca ggtcagccaa aattacccta tagtgcagaa catccagggg 3480
caaatggtac atcaggccat atcacctaga actttaaatg catgggtaaa agtagtagaa 3540
gagaaggctt tcagcccaga agtgataccc atgttttcag cattatcaga aggagccacc 3600
ccacaagatt taaacaccat gctaaacaca gtggggggac atcaagcagc catgcaaatg 3660
ttaaaagaga ccatcaatga ggaagctgca gaatgggata gagtgcatcc agtgcatgca 3720
gggcctattg caccaggcca gatgagagaa ccaaggggaa gtgacatagc aggaactact 3780
agtacccttc aggaacaaat aggatggatg acaaataatc cacctatccc agtaggagaa 3840
atttataaaa gatggataat cctgggatta aataaaatag taagaatgta tagccctacc 3900
agcattctgg acataagaca aggaccaaaa gaacccttta gagactatgt agaccggttc 3960
tataaaactc taagagccga gcaagcttca caggaggtaa aaaattggat gacagaaacc 4020
ttgttggtcc aaaatgcgaa cccagattgt aagactattt taaaagcatt gggaccagcg 4080
gctacactag aagaaatgat gacagcatgt cagggagtag gaggacccgg ccataaggca 4140
agagttttgt aggtttaaac taagccgaat tctgcagatc gcgccgagct cgctgatcag 4200
cctcgactgt gccttctagt tgccagccat ctgttgtttg cccctccccc gtgccttcct 4260
tgaccctgga aggtgccact cccactgtcc tttcctaata aaatgaggaa attgcatcgc 4320
attgtctgag taggtgtcat tctattctgg ggggtggggt ggggcaggac agcaaggggg 4380
aggattggga agacaatagc aggcatgctg gggaattt 4418

SEQ ID NO:8

aaatgggggc gctgaggtct gcctcgtgaa gaaggtgttg ctgactcata ccaggcctga 60
atcgccccat catccagcca gaaagtgagg gagccacggt tgatgagagc tttgttgtag 120
gtggaccagt tggtgatttt gaacttttgc tttgccacgg aacggtctgc gttgtcggga 180
agatgcgtga tctgatcctt caactcagca aaagttcgat ttattcaaca aagccgccgt 240
cccgtcaagt cagcgtaatg ctctgccagt gttacaacca attaaccaat tctgcgttca 300
aaatggtatg cgttttgaca catccactat atatccgtgt cgttctgtcc actcctgaat 360
cccattccag aaattctcta gcgattccag aagtttctca gagtcggaaa gttgaccaga 420
cattacgaac tggcacagat ggtcataacc tgaaggaaga tctgattgct taactgcttc 480
agttaagacc gacgcgctcg tcgtataaca gatgcgatga tgcagaccaa tcaacatggc 540
acctgccatt gctacctgta cagtcaagga tggtagaaat gttgtcggtc cttgcacacg 600
aatattacgc catttgcctg catattcaaa cagctcttct acgataaggg cacaaatcgc 660
atcgtggaac gtttgggctt ctaccgattt agcagtttga tacactttct ctaagtatcc 720
acctgaatca taaatcggca aaatagagaa aaattgacca tgtgtaagcg gccaatctga 780

ttccacctga gatgcataat ctagtagaat ctcttcgcta tcaaaattca cttccacctt 840
ccactcaccg gttgtccatt catggctgaa ctctgcttcc tctgttgaca tgacacacat 900
catctcaata tccgaatacg gaccatcagt ctgacgacca agagagccat aaacaccaat 960
agccttaaca tcatccccat atttatccaa tattcgttcc ttaatttcat gaacaatctt 1020
cattctttct tctctagtca ttattattgg tccgttcata acaccccttg tattactgtt 1080
tatgtaagca gacagtttta ttgttcatga tgatatattt ttatcttgtg caatgtaaca 1140
tcagagattt tgagacacaa cgtggctttc cccggcccat gaccaaaatc ccttaacgtg 1200
agttttcgtt ccactgagcg tcagaccccg tagaaaagat caaaggatct tcttgagatc 1260
cttttttct gcgcgtaatc tgctgcttgc aaacaaaaaa accaccgcta ccagcggtgg 1320
tttgtttgcc ggatcaagag ctaccaactc tttttccgaa ggtaactggc ttcagcagag 1380
cgcagatacc aaatactgtc cttctagtgt agccgtagtt aggccaccac ttcaagaact 1440
ctgtagcacc gcctacatac ctcgctctgc taatcctgtt accagtggct gctgccagtg 1500
gcgataagtc gtgtcttacc gggttggact caagacgata gttaccggat aaggcgcagc 1560
ggtcgggctg aacggggggt tcgtgcacac agcccagctt ggagcgaacg acctacaccg 1620
aactgagata cctacagcgt gagctatgag aaagcgccac gcttcccgaa gggagaaagg 1680
cggacaggta tccggtaagc ggcagggtcg gaacaggaga gcgcacgagg gagcttccag 1740
ggggaaacgc ctggtatctt tatagtcctg tcgggtttcg ccacctctga cttgagcgtc 1800
gatttttgtg atgctcgtca ggggggcgga gcctatggaa aaacgccagc aacgcggcct 1860
ttttacggtt cctggccttt tgctggcctt ttgctcacat gttctttcct gcgttatccc 1920
ctgattctgt ggataaccgt attaccgcct ttgagtgagc tgataccgct cgccgcagcc 1980
gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga agagcgcctg atgcggtatt 2040
ttctccttac gcatctgtgc ggtatttcac accgcatatg gtgcactctc agtacaatct 2100
gctctgatgc cgcatagtta agccagtatc tgctccctgc ttgtgtgttg gaggtcgctg 2160
agtagtgcgc gagcaaaatt taagctacaa caaggcaagg cttgaccgac aattgcatga 2220
agaatctgct tagggttagg cgttttgcgc tgcttcgcga tgtacgggcc agatatagcc 2280
gcggcatcga tgatatccat tgcatacgtt gtatctatat cataatatgt acatttatat 2340
tggctcatgt ccaatatgac cgccatgttg acattgatta ttgactagtt attaatagta 2400
atcaattacg gggtcattag ttcatagccc atatatggag ttccgcgtta cataacttac 2460
ggtaaatggc ccgcctggct gaccgcccaa cgacccccgc ccattgacgt caataatgac 2520
gtatgttccc atagtaacgc caatagggac tttccattga cgtcaatggg tggagtattt 2580
acggtaaact gcccacttgg cagtacatca agtgtatcat atgccaagtc cgccccctat 2640
tgacgtcaat gacggtaaat ggcccgcctg gcattatgcc cagtacatga ccttacggga 2700
ctttcctact tggcagtaca tctacgtatt agtcatcgct attaccatgg tgatgcggtt 2760
ttggcagtac atcaatgggc gtggatagcg gtttgactca cggggatttc caagtctcca 2820
ccccattgac gtcaatggga gtttgttttg gcaccaaaat caacgggact ttccaaaatg 2880
tcgtaacaac tccgccccat tgacgcaaat gggcggtagg cgtgtacggt gggaggtcta 2940
tataagcaga gctcgtttag tgaaccgtca gatcgcctgg agacgccatc cacgctgttt 3000
tgacctccat agaagacacc gggaccgatc cagcctccgc gggcgcgcgt cgacgccacc 3060
atgggggcgc gggcgtccgt cctctccggg ggggagctcg atcggtggga gaaaattcgg 3120
ctccggccgg gggggaagaa aaaatataaa ctcaaacata ttgtctgggc gtcccgggag 3180
ctcgagcggt tcgcggtcaa tccggggctg ctcgagacgt ccgagggctg tgcgcaaatt 3240
ctcgggcagc tccaaccgtc cctccagacg gggtccgagg agctccggtc cctctataat 3300
acggtcgcga cgctctattg tgtccatcaa cggattgaga ttaaagacac gaaggaggcg 3360
ctcgacaaga ttgaggagga gcaaaacaaa tccaagaaaa aagcgcagca agcggcggcg 3420
gacacggggc actccaatca ggtctcccaa aattacccga ttgtccagaa cattcagggg 3480
caaatggtcc atcaggcgat ttccccgcgg acgctcaatg cgtgggtcaa agtcgtcgag 3540
gagaaggcgt tctccccgga ggtcattccg atgttttcag cgctctccga gggggcgacg 3600
ccgcaagatc tcaacacgat gctcaacacg gtcgggggc atcaagcggc gatgcaaatg 3660

ctcaaagaga cgattaatga ggaggcggcg gagtgggatc gggtccatcc ggtccatgcg 3720
gggccgattg cgccggggca gatgcgggag ccgcgggggt ccgacattgc ggggacgacg 3780
tccacgctcc aggagcaaat tgggtggatg acgaataatc cgccgattcc ggtcggggag 3840
atttataaac ggtggattat tctcgggctc aataaaattg tccggatgta ttccccgacg 3900
tccattctcg acattcggca agggccgaag gagccgtttc gggactatgt agaccggttc 3960
tataaaacgc tccgggcgga gcaagcgtcc caggaggtca aaaattggat gacggagacg 4020
ctcctcgtcc aaaatgcgaa cccggattgt aagacgattc tcaaagcgct cgggccggcg 4080
gctacgctcg aggagatgat gacggcgtgt cagggggtcg ggggggccggg gcataaggcg 4140
cgggtcctct aatgaggcgc gccgagctcg ctgatcagcc tcgactgtgc cttctagttg 4200
ccagccatct gttgtttgcc cctcccccgt gccttccttg accctggaag gtgccactcc 4260
cactgtcctt tcctaataaa atgaggaaat tgcatcgcat tgtctgagta ggtgtcattc 4320
tattctgggg ggtggggtgg ggcaggacag caagggggag gattgggaag acaatagcag 4380
gcatgctggg gaattt 4396

[0179] A plasmid map of the plasmid construct set forth in SEQ ID NO:7 is provided as Figure 2 and a plasmid map of the plasmid construct as set forth in SEQ ID NO:8 is provided as Figure 3.

[0180] The results of two experiments to compare the levels of gag expression of the wild-type to the modified gene are provided in Table III.

Table III.

Experiment 1:

Expression from wild-type gag (plasmid construct of SEQ ID NO:7) = 8 ng/ml

Expression from modified gag (plasmid construct of SEQ ID NO:8) = 88 ng/ml

Experiment 2:

Expression from wild-type gag (SEQ ID NO:7) = 0.6 ng/ml

Expression from modified gag (SEQ ID NO:8) = 10 ng/ml

[0181] As indicated by the experimental results provided in Table III, the modified polynucleotide prepared in accordance with an embodiment of the present invention provided at least a ten fold increase in expression over its corresponding wild-type polynucleotide.

Example 3-Enhancement of Expression of HIV-1 gp160 envelope primary isolate 6101.

[0182] a. A third example illustrating the unexpected benefits of using “surrogate” codons in lieu of wild-type codons in a nucleic acid sequence was found for an HIV-1 gp160 envelope gene derived from a primary isolate 6101. The sequences (SEQ ID NO:5, the modified polynucleotide, and SEQ ID NO:6, the protein) are provided below.

SEQ ID NO:5 (polypeptide) and SEQ ID NO:6 (protein)

```
1    ATGCGGGCGAAGGAGATGCGGAAGTCCTGTCAGCACCTCCGGAAATGGGGGATTCTCCTCTTTGGGGTCCTCATGATTTGT
1    M  R  A  K  E  M  R  K  S  C  Q  H  L  R  K  W  G  I  L  L  F  G  V  L  M  I  C
82   TCCGCGGAGGAGAAGCTCTGGGTCACGGTCTATTATGGGGTCCCGGTCTGGAAAGAGGCGACGACGACGCTCTTTTGTGCG
28   S  A  E  E  K  L  W  V  T  V  Y  Y  G  V  P  V  W  K  E  A  T  T  T  L  F  C  A
163  TCCGATGCGAAGGCGCATCATGCGGAGGCGCATAATGTCTGGGCGACGCATGCGTGTGTCCCGACGGACCCGAACCCGCAA
55   S  D  A  K  A  H  H  A  E  A  H  N  V  W  A  T  H  A  C  V  P  T  D  P  N  P  Q
244  GAGGTCATTCTCGAGAATGTCACGGAGAAATATAACATGTGGAAAAATAACATGGTAGACCAGATGCATGAGGATATTATT
82   E  V  I  L  E  N  V  T  E  K  Y  N  M  W  K  N  N  M  V  D  Q  M  H  E  D  I  I
325  TCCCTCTGGGATCAATCCCTCAAGCCGTGTGTCAAACTCACGCCGCTCTGTGTCACGCTCAATTGCACGAATGCGACGTAT
109  S  L  W  D  Q  S  L  K  P  C  V  K  L  T  P  L  C  V  T  L  N  C  T  N  A  T  Y
406  ACGAATTCCGACTCCAAGAATTCCACTAGTAATTCCTCCCTCGAGGACTCCGGGAAAGGGGACATGAACTGCTCCTTCGAT
136  T  N  S  D  S  K  N  S  T  S  N  S  S  L  E  D  S  G  K  G  D  M  N  C  S  F  D
487  GTCACGACGTCCATTGATAAAAAGAAGAAGACGGAGTATGCGATTTTTGATAAACTCGATGTCATGAATATTGGGAATGGG
163  V  T  T  S  I  D  K  K  K  K  T  E  Y  A  I  F  D  K  L  D  V  M  N  I  G  N  G
568  CGGTATACGCTCCTCAATTGTAACACGTCCGTCATTACGCAGGCGTGTCCGAAGATGTCCTTTGAGCCGATTCCGATTCAT
190  R  Y  T  L  L  N  C  N  T  S  V  I  T  Q  A  C  P  K  M  S  F  E  P  I  P  I  H
649  TATTGTACGCCGGCGGGGTATGCGATTCTCAAGTGTAATGATAATAAGTTCAATGGGACGGGGCCGTGTACGAATGTCTCC
217  Y  C  T  P  A  G  Y  A  I  L  K  C  N  D  N  K  F  N  G  T  G  P  C  T  N  V  S
730  ACGATTCAATGTACGCATGGGATTAAGCCGGTCGTCTCCACGCAACTCCTCCTCAATGGATCCCTCGCGGAGGGGGGGGAG
244  T  I  Q  C  T  H  G  I  K  P  V  V  S  T  Q  L  L  L  N  G  S  L  A  E  G  G  E
811  GTCATTATTCGGTCCGAGAATCTCACGGACAATGCGAAAACGATTATTGTCCAGCTCAAGGAGCCGGTCGAGATTAATTGT
271  V  I  I  R  S  E  N  L  T  D  N  A  K  T  I  I  V  Q  L  K  E  P  V  E  I  N  C
892  ACGCGGCCGAACAACAATACGCGGAAATCCATTCATATGGGGCCGGGGGCGGCGTTTTATGCGCGGGGGGAGGTCATTGGG
298  T  R  P  N  N  N  T  R  K  S  I  H  M  G  P  G  A  A  F  Y  A  R  G  E  V  I  G
973  GATATTCGGCAAGCGCATTGCAACATTTCCCGGGGGCGGTGGAATGACACGCTCAAACAGATTGCGAAAAAGCTCCGGGAG
325  D  I  R  Q  A  H  C  N  I  S  R  G  R  W  N  D  T  L  K  Q  I  A  K  K  L  R  E
1054 CAATTTAATAAAACGATTTCCCTCAACCAATCCTCCGGGGGGGACCTCGAGATTGTCATGCACACGTTTAATTGTGGGGGG
352  Q  F  N  K  T  I  S  L  N  Q  S  S  G  G  D  L  E  I  V  M  H  T  F  N  C  G  G
1135 GAGTTTTTCTACTGTAATACGACGCAGCTCTTTAATTCCACGTGGAATGAGAATGATACGACGTGGAATAATACGGCGGGG
379  E  F  F  Y  C  N  T  T  Q  L  F  N  S  T  W  N  E  N  D  T  T  W  N  N  T  A  G
1216 TCCAATAACAATGAGACGATTACGCTCCCGTGTCGGATTAAACAAATTATTAACCGGTGGCAGGAGGTCGGGAAAGCGATG
406  S  N  N  N  E  T  I  T  L  P  C  R  I  K  Q  I  I  N  R  W  Q  E  V  G  K  A  M
1297 TATGCGCCGCCGATTTCCGGGCCGATTAATTGTCTCTCCAATATTACGGGGCTCCTCCTCACGCGTGATGGGGGGGACAAC
433  Y  A  P  P  I  S  G  P  I  N  C  L  S  N  I  T  G  L  L  L  T  R  D  G  G  D  N
1378 AATAATACGATTGAGACGTTCCGGCCGGGGGGGGGGGATATGCGGGACAATTGGCGGTCCGAGCTCTATAAATATAAAGTC
460  N  N  T  I  E  T  F  R  P  G  G  G  D  M  R  D  N  W  R  S  E  L  Y  K  Y  K  V
1459 GTCCGGATTGAGCCGCTCGGGATTGCGCCGACGAAGGCGAAGCGGCGGGTCGTCCAACGGGAGAAACGGGCGGTCGGGATT
487  V  R  I  E  P  L  G  I  A  P  T  K  A  K  R  R  V  V  Q  R  E  K  R  A  V  G  I
1540 GGGGCGATGTTCCTCGGGTTCCTCGGGGCGGCGGGGTCCACGATGGGGGCGGCGTCCGTCACGCTCACGGTCCAGGCGCGG
514  G  A  M  F  L  G  F  L  G  A  A  G  S  T  M  G  A  A  S  V  T  L  T  V  Q  A  R
1621 CTCCTCCTCTCCGGGATTGTCCAACAGCAAAACAATCTCCTCCGGGCGATTGAGGCGCAACAGCATCTCCTCCAACTCACG
541  L  L  L  S  G  I  V  Q  Q  Q  N  N  L  L  R  A  I  E  A  Q  Q  H  L  L  Q  L  T
1702 GTCTGGGGGATTAAGCAGCTCCAGGCGCGGGTCCTCGCGATGGAGCGGTACCTCAAGGATCAACAGCTCCTCGGGATTTGG
568  V  W  G  I  K  Q  L  Q  A  R  V  L  A  M  E  R  Y  L  K  D  Q  Q  L  L  G  I  W
1783 GGGTGCTCCGGGAAACTCATTTGCACGACGAATGTCCCGTGGAATGCGTCCTGGTCCAATAAATCCCTCGACAAGATTTGG
595  G  C  S  G  K  L  I  C  T  T  N  V  P  W  N  A  S  W  S  N  K  S  L  D  K  I  W
1864 CATAACATGACGTGGATGGAGTGGGACCGGGAGATTGACAATTACACGAAACTCATTTACACGCTCATTGAGGCGTCCCAG
622  H  N  M  T  W  M  E  W  D  R  E  I  D  N  Y  T  K  L  I  Y  T  L  I  E  A  S  Q
1945 ATTCAGCAGGAGAAGAATGAGCAAGAGCTCCTCGAGCTCGATTCCTGGGCGTCCCTCTGGTCCTGGTTTGACATTTCCAAA
649  I  Q  Q  E  K  N  E  Q  E  L  L  E  L  D  S  W  A  S  L  W  S  W  F  D  I  S  K
2026 TGGCTCTGGTATATTGGGGTCTTCATTATTGTCATTGGGGGGCTCGTCGGGCTCAAAATTGTCTTTGCGGTCCTCTCCATT
676  W  L  W  Y  I  G  V  F  I  I  V  I  G  G  L  V  G  L  K  I  V  F  A  V  L  S  I
2107 GTCAATCGGGTCCGGCAGGGGTACTCCCCGCTCTCCTTTCAGACGCGGCTCCCGGCGCCGCGGGGGCCGGACCGGCCGGAG
703  V  N  R  V  R  Q  G  Y  S  P  L  S  F  Q  T  R  L  P  A  P  R  G  P  D  R  P  E
2188 GGGATTGAGGAGGGGGGGGGGGAGCGGGACCGGGACAGATCTGATCAACTCGTCACGGGGTTCCTCGCGCTCATTTGGGAC
730  G  I  E  E  G  G  G  E  R  D  R  D  R  S  D  Q  L  V  T  G  F  L  A  L  I  W  D
2269 GATCTCCGGTCCCTCTGCCTCTTCTCCTACCACCGGCTCCGGGACCTCCTCCTCATTGTCGCGCGGATTGTCGAGCTCCTC
757  D  L  R  S  L  C  L  F  S  Y  H  R  L  R  D  L  L  L  I  V  A  R  I  V  E  L  L
2350 GGGCGGCGGGGGTGGGAGGCGCTCAAGTATTGGTGGAATCTCCTCCAATATTGGATTCAGGAGCTCAAGAATTCCGCGGTC
784  G  R  R  G  W  E  A  L  K  Y  W  W  N  L  L  Q  Y  W  I  Q  E  L  K  N  S  A  V
2431 TCCCTCCTCAACGCGACGGCGATTGCGGTCGCGGAGGGGACGGATCGGATTATTGAGGTCGTCCAACGGATTGGGCGGGCG
811  S  L  L  N  A  T  A  I  A  V  A  E  G  T  D  R  I  I  E  V  V  Q  R  I  G  R  A
2512 ATTCTCCACATTCCGCGGCGGATTCGGCAGGGGCTCGAGCGGGCGCTCCTCTAATGA
838  I  L  H  I  P  R  R  I  R  Q  G  L  E  R  A  L  L
```

[0183] Gene assembly methods were identical to those employed above for HIV-1 gag. Since this gp160 gene exceeds 2.5kb, it was assembled in 3 segments (each of approximately 800bp-900bp). A person skilled in the art would readily be able to select and assemble suitable segments.

[0184] The plasmid construct set forth in SEQ ID NO:9 was used as the vector for transfection of the modified polynucleotide prepared in accordance with an embodiment of the present invention.

SEQ ID NO:9:

aaatgggggc gctgaggtct gcctcgtgaa gaaggtgttg ctgactcata ccaggcctga 60
atcgccccat catccagcca gaaagtgagg gagccacggt tgatgagagc tttgttgtag 120
gtggaccagt tggtgatttt gaacttttgc tttgccacgg aacggtctgc gttgtcggga 180
agatgcgtga tctgatcctt caactcagca aaagttcgat ttattcaaca aagccgccgt 240
cccgtcaagt cagcgtaatg ctctgccagt gttacaacca attaaccaat tctgcgttca 300
aaatggtatg cgttttgaca catccactat atatccgtgt cgttctgtcc actcctgaat 360
cccattccag aaattctcta gcgattccag aagtttctca gagtcggaaa gttgaccaga 420
cattacgaac tggcacagat ggtcataacc tgaaggaaga tctgattgct taactgcttc 480
agttaagacc gacgcgctcg tcgtataaca gatgcgatga tgcagaccaa tcaacatggc 540
acctgccatt gctacctgta cagtcaagga tggtagaaat gttgtcggtc cttgcacacg 600
aatattacgc catttgcctg catattcaaa cagctcttct acgataaggg cacaaatcgc 660
atcgtggaac gtttgggctt ctaccgattt agcagtttga tacactttct ctaagtatcc 720
acctgaatca taaatcggca aaatagagaa aaattgacca tgtgtaagcg gccaatctga 780
ttccacctga gatgcataat ctagtagaat ctcttcgcta tcaaaattca cttccacctt 840
ccactcaccg gttgtccatt catggctgaa ctctgcttcc tctgttgaca tgacacacat 900
catctcaata tccgaatacg gaccatcagt ctgacgacca agagagccat aaacaccaat 960
agccttaaca tcatccccat atttatccaa tattcgttcc ttaatttcat gaacaatctt 1020
cattctttct tctctagtca ttattattgg tccgttcata acaccccttg tattactgtt 1080
tatgtaagca gacagtttta ttgttcatga tgatatattt ttatcttgtg caatgtaaca 1140
tcagagattt tgagacacaa cgtggctttc cccggcccat gaccaaaatc ccttaacgtg 1200
agttttcgtt ccactgagcg tcagaccccg tagaaaagat caaaggatct tcttgagatc 1260
cttttttct gcgcgtaatc tgctgcttgc aaacaaaaaa accaccgcta ccagcggtgg 1320
tttgtttgcc ggatcaagag ctaccaactc tttttccgaa ggtaactggc ttcagcagag 1380
cgcagatacc aaatactgtc cttctagtgt agccgtagtt aggccaccac ttcaagaact 1440
ctgtagcacc gcctacatac ctcgctctgc taatcctgtt accagtggct gctgccagtg 1500
gcgataagtc gtgtcttacc gggttggact caagacgata gttaccggat aaggcgcagc 1560
ggtcgggctg aacggggggt tcgtgcacac agcccagctt ggagcgaacg acctacaccg 1620
aactgagata cctacagcgt gagctatgag aaagcgccac gcttcccgaa gggagaaagg 1680
cggacaggta tccggtaagc ggcagggtcg gaacaggaga gcgcacgagg gagcttccag 1740
ggggaaacgc ctggtatctt tatagtcctg tcgggtttcg ccacctctga cttgagcgtc 1800
gatttttgtg atgctcgtca ggggggcgga gcctatggaa aaacgccagc aacgcggcct 1860
ttttacggtt cctggccttt tgctggcctt ttgctcacat gttctttcct gcgttatccc 1920
ctgattctgt ggataaccgt attaccgcct ttgagtgagc tgataccgct cgccgcagcc 1980
gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga agagcgcctg atgcggtatt 2040
ttctccttac gcatctgtgc ggtatttcac accgcatatg gtgcactctc agtacaatct 2100
gctctgatgc cgcatagtta agccagtatc tgctccctgc ttgtgtgttg gaggtcgctg 2160
agtagtgcgc gagcaaaatt taagctacaa caaggcaagg cttgaccgac aattgcatga 2220

agaatctgct tagggttagg cgttttgcgc tgcttcgcga tgtacgggcc agatatagcc 2280
gcggcatcga tgatatccat tgcatacgtt gtatctatat cataatatgt acatttatat 2340
tggctcatgt ccaatatgac cgccatgttg acattgatta ttgactagtt attaatagta 2400
atcaattacg gggtcattag ttcatagccc atatatggag ttccgcgtta cataacttac 2460
ggtaaatggc ccgcctggct gaccgcccaa cgacccccgc ccattgacgt caataatgac 2520
gtatgttccc atagtaacgc caatagggac tttccattga cgtcaatggg tggagtattt 2580
acggtaaact gcccacttgg cagtacatca agtgtatcat atgccaagtc cgccccctat 2640
tgacgtcaat gacggtaaat ggcccgcctg gcattatgcc cagtacatga ccttacggga 2700
ctttcctact tggcagtaca tctacgtatt agtcatcgct attaccatgg tgatgcggtt 2760
ttggcagtac atcaatgggc gtggatagcg gtttgactca cggggatttc caagtctcca 2820
ccccattgac gtcaatggga gtttgttttg gcaccaaaat caacgggact ttccaaaatg 2880
tcgtaacaac tccgccccat tgacgcaaat gggcggtagg cgtgtacggt gggaggtcta 2940
tataagcaga gctcgtttag tgaaccgtca gatcgcctgg agacgccatc cacgctgttt 3000
tgacctccat agaagacacc gggaccgatc cagcctccgc gggcgcgcgt cgacgccacc 3060
atgcgggcga aggagatgcg gaagtcctgt cagcacctcc ggaaatgggg gattctcctc 3120
tttggggtcc tcatgatttg ttccgcggag gagaagctct gggtcacggt ctattatggg 3180
gtcccggtct ggaaagaggc gacgacgacg ctcttttgtg cgtccgatgc gaaggcgcat 3240
catgcggagg cgcataatgt ctgggcgacg catgcgtgtg tcccgacgga cccgaacccg 3300
caagaggtca ttctcgagaa tgtcacggag aaatataaca tgtggaaaaa taacatggta 3360
gaccagatgc atgaggatat tatttccctc tgggatcaat ccctcaagcc gtgtgtcaaa 3420
ctcacgccgc tctgtgtcac gctcaattgc acgaatgcga cgtatacgaa ttccgactcc 3480
aagaattcca ctagtaattc ctccctcgag gactccggga aaggggacat gaactgctcc 3540
ttcgatgtca cgacgtccat tgataaaaag aagaagacgg agtatgcgat ttttgataaa 3600
ctcgatgtca tgaatattgg gaatgggcgg tatacgctcc tcaattgtaa cacgtccgtc 3660
attacgcagg cgtgtccgaa gatgtccttt gagccgattc cgattcatta ttgtacgccg 3720
gcggggtatg cgattctcaa gtgtaatgat aataagttca atgggacggg gccgtgtacg 3780
aatgtctcca cgattcaatg tacgcatggg attaagccgg tcgtctccac gcaactcctc 3840
ctcaatggat ccctcgcgga ggggggggag gtcattattc ggtccgagaa tctcacggac 3900
aatgcgaaaa cgattattgt ccagctcaag gagccggtcg agattaattg tacgcggccg 3960
aacaacaata cgcggaaatc cattcatatg ggccggggg cggcgtttta tgcgcggggg 4020
gaggtcattg gggatattcg gcaagcgcat tgcaacattt cccgggggcg gtggaatgac 4080
acgctcaaac agattgcgaa aaaactccgg gagcaattta ataaaacgat ttccctcaac 4140
caatcctccg ggggggacct cgagattgtc atgcacacgt ttaattgtgg gggggagttt 4200
ttctactgta atacgacgca gctctttaat tccacgtgga atgagaatga tacgacgtgg 4260
aataatacgg cggggtccaa taacaatgag acgattacgc tcccgtgtcg gattaaacaa 4320
attattaacc ggtggcagga ggtcgggaaa gcgatgtatg cgccgccgat ttccgggccg 4380
attaattgtc tctccaatat tacggggctc ctcctcacgc gtgatggggg ggacaacaat 4440
aatacgattg agacgttccg gccggggggg ggggatatgc gggacaattg gcggtccgag 4500
ctctataaat ataaagtcgt ccggattgag ccgctcggga ttgcgccgac gaaggcgaag 4560
cggcgggtcg tccaacggga gaaacgggcg gtcgggattg gggcgatgtt cctcgggttc 4620
ctcggggcgg cggggtccac gatgggggcg gcgtccgtca cgctcacggt ccaggcgcgg 4680
ctcctcctct ccgggattgt ccaacagcaa aacaatctcc tccgggcgat tgaggcgcaa 4740
cagcatctcc tccaactcac ggtctggggg attaagcagc tccaggcgcg ggtcctcgcg 4800
atggagcggt acctcaagga tcaacagctc ctcgggattt gggggtgctc cgggaaactc 4860
atttgcacga cgaatgtccc gtggaatgcg tcctggtcca ataaatccct cgacaagatt 4920
tggcataaca tgacgtggat ggagtgggac cgggagattg acaattacac gaaactcatt 4980
tacacgctca ttgaggcgtc ccagattcag caggagaaga atgagcaaga gctcctcgag 5040
ctcgattcct gggcgtccct ctggtcctgg tttgacattt ccaaatggct ctggtatatt 5100
ggggtcttca ttattgtcat tgggggggctc gtcgggctca aaattgtctt tgcggtcctc 5160
tccattgtca atcgggtccg gcaggggtac tccccgctct cctttcagac gcggctcccg 5220
gcgccgcggg ggccggaccg gccggagggg attgaggagg gggggggga gcgggaccgg 5280
gacagatctg atcaactcgt cacggggttc ctcgcgctca tttgggacga tctccggtcc 5340

ctctgcctct tctcctacca ccggctccgg gacctcctcc tcattgtcgc gcggattgtc 5400
gagctcctcg ggcggcgggg gtgggaggcg ctcaagtatt ggtggaatct cctccaatat 5460
tggattcagg agctcaagaa ttccgcggtc tccctcctca acgcgacggc gattgcggtc 5520
gcggaggga cggatcggat tattgaggtc gtccaacgga ttgggcgggc gattctccac 5580
attccgcggc ggattcggca ggggctcgag cgggcgctcc tctaatgagg cgcgccgagc 5640
tcgctgatca gcctcgactg tgccttctag ttgccagcca tctgttgttt gcccctcccc 5700
cgtgccttcc ttgaccctgg aaggtgccac tcccactgtc ctttcctaat aaaatgagga 5760
aattgcatcg cattgtctga gtaggtgtca ttctattctg gggggtgggg tggggcagga 5820
cagcaagggg gaggattggg aagacaatag caggcatgct ggggaattt 5869

The plasmid construct set forth in SEQ ID NO:10 is the vector for the transfection of the wild-type gene.

SEQ ID NO:10:

aaatgggggc gctgaggtct gcctcgtgaa gaaggtgttg ctgactcata ccaggcctga 60
atcgccccat catccagcca gaaagtgagg gagccacggt tgatgagagc tttgttgtag 120
gtggaccagt tggtgatttt gaacttttgc tttgccacgg aacggtctgc gttgtcggga 180
agatgcgtga tctgatcctt caactcagca aaagttcgat ttattcaaca aagccgccgt 240
cccgtcaagt cagcgtaatg ctctgccagt gttacaacca attaaccaat tctgcgttca 300
aaatggtatg cgttttgaca catccactat atatccgtgt cgttctgtcc actcctgaat 360
cccattccag aaattctcta gcgattccag aagtttctca gagtcggaaa gttgaccaga 420
cattacgaac tggcacagat ggtcataacc tgaaggaaga tctgattgct taactgcttc 480
agttaagacc gacgcgctcg tcgtataaca gatgcgatga tgcagaccaa tcaacatggc 540
acctgccatt gctacctgta cagtcaagga tggtagaaat gttgtcggtc cttgcacacg 600
aatattacgc catttgcctg catattcaaa cagctcttct acgataaggg cacaaatcgc 660
atcgtggaac gtttgggctt ctaccgattt agcagtttga tacactttct ctaagtatcc 720
acctgaatca taaatcggca aaatagagaa aaattgacca tgtgtaagcg gccaatctga 780
ttccacctga gatgcataat ctagtagaat ctcttcgcta tcaaaattca cttccacctt 840
ccactcaccg gttgtccatt catggctgaa ctctgcttcc tctgttgaca tgacacacat 900
catctcaata tccgaatacg gaccatcagt ctgacgacca agagagccat aaacaccaat 960
agccttaaca tcatccccat atttatccaa tattcgttcc ttaatttcat gaacaatctt 1020
cattctttct tctctagtca ttattattgg tccgttcata acaccccttg tattactgtt 1080
tatgtaagca gacagtttta ttgttcatga tgatatattt ttatcttgtg caatgtaaca 1140
tcagagattt tgagacacaa cgtggctttc cccggcccat gaccaaaatc ccttaacgtg 1200
agttttcgtt ccactgagcg tcagaccccg tagaaaagat caaaggatct tcttgagatc 1260
cttttttct gcgcgtaatc tgctgcttgc aaacaaaaaa accaccgcta ccagcggtgg 1320
tttgtttgcc ggatcaagag ctaccaactc tttttccgaa ggtaactggc ttcagcagag 1380
cgcagatacc aaatactgtc cttctagtgt agccgtagtt aggccaccac ttcaagaact 1440
ctgtagcacc gcctacatac ctcgctctgc taatcctgtt accagtggct gctgccagtg 1500
gcgataagtc gtgtcttacc gggttggact caagacgata gttaccggat aaggcgcagc 1560
ggtcgggctg aacggggggt tcgtgcacac agcccagctt ggagcgaacg acctacaccg 1620
aactgagata cctacagcgt gagctatgag aaagcgccac gcttcccgaa gggagaaagg 1680
cggacaggta tccggtaagc ggcagggtcg gaacaggaga gcgcacgagg gagcttccag 1740
ggggaaacgc ctggtatctt tatagtcctg tcgggtttcg ccacctctga cttgagcgtc 1800
gatttttgtg atgctcgtca ggggggcgga gcctatggaa aaacgccagc aacgcggcct 1860
ttttacggtt cctggccttt tgctggcctt ttgctcacat gttctttcct gcgttatccc 1920
ctgattctgt ggataaccgt attaccgcct ttgagtgagc tgataccgct cgccgcagcc 1980
gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga agagcgcctg atgcggtatt 2040
ttctccttac gcatctgtgc ggtatttcac accgcatatg gtgcactctc agtacaatct 2100
gctctgatgc cgcatagtta agccagtatc tgctccctgc ttgtgtgttg gaggtcgctg 2160

agtagtgcgc gagcaaaatt taagctacaa caaggcaagg cttgaccgac aattgcatga 2220
agaatctgct tagggttagg cgttttgcgc tgcttcgcga tgtacgggcc agatatagcc 2280
gcggcatcga tgatatcgcg gctatctgag ggactaggg tgtgtttagg cgaaaagcgg 2340
ggcttcggtt gtacgcggtt aggagtcccc tcaccattgc atacgttgta tctatatcat 2400
aatatgtaca tttatattgg ctcatgtcca atatgaccgc catgttgaca ttgattattg 2460
actagttatt aatagtaatc aattacgggg tcattagttc atagcccata tatggagttc 2520
cgcgttacat aacttacggt aaatggcccg cctggctgac cgcccaacga cccccgccca 2580
ttgacgtcaa taatgacgta tgttcccata gtaacgccaa tagggacttt ccattgacgt 2640
caatgggtgg agtatttacg gtaaactgcc cacttggcag tacatcaagt gtatcatatg 2700
ccaagtccgc cccctattga cgtcaatgac ggtaaatggc ccgcctggca ttatgcccag 2760
tacatgacct tacgggactt tcctacttgg cagtacatct acgtattagt catcgctatt 2820
accatggtga tgcggttttg gcagtacatc aatgggcgtg gatagcggtt tgactcacgg 2880
ggatttccaa gtctccaccc cattgacgtc aatgggagtt tgttttggca ccaaaatcaa 2940
cgggactttc caaaatgtcg taacaactcc gccccattga cgcaaatggg cggtaggcgt 3000
gtacggtggg aggtctatat aagcagagct cgtttagtga accgtcagat cgcctggaga 3060
cgccatccac gctgttttga cctccataga agacaccggg accgatccag cctccgcggg 3120
cgcgcgtcga cgccaccatg agagcgaagg agatgaggaa gagttgtcag cacttgagga 3180
aatggggcat cttgctcttt ggagtgttga tgatctgtag tgctgaagaa aagttgtggg 3240
tcacagtcta ttatggggta cctgtgtgga aagaagcaac caccactcta ttttgtgcat 3300
cagatgctaa ggcacatcat gcagaggcac ataatgtttg ggccacacat gcctgtgtac 3360
ccacagaccc taacccacaa gaagtaatat tggaaaatgt gacagaaaaa tataacatgt 3420
ggaaaaataa catggtagac cagatgcatg aggatataat cagtttatgg gatcaaagcc 3480
taaagccatg tgtaaaatta accccactct gtgttacttt aaattgcact aatgcgacgt 3540
atactaatag tgacagtaag aatagtacca gtaatagtag tttggaagac agtgggaaag 3600
gagacatgaa ctgctctttc gatgtcacca caagcataga taaaaagaag aagacagaat 3660
atgcaatttt tgataaactt gatgtaatga atataggtaa tggaagatat acattactaa 3720
attgtaacac ctcagtcatt acacaggcct gtccaaagat gtcctttgaa ccaattccca 3780
tacattattg taccccggct ggttatgcga ttctaaagtg taatgataat aagttcaatg 3840
gaacaggacc atgtacaaat gtcagcacaa tacaatgtac acatggaatt aagccagtag 3900
tgtcaactca actgctgtta aatggcagtc tagcagaagg aggagaggta ataattagat 3960
ctgaaaatct cacagacaat gctaaaacca taatagtaca gctcaaggaa cctgtagaaa 4020
tcaattgtac aagacccaac aacaatacaa gaaaaagtat acatatggga ccaggagcag 4080
cattttatgc aagaggagaa gtaataggag atataagaca agcacattgc aacattagta 4140
gaggaagatg gaatgacact ttaaaacaga tagctaaaaa attaagagaa caatttaata 4200
aaacaataag ccttaaccaa tcctcaggag gggacctaga aattgtaatg cacactttta 4260
attgtggagg ggaatttttc tactgtaata caacacagct gtttaatagt acttggaatg 4320
agaatgatac tacctggaat aatacagcag ggtcaaataa caatgaaact atcacactcc 4380
catgtagaat aaaacaaatt ataaacaggt ggcaggaagt aggaaaagca atgtatgccc 4440
ctcccatcag tggaccaatt aattgtttat caaatatcac agggctatta ttaacaagag 4500
atggtggtga caacaataat acaatagaga ccttcagacc tggaggagga gatatgaggg 4560
acaattggag aagtgaatta tataaatata aagtagtaag aattgagcca ttaggaatag 4620
cacccaccaa ggcaaagaga agagtggtgc aaagagaaaa aagagcagtg ggaataggag 4680
ctatgttcct tgggttcttg ggagcagcag gaagcactat gggcgcagcg tcagtgacgc 4740
tgacggtaca ggccagacta ttattgtctg gtatagtgca acagcaaaac aatttgctga 4800
gagctatcga ggcgcaacag catctgttgc aactcacagt ctggggcatc aagcagctcc 4860
aggctagagt cctggctatg gaaagatacc taaaggatca acagctccta gggatttggg 4920
gttgctctgg aaaactcatt tgcaccacta atgtgccttg gaatgctagt tggagtaata 4980
aatctctgga caagatttgg cataacatga cctggatgga gtgggacaga gaaattgaca 5040
attacacaaa attaatatac accttaattg aagcatcgca gatccagcag gaaaagaatg 5100
aacaagaatt attggaattg gatagttggg caagtttgtg gagttggttt gacatctcaa 5160
aatggctgtg gtatatagga gtattcataa tagtaatagg aggtttagta ggtttaaaaa 5220
tagtttttgc tgtactttct atagtaaata gagttaggca gggatactca ccattatcat 5280

ttcagacccg cctcccagcc ccgcggggac ccgacaggcc cgaaggaatc gaagaaggag 5340
gtggagagag agacagagac agatccgatc aattagtgac tggattctta gcactcatct 5400
gggacgatct gcggagcctg tgcctcttca gctaccaccg cttgagagac ttactcttga 5460
ttgtagcgag gattgtggaa cttctgggac gcagggggtg ggaagccctg aagtattggt 5520
ggaatctcct gcaatattgg attcaggaac taaagaatag tgctgttagt ttgcttaacg 5580
ccacagctat agcagtagcc gaggggacag ataggattat agaagtagta caaaggattg 5640
gtagagctat tctccacata cctagaagaa taagacaggg cttagaaagg gctttgctat 5700
aatagggcgc gccgagctcg ctgatcagcc tcgactgtgc cttctagttg ccagccatct 5760
gttgtttgcc cctcccccgt gccttccttg accctggaag gtgccactcc cactgtcctt 5820
tcctaataaa atgaggaaat tgcatcgcat tgtctgagta ggtgtcattc tattctgggg 5880
ggtggggtgg ggcaggacag caagggggag gattgggaag acaatagcag gcatgctggg 5940
gaattt 5946

[0185] A plasmid map of the plasmid construct set forth in SEQ ID NO:9 is provided as Figure 4 and a plasmid map of the plasmid construct set forth in SEQ ID NO:10 is provided as Figure 5.

[0186] Western blot detection and ELISA methods were employed to compare transfected cells expressing the wild type or the modified gp160 genes.

[0187] Two Western blots confirmed gp160 antigen specificity from SEQ ID NO:9 plasmid construct-transfected 293 cells forty eight hours later (data not shown). Initial studies tested two SEQ ID NO:9 plasmid construct clones with later focus on clone 6, hereafter just denoted SEQ ID NO:9. These Western blots demonstrated recognition of SEQ ID NO: 9 plasmid construct-transfected lysates by both an anti IIIB gp120 polyclonal rabbit serum as well as an anti-MN gp41 monoclonal antibody (data not shown). Each blot revealed reactivity with their respective positive control recombinant proteins (451 for gp160 and MN expressed in E. coli for gp41. Since the amino acid sequences differ between the 6101 primary isolate (encoded by the SEQ ID NO:9 plasmid construct) and the MN strain, no direct quantitative comparisons can be made between these envelopes in these Western blots or in the ELISA assays listed below.

[0188] Enhanced expression levels of the 6101 gp160 envelope geneaccording to an embodiment of the present invention was observed. The plasmid construct for the gene modified in accordance with an embodiment of the present invention (SEQ ID NO:9) expressed substantially higher levels of gp160 compared to the wild-type 6101 gene (which was undetectable by Western blot). Envelope 6101 gp160 expression levels were quantified

for 293 as well as for COS-7, Hela, and RD cell lines after transient transfection from total cell lysates using an anti-gp120 ELISA capture kit (ABI, Cat No. 15-102-000).

Table IV. HIV-1 Gp160 6101 protein levels (in ng/ml) from total cell lysates

	Cells			
Constructs	COS-7	Hela	RD	293
construct for modified polynucleotide (SEQ ID NO:9)	4	5.4	0.8	80
construct for wild-type (SEQ ID NO:10)	**	**	**	**

*Lower limit of standard curve = 78pg/ml
**not detected

[0189] From these studies it can be concluded that the construct for the modified gene (SEQ ID NO: 9) expresses the altered 6101 gp160 protein at levels far superior (almost 100 times) to its wild-type counterpart (SEQ ID NO:10) in several cell lines (as shown in Table IV). Quantification of this primary isolate can be achieved by an ABI anti-gp120 ELISA kit and is at substantially lower levels than observed for p37 gag (in the ug/ml range in cell lysates).

Example 4-Modification of the env gene increased gp160 protein levels relative to wild-type

[0190] A further study comparing the expression of a modified polynucleotide of an embodiment of the present invention for gp160 to the wild-type version of the gene was conducted.

[0191] For the purposes of the study, a modified polynucleotide of an embodiment of the present invention for gp160 was prepared as described in Example 3 above. A wild-type gp160 polynucleotide for the gene was also obtained for the study.

[0192] Expression of the two types of polynucleotides was measured using the systems described in Examples 1-3 above.

[0193] Referring to Figure 1, the results of the study are illustrated by the graph. As is clearly shown, the modified polynucleotide of an embodiment of the present invention for the gp160 ("optimized") gene provides substantially better expression than the wild-type gene.

Example 5-Enhanced expression of human IL-15

[0194] A study was conducted to compare IL-15 expression by various IL-15 constructs in accordance with embodiments of the present invention, such as an IL-15 recombinant construct (modified with surrogate codons) with a human IgE leader sequence or with the long leader sequence, unmodified IL-15 with an IgE leader, and two alternative optimized IL-15 constructs with IgE leader against expression by other IL-15 constructs. The results of the study show that the constructs of the present invention provide unexpectedly improved expression of IL-15. In particular, the IgE leader sequence in combination with the less intensive modified surrogate codon approach provides synergistically improved expression over currently used IL-15 constructs and comparable results to codon optimized or "preferred codon" approaches with a lower intensive and thus highly efficient and accurate approach. The experimental procedures and results are described below and illustrated in the following Tables and in Figs. 6-10.

[0195] Various constructs were used for comparative purposes, as follows:

1. IL-15 constructs with the native IL-15 signal peptide replaced by the human IgE leader sequence.

2. IL-15 constructs with optimized codons (codon optimization alternative 1.

3. IL-15 constructs with the IL-15 nucleotide sequence optimized to reduce mRNA secondary structure (codon optimization alternative 2).

4. IL-15 constructs with combinations of IgE leader sequences and gene optimization techniques.

Cloning:

[0196] All gene sequences were designed based upon published codon tables and synthesized from Blue Heron Technologies. Genes were then subcloned into the DNA vaccine vector backbone.

Cell Culture and Transfection:

[0197] RD, 293, Hela and COS-7 cells were used in transient transfections. All transfections were carried out using Fugene-6 (Roche) according to the manufacturer's instructions. A total of 0.25 mg of human IL-15 plasmid and 0.5 mg of SEAP (a secreted form of human placental alkaline phosphatase) control vector with 4 ml of Fugene-6 was used for each transfection. For dose titration, 0.25-2.0 mg of the test plasmid was used along with the control DNA and the total DNA was made up to a final concentration of 2.0 mg per trannsfection. Dose titration was performed to identify an appropriate concentration of plasmid to be used for comparative analysis. Forty-eight hours after transfection, cell culture media and cells were harvested and analyzed for IL-15 by ELISA (R&D Systems) and CTLL2 proliferation assay. The cell lysates were tested for total protein concentration by Micro BCA protein assay. Data is depicted as pg of IL-15 per mg of protein in cell lysates and pg of IL-15 per 10,000 units of seap activity.

Intramuscular Immunization of Mice:

[0198] Six to eight-week-old female BALB/c mice were used in this study. Each group consisted of 2 animals and mice were immunized intramuscularly in both quadriceps muscles with a total of 200 mg plasmid DNA (formulated with 0.25% bupivacaine) in a 50 ml volume using a 28-gauge needle. In all 4 muscles were analysed at each time point. The quadriceps muscles were taken at 2, 5, 9 and 15 days post-immunization and homogenized in cell lysis buffer (50mM Tris, pH8.0 – 50mM NaCl – 1%Triton-X100) containing proteinase inhibitor mixture (Roche). The cell lysates were subjected to three freeze and thaw cycles, centrifuged and supernatants were evaluated for IL-15 protein by ELISA (R&D Systems). Data are represented as average expression in 4 muscle samples per group.

CTLL2 cell proliferation assay

[0199] Mouse CTLL2 cells were washed twice with PBS and incubated in a 96 well-plate at a density of 100000 cells /well in complete medium with either different amounts of human recombinant IL-15 (R&D Systems) as standard controls or indicated media of cells transfected with hIL-15 expression construct. Forty eight hours post-incubation, MTT reagent (3-(4,5-dimethylthiazolyl-2)-2, 5-diphenyltetrazolium bromide) was added and further incubated for four hours. Conversion of the tetrazolium salt to the purple formazon product by mitochondrial enzymes in viable cells allows a visual assessment of the reaction. When the purple formazon precipitate was clearly visible in the microscope the cells were lysed with the detergent and absorbances read at 570 nm. Final concentration is based upon the

known standards used in the assay and data are represented as pg of IL-15 per ml of supernatant from transfected cells.

Results:

Human IL-15 Constructs:

[0200] The following seven human IL-15 inserts were subcloned into a vector backbone, which contains human CMV promoter. All the constructs were confirmed by sequencing and used for *in vitro* and *in vivo* human IL-15 expression assays.

Construct	Description
+++++______________	LP-IL-15-IgE leader (surrogate codons)
---------______________	Current clinical IL-15 (native IL-15 with long signal peptide)
+++++______________	Native IL-15 with IL-15-IgE leader that replaces the long signal peptide
+++++______________	O-IL-15-IgE leader (preffered codons)
+++++______________	BH-IL-15-IgE leader (secondary structure optimization)
---------______________	O-15 with a long signal peptide
---------______________	LP-15 with a long signal peptide
---------______________	RNA optimization with a long signal peptide

------ **Native Leader Sequence** +++++ **IgE Leader Sequence**

[0201] As shown in Table V(A) and V(B), constructs according to embodiments of the present invention significantly improve IL-15 expression *in vitro*. In particular, Table V(A) shows expression in cells and supernatants of 293 cells. Table V(B) shows expression in cells and supernatants of RD cells

(A)

Human IL-15 expression in 293 cell lysates (ELISA)		
Group	human IL15 (pg/mg protein)	Fold increase compared to WLV125M
WLV125M	7139.83	1.00
WLV134M	23893.23	3.35
WLV186M	123002.31	17.23
WLV187M	80523.75	11.28
WLV188M	29772.71	4.17
WLV211M	33000.66	4.62
WLV217M	11403.65	1.60
WLV225M	29103.13	4.08
WLV001AM	0.00	0

Human IL-15 expression in 293 cell supernatants (ELISA)		
Group	human IL15 (pg/ml/10000unit SEAP)	Fold increase compared to WLV125M
WLV125M	64.24	1.00
WLV134M	928.76	14.46
WLV186M	6807.04	105.96
WLV187M	4389.32	68.33
WLV188M	1327.20	20.66
WLV211M	967.94	15.07
WLV217M	217.81	3.39
WLV225M	1556.50	24.23
WLV001AM	0.00	0

(B)

Human IL-15 expression in RD cell supernatants (ELISA)		
Group	human IL15 (pg/ml/10000unit SEAP)	Fold increase compared to WLV125M
WLV125M	72.97	1.00
WLV134M	528.40	7.24
WLV186M	9544.01	130.79
WLV187M	4102.73	56.22
WLV188M	1548.02	21.21
WLV211M	6287.93	86.17
WLV217M	407.16	5.58
WLV225M	1958.41	26.84
WLV001AM	0.00	0

Human IL-15 expression in RD cell lysates (EL		
Group	human IL15 (pg/mg protein)	Fold incre compare WLV12
WLV125M	1056.64	
WLV134M	2786.32	
WLV186M	20877.53	
WLV187M	7287.57	
WLV188M	3275.43	
WLV211M	6183.53	
WLV217M	1409.34	
WLV225M	4443.84	
WLV001AM	0.00	

[0202] Table VI shows *in vivo* gene expression from IL-15 constructs in accordance with the invention as well as previously used IL-15 constructs for purposes of comparison. Codon engineering in addition to the replacement of the native signal peptide with human IgE leader significantly improved IL-15 expression *in vivo*. Four mice per group received 200 mg of plasmid DNA. Animals were sacrificed and analyzed at 2, 5, 9 and 15 days after immunization. Data summarized are an average IL-15 protein expression from a group of 4 muscles per time point.

Groups	Human IL-15 expression in the mouse muscles(pg/10 mg of protein)			
	Day 2	Day 5	Day 9	Day 15
WLV125M	2.959	2.714	2.889	0.845
WLV134M	4.134	3.028	2.927	0.811
WLV186M	25.846	31.830	3.403	1.220
WLV187M	15.072	4.826	2.499	0.829

[0203] Table VII shows the results of the CTLL2 assay. Supernatants from RD cells transfected with optimized constructs induced 5-30 fold higher functional IL-15 than the native plasmid in a MTT cell proliferation bioassay (see materials and methods for details). The proliferation rate was estimated from a standard curve obtained with purified recombinant human IL-15 (pg/ml).

Human IL-15 expression in 293 cell lysates (CTLL2 Assay)		
Group	human IL15 (ng/ml of supernatant)	Fold increase compared to WLV125M
WLV125M	3.12	1.00
WLV134M	16.22	5.19
WLV186M	98.95	31.69
WLV187M	71.42	22.87
WLV188M	34.36	11.01
WLV001AN	0.00	0.00

[0204] The foregoing study demonstrates that various gene modification strategies significantly improve human IL-15 expression. Replacement of native IL-15 signal peptide sequence with that of human IgE leader up-regulated its expression by 5-8 fold demonstrating the negative regulatory feature of the IL-15 leader. Not only did optimized further enhance the expression by 4-15 fold, but even more suprisingly, the less intensive surrogate codon approach as described herein did so as well.

[0205] Codon engineering in addition to secretary signal substitution resulted in as much as 40-100 fold increase in IL-15 gene expression in various cell lines tested. The functionality of IL-15 produced from constructs was demonstrated by CTLL2 cell proliferation assay.

[0206] Consistent with 'in vitro' data, 'in vivo' gene expression from the IL-15 constructs according to embodiments of the invention was considerably elevated. Taken together, this data suggest that this combined method represents a novel and unexpected approach for enhancing IL-15 gene expression.

[0207] The IgE leader sequence for use in certain embodiments of the invention is provided below.

IgE Leader Sequence (SEQ ID NO:11)

ATGGATTGGACTTGGATCTTATTTTTAGTTGCTGCTGCTACTAGAGTTCATT
CT

[0208] The following are the nucleic acid sequences of constructs in accordance with embodiments of the present invention. Leader sequences are indicated by underlining.

Surrogate codon usage HuIL-15 sequence (SEQ ID NO:12)

ATGCGGATTTCCAAACCTCATCTCAGGTCCATTTCCATCCAGTGCTACCTCT
GTCTCCTCCTCAACTCCCATTTTCTCACGGAAGCTGGCATTCATGTCTTCATT
GTCGGCTGTTTCTCCGCGGGGCTCCCTAAAACGGAAGCCAACTGGGTGAATG
TCATTTCCGATCTCAAAAAAATTGAAGATCTCATTCAATCCATGCATATTGATGC
GACGCTCTATACGGAATCCGATGTCCACCCCTCCTGCAAAGTCACCGCGATGAAG
TGCTTTCTCCTCGAGCTCCAAGTCATTTCCCTCGAGTCCGGGGATGCGTCCATTCA

TGATACGGTCGAAAATCTGATCATCCTCGCGAACAACTCCCTCTCCTCCAATGGG
AATGTCACGGAATCCGGGTGCAAAGAATGTGAGGAACTGGAGGAAAAAAATATT
AAAGAATTTCTCCAGTCCTTTGTCCATATTGTCCAAATGTTCATCAACACGTCCTA
G

IgE leader Human IL-15 sequence (SEQ ID NO:13)

ATGGATTGGACTTGGATCTTATTTTTAGTTGCTGCTGCTACTAGAGTTCATT
CTAACTGGGTGAATGTAATAAGTGATTTGAAAAAAATTGAAGATCTTATTCAATC
TATGCATATTGATGCTACTTTATATACGGAAAGTGATGTTCACCCCAGTTGCAAA
GTAACAGCAATGAAGTGCTTTCTCTTGGAGTTACAAGTTATTTCACTTGAGTCCG
GAGATGCAAGTATTCATGATACAGTAGAAAATCTGATCATCCTAGCAAACAACA
GTTTGTCTTCTAATGGGAATGTAACAGAATCTGGATGCAAAGAATGTGAGGAACT
GGAGGAAAAAAATATTAAAGAATTTTTGCAGAGTTTTGTACATATTGTCCAAATG
TTCATCAACACTTCTTGA

IgE leader + surrogate codon usage HuIL-15 sequence (SEQ ID NO:14)

ATGGATTGGACGTGGATCCTCTTTCTCGTCGCGGCGGCGACGCGGGTCCAT
TCCAACTGGGTGAATGTCATTTCCGATCTCAAAAAAATTGAAGATCTCATTCAAT
CCATGCATATTGATGCGACGCTCTATACGGAATCCGATGTCCACCCCTCCTGCAA
AGTCACCGCGATGAAGTGCTTTCTCCTCGAGCTCCAAGTCATTTCCCTCGAGTCC
GGGGATGCGTCCATTCATGATACGGTCGAAAATCTGATCATCCTCGCGAACAACT
CCCTCTCCTCCAATGGGAATGTCACGGAATCCGGGTGCAAAGAATGTGAGGAAC
TGGAGGAAAAAAATATTAAAGAATTTCTCCAGTCCTTTGTCCATATTGTCCAAAT
GTTCATCAACACGTCCTAG

IgE leader + optimized HuIL-15 sequence (optimized alternative 1) (SEQ ID NO:15)

ATGGACTGGACCTGGATCCTGTTCCTGGTGGCCGCCGCCACCCGCGTGCAC
TCCAACTGGGTGAACGTGATCAGCGACCTGAAGAAGATCGAGGACCTGATCCAG
AGCATGCACATCGACGCCACCCTGTACACCGAGAGCGACGTGCACCCCAGCTGC
AAGGTGACCGCCATGAAGTGCTTCCTGCTGGAGCTGCAGGTGATCAGCCTGGAG
AGCGGCGACGCCAGCATCCACGACACCGTGGAGAACCTGATCATCCTGGCCAAC
AACAGCCTGAGCAGCAACGGCAACGTGACCGAGAGCGGCTGCAAGGAGTGCGA
GGAGCTGGAGGAGAAGAACATCAAGGAGTTCCTGCAGAGCTTCGTGCACATCGT
GCAGATGTTCATCAACACCAGCTAG

IgE leader + Secondary structure optimized HuIL-15 sequence (Optimized Alternative 2) (SEQ ID NO: 16)

ATGGATTGGACCTGGATCCTCTTTCTTGTCGCCGCTGCCACTCGAGTACATT
CAAACTGGGTAAATGTGATTTCCGACCTTAAAAAAATTGAAGACCTTATCCAAA
GCATGCACATAGACGCCACCCTTTATACTGAATCCGACGTACACCCCTCCTGCAA
AGTTACCGCCATGAAATGTTTTCTCCTCGAACTCCAAGTAATTAGCCTCGAATCC

GGAGACGCCTCTATCCACGACACAGTTGAAAACCTCATAATCCTTGCAAATAACT
CTCTTAGCTCAAACGGAAATGTTACTGAATCTGGTTGTAAAGAATGCGAAGAACT
TGAAGAAAAAAATATAAAAGAATTTCTGCAATCATTTGTCCACATCGTTCAAATG
TTTATCAATACCTCTTAG

[0209] The following is the sequence of naturally-occurring human IL-15 sequence provided herein for comparative purposes.

Human IL-15 sequence (SEQ ID NO:17)

ATGAGAATTTCGAAACCACATTTGAGAAGTATTTCCATCCAGTGCTACTTGT
GTTTACTTCTAAACAGTCATTTTCTAACTGAAGCTGGCATTCATGTCTTCATT
TTGGGCTGTTTCAGTGCAGGGCTTCCTAAAACAGAAGCCAACTGGGTGAATG
TAATAAGTGATTTGAAAAAAATTGAAGATCTTATTCAATCTATGCATATTGATGC
TACTTTATATACGGAAAGTGATGTTCACCCCAGTTGCAAAGTAACAGCAATGAAG
TGCTTTCTCTTGGAGTTACAAGTTATTTCACTTGAGTCTGGAGATGCAAGTATTCA
TGATACAGTAGAAAATCTGATCATCCTAGCAAACAACAGTTTGTCTTCTAATGGG
AATGTAACAGAATCTGGATGCAAAGAATGTGAGGAACTGGAGGAAAAAAATATT
AAAGAATTTTTGCAGAGTTTTGTACATATTGTCCAAATGTTCATCAACACTTCTTG
A

The following is the nucleic acid sequence for the O-IL-15-IgE leader plasmid construct (SEQ ID NO:18):

AAATGGGGGCGCTGAGGTCTGCCTCGTGAAGAAGGTGTTGCTGACTCATACCAGGCCT
GAATCGCCCCATCATCCAGCCAGAAAGTGAGGGAGCCACGGTTGATGAGAGCTTTGTT
GTAGGTGGACCAGTTGGTGATTTTGAACTTTTGCTTTGCCACGGAACGGTCTGCGTTGT
CGGGAAGATGCGTGATCTGATCCTTCAACTCAGCAAAAGTTCGATTTATTCAACAAAGC
CGCCGTCCCGTCAAGTCAGCGTAATGCTCTGCCAGTGTTACAACCAATTAACCAATTCT
GCGTTCAAAATGGTATGCGTTTTGACACATCCACTATATATCCGTGTCGTTCTGTCCACT
CCTGAATCCCATTCCAGAAATTCTCTAGCGATTCCAGAAGTTTCTCAGAGTCGGAAAGT
TGACCAGACATTACGAACTGGCACAGATGGTCATAACCTGAAGGAAGATCTGATTGCTT
AACTGCTTCAGTTAAGACCGACGCGCTCGTCGTATAACAGATGCGATGATGCAGACCA
ATCAACATGGCACCTGCCATTGCTACCTGTACAGTCAAGGATGGTAGAAATGTTGTCGG
TCCTTGCACACGAATATTACGCCATTTGCCTGCATATTCAAACAGCTCTTCTACGATAAG
GGCACAAATCGCATCGTGGAACGTTTGGGCTTCTACCGATTTAGCAGTTTGATACACTT
TCTCTAAGTATCCACCTGAATCATAAATCGGCAAAATAGAGAAAAATTGACCATGTGTAA
GCGGCCAATCTGATTCCACCTGAGATGCATAATCTAGTAGAATCTCTTCGCTATCAAAAT
TCACTTCCACCTTCCACTCACCGGTTGTCCATTCATGGCTGAACTCTGCTTCCTCTGTT
GACATGACACACATCATCTCAATATCCGAATACGGACCATCAGTCTGACGACCAAGAGA
GCCATAAACACCAATAGCCTTAACATCATCCCCATATTTATCCAATATTCGTTCCTTAATT
TCATGAACAATCTTCATTCTTTCTTCTCTAGTCATTATTATTGGTCCGTTCATAACACCCC
TTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCATGATGATATATTTTTATCTTGT

GCAATGTAACATCAGAGATTTTGAGACACAACGTGGCTTTCCCCGGCCCATGACCAAAA
TCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGA
TCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCG
CTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAC
TGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCC
ACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCA
GTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGT
TACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCT
TGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGC
CACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAAC
AGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTC
GGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGA
GCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCC
TTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGC
CTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGT
GAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGT
ATTTCACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAG
CCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATT
TAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAG
GCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATAGCCGCGGCATCGATGATATC
CATTGCATACGTTGTATCTATATCATAATATGTACATTTATATTGGCTCATGTCCAATATG
ACCGCCATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATT
AGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTG
GCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGT
AACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCC
ACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGAC
GGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTG
GCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACA
TCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGA
CGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACA
ACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAG
CAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGAC
CTCCATAGAAGACACCGGGACCGATCCAGCCTCCGCGGGCGCGCGTCGACCACCATG
GACTGGACCTGGATCCTGTTCCTGGTGGCCGCCGCCACCCGCGTGCACTCCAACTGG
GTGAACGTGATCAGCGACCTGAAGAAGATCGAGGACCTGATCCAGAGCATGCACATCG
ACGCCACCCTGTACACCGAGAGCGACGTGCACCCCAGCTGCAAGGTGACCGCCATGA
AGTGCTTCCTGCTGGAGCTGCAGGTGATCAGCCTGGAGAGCGGCGACGCCAGCATCC
ACGACACCGTGGAGAACCTGATCATCCTGGCCAACAACAGCCTGAGCAGCAACGGCAA
CGTGACCGAGAGCGGCTGCAAGGAGTGCGAGGAGCTGGAGGAGAAGAACATCAAGGA
GTTCCTGCAGAGCTTCGTGCACATCGTGCAGATGTTCATCAACACCAGCTAGTGAGTC
GACGGGCGACGCGAAACTTGGGCCCACTCGAGAGGCGCGCCGAGCTCGCTGATCAG
CCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCC
TTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATC
GCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAA
GGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGAATTT

The following is the nucleic acid sequence for the :LP-IL-15-IgE leader plasmid construct (SEQ ID NO:19)

AAATGGGGGCGCTGAGGTCTGCCTCGTGAAGAAGGTGTTGCTGACTCATACCAGGCCT
GAATCGCCCCATCATCCAGCCAGAAAGTGAGGGAGCCACGGTTGATGAGAGCTTTGTT
GTAGGTGGACCAGTTGGTGATTTTGAACTTTTGCTTTGCCACGGAACGGTCTGCGTTGT
CGGGAAGATGCGTGATCTGATCCTTCAACTCAGCAAAAGTTCGATTTATTCAACAAAGC
CGCCGTCCCGTCAAGTCAGCGTAATGCTCTGCCAGTGTTACAACCAATTAACCAATTCT
GCGTTCAAAATGGTATGCGTTTTGACACATCCACTATATATCCGTGTCGTTCTGTCCACT
CCTGAATCCCATTCCAGAAATTCTCTAGCGATTCCAGAAGTTTCTCAGAGTCGGAAAGT
TGACCAGACATTACGAACTGGCACAGATGGTCATAACCTGAAGGAAGATCTGATTGCTT
AACTGCTTCAGTTAAGACCGACGCGCTCGTCGTATAACAGATGCGATGATGCAGACCA
ATCAACATGGCACCTGCCATTGCTACCTGTACAGTCAAGGATGGTAGAAATGTTGTCGG
TCCTTGCACACGAATATTACGCCATTTGCCTGCATATTCAAACAGCTCTTCTACGATAAG
GGCACAAATCGCATCGTGGAACGTTTGGGCTTCTACCGATTTAGCAGTTTGATACACTT
TCTCTAAGTATCCACCTGAATCATAAATCGGCAAAATAGAGAAAAATTGACCATGTGTAA
GCGGCCAATCTGATTCCACCTGAGATGCATAATCTAGTAGAATCTCTTCGCTATCAAAAT
TCACTTCCACCTTCCACTCACCGGTTGTCCATTCATGGCTGAACTCTGCTTCCTCTGTT
GACATGACACACATCATCTCAATATCCGAATACGGACCATCAGTCTGACGACCAAGAGA
GCCATAAACACCAATAGCCTTAACATCATCCCCATATTTATCCAATATTCGTTCCTTAATT
TCATGAACAATCTTCATTCTTTCTTCTCTAGTCATTATTATTGGTCCGTTCATAACACCCC
TTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCATGATGATATATTTTTATCTTGT
GCAATGTAACATCAGAGATTTTGAGACACAACGTGGCTTTCCCCGGCCCATGACCAAAA
TCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGA
TCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCG
CTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAC
TGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCC
ACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCA
GTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGT
TACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCT
TGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGC
CACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAAC
AGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTC
GGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGA
GCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCC
TTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGC
CTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGT
GAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGT
ATTTCACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAG
CCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATT
TAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAG
GCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATAGCCGCGGCATCGATGATATC
CATTGCATACGTTGTATCTATATCATAATATGTACATTTATATTGGCTCATGTCCAATATG
ACCGCCATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATT
AGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTG
GCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGT
AACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCC
ACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGAC
GGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTG
GCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACA
TCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGA

CGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACA
ACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAG
CAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGAC
CTCCATAGAAGACACCGGGACCGATCCAGCCTCCGCGGGCGCGCGTCGACCACCATG
GATTGGACGTGGATCCTCTTTCTCGTCGCGGCGGCGACGCGGGTCCATTCCAACTGG
GTGAATGTCATTTCCGATCTCAAAAAAATTGAAGATCTCATTCAATCCATGCATATTGAT
GCGACGCTCTATACGGAATCCGATGTCCACCCCTCCTGCAAAGTCACCGCGATGAAGT
GCTTTCTCCTCGAGCTCCAAGTCATTTCCCTCGAGTCCGGGGATGCGTCCATTCATGAT
ACGGTCGAAAATCTGATCATCCTCGCGAACAACTCCCTCTCCTCCAATGGGAATGTCAC
GGAATCCGGGTGCAAAGAATGTGAGGAACTGGAGGAAAAAAATATTAAAGAATTTCTCC
AGTCCTTTGTCCATATTGTCCAAATGTTCATCAACACGTCCTAGTGAGTCGACGGGCGA
CGCGAAACTTGGGCCCACTCGAGAGGCGCGCCGAGCTCGCTGATCAGCCTCGACTGT
GCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGG
AAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTG
AGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGAT
TGGGAAGACAATAGCAGGCATGCTGGGGAATTT

The following is the nucleic acid sequence for the BH-IL-15-IgE leader plasmid construct (SEQ ID NO:20)

AAATGGGGGCGCTGAGGTCTGCCTCGTGAAGAAGGTGTTGCTGACTCATACCAGGCCT
GAATCGCCCCATCATCCAGCCAGAAAGTGAGGGAGCCACGGTTGATGAGAGCTTTGTT
GTAGGTGGACCAGTTGGTGATTTTGAACTTTTGCTTTGCCACGGAACGGTCTGCGTTGT
CGGGAAGATGCGTGATCTGATCCTTCAACTCAGCAAAAGTTCGATTTATTCAACAAAGC
CGCCGTCCCGTCAAGTCAGCGTAATGCTCTGCCAGTGTTACAACCAATTAACCAATTCT
GCGTTCAAAATGGTATGCGTTTTGACACATCCACTATATATCCGTGTCGTTCTGTCCACT
CCTGAATCCCATTCCAGAAATTCTCTAGCGATTCCAGAAGTTTCTCAGAGTCGGAAAGT
TGACCAGACATTACGAACTGGCACAGATGGTCATAACCTGAAGGAAGATCTGATTGCTT
AACTGCTTCAGTTAAGACCGACGCGCTCGTCGTATAACAGATGCGATGATGCAGACCA
ATCAACATGGCACCTGCCATTGCTACCTGTACAGTCAAGGATGGTAGAAATGTTGTCGG
TCCTTGCACACGAATATTACGCCATTTGCCTGCATATTCAAACAGCTCTTCTACGATAAG
GGCACAAATCGCATCGTGGAACGTTTGGGCTTCTACCGATTTAGCAGTTTGATACACTT
TCTCTAAGTATCCACCTGAATCATAAATCGGCAAAATAGAGAAAAATTGACCATGTGTAA
GCGGCCAATCTGATTCCACCTGAGATGCATAATCTAGTAGAATCTCTTCGCTATCAAAAT
TCACTTCCACCTTCCACTCACCGGTTGTCCATTCATGGCTGAACTCTGCTTCCTCTGTT
GACATGACACACATCATCTCAATATCCGAATACGGACCATCAGTCTGACGACCAAGAGA
GCCATAAACACCAATAGCCTTAACATCATCCCCATATTTATCCAATATTCGTTCCTTAATT
TCATGAACAATCTTCATTCTTTCTTCTCTAGTCATTATTATTGGTCCGTTCATAACACCCC
TTGTATTACTGTTTATGTAAGCAGACAGTTTTATTGTTCATGATGATATATTTTTATCTTGT
GCAATGTAACATCAGAGATTTTGAGACACAACGTGGCTTTCCCCGGCCCATGACCAAAA

TCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGA
TCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCG
CTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAAC
TGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCC
ACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCA
GTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGT
TACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCT
TGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGC
CACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAAC
AGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTC
GGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGA
GCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCC
TTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGC
CTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGT
GAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCGGT
ATTTCACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAG
CCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATT
TAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAG
GCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATAGCCGCGGCATCGATGATATC
CATTGCATACGTTGTATCTATATCATAATATGTACATTTATATTGGCTCATGTCCAATATG
ACCGCCATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATT
AGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTG
GCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGT
AACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCC
ACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGAC
GGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTG
GCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACA
TCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGA
CGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACA
ACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAG
CAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGAC
CTCCATAGAAGACACCGGGACCGATCCAGCCTCCGCGGGCGCGCGTCGACCACCATG
GATTGGACCTGGATCCTCTTTCTTGTCGCCGCTGCCACTCGAGTACATTCAAACTGGGT
AAATGTGATTTCCGACCTTAAAAAAATTGAAGACCTTATCCAAAGCATGCACATAGACGC
CACCCTTTATACTGAATCCGACGTACACCCCTCCTGCAAAGTTACCGCCATGAAATGTT

TTCTCCTCGAACTCCAAGTAATTAGCCTCGAATCCGGAGACGCCTCTATCCACGACACA
GTTGAAAACCTCATAATCCTTGCAAATAACTCTCTTAGCTCAAACGGAAATGTTACTGAA
TCTGGTTGTAAAGAATGCGAAGAACTTGAAGAAAAAAATATAAAAGAATTTCTGCAATCA
TTTGTCCACATCGTTCAAATGTTTATCAATACCTCTTAGTGAGTCGACGGGCGACGCGA
AACTTGGGCCCACTCGAGAGGCGCGCCGAGCTCGCTGATCAGCCTCGACTGTGCCTT
CTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGT
GCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAG
GTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGA
AGACAATAGCAGGCATGCTGGGGAATTT

[0210] The invention now being fully described, it will be apparent to one of ordinary skill in the art that many changes and modifications can be made thereto without departing from the spirit or scope of the invention as set forth herein. The foregoing describes the preferred embodiments of the present invention along with a number of possible alternatives. These embodiments, however, are merely for example and the invention is not restricted thereto.

WHAT IS CLAIMED IS:

1. A method for preparing a polynucleotide that provides enhanced expression of a gene comprising:

assembling oligonucleotides comprising surrogate codons to form a modified polynucleotide comprising a predetermined nucleic acid sequence encoding the same protein or polypeptide as a wild-type polynucleotide, said nucleic acid sequence comprising surrogate codons in place of naturally-occurring codons that encode the amino acids alanine, arginine, leucine, proline, glutamic acid, glycine, isoleucine, serine, threonine and valine and that have adenine (A) or uracil (U) or thymine (T) in the wobble position.

2. The method of claim 1, additionally comprising adding a non-native leader sequence to the modified polynucleotide.

3. The method of claim 1, additionally comprising adding a human non-native leader sequence to the modified polynucleotide.

4. The method of claim 1, additionally comprising adding an immunoglobulin leader sequence to the modified polynucleotide.

5. The method of claim 1, additionally comprising adding to the modified polynucleotide: (a) an IgE leader sequence or (b) a leader sequence that hybridizes to an IgE leader sequence under stringent conditions.

6. The method of claim 1, additionally comprising adding to the modified polynucleotide: (a) a leader sequence having SEQ ID NO:11; or (b) a leader sequence that hybridizes to SEQ ID NO:11 under stringent conditions.

7. The method of claim 1, additionally comprising adding to the modified polynucleotide a leader sequence having at least 90% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

8. The method of claim 1, additionally comprising adding to the modified polynucleotide a leader sequence having the nucleic acid sequence of SEQ ID NO:11.

9. The method of claim 1, wherein the modified polynucleotide expresses a viral, bacterial, protist, fungal, plant, or animal polypeptide

10. The method of claim 1, wherein the modified polynucleotide expresses a mammalian polypeptide.

11. The method of claim 1, wherein the modified polynucleotide expresses a bacterial polypeptide.

12. The method of claim 1, wherein the modified polynucleotide expresses a viral polypeptide.

13. The method of claim 1, wherein the viral polypeptide is a human papillomavirus 16 (HPV16) polypeptide, human immunodeficiency virus-1 (HIV-1) polypeptide or gp160 polypeptide.

14. The method of claim 1, wherein the modified polynucleotide comprises the open reading frame (ORF) for the HPV16 E7 gene (E7).

15. The method of claim 1, wherein the modified polynucleotide comprises the ORF for the HIV-1 gag gene (gag).

16. The method of claim 1, wherein the modified polynucleotide comprises the ORF for the gp160 envelope gene (env).

17. The method of any of claims 1-8, wherein the modified polynucleotide encodes for a cytokine.

18. The method of any of claims 1-8, wherein the modified polynucleotide encodes for an interleukin.

19. The method of any of claims 1-8, wherein the modified polynucleotide encodes for IL-15 or a peptide or polypeptide having the activity of IL-15.

20. The method of any of claims 1-8, wherein the modified polynucleotide encodes for IL-15.

21. The method of claim 1-8, wherein the modified polynucleotide a leader sequence has at least 90% sequence identity to the nucleic acid sequence of any of SEQ ID NOS:12-16.

22. The method of claim 1-8, wherein the modified polynucleotide has at least 95% sequence identity to the nucleic acid sequence of any of SEQ ID NOS:12-16.

23. The method of any of claims 1-8, wherein the modified polynucleotide comprises the nucleic acid sequence of any of SEQ ID NOS: 12-16.

24. The method of any of claims 1-8, wherein the modified polynucleotide comprises the nucleic acid sequence of SEQ ID NO:14.

25. The method of claim 1, wherein the surrogate codons are a randomized selection of at least about 10% of the codons in said modified polynucleotide that encode for any of the amino acids alanine, arginine, leucine, proline, glutamic acid, glycine, isoleucine, serine, threonine and valine.

26. The method of claim 1, wherein the surrogate codons are a randomized selection of at least about 50% of the codons in said modified polynucleotide that encode for any of the amino acids alanine, arginine, leucine, proline, glutamic acid, glycine, isoleucine, serine, threonine and valine.

27. The method of claim 1, wherein the surrogate codons are a randomized selection of at least about 90% of the codons in said modified polynucleotide that encode for any of the amino acids alanine, arginine, leucine, proline, glutamic acid, glycine, isoleucine, serine, threonine and valine.

28. The method of claim 1, wherein the surrogate codons are each of the codons in said modified polynucleotide that encode for the amino acids alanine, arginine, leucine, proline, glutamic acid, glycine, isoleucine, serine, threonine and valine.

29. The method of claim 1, wherein the modified polynucleotide is a DNA molecule.

30. The method of claim 1, wherein the modified polynucleotide is an RNA molecule.

31. The method of claim 1, wherein the surrogate codons are selected independently of results from sequence specific experimental studies of codon optimization or inhibitory sequence location.

32. A method for preparing a modified polynucleotide that provides enhanced expression of a gene comprising:

determining for said gene a modified nucleic acid sequence comprising surrogate codons in place of the corresponding naturally-occurring codons having the nucleotides adenine (A) or uracil (U) or thymine (T) in the wobble position;

selecting oligonucleotides having nucleotide sequences corresponding to portions of said determined modified nucleic acid sequence; and

assembling the oligonucleotides to form a modified polynucleotide comprising the determined modified nucleic acid sequence.

33. The method of claim 32, additionally comprising adding a non-native leader sequence to the modified polynucleotide.

34. The method of claim 32, additionally comprising adding a human non-native leader sequence to the modified polynucleotide.

35. The method of claim 32, additionally comprising adding an immunoglobulin leader sequence to the modified polynucleotide.

36. The method of claim 32, additionally comprising adding to the modified polynucleotide: (a) an IgE leader sequence or (b) a leader sequence that hybridizes to an IgE leader sequence under stringent conditions.

37. The method of claim 32, additionally comprising adding to the modified polynucleotide: (a) a leader sequence having SEQ ID NO:11; or (b) a leader sequence that hybridizes to SEQ ID NO:11 under stringent conditions.

38. The method of claim 32, additionally comprising adding to the modified polynucleotide a leader sequence having at least 90% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

39. The method of claim 32, additionally comprising adding to the modified polynucleotide a leader sequence having the nucleic acid sequence of SEQ ID NO:11.

40. The method of any of claims 32-39, wherein the modified polynucleotide encodes for a cytokine.

41. The method of any of claims 32-39, wherein the modified polynucleotide encodes for an interleukin.

42. The method of claims 32-39, wherein the modified polynucleotide encodes for IL-15 or a peptide or polypeptide having the activity of IL-15.

43. The method of any of claims 32-39, wherein the modified polynucleotide encodes for IL-15.

44. The method of claim 32-39, wherein the modified polynucleotide has at least 90% sequence identity to the nucleic acid sequence of any of SEQ ID NOS:12-16.

45. The method of claim 32-39, additionally comprising adding to the modified polynucleotide has at least 95% sequence identity to the nucleic acid sequence of any of SEQ ID NOS:12-16.

46. The method of any of claims 32-39, wherein the modified polynucleotide comprises the nucleic acid sequence of any of SEQ ID NOS:12-16.

47. The method of any of claims 32-39, wherein the modified polynucleotide comprises the nucleic acid sequence of SEQ ID NO:14.

48. The method of claim 32, wherein the modified nucleic acid sequence is determined by a computer.

49. The method of claim 32, wherein the modified nucleic acid sequence is determined manually.

50. A method for enhancing expression of a gene comprising:

assembling a modified polynucleotide having a nucleic acid sequence in which each naturally-occurring codon having adenine (A) or uracil (U) or thymine (T) in the wobble position is replaced in the sequence by a surrogate codon; and

expressing the gene.

51. The method of claim 50, additionally comprising adding a non-native leader sequence to the modified polynucleotide.

52. The method of claim 50, additionally comprising adding a human non-native leader sequence to the modified polynucleotide.

53. The method of claim 50, additionally comprising adding an immunoglobulin leader sequence to the modified polynucleotide.

54. The method of claim 50, additionally comprising adding to the modified polynucleotide: (a) an IgE leader sequence or (b) a leader sequence that hybridizes to an IgE leader sequence under stringent conditions.

55. The method of claim 50, additionally comprising adding to the modified polynucleotide: (a) a leader sequence having SEQ ID NO:11; or (b) a leader sequence that hybridizes to SEQ ID NO:11 under stringent conditions.

56. The method of claim 50, additionally comprising adding to the modified polynucleotide a leader sequence having at least 90% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

57. The method of claim 50, additionally comprising adding to the modified polynucleotide a leader sequence having at least 95% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

58. The method of claim 50, additionally comprising adding to the modified polynucleotide a leader sequence having at least 97% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

59. The method of claim 50, additionally comprising adding to the modified polynucleotide a leader sequence having the nucleic acid sequence of SEQ ID NO:11.

60. The method of claim 50, additionally comprising adding an IgE leader sequence to the modified polynucleotide.

61. The method of claim 50, comprising replacing each naturally-occuring codon for alanine with GCG, each disfavored codon encoding arginine with CGG or AGG, each disfavored codon encoding leucine with CTC, each disfavored codon encoding proline with CCT or CCG, each disfavored codon encoding glutamic acid with GAG, each disfavored codon encoding glycine with GGG, each disfavored codon encoding isoleucine with ATT, each disfavored codon encoding serine with TCC, each disfavored codon encoding threonine with ACG, and each disfavored codon encoding valine with GTC.

62. The method of claim 50, wherein expression is enhanced *in vitro.*

63. The method of claim 50, wherein expression is enhanced *in vivo.*

64. The method of claim 50, wherein the gene expresses a viral, bacterial, protist, fungal, plant, or animal polypeptide.

65. The method of claim 50, wherein the gene expresses a mammalian polypeptide.

66. The method of claim 50, wherein the gene expresses a bacterial polypeptide.

67. The method of claim 50, wherein the gene expresses a viral polypeptide.

68. The method of claim 50, wherein the viral polypeptide is an HPV16, HIV-1 or gp160 protein.

69. The method of claim 50, wherein the gene is HPV16 E7.

70. The method of claim 50, wherein the gene is HIV-1 gag.

71. The method of claim 50, wherein the gene is gp160 env.

72. A method for enhancing expression of a gene comprising: expressing *in vivo* or *in vitro* a modified polynucleotide according to any of claims 80-155 or any combination thereof.

73. A method of preventing or treating a disease in a mammal comprising:

administering to the mammal an effective amount of one or more composition of any of claims 156-374.

74. The method of claim 73, wherein the composition is administered parenterally.

75. The method of claim 73, wherein the composition is administered mucosally.

76. The method of claim 73, wherein the composition is administered by subcutaneous or intramuscular injection.

77. A method for preparing a modified polynucleotide that provides enhanced expression of a gene comprising:

attaching an IgE leader sequence to a nucleic acid sequence encoding the same protein or polypeptide as a naturally-occurring polynucleotide, said nucleic acid sequence comprising a surrogate codon in place of a naturally-occurring codon that encodes for any of the amino acids alanine, arginine, leucine, proline, glutamic acid, glycine, isoleucine, serine, threonine and valine and that have adenine (A) or uracil (U) or thymine (T) in the wobble position.

78. The method of claim 77, wherein the nucleic acid sequence comprises any of SEQ ID NOS:12-16.

79. The method of claim 77, wherein the modified polynucleotide is SEQ ID NO:14.

80. A modified polynucleotide sequence comprising:

a nucleic acid sequence encoding the same protein as a wild-type polynucleotide, said nucleic acid sequence comprising surrogate codons in place of naturally-occurring codons that encode the amino acids alanine, arginine, leucine, proline, glutamic acid, glycine, isoleucine, serine, threonine and valine and that have adenine (A) or uracil (U) or thymine (T) in the wobble position.

81. The modified polynucleotide of claim 80, additionally comprising a non-native leader sequence.

82. The modified polynucleotide of claim 80, additionally comprising a human non-native leader sequence.

83. The modified polynucleotide of claim 80, additionally comprising an immunoglobulin leader sequence.

84. The modified polynucleotide of claim 80, additionally comprising (a) an IgE leader sequence or (b) a leader sequence that hybridizes to an IgE leader sequence under stringent conditions.

85. The modified polynucleotide of claim 80, additionally comprising (a) a leader sequence having SEQ ID NO:11 or (b) a leader sequence that hybridizes to SEQ ID NO:11 under stringent conditions.

86. The modified polynucleotide of claim 80, additionally comprising adding to the modified polynucleotide a leader sequence having at least 90% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

87. The modified polynucleotide of claim 80, additionally comprising a leader sequence having at least 95% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

88. The modified polynucleotide of claim 80, additionally comprising a leader sequence having at least 97% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

89. The modified polynucleotide of claim 80, additionally comprising a leader sequence having the nucleic acid sequence of SEQ ID NO:11.

90. The modified polynucleotide of claim 80, wherein the modified polynucleotide expresses a viral, bacterial, protist, fungal, plant, or animal polypeptide.

91. The modified polynucleotide of claim 80, wherein the modified polynucleotide expresses a mammalian polypeptide.

92. The modified polynucleotide of claim 80, wherein the modified polynucleotide expresses a bacterial polypeptide.

93. The modified polynucleotide of claim 80, wherein the modified polynucleotide expresses a viral polypeptide.

94. The modified polynucleotide of claim 80, wherein the viral polypeptide is an HPV16 polypeptide or an HIV-1 polypeptide.

95. The modified polynucleotide of claim 80, wherein the modified polynucleotide comprises the ORF for the HPV16 E7 gene.

96. The modified polynucleotide of claim 80, wherein the modified polynucleotide comprises the ORF for the HIV-1 gag gene.

97. The modified polynucleotide of claim 80, wherein the modified polynucleotide comprises the ORF for the gp160 envelope gene.

98. The modified polynucleotide of claim 80, wherein the surrogate codons are a randomized selection of at least about 10% of the codons in said modified polynucleotide that encode for any of the amino acids alanine, arginine, leucine, proline, glutamic acid, glycine, isoleucine, serine, threonine and valine.

99. The modified polynucleotide of claim 80, wherein the surrogate codons are a randomized selection of at least about 50% of the codons in said modified polynucleotide that encode for any of the amino acids alanine, arginine, leucine, proline, glycine, isoleucine, serine, threonine and valine.

100. The modified polynucleotide of claim 80, wherein the surrogate codons are a randomized selection of at least about 90% of the codons in said recomb1inant polynucleotide that encode for any of the amino acids alanine, arginine, leucine, proline, glycine, isoleucine, serine, threonine and valine.

101. The modified polynucleotide of claim 80, wherein the surrogate codons are each of the codons in said modified polynucleotide that encode for the amino acids alanine, arginine, leucine, proline, glycine, isoleucine, serine, threonine and valine.

102. The modified polynucleotide of claim 80, wherein the modified polynucleotide is a DNA molecule.

103. The modified polynucleotide of claim 80, wherein the modified polynucleotide is an RNA molecule.

104. A modified polynucleotide comprising:

a nucleic acid sequence in which each codon, corresponding to a naturally-occurring codon having A, T or U in the wobble position, encoding alanine is GCG, encoding arginine is CGG or AGG, encoding leucine is CTC, encoding proline is CCT or CCG, encoding glutamic acid is GAG, encoding glycine is GGG, encoding isoleucine is ATT, encoding serine is TCC, encoding threonine is ACG, and encoding valine is GTC.

105. The modified polynucleotide of claim 104, additionally comprising a non-native leader sequence.

106. The modified polynucleotide of claim 104, additionally comprising a human non-native leader sequence.

107. The modified polynucleotide of claim 104, additionally comprising an immunoglobulin leader sequence.

108. The modified polynucleotide of claim 104, additionally comprising (a) an IgE leader sequence or (b) a leader sequence that hybridizes to an IgE leader sequence under stringent conditions.

109. The modified polynucleotide of claim 104, additionally comprising (a) a leader sequence having SEQ ID NO:11 or (b) a leader sequence that hybridizes to SEQ ID NO:11 under stringent conditions.

110. The modified polynucleotide of claim 104, additionally comprising adding to the modified polynucleotide a leader sequence having at least 90% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

111. The modified polynucleotide of claim 104, additionally comprising a leader sequence having at least 95% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

112. The modified polynucleotide of claim 104, additionally comprising a leader sequence having at least 97% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

113. The modified polynucleotide of claim 104, additionally comprising a leader sequence having the nucleic acid sequence of SEQ ID NO:11.

114. The modified polynucleotide of claim 104, wherein the expression levels of the modified polynucleotide are at least two times more than the expression levels of a wild-type polynucleotide encoding the same polypeptide, when expressed in a mammalian host cell.

115. The modified polynucleotide of claim 104, wherein the expression levels of the modified polynucleotide are at least ten times more than the expression

levels of a wild-type polynucleotide encoding the same polypeptide, when expressed in a mammalian host cell.

116. The modified polynucleotide of claim 104, wherein the modified polynucleotide expresses a mammalian polypeptide.

117. The modified polynucleotide of claim 104, wherein the modified polynucleotide expresses a bacterial polypeptide.

118. The modified polynucleotide of claim 104, wherein the modified polynucleotide expresses a viral polypeptide.

119. The modified polynucleotide of claim 104, wherein the viral polypeptide is an HPV16 polypeptide or an HIV-1 polypeptide.

120. The modified polynucleotide of claim 104, wherein the modified polynucleotide comprises the ORF for the HPV16 E7 gene.

121. The modified polynucleotide of claim 104, wherein the modified polynucleotide comprises the ORF for the HIV-1 gag gene.

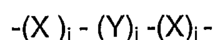
122. The modified polynucleotide of claim 104, wherein the modified polynucleotide comprises the ORF for the env gene.

123. The modified polynucleotide of claim 104, wherein the modified polynucleotide is a DNA molecule.

124. The modified polynucleotide of claim 104, wherein the modified polynucleotide is an RNA molecule.

125. A modified polynucleotide comprising:

a nucleic acid sequence having the general formula:



$$-(X)_i - (Y)_j - (X)_i -$$

wherein X represents non-surrogate codons having the nucleic acid sequence of any of the corresponding wild-type codons in the naturally-occurring polynucleotide that encode the same polypeptide as said modified polynucleotide, said wild-type codons having cytosine (C) or guanine (G) in the wobble position;

wherein Y represents surrogate codons having a nucleic acid sequence that is different from the corresponding wild-type codons in the naturally-occurring polynucleotide that encode the same polypeptide as said modified polynucleotide, said wild-type codons having adenine (A) or uracil (U) or thymine (T) in the wobble position, said surrogate codons encoding the same amino acid as the corresponding wild-type codons in the naturally-occurring polypeptide that encodes the same protein or polypeptide as said modified polynucleotide;

wherein i is any positive integer of at least 0; and

wherein j is any positive integer of at least 1.

126. The modified polynucleotide of claim 125, additionally comprising a non-native leader sequence.

127. The modified polynucleotide of claim 125, additionally comprising a human non-native leader sequence.

128. The modified polynucleotide of claim 125, additionally comprising an immunoglobulin leader sequence.

129. The modified polynucleotide of claim 125, additionally comprising (a) an IgE leader sequence or (b) a leader sequence that hybridizes to an IgE leader sequence under stringent conditions.

130. The modified polynucleotide of claim 125, additionally comprising (a) a leader sequence having SEQ ID NO:11 or (b) a leader sequence that hybridizes to SEQ ID NO:11 under stringent conditions.

131. The modified polynucleotide of claim 125, additionally comprising adding to the modified polynucleotide a leader sequence having at least 90% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

132. The modified polynucleotide of claim 125, additionally comprising a leader sequence having at least 95% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

133. The modified polynucleotide of claim 125, additionally comprising a leader sequence having at least 97% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

134. The modified polynucleotide of claim 125, additionally comprising a leader sequence having the nucleic acid sequence of SEQ ID NO:11.

135. The modified polynucleotide of claim 125, wherein the modified polynucleotide expresses a viral, bacterial, protist, fungal, plant, or animal polypeptide.

136. The modified polynucleotide of claim 125, wherein the modified polynucleotide expresses a mammalian polypeptide.

137. The modified polynucleotide of claim 125, wherein the modified polynucleotide expresses a bacterial polypeptide.

138. The modified polynucleotide of claim 125, wherein the modified polynucleotide expresses a viral polypeptide.

139. The modified polynucleotide of claim 125, wherein the viral protein is an HPV16 or HIV-1 polypeptide.

140. The modified polynucleotide of claim 125, wherein the modified polynucleotide comprises the ORF for the HPV16 E7 gene.

141. The modified polynucleotide of claim 125, wherein the modified polynucleotide comprises the ORF for the HIV-1 gag gene.

142. The modified polynucleotide of claim 125, wherein the modified polynucleotide comprises the ORF for the gp160 envelope gene.

143. The modified polynucleotide of claim 125, wherein the surrogate codons are a randomized selection of at least about 10% of the codons in said modified polynucleotide that encode for any of the amino acids alanine, arginine, leucine, proline, glutamic acid, glycine, isoleucine, serine, threonine and valine.

144. The modified polynucleotide of claim 125, wherein the surrogate codons are a randomized selection of at least about 50% of the codons in said modified

polynucleotide that encode for any of the amino acids alanine, arginine, leucine, proline, glycine, isoleucine, serine, threonine and valine.

145. The modified polynucleotide of claim 125, wherein the surrogate codons are a randomized selection of at least about 90% of the codons in said modified polynucleotide that encode for any of the amino acids alanine, arginine, leucine, proline, glycine, isoleucine, serine, threonine and valine.

146. The modified polynucleotide of claim 125, wherein the surrogate codons are each of the codons in said modified polynucleotide that encode for the amino acids alanine, arginine, leucine, proline, glycine, isoleucine, serine, threonine and valine.

147. The modified polynucleotide of claim 125, wherein the modified polynucleotide is a DNA molecule.

148. The modified polynucleotide of claim 125, wherein the modified polynucleotide is an RNA molecule.

149. A modified polynucleotide comprising:

the nucleic acid sequence of any of SEQ ID NOS: 1, 3 or 5;

the nucleic acid sequence encoding any of SEQ ID NOS:2, 4 or 6;

an immunogenic encoding portion of (a) or (b); or

a nucleic acid sequence that hybridizes under stringent conditions to any of (a), (b) or (c).

150. The modified polynucleotide of claim 149, wherein the modified polynucleotide comprises the nucleic acid sequence of any of SEQ ID NOS: 1, 3 or 5.

151. The modified polynucleotide of claim 149, wherein the modified polynucleotide comprises the nucleic acid sequence of SEQ ID NO:1.

152. The modified polynucleotide of claim 149, wherein the modified polynucleotide comprises the nucleic acid sequence of SEQ ID NO:3.

153. The modified polynucleotide of claim 149, wherein the modified polynucleotide comprises the nucleic acid sequence of SEQ ID NO:5.

154. The modified polynucleotide of claim 149, wherein the stringent conditions are high stringency southern hybridization conditions.

155. A modified polynucleotide comprising:

(a) the nucleic acid sequence of any of SEQ ID NOS: 12-16;

(b) an immunogenic encoding portion of (a); or

(c) a nucleic acid sequence that hybridizes under stringent conditions to any of (a) or (b).

156. A composition comprising:

a modified polynucleotide comprising a nucleic acid sequence with surrogate codons in place of the corresponding naturally-occurring codons having adenine (A), thymine (T) or uracil (U) in the wobble position, of a wild-type polypeptide, said modified polynucleotide expressing the same protein or polypeptide as said wild-type polypeptide; and

a pharmaceutically acceptable vector.

157. The compostion of claim 156, wherein the modified polynucleotide additionally comprises a non-native leader sequence.

158. The composition of claim 156, wherein the modified polynucleotide additionally comprises a human non-native leader sequence.

159. The composition of claim 156, wherein the modified polynucleotide additionally comprises an immunoglobulin leader sequence.

160. The composition of claim 156, wherein the modified polynucleotide additionally comprises (a) an IgE leader sequence or (b) a leader sequence that hybridizes to an IgE leader sequence under stringent conditions.

161. The composition of claim 156, wherein the modified polynucleotide additionally comprises (a) a leader sequence having SEQ ID NO:11 or (b) a leader sequence that hybridizes to SEQ ID NO:11 under stringent conditions.

162. The composition of claim 156, wherein the modified polynucleotide additionally comprises adding to the modified polynucleotide a leader sequence having at least 90% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

163. The composition of claim 156, wherein the modified polynucleotide additionally comprises a leader sequence having at least 95% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

164. The composition of claim 156, wherein the modified polynucleotide additionally comprises a leader sequence having at least 97% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

165. The composition of claim 156, wherein the modified polynucleotide additionally comprises a leader sequence having the nucleic acid sequence of SEQ ID NO:11.

166. The composition of claim 156, wherein the pharmaceutically acceptable vector is a plasmid.

167. The composition of claim 156, wherein the pharmaceutically acceptable vector is a phage.

168. The composition of claim 156, additionally comprising a nucleic acid sequence encoding for an additional polypeptide.

169. The composition of claim 156, wherein the modified polynucleotide expresses a mammalian polypeptide.

170. The composition of claim 156, wherein the modified polynucleotide expresses a bacterial polypeptide.

171. The composition of claim 156, wherein the modified polynucleotide expresses a viral polypeptide.

172. The composition of claim 171, wherein the viral polypeptide is an HPV16, HIV-1 or gp160 polypeptide.

173. The composition of claim 171, wherein the modified polynucleotide comprises the ORF for the HPV16 E7 gene.

174. The composition of claim 171, wherein the modified polynucleotide comprises the ORF for the HIV-1 gag gene.

175. The composition of claim 171, wherein the modified polynucleotide comprises the ORF for the gp160 envelope gene.

176. The composition of claim 156, wherein the surrogate codons encode each of the codons in said modified polynucleotide that encode for the amino acids alanine, arginine, leucine, proline, glycine, isoleucine, serine, threonine and valine.

177. A composition comprising:

a modified polynucleotide comprising a nucleic acid sequence in which each codon, corresponding to a naturally-occurring codon having A, T or U in the wobble position, encoding alanine is GCG, encoding arginine is CGG or AGG, encoding leucine is CTC, encoding proline is CCT or CCG, encoding glutamic acid is GAG, encoding glycine is GGG, encoding isoleucine is ATT, encoding serine is TCC, encoding threonine is ACG, and encoding valine is GTC; and

a pharmaceutically acceptable vector.

178. The compostion of claim 177, wherein the modified polynucleotide additionally comprises a non-native leader sequence.

179. The composition of claim 177, wherein the modified polynucleotide additionally comprises a human non-native leader sequence.

180. The composition of claim 177, wherein the modified polynucleotide additionally comprises an immunoglobulin leader sequence.

181. The composition of claim 177, wherein the modified polynucleotide additionally comprises (a) an IgE leader sequence or (b) a leader sequence that hybridizes to an IgE leader sequence under stringent conditions.

182. The composition of claim 177, wherein the modified polynucleotide additionally comprises (a) a leader sequence having SEQ ID NO:11 or (b) a leader sequence that hybridizes to SEQ ID NO:11 under stringent conditions.

183. The composition of claim 177, wherein the modified polynucleotide additionally comprises adding to the modified polynucleotide a leader sequence having at least 90% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

184. The composition of claim 177, wherein the modified polynucleotide additionally comprises a leader sequence having at least 95% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

185. The composition of claim 177, wherein the modified polynucleotide additionally comprises a leader sequence having at least 97% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

186. The composition of claim 177, wherein the modified polynucleotide additionally comprises a leader sequence having the nucleic acid sequence of SEQ ID NO:11.

187. The composition of claim 177, wherein the pharmaceutically acceptable vector is a plasmid.

188. The composition of claim 177, wherein the pharmaceutically acceptable vector is a phage.

189. The composition of claim 177, additionally comprising a nucleic acid sequence encoding for an additional polypeptide.

190. The composition of claim 177, wherein the modified polynucleotide expresses a mammalian polypeptide.

191. The composition of claim 177, wherein the modified polynucleotide expresses a bacterial polypeptide.

192. The composition of claim 177, wherein the modified polynucleotide expresses a viral polypeptide.

193. The composition of claim 177, wherein the viral polypeptide is HPV16, HIV-1 or gp160 polypeptide.

194. The composition of claim 177, wherein the modified polynucleotide comprises the ORF for the HPV16 E7 gene.

195. The composition of claim 177, wherein the modified polynucleotide comprises the ORF for the HIV-1 gag gene.

196. The composition of claim 177, wherein the modified polynucleotide comprises the ORF for the gp160 envelope gene.

197. The composition of claim 177, wherein the modified polynucleotide is a DNA molecule.

198. The composition of claim 177, wherein the modified polynucleotide is an RNA molecule.

199. A composition comprising:

a pharmaceutically acceptable buffer, diluent, adjuvant, carrier or vector; and

a polynucleotide comprising a nucleic acid sequence having the general formula:

$$-(X)_i - (Y)_j - (X)_i -$$

wherein X represents non-surrogate codons having the nucleic acid sequence of any of the corresponding wild-type codons in the naturally-occurring polypeptide that encode the same protein or polypeptide as said modified polynucleotide, said wild-type codons having cytosine (C) or guanine (G) in the wobble position;

wherein Y represents surrogate codons having a nucleic acid sequence that is different from the corresponding wild-type codons in the naturally-occurring polypeptide that encode the same protein or polypeptide as said modified polynucleotide, said wild-type codons having adenine (A) or uracil (U) or thymine (T) in the wobble position, said surrogate codons encoding the same amino acid as the corresponding wild-type codons in the naturally-occurring polypeptide that encodes the same protein or polypeptide as said modified polynucleotide;

wherein i is any positive integer of at least 0; and

wherein j is any positive integer of at least 1.

200. The composition of claim 199, wherein the modified polynucleotide additionally comprises a non-native leader sequence.

201. The composition of claim 199, wherein the modified polynucleotide additionally comprises a human non-native leader sequence.

202. The composition of claim 199, wherein the modified polynucleotide additionally comprises an immunoglobulin leader sequence.

203. The composition of claim 199, wherein the modified polynucleotide additionally comprises (a) an IgE leader sequence or (b) a leader sequence that hybridizes to an IgE leader sequence under stringent conditions.

204. The composition of claim 199, wherein the modified polynucleotide additionally comprises (a) a leader sequence having SEQ ID NO:11 or (b) a leader sequence that hybridizes to SEQ ID NO:11 under stringent conditions.

205. The composition of claim 199, wherein the modified polynucleotide additionally comprises adding to the modified polynucleotide a leader sequence having at least 90% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

206. The composition of claim 199, wherein the modified polynucleotide additionally comprises a leader sequence having at least 95% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

207. The composition of claim 199, wherein the modified polynucleotide additionally comprises a leader sequence having at least 97% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

208. The composition of claim 199, wherein the modified polynucleotide additionally comprises a leader sequence having the nucleic acid sequence of SEQ ID NO:11.

209. The composition of claim 199, wherein the pharmaceutically acceptable vector is a plasmid.

210. The composition of claim 199, wherein the pharmaceutically acceptable vector is a plasmid.

211. The composition of claim 199, wherein the pharmaceutically acceptable vector is a phage.

212. The composition of claim 199, additionally comprising a nucleic acid sequence encoding for an additional peptide, polypeptide or protein.

213. The composition of claim 199, wherein the modified polynucleotide expresses a mammalian polypeptide.

214. The composition of claim 199, wherein the modified polynucleotide expresses a bacterial polypeptide.

215. The composition of claim 199, wherein the modified polynucleotide expresses a viral polypeptide.

216. The composition of claim 199, wherein the viral polypeptide is HPV16, HIV-1 or gp160 polypeptide.

217. The composition of claim 199, wherein the modified polynucleotide comprises the ORF for the HPV16 E7 gene.

218. The composition of claim 199, wherein the modified polynucleotide comprises the ORF for the HIV-1 gag gene.

219. The composition of claim 199, wherein the modified polynucleotide comprises the ORF for the gp160 envelope gene.

220. The composition of claim 199, wherein the surrogate codons are each of the codons in said modified polynucleotide that encode for the amino acids alanine, arginine, leucine, proline, glycine, isoleucine, serine, threonine and valine.

221. A composition comprising:

the nucleic acid sequence of any of SEQ ID NOS: 1, 3 or 5;

an immunogenic encoding portion of (a); or

a nucleic acid sequence that hybridizes under stringent conditions to any of (a) or (b).

222. The composition of claim 221, wherein the composition additionally comprises one or more of a pharmaceutically acceptable diluent, buffer, adjuvant, carrier or vector.

223. The composition of claim 221, wherein the modified polynucleotide comprises the nucleic acid sequence of any of SEQ ID NOS: 1, 3 or 5.

224. The composition of claim 221, wherein the pharmaceutically acceptable vector is a plasmid.

225. The composition of claim 221, wherein the pharmaceutically acceptable vector is a phage.

226. The composition of claim 221, additionally comprising a nucleic acid sequence encoding for an additional peptide, polypeptide or protein.

227. A composition comprising:

a polynucleotide comprising the nucleic acid sequence of any of SEQ ID NOS: 1, 3 or 5; and

a vector.

228. The composition of claim 227, wherein the modified polynucleotide comprises the nucleic acid sequence of SEQ ID NO:1.

229. The composition of claim 227, wherein the modified polynucleotide comprises the nucleic acid sequence of SEQ ID NO:3.

230. The composition of claim 227, wherein the modified polynucleotide comprises the nucleic acid sequence of SEQ ID NO:5.

231. A composition comprising:

a recombinantly prepared protein or polypeptide encoded by a modified polynucleotide comprising any of:

the nucleic acid sequence of any of SEQ ID NOS: 1, 3 or 5;

an immunogenic encoding portion of (a); or

a nucleic acid sequence that hybridizes under stringent conditions to any of (a) or (b).

232. The composition of claim 231, wherein the composition additionally comprises a pharmaceutically acceptable carrier.

233. The composition of claim 231, wherein the composition additionally comprises an adjuvant.

234. The composition of claim 231, wherein the composition additionally comprises a polysaccharide.

235. The composition of claim 231, additionally comprises an additional peptide, polypeptide or protein.

236. The composition of claim 231, wherein the additional peptide, polypeptide or protein forms a conjugate with said modified polypeptide.

237. The composition of claim 231, wherein the modifiedly prepared protein or polypeptide is an immunogenic polypeptide.

238. A composition comprising:

a recombinantly prepared protein or polypeptide encoded by a modified polynucleotide comprising a nucleic acid sequence comprising surrogate codons in which the nucleotides cytosine (C) or guanine (G) occupy the wobble position in place of the corresponding nucleotides adenine (A) or uracil (U) or thymine (T) of a naturally-occurring polynucleotide that expresses the same protein or polypeptide as said modified polynucleotide.

239. The composition of claim 238, wherein the modified polynucleotide additionally comprises a non-native leader sequence.

240. The composition of claim 238, wherein the modified polynucleotide additionally comprises a human non-native leader sequence.

241. The composition of claim 238, wherein the modified polynucleotide additionally comprises an immunoglobulin leader sequence.

242. The composition of claim 238, wherein the modified polynucleotide additionally comprises (a) an IgE leader sequence or (b) a leader sequence that hybridizes to an IgE leader sequence under stringent conditions.

243. The composition of claim 238, wherein the modified polynucleotide additionally comprises (a) a leader sequence having SEQ ID NO:11 or (b) a leader sequence that hybridizes to SEQ ID NO:11 under stringent conditions.

244. The composition of claim 238, wherein the modified polynucleotide additionally comprises adding to the modified polynucleotide a leader sequence having at least 90% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

245. The composition of claim 238, wherein the modified polynucleotide additionally comprises a leader sequence having at least 95% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

246. The composition of claim 238, wherein the modified polynucleotide additionally comprises a leader sequence having at least 97% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

247. The composition of claim 238, wherein the modified polynucleotide additionally comprises a leader sequence having the nucleic acid sequence of SEQ ID NO:11.

248. The composition of claim 238, wherein the composition additionally comprises a pharmaceutically acceptable carrier.

249. The composition of claim 238, wherein the composition additionally comprises an adjuvant.

250. The composition of claim 238, wherein the composition additionally comprises a polysaccharide.

251. The composition of claim 238, additionally comprises an additional peptide, polypeptide or protein.

252. The composition of claim 238, wherein the additional peptide, polypeptide or protein forms a conjugate with said modified polypeptide.

253. The composition of claim 238, wherein the modifiedly prepared protein or polypeptide is an immunogenic polypeptide.

254. A composition that comprises a modified polynucleotide comprising:

(a) a non-native leader sequence; and

(b) a nucleic acid sequence comprising cytosine (C) or guanine (G) at the wobble position of at least one codon that encodes any of the amino acids alanine, arginine, leucine, proline, glutamic acid, glycine, isoleucine, serine, threonine, or valine where adenine (A), uracil (U) or thymine (T) occupy the wobble position of the corresponding codon of the naturally-occuring nucleic acid sequence.

255. The compostion of claim 254, wherein the modified polynucleotide additionally comprises a non-native leader sequence.

256. The composition of claim 254, wherein the modified polynucleotide additionally comprises a human non-native leader sequence.

257. The composition of claim 254, wherein the modified polynucleotide additionally comprises an immunoglobulin leader sequence.

258. The composition of claim 254, wherein the modified polynucleotide additionally comprises (a) an IgE leader sequence or (b) a leader sequence that hybridizes to an IgE leader sequence under stringent conditions.

259. The composition of claim 254, wherein the modified polynucleotide additionally comprises (a) a leader sequence having SEQ ID NO:11 or (b) a leader sequence that hybridizes to SEQ ID NO:11 under stringent conditions.

260. The composition of claim 254, wherein the modified polynucleotide additionally comprises adding to the modified polynucleotide a leader sequence having at least 90% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

261. The composition of claim 254, wherein the modified polynucleotide additionally comprises a leader sequence having at least 95% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

262. The composition of claim 254, wherein the modified polynucleotide additionally comprises a leader sequence having at least 97% sequence identity to the nucleic acid sequence of SEQ ID NO:11.

263. The composition of claim 254, wherein the modified polynucleotide additionally comprises a leader sequence having the nucleic acid sequence of SEQ ID NO:11.

264. A composition comprising:

an antibody that immunospecifically binds to the modifiedly prepared protein of any of claims 132-145.

265. The composition of claim 264, wherein the antibody is a monoclonal antibody.

266. The composition of claim 264, wherein the antibody is a polyclonal antibody.

267. A composition prepared by a process comprising:

inserting into a vector a modified nucleic acid sequence comprising surrogate codons in place of the corresponding naturally-occurring codons having adenine (A) or uracil (U) or thymine (T) in the wobble position.

268. The composition of claim 267, wherein a non-native leader sequence is additionally inserted into the vector along with the modified nucleic acid sequence.

269. The composition of claim 267, wherein a human non-native leader sequence is additionally inserted into the vector along with the modified nucleic acid sequence.

270. The composition of claim 267, wherein an immunoglobulin leader sequence is additionally inserted into the vector along with the modified nucleic acid sequence.

271. The composition of claim 267, wherein an IgE leader sequence or a leader sequence that hybridizes to an IgE leader sequence under stringent conditions is additionally inserted into the vector along with the modified nucleic acid sequence.

272. The composition of claim 267, wherein a leader sequence having SEQ ID NO:11 or a leader sequence that hybridizes to SEQ ID NO:11 under stringent conditionsis additionally inserted into the vector along with the modified nucleic acid sequence.

273. The composition of claim 267, wherein a leader sequence having at least 90% sequence identity to the nucleic acid sequence of SEQ ID NO:11 is additionally inserted into the vector along with the modified nucleic acid sequence.

274. The composition of claim 267, a leader sequence having at least 95% sequence identity to the nucleic acid sequence of SEQ ID NO:11 is additionally inserted into the vector along with the modified nucleic acid sequence.

275. The composition of claim 267, wherein a leader sequence having at least 97% sequence identity to the nucleic acid sequence of SEQ ID NO:11 is additionally inserted into the vector along with the modified nucleic acid sequence.

276. The composition of claim 267, wherein a leader sequence having the nucleic acid sequence of SEQ ID NO:11 is additionally inserted into the vector along with the modified nucleic acid sequence.

277. The composition of claim 267, wherein the modified polynucleotide expresses a viral, bacterial, protist, fungal, plant, or animal polypeptide.

278. The composition of claim 267, wherein the modified polynucleotide expresses a mammalian polypeptide.

279. The composition of claim 267, wherein the modified polynucleotide expresses a bacterial polypeptide.

280. The composition of claim 267, wherein the modified polynucleotide expresses a viral polypeptide.

281. The composition of claim 267, wherein the viral polypeptide is HPV16 polypeptide, HIV-1 polypeptide or gp160 polypetide.

282. The composition of claim 267, wherein the modified polynucleotide comprises the ORF for the HPV16 E7 gene.

283. The composition of claim 267, wherein the modified polynucleotide comprises the ORF for the HIV-1 gag gene.

284. The composition of claim 267, wherein the modified polynucleotide comprises the ORF for the gp160 envelope gene.

285. The composition of claim 267, wherein the surrogate codons are a randomized selection of at least about 10% of the codons in said modified polynucleotide that encode for any of the amino acids alanine, arginine, leucine, proline, glutamic acid, glycine, isoleucine, serine, threonine and valine.

286. The composition of claim 267, wherein the surrogate codons are a randomized selection of at least about 50% of the codons in said modified polynucleotide that encode for any of the amino acids alanine, arginine, leucine, proline, glutamic acid, glycine, isoleucine, serine, threonine and valine.

287. The composition of claim 267, wherein the surrogate codons are a randomized selection of at least about 90% of the codons in said modified polynucleotide that encode for any of the amino acids alanine, arginine, leucine, proline, glutamic acid, glycine, isoleucine, serine, threonine and valine.

288. The composition of claim 267, wherein the surrogate codons are each of the codons in said modified polynucleotide that encode for the amino acids alanine, arginine, leucine, proline, glutamic acid, glycine, isoleucine, serine, threonine and valine.

289. The composition of claim 267, wherein the modified polynucleotide is a DNA molecule.

290. The composition of claim 267, wherein the modified polynucleotide is an RNA molecule.

291. A composition prepared by a process comprising:

inserting into a vector a nucleic acid sequence in which each codon encoding alanine is GCG, each codon encoding arginine is CGG or AGG, each codon encoding leucine is CTC, each codon encoding proline is CCT or CCG, each codon encoding glutamic acid is GAG, each codon encoding glycine is GGG, each codon

encoding isoleucine is ATT, each codon encoding serine is TCC, each codon encoding threonine is ACG, and each codon encoding valine is GTC.

292. The compostion of claim 291, wherein a non-native leader sequence is additionally inserted into the vector along with the modified nucleic acid sequence.

293. The composition of claim 291, wherein a human non-native leader sequence is additionally inserted into the vector along with the modified nucleic acid sequence.

294. The composition of claim 291, wherein an immunoglobulin leader sequence is additionally inserted into the vector along with the modified nucleic acid sequence.

295. The composition of claim 291, wherein an IgE leader sequence or a leader sequence that hybridizes to an IgE leader sequence under stringent conditions is additionally inserted into the vector along with the modified nucleic acid sequence.

296. The composition of claim 291, wherein a leader sequence having SEQ ID NO:11 or a leader sequence that hybridizes to SEQ ID NO:11 under stringent conditionsis additionally inserted into the vector along with the modified nucleic acid sequence.

297. The composition of claim 291, wherein a leader sequence having at least 90% sequence identity to the nucleic acid sequence of SEQ ID NO:11 is additionally inserted into the vector along with the modified nucleic acid sequence.

298. The composition of claim 291, a leader sequence having at least 95% sequence identity to the nucleic acid sequence of SEQ ID NO:11 is additionally inserted into the vector along with the modified nucleic acid sequence.

299. The composition of claim 291, wherein a leader sequence having at least 97% sequence identity to the nucleic acid sequence of SEQ ID NO:11 is additionally inserted into the vector along with the modified nucleic acid sequence.

300. The composition of claim 291, wherein a leader sequence having the nucleic acid sequence of SEQ ID NO:11 is additionally inserted into the vector along with the modified nucleic acid sequence.

301. The composition of claim 291, wherein the modified polynucleotide expresses a viral, bacterial, protist, fungal, plant, or animal polypeptide.

302. The composition of claim 291, wherein the modified polynucleotide expresses a mammalian polypeptide.

303. The composition of claim 291, wherein the modified polynucleotide expresses a bacterial polypeptide.

304. The composition of claim 291, wherein the modified polynucleotide expresses a viral polypeptide.

305. The composition of claim 291, wherein the viral polypeptide is HPV16, HIV-1 or gp160 polypeptide.

306. The composition of claim 291, wherein the modified polynucleotide comprises the ORF for the HPV16 E7 gene.

307. The composition of claim 291, wherein the modified polynucleotide comprises the ORF for the HIV-1 gag gene.

308. The composition of claim 291, wherein the modified polynucleotide comprises the ORF for the gp160 envelope gene.

309. The composition of claim 291, wherein the surrogate codons are a randomized selection of at least about 10% of the codons in said modified polynucleotide that encode for any of the amino acids alanine, arginine, leucine, proline, glutamic acid, glycine, isoleucine, serine, threonine and valine.

310. The composition of claim 291, wherein the surrogate codons are a randomized selection of at least about 50% of the codons in said modified polynucleotide that encode for any of the amino acids alanine, arginine, leucine, proline, glutamic acid, glycine, isoleucine, serine, threonine and valine.

311. The composition of claim 291, wherein the surrogate codons are a randomized selection of at least about 90% of the codons in said modified polynucleotide that encode for any of the amino acids alanine, arginine, leucine, proline, glutamic acid, glycine, isoleucine, serine, threonine and valine.

312. The composition of claim 291, wherein the surrogate codons are each of the codons in said modified polynucleotide that encode for the amino acids alanine, arginine, leucine, proline, glutamic acid, glycine, isoleucine, serine, threonine and valine.

313. The composition of claim 291, wherein the modified polynucleotide is a DNA molecule.

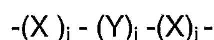
314. The composition of claim 291, wherein the modified polynucleotide is an RNA molecule.

315. The composition of claim 291, wherein the vector is a plasmid.

316. The composition of claim 291, wherein the vector is a phage.

317. A composition prepared by a process comprising:

inserting into a vector a polynucleotide comprising a nucleic acid sequence having the general formula:



$$-(X)_i - (Y)_j - (X)_i -$$

wherein X represents non-surrogate codons having the nucleic acid sequence of any of the corresponding wild-type codons in the naturally-occurring polypeptide that encode the same protein or polypeptide as said modified polynucleotide, said wild-type codons having cytosine (C) or guanine (G) in the wobble position;

wherein Y represents surrogate codons having a nucleic acid sequence that is different from the corresponding wild-type codons in the naturally-occurring polypeptide that encode the same protein or polypeptide as said modified polynucleotide, said wild-type codons having adenine (A) or uracil (U) or thymine (T) in the wobble position, said surrogate codons having cytosine (C) or guanine (G) in the wobble position and encoding the same amino acid as the corresponding wild-type codons in the naturally-occurring polypeptide that encodes the same protein or polypeptide as said modified polynucleotide;

wherein i is any positive integer of at least 0; and

wherein j is any positive integer of at least 1.

318. The composition of claim 317, wherein a non-native leader sequence is additionally inserted into the vector along with the modified nucleic acid sequence.

319. The composition of claim 317, wherein a human non-native leader sequence is additionally inserted into the vector along with the modified nucleic acid sequence.

320. The composition of claim 317, wherein an immunoglobulin leader sequence is additionally inserted into the vector along with the modified nucleic acid sequence.

321. The composition of claim 317, wherein an IgE leader sequence or a leader sequence that hybridizes to an IgE leader sequence under stringent conditions is additionally inserted into the vector along with the modified nucleic acid sequence.

322. The composition of claim 317, wherein a leader sequence having SEQ ID NO:11 or a leader sequence that hybridizes to SEQ ID NO:11 under stringent conditionsis additionally inserted into the vector along with the modified nucleic acid sequence.

323. The composition of claim 317, wherein a leader sequence having at least 90% sequence identity to the nucleic acid sequence of SEQ ID NO:11 is additionally inserted into the vector along with the modified nucleic acid sequence.

324. The composition of claim 317, a leader sequence having at least 95% sequence identity to the nucleic acid sequence of SEQ ID NO:11 is additionally inserted into the vector along with the modified nucleic acid sequence.

325. The composition of claim 317, wherein a leader sequence having at least 97% sequence identity to the nucleic acid sequence of SEQ ID NO:11 is additionally inserted into the vector along with the modified nucleic acid sequence.

326. The composition of claim 317, wherein a leader sequence having the nucleic acid sequence of SEQ ID NO:11 is additionally inserted into the vector along with the modified nucleic acid sequence.

327. The composition of claim 317, wherein the modified polynucleotide expresses a viral, bacterial, protist, fungal, plant, or animal polypeptide.

328. The composition of claim 317, wherein the modified polynucleotide expresses a mammalian polypeptide.

329. The composition of claim 317, wherein the modified polynucleotide expresses a bacterial polypeptide.

330. The composition of claim 317, wherein the modified polynucleotide expresses a viral polypeptide.

331. The composition of claim 317, wherein the viral polypeptide is human HPV16, HIV-1 or gp160 polypeptide.

332. The composition of claim 317, wherein the modified polynucleotide comprises the ORF for the HPV16 E7 gene.

333. The composition of claim 317, wherein the modified polynucleotide comprises the ORF for the HIV-1 gag gene.

334. The composition of claim 317, wherein the modified polynucleotide comprises the ORF for the gp160 envelope gene.

335. The composition of claim 317, wherein the surrogate codons are a randomized selection of at least about 10% of the codons in said modified polynucleotide that encode for any of the amino acids alanine, arginine, leucine, proline, glutamic acid, glycine, isoleucine, serine, threonine and valine.

336. The composition of claim 317, wherein the surrogate codons are a randomized selection of at least about 50% of the codons in said modified polynucleotide that encode for any of the amino acids alanine, arginine, leucine, proline, glycine, isoleucine, serine, threonine and valine.

337. The composition of claim 317, wherein the surrogate codons are a randomized selection of at least about 90% of the codons in said modified polynucleotide that encode for any of the amino acids alanine, arginine, leucine, proline, glycine, isoleucine, serine, threonine and valine.

338. The composition of claim 317, wherein the surrogate codons are each of the codons in said modified polynucleotide that encode for the amino acids alanine, arginine, leucine, proline, glycine, isoleucine, serine, threonine and valine.

339. The composition of claim 317, wherein the modified polynucleotide is a DNA molecule.

340. The composition of claim 317, wherein the modified polynucleotide is an RNA molecule.

341. A composition prepared by a process comprising:

inserting into a vector any of:

the nucleic acid sequence of any of SEQ ID NOS: 1, 3 or 5;

an immunogenic encoding portion of (a); or

a nucleic acid sequence that hybridizes under stringent conditions to any of (a) or (b).

342. The composition of claim 341, wherein a non-native leader sequence is additionally inserted into the vector along with the modified nucleic acid sequence.

343. The composition of claim 341, wherein a human non-native leader sequence is additionally inserted into the vector along with the modified nucleic acid sequence.

344. The composition of claim 341, wherein an immunoglobulin leader sequence is additionally inserted into the vector along with the modified nucleic acid sequence.

345. The composition of claim 341, wherein an IgE leader sequence or a leader sequence that hybridizes to an IgE leader sequence under stringent conditions is additionally inserted into the vector along with the modified nucleic acid sequence.

346. The composition of claim 341, wherein a leader sequence having SEQ ID NO:11 or a leader sequence that hybridizes to SEQ ID NO:11 under stringent conditionsis additionally inserted into the vector along with the modified nucleic acid sequence.

347. The composition of claim 341, wherein a leader sequence having at least 90% sequence identity to the nucleic acid sequence of SEQ ID NO:11 is additionally inserted into the vector along with the modified nucleic acid sequence.

348. The composition of claim 341, a leader sequence having at least 95% sequence identity to the nucleic acid sequence of SEQ ID NO:11 is additionally inserted into the vector along with the modified nucleic acid sequence.

349. The composition of claim 341, wherein a leader sequence having at least 97% sequence identity to the nucleic acid sequence of SEQ ID NO:11 is additionally inserted into the vector along with the modified nucleic acid sequence.

350. The composition of claim 341, wherein a leader sequence having the nucleic acid sequence of SEQ ID NO:11 is additionally inserted into the vector along with the modified nucleic acid sequence.

351. The composition of claim 341, wherein the modified polynucleotide expresses a viral, bacterial, protist, fungal, plant, or animal polypeptide.

352. The composition of claim 341, wherein the modified polynucleotide expresses a mammalian polypeptide.

353. The composition of claim 341, wherein the modified polynucleotide expresses a bacterial polypeptide.

354. The composition of claim 341, wherein the modified polynucleotide expresses a viral polypeptide.

355. The composition of claim 341, wherein the viral polypeptide is HPV16, HIV-1 or gp160 polypeptides.

356. The composition of claim 341, wherein the modified polynucleotide comprises the ORF for HPV16 E7 gene.

357. The composition of claim 341, wherein the modified polynucleotide comprises the ORF for the HIV-1 gag gene.

358. The composition of claim 341, wherein the modified polynucleotide comprises the ORF for the gp160 envelope gene.

359. The composition of claim 341, wherein the surrogate codons are a randomized selection of at least about 10% of the codons in said modified polynucleotide that encode for any of the amino acids alanine, arginine, leucine, proline, glutamic acid, glycine, isoleucine, serine, threonine and valine.

360. The composition of claim 341, wherein the surrogate codons are a randomized selection of at least about 50% of the codons in said modified polynucleotide that encode for any of the amino acids alanine, arginine, leucine, proline, glycine, isoleucine, serine, threonine and valine.

361. The composition of claim 341, wherein the surrogate codons are a randomized selection of at least about 90% of the codons in said modified polynucleotide that encode for any of the amino acids alanine, arginine, leucine, proline, glycine, isoleucine, serine, threonine and valine.

362. The composition of claim 341, wherein the surrogate codons are each of the codons in said modified polynucleotide that encode for the amino acids alanine, arginine, leucine, proline, glycine, isoleucine, serine, threonine and valine.

363. The composition of claim 341, wherein the modified polynucleotide is a DNA molecule.

364. The composition of claim 341, wherein the modified polynucleotide is an RNA molecule.

365. A composition comprising:

a compound for preventing or treating a condition in an animal; and

an adjuvant for the compound for preventing or treating a condition in an animal,

wherein the adjuvant comprises a modified polynucleotide of any of claims 80-155.

366. A composition comprising:

a modified polynucleotide of any of claims 80-155; and

an immunogenic compound or composition.

367. The composition of claim 366, wherein the modified polynucleotide and the immunogenic compound or composition are contained within a vector.

368. A composition that comprises a recombinant polynucleotide comprising:

(a) an IgE leader sequence; and

(b) a nucleic acid sequence comprising cytosine (C) or guanine (G) at the wobble position of at least one codon that encodes any of the amino acids alanine, arginine, leucine, proline, glutamic acid, glycine, isoleucine, serine, threonine, or valine where adenine (A), uracil (U) or thymine (T) occupy the wobble position of the corresponding codon of the naturally-occuring nucleic acid sequence.

369. A composition comprising:

a polynucleotide comprising (a) a nucleic acid sequence having at least about 70% sequence identity to the nucleic acid sequence of SEQ ID NO:14; or (b) a nucleic acid sequence that hybridizes to SEQ ID NO:14 under stringent conditions.

370. The composition of claim 369, wherein the nucleic acid sequence has at least about 80% sequence identity to SEQ ID NO:14.

371. The composition of claim 369, wherein the nucleic acid sequence has at least about 90% sequence identity to SEQ ID NO:14.

372. The composition of claim 369, wherein the nucleic acid sequence has at least about 95% sequence identity to SEQ ID NO:14.

373. The composition of claim 369, wherein the nucleic acid sequence has at least about 97% sequence identity to SEQ ID NO:14.

374. The composition of claim 369, wherein the nucleic acid sequence comprises the nucleic acid sequence of SEQ ID NO:14.

375. A vector comprising any of the modified polynucleotides of any of claims 80-155.

376. The vector of claim 375, wherein the vector is a plasmid.

377. The vector of claim 375, wherein the modified polynucleotide comprises the nucleic acid sequence of SEQ ID NO:14.

378. The vector of claim 375, additionally comprising one or more immunogenic compounds and/or compositions.

379. The vector of claim 378, wherein the immunogenic compound or composition comprises DNA or RNA.

380. The vector of claim 278, wherein the immunogenic compound or composition provides immunogenicity against one or more retrovirus.

381. The vector of claim 278, wherein the immunogenic compound or composition provides protection against HIV and/or AIDS when administered to a human.

382. Use of the composition of any of claims 156-374 in the preparation of a medicament for inducing an immune response in a mammal.

383. The use according to claim 382, wherein said composition is administered parenterally.

384. The use according to claim 382, wherein said composition is administered mucosally.

385. The use according to claim 382, wherein said composition is administered by subcutaneous or intramuscular injection.

386. The use according to claim 382, wherein said composition is effective against infection in the mammal.

387. The use according to claim 382, wherein said infection is infection by HIV.

388. Use of the composition of any of claims 156-374 in the preparation of a medicament for treating a condition in a mammal.

389. The use according to claim 388, wherein said condition is HIV infection and/or AIDS.

390. The use according to claim 388, wherein said composition is administered parenterally.

391. The use according to claim 388, wherein said composition is administered mucosally.

392. The use according to claim 388, wherein said composition is administered by subcutaneous or intramuscular injection.

393. The use according to claim 388, wherein said composition is effective against infection in the mammal.

394. A transformed, transfected, lipofected or infected cell line comprising:

a modified cell that expresses any of:

the nucleic acid sequence of any of SEQ ID NOS: 1, 3, 5, 12, 13, 14, 15 or 16;

an immunogenic encoding portion of (a); or

a nucleic acid sequence that hybridizes under stringent conditions to any of (a) or (b).

395. The cell line of claim 394, wherein the cell is a hybridoma.

396. The cell line of claim 394, wherein the cell is a trioma.

397. A transformed, transfected, lipofected or infected cell line comprising:

a modified cell that expresses any of the modified polynucleotides described herein.

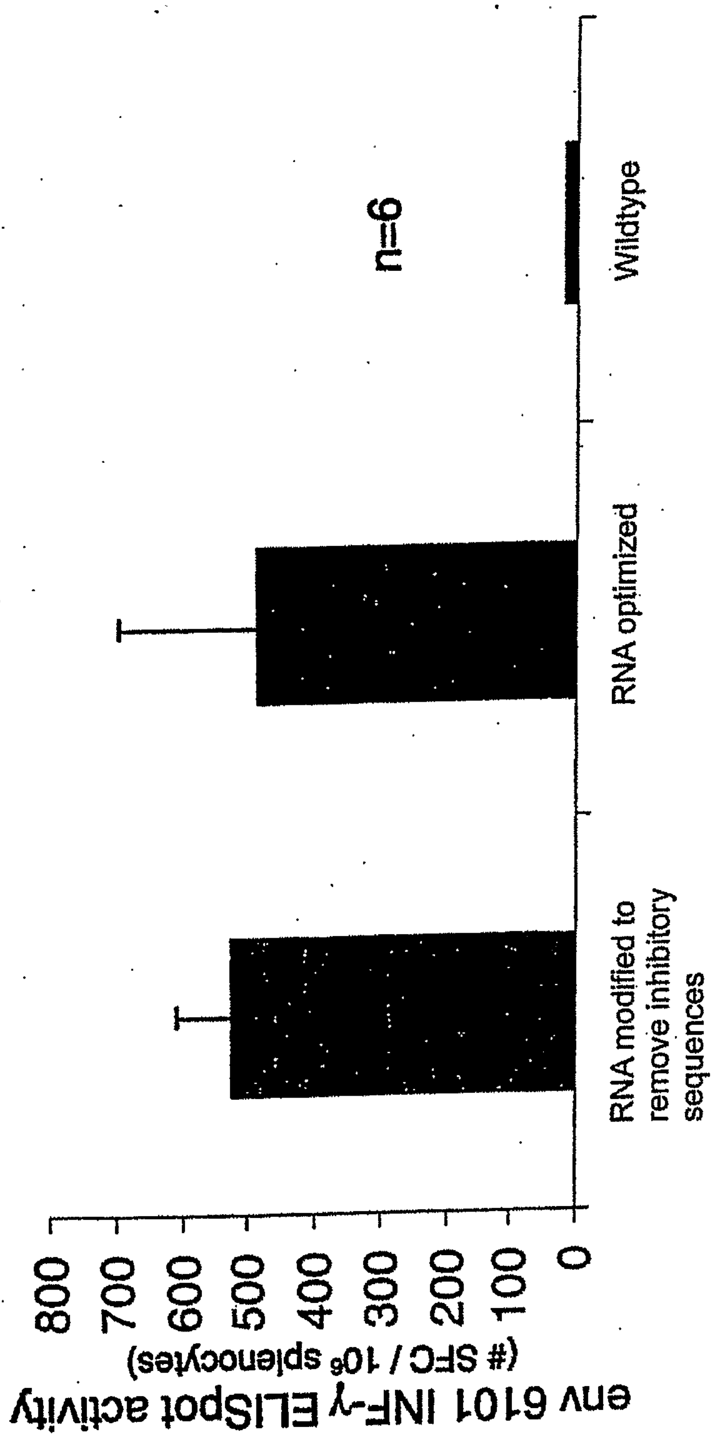


Fig. 1

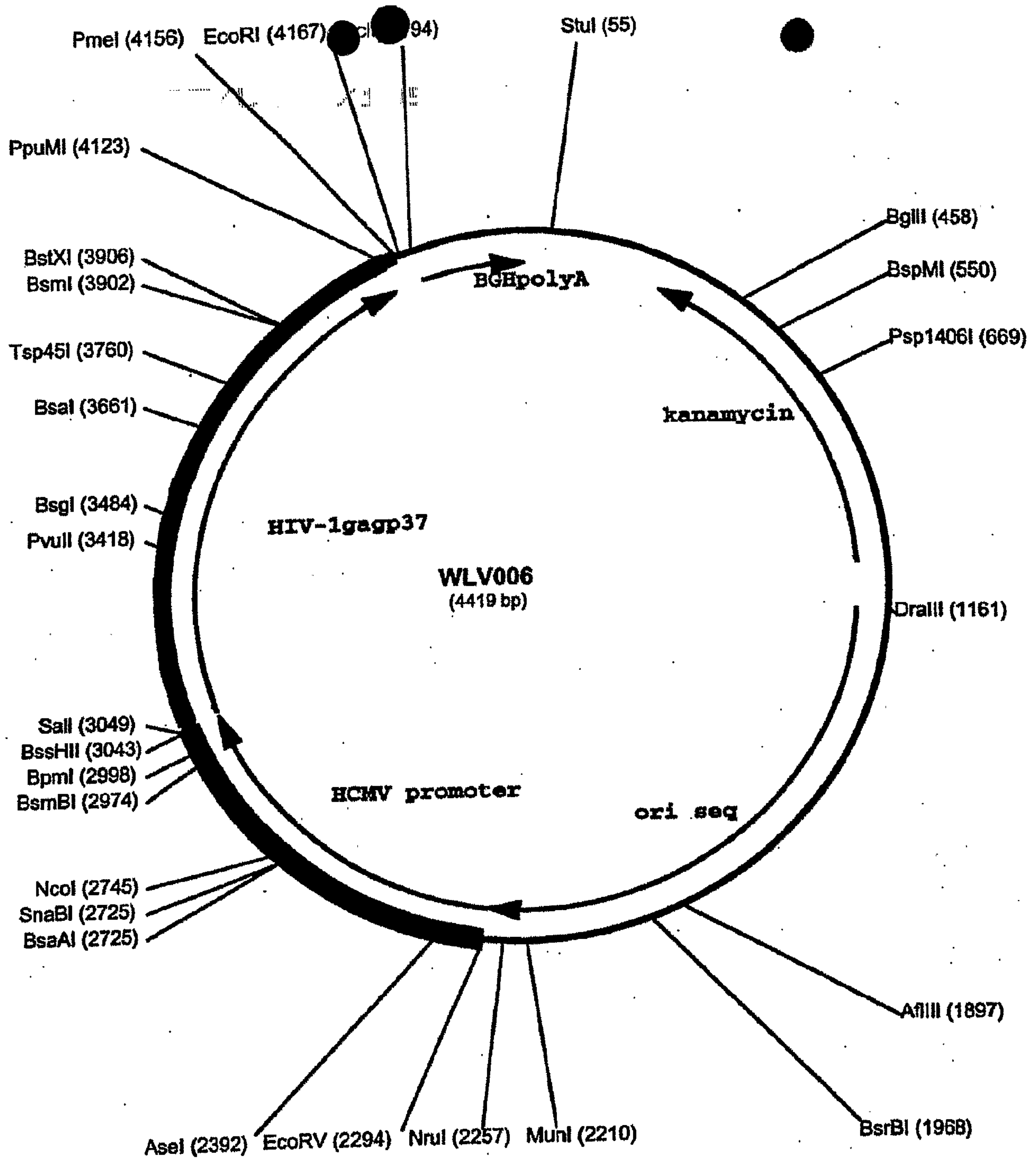


Fig. 2

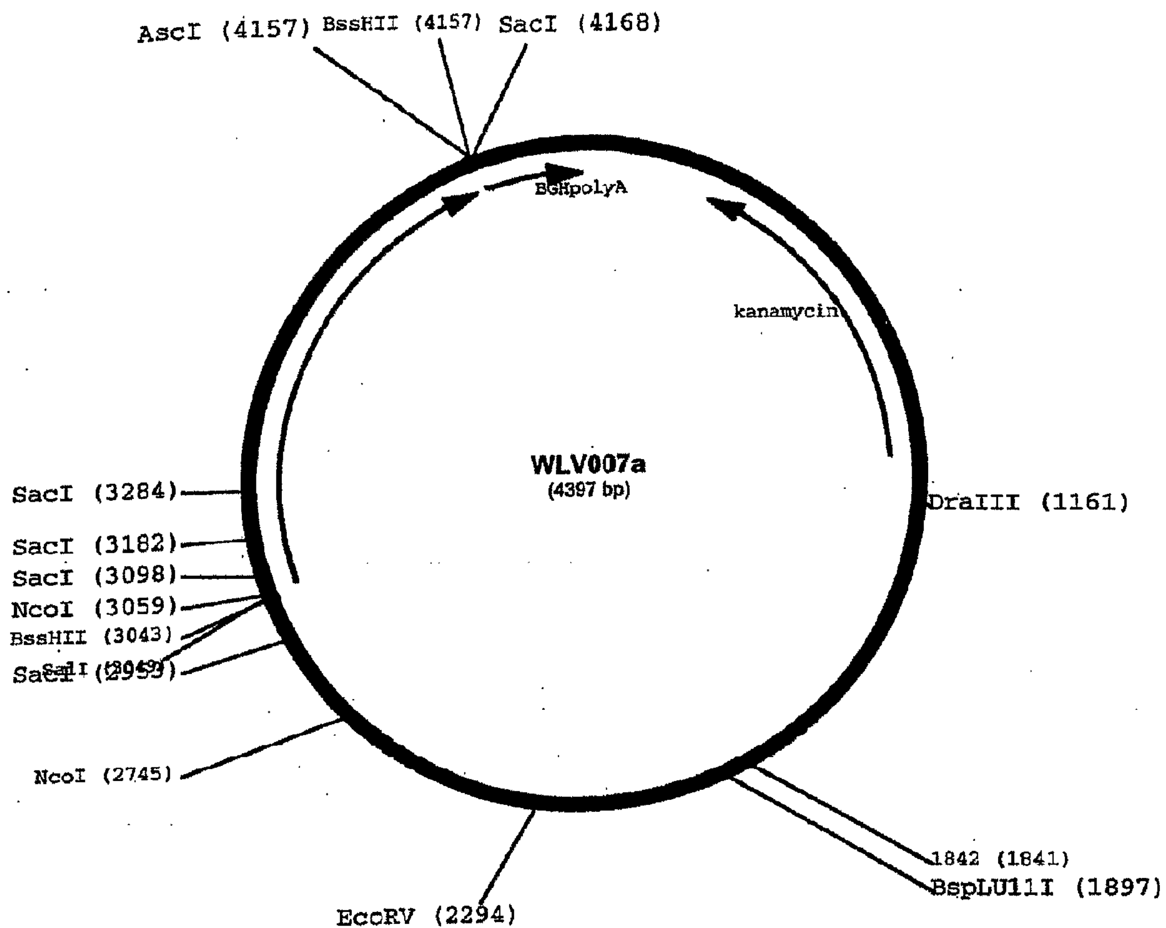


Fig. 3

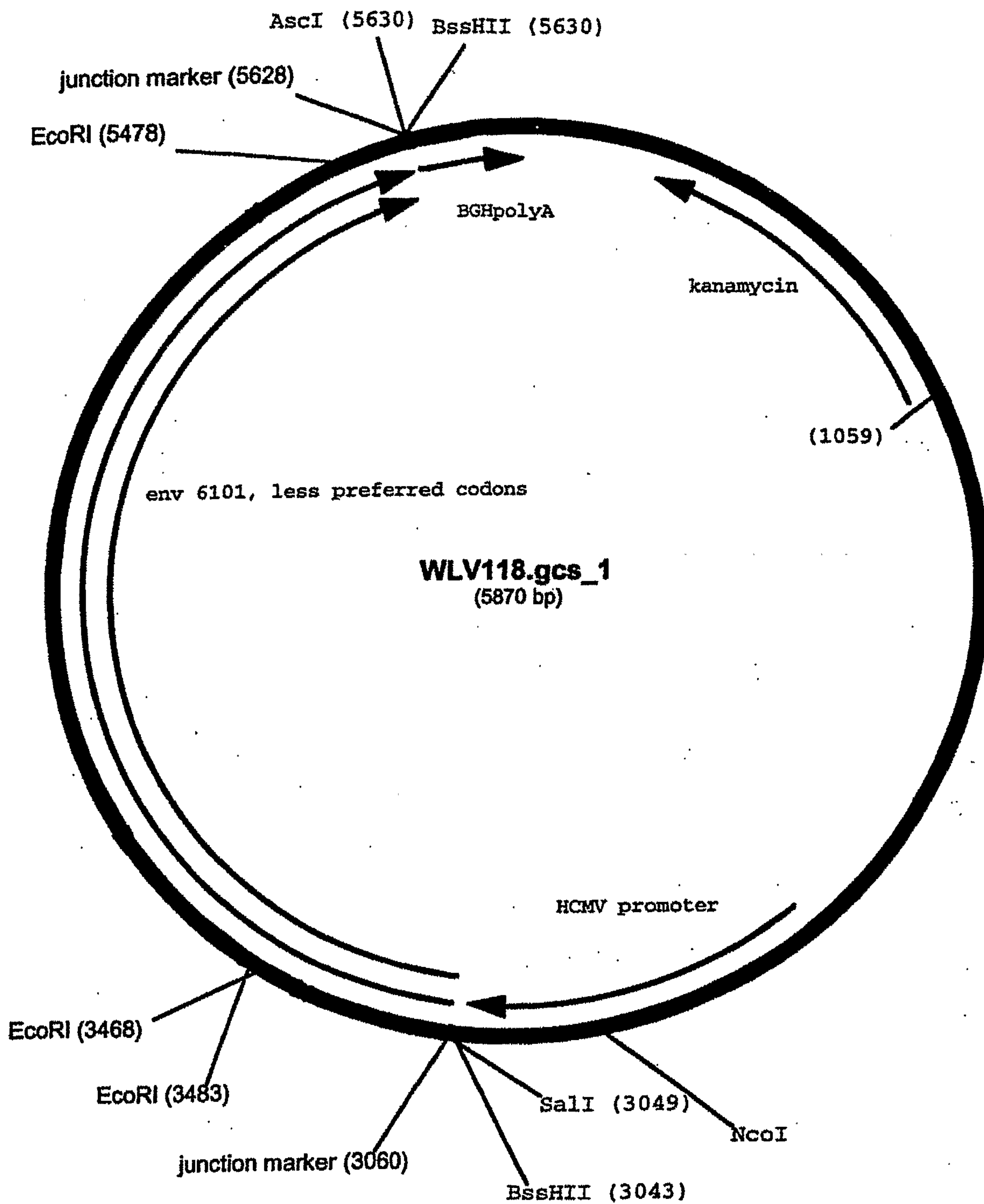


Fig. 4

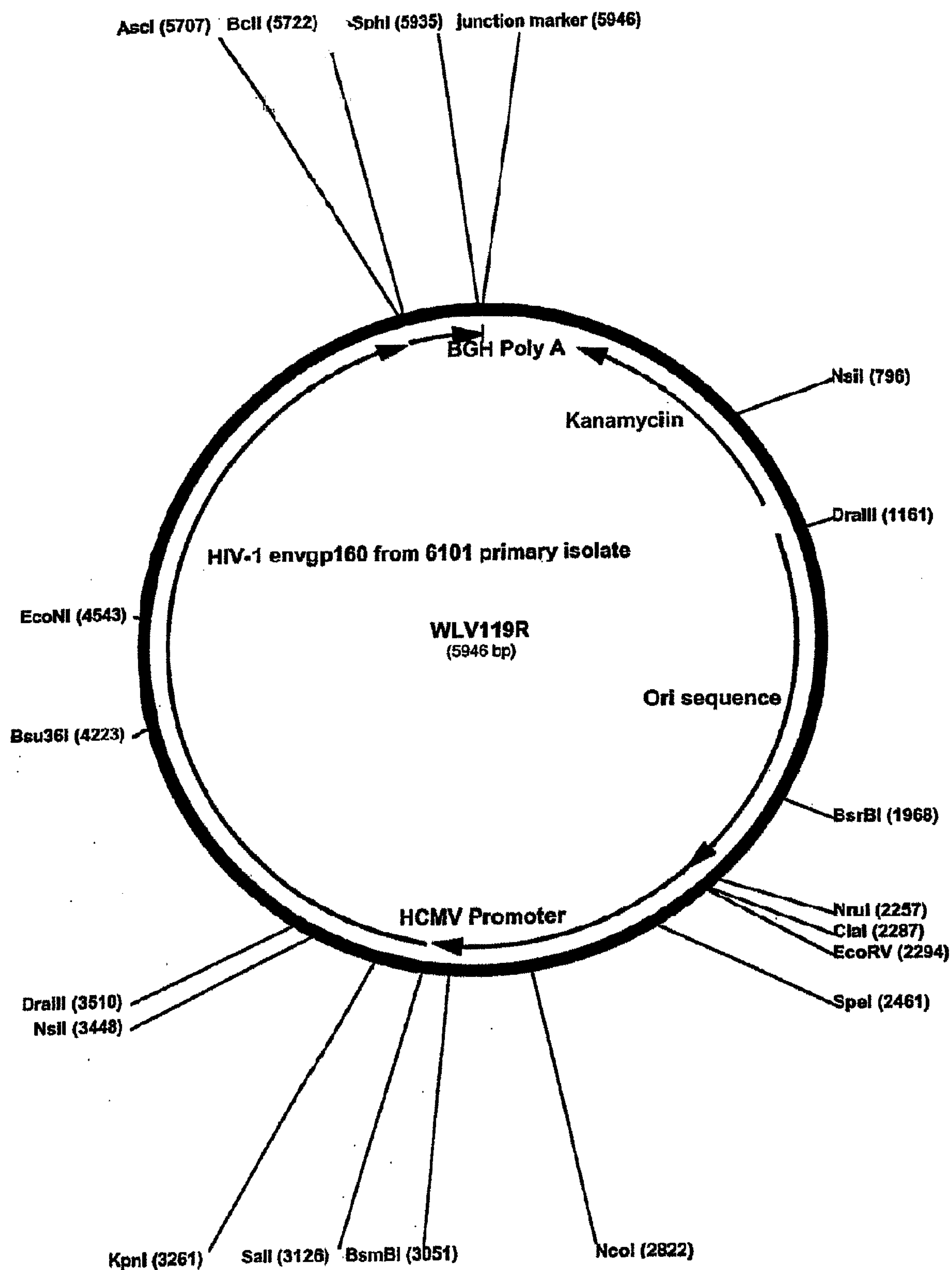
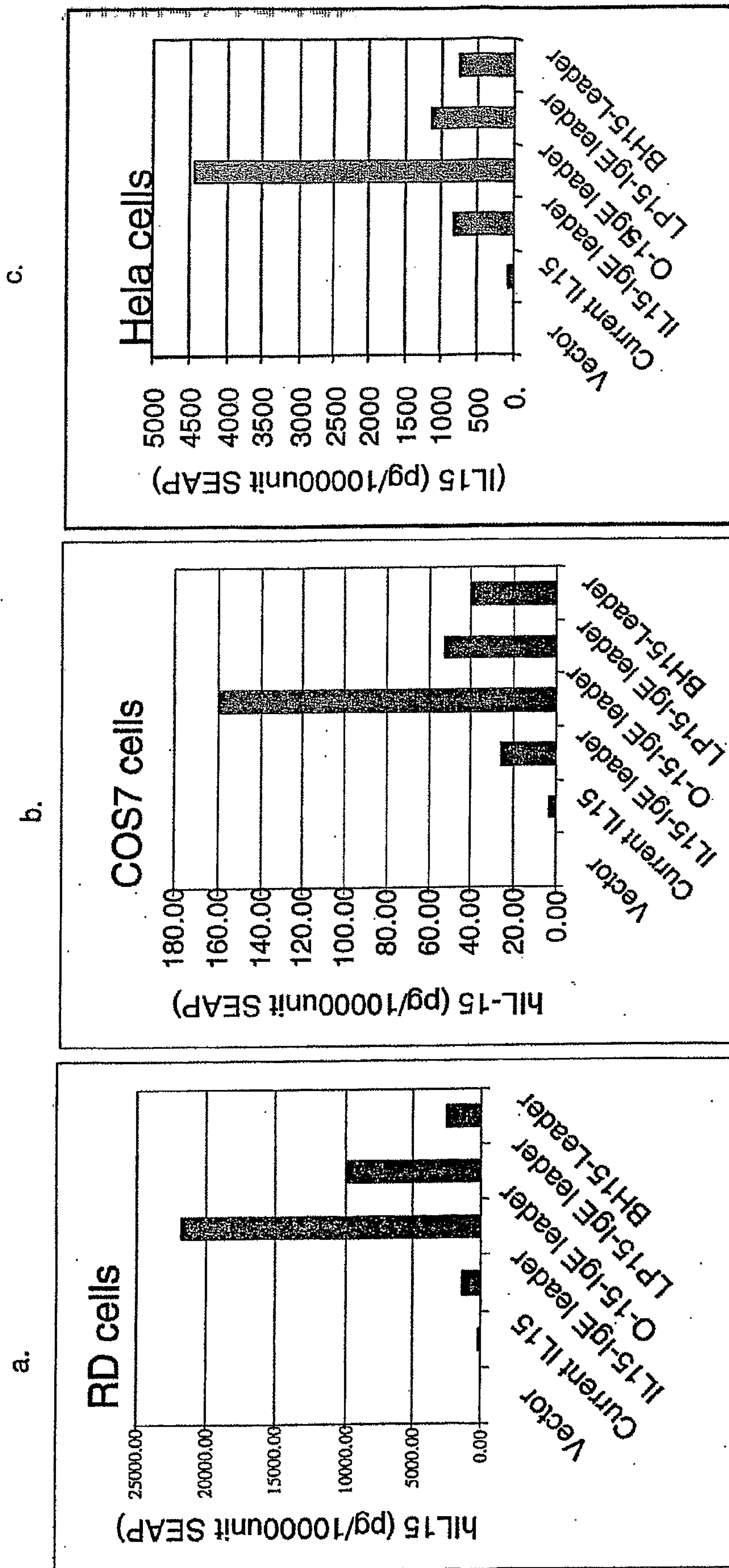


Fig. 5



C: current clinical construct, LP: modified IL-15, O: optimized alternative 1, BH: optimized alternative 2

FIG. 6

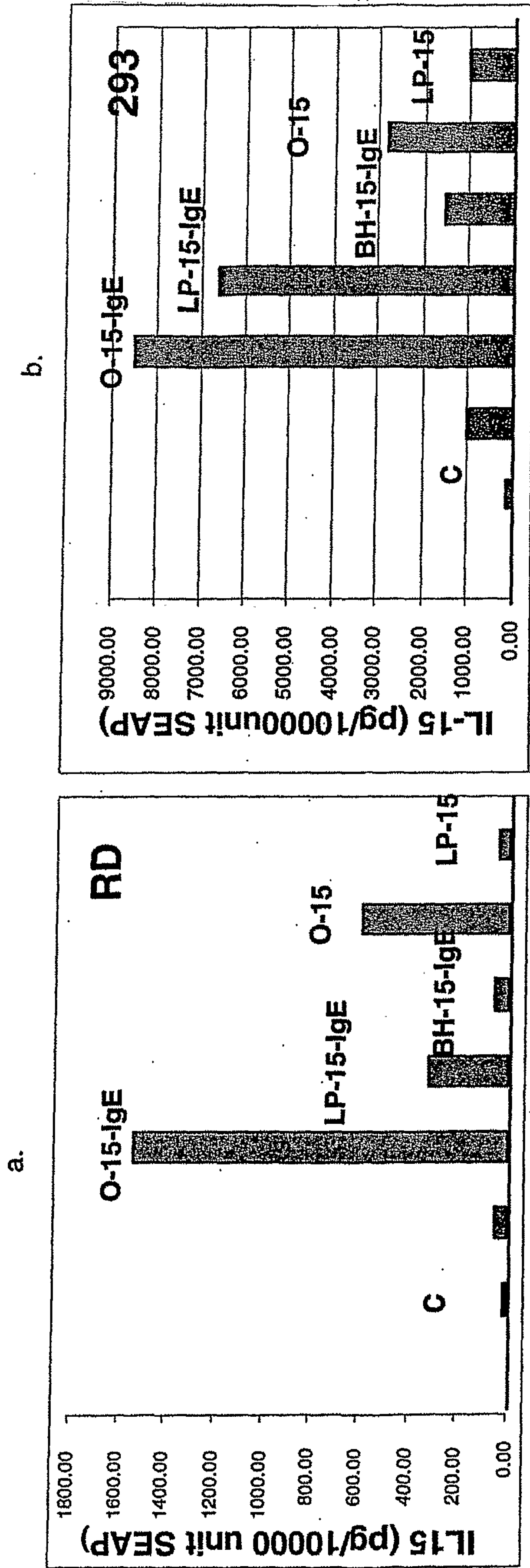


FIG. 7

IL-15 Expression (Fold Increase)

Cell type	Current IL-15	IL-15IgE leader	O-15IgE leader	LP-15IgE leader	BH-15IgE leader	O-15	LP-15
RD	1.00	3.75	102.00	20.00	3.80	41.00	2.70
COS7	1.00	6.61	41.06	13.69	10.32	nd	nd
Hela	1.00	12.60	66.41	17.16	11.43	nd	nd
293	1.00	7.20	61.00	47.00	11.00	20.00	7.30

nd: not done

FIG. 8

CTLL2 Assay

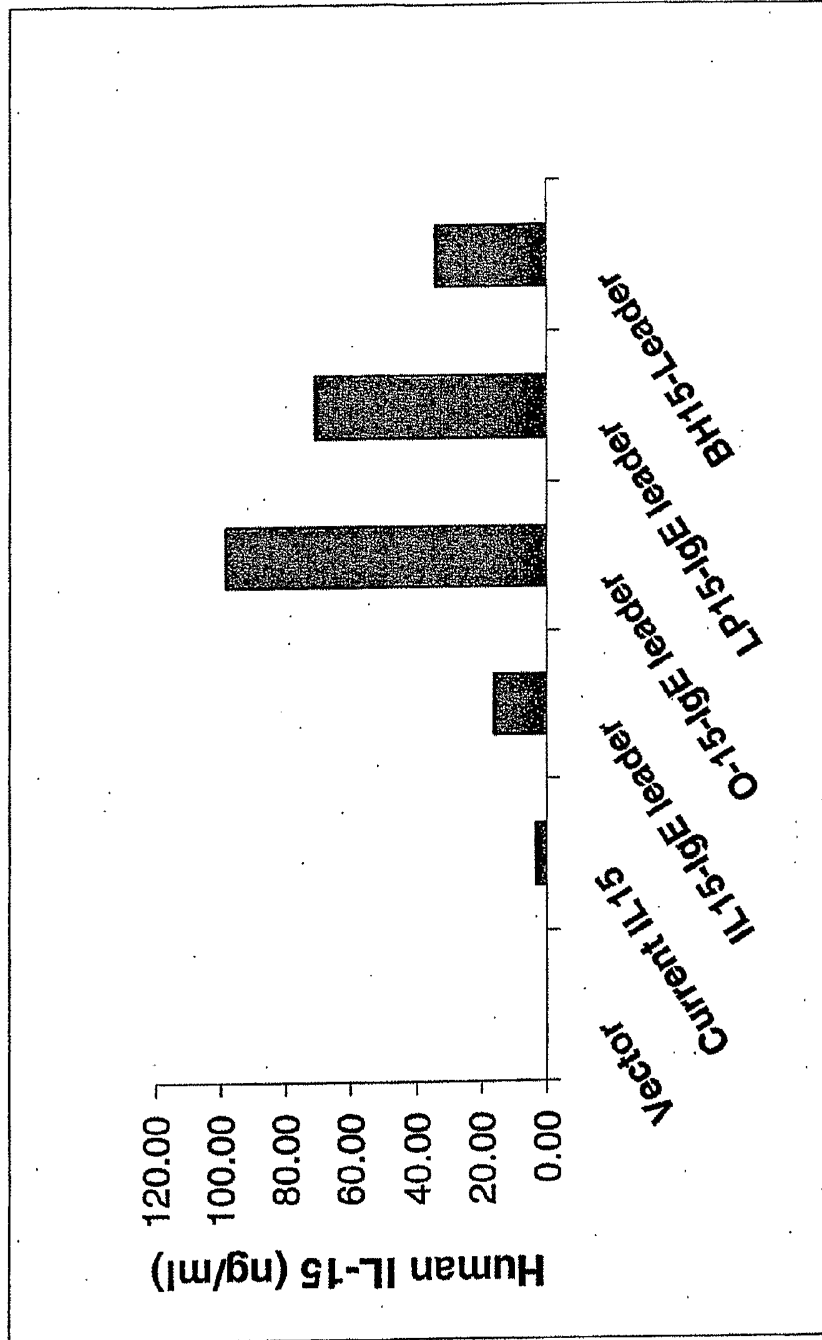


FIG. 9

In Vivo Gene Expression From IL-15 Plasmids

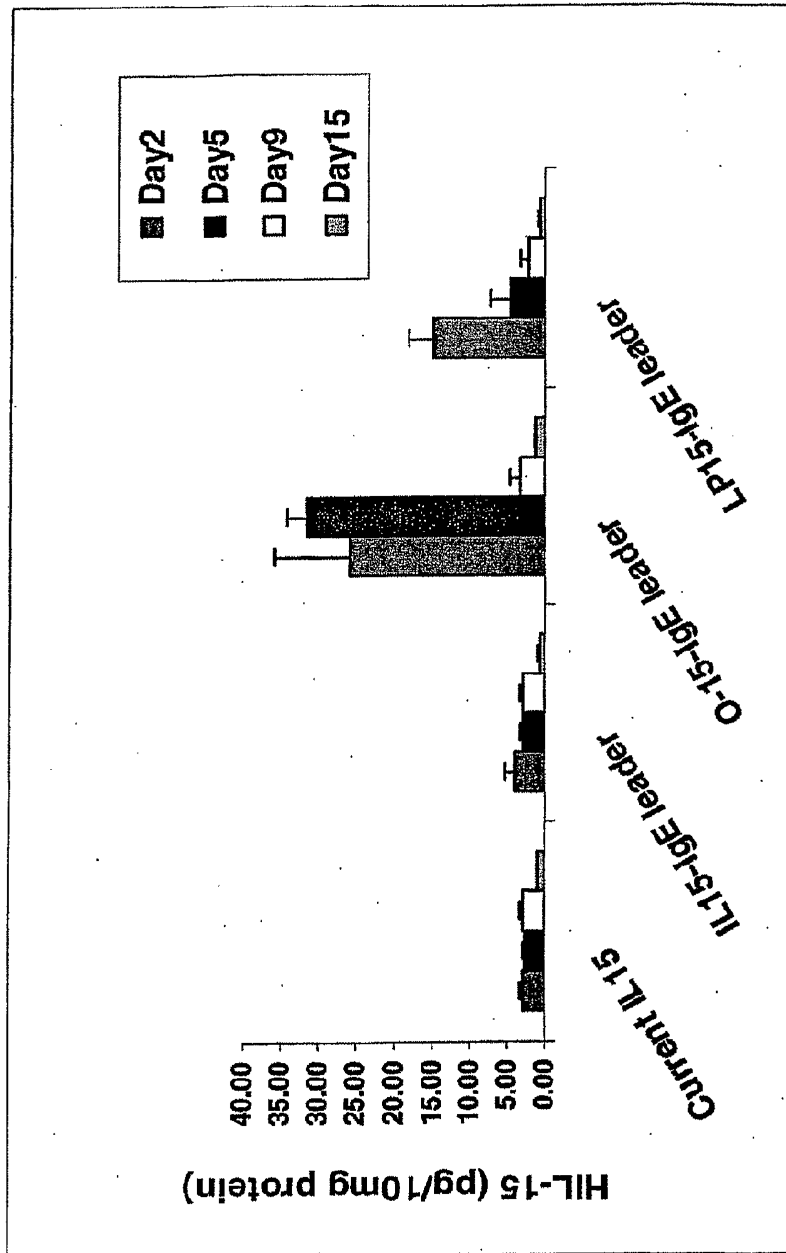


FIG. 10

FIGURE 11

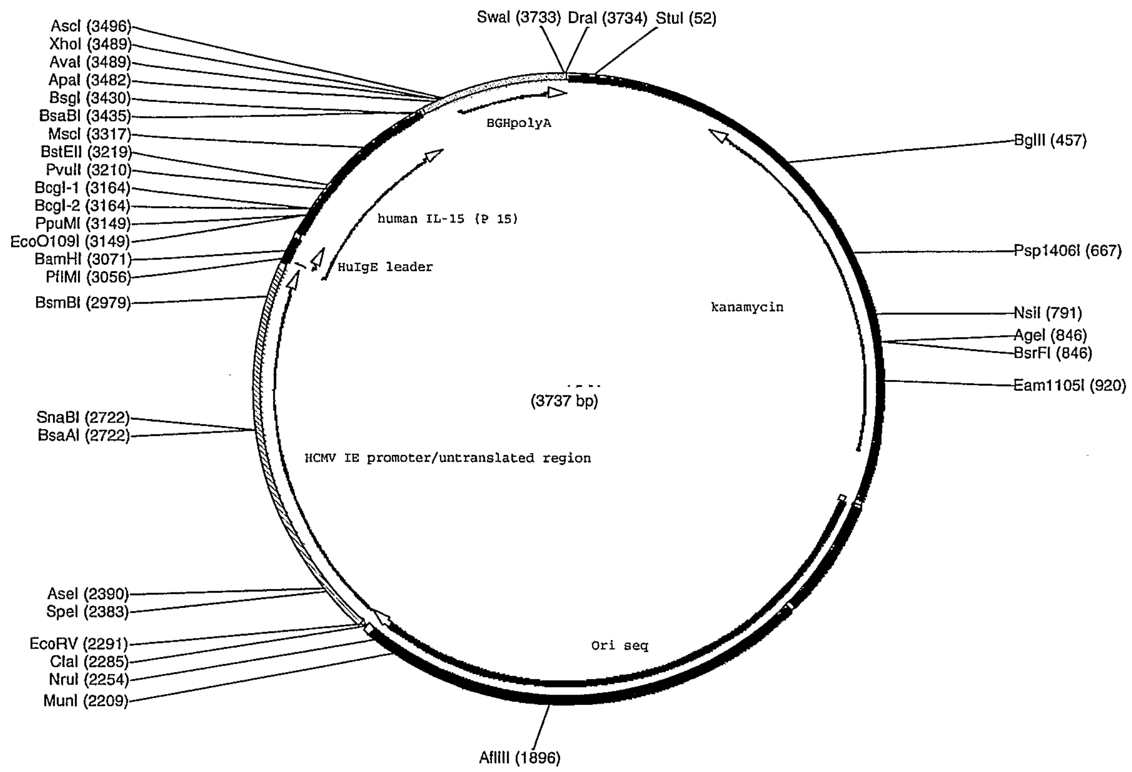


FIGURE 12

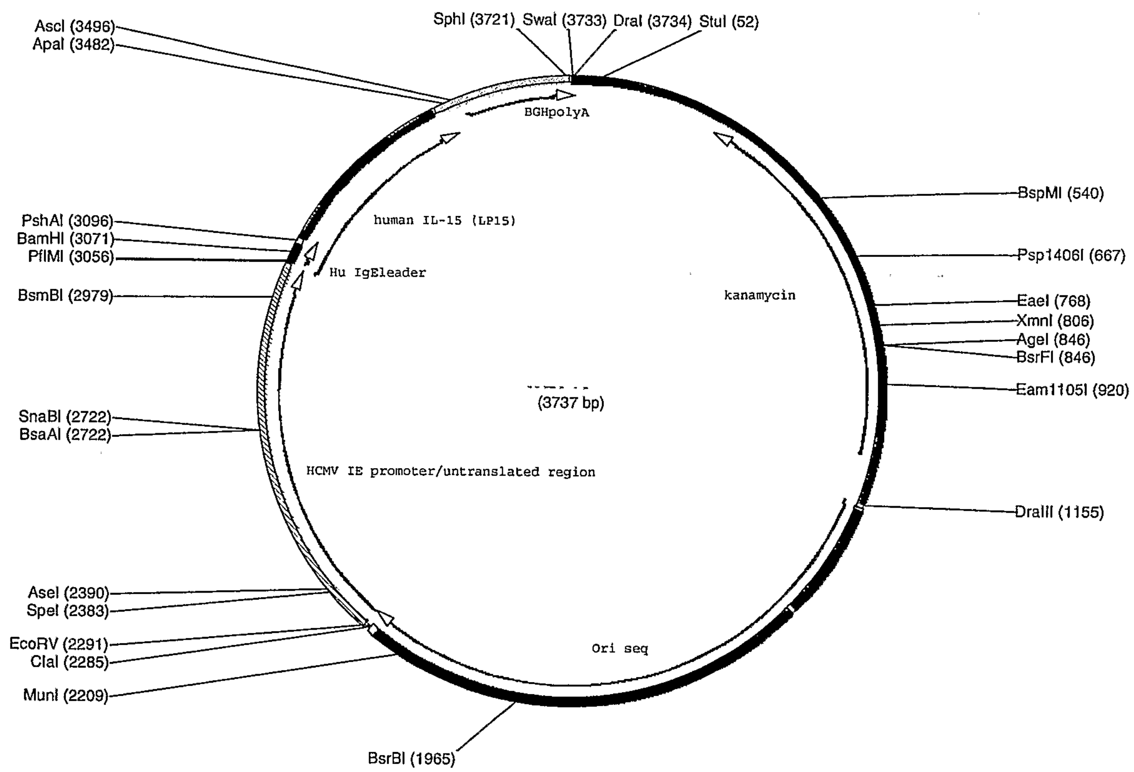
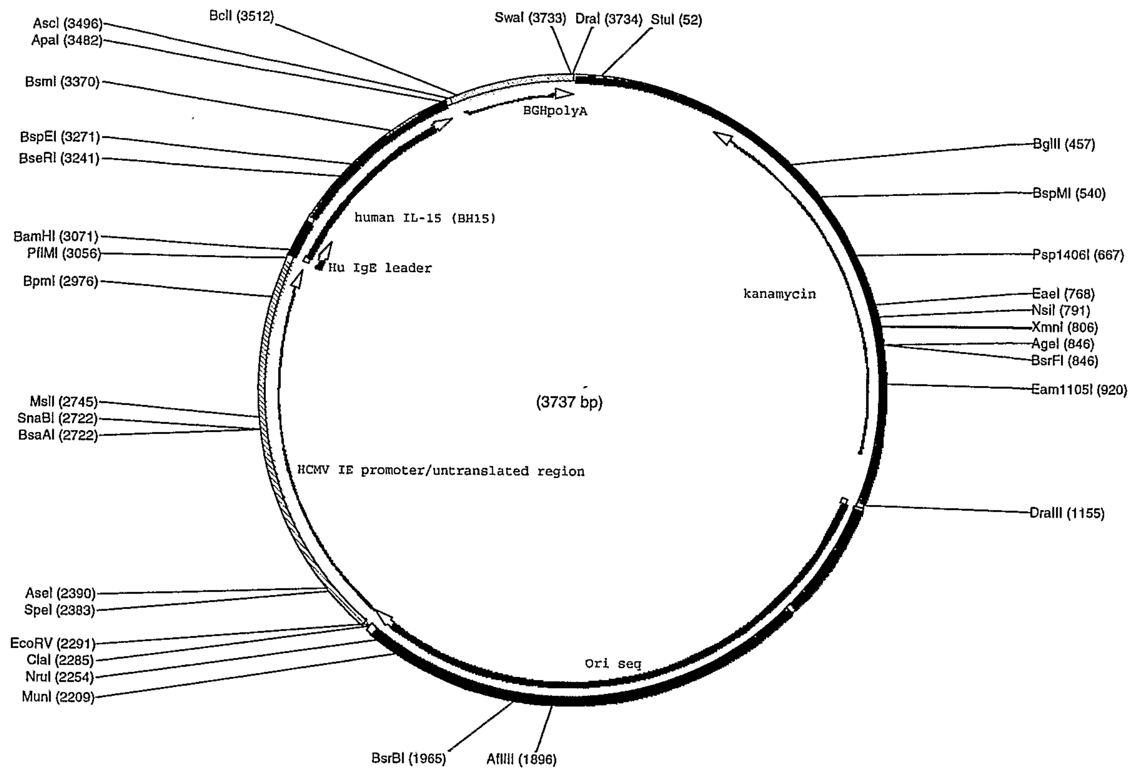


FIGURE 13



INTERNATIONAL SEARCH REPORT

International application No.
PCT/US05/19592

A. CLASSIFICATION OF SUBJECT MATTER
IPC(7) : C12Q 1/68; C12N 1/00, 15/09, 15/63, 15/70, 15/74, 5/00, 5/02; C07H 21/02, 21/04
US CL : 435/6; 435/350; 435/325; 435/320.1; 536/23.1
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 435/6; 435/350; 435/325; 435/320.1; 536/23.1

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Continuation Sheet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	HAAS et al. Codon Usage Limitation in the Expression of HIV-1 Envelope Glycoprotein Current biology 1996 Vol.6 no.3 pp315-324	1-397
A	US 6,414,132 B1 (PAVLAKIS et al.) 02 June 2002 (02.06.2002).	1-397

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search
02 September 2005 (02.09.2005)

Date of mailing of the international search report
23 SEP 2005

Name and mailing address of the ISA/US
Mail Stop PCT, Attn: ISA/US
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450
Facsimile No. (703) 305-3230

Authorized officer
Konstantina Katcheves
Telephone No. (703) 305-3388

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US05/19592

Continuation of B. FIELDS SEARCHED Item 3:
EAST (PGPUB, JPO, EPO, DERWENT, USPAT); STN; STIC
wobble codon, mutant, seq id nos