Title: IMPROVED RECOMBINANT PROTEIN EXPRESSION USING A HYBRID CHEF1 PROMOTER

Abstract: The invention provides expression vectors and host cells for high-level expression of recombinant proteins. The expression vectors comprise Chinese hamster ovary elongation factor 1-α (CHEF1) transcriptional regulatory DNA elements and a cytomegalovirus (CMV) promoter and/or a human adenovirus tripartite leader (AdTPL) sequence. The invention achieves increased protein expression and better productivity of host cells compared to previously described expression systems.
IMPROVED RECOMBINANT PROTEIN EXPRESSION USING A HYBRID CHEF1 PROMOTER

[0001] This application contains, as a separate part of the disclosure, a sequence listing in computer-readable form (Filename: 44744A_SeqListing.txt; Size: 37,528 bytes; Created: March 11, 2014) which is incorporated by reference in its entirety.

CROSS-REFERENCE TO RELATED APPLICATION

[0002] This application claims the benefit of priority of U.S. Provisional Patent Application Serial No. 61/777,603, filed March 12, 2013. The disclosure of the priority application is incorporated herein by reference.

FIELD OF THE INVENTION

[0003] This invention is directed to expression vectors comprising a novel promoter-enhancer combination that increases heterologous protein expression and has practical application in the field of recombinant protein production.

BACKGROUND OF THE INVENTION

[0004] Increasing recombinant protein expression through improvements in transcription, translation, protein folding and/or secretion is a fundamental priority for optimizing yield during cell line development. The Chinese hamster ovary elongation factor 1-α (CHEF1) expression system has been used extensively to create clinical cell lines for producing recombinant proteins. The elongation factor 1-α (EF-1α) gene is highly expressed in most tissue types, and EF-1 is one of the most abundant proteins in human cells (Beck et al., Molecular Systems Biology 7: 549; 2011). CHEF1 expression vectors achieve high-level recombinant protein expression in Chinese hamster ovary (CHO) cells, as well as in non-hamster cells.

[0005] CHEF1 expression is coordinated with growth such that titer increases are driven by increased volumetric productivity. Typically, protein expression initiates early in the exponential phase of growth and drops off during stationary phase and decline. The linkage between protein expression and cell growth is consistent with the regulation of the native EF-1α gene, which is constitutively expressed to function in ribosomal protein complexes. Expression of EF-1α has been documented to increase in transformed (Sanders et al., Nucleic Acids Research 20: 5907; 1992) and mitogen-stimulated cells (Thomas and Thomas, Journal of Cell}
Biology 103: 2137; 1986), consistent with constitutive expression of EF-1α in actively growing cells. In addition to transcriptional control in growing cells, the growth factor insulin regulates the translation of EF-1α through the mRNA 5′ untranslated region (5′UTR) (Hammond and Bowman, *Journal of Biological Chemistry* 25: 17785; 1988; Proud and Denton, *Biochemical Journal* 328: 329; 1997). This translational control is achieved through the Tract of Polypyrimidine (TOP) sequence found in the 5′UTR (Mariottini and Amaldi, *Molecular and Cellular Biology* 10: 816; 1990).

**0006** CHEF1 expression systems have been shown to be capable of achieving higher levels of protein expression than vectors employing other commonly used promoters, such as the cytomegalovirus (CMV), human EF-1α, and Simian virus 40 (SV40) promoters (Running Deer and Allison, *Biotechnology Progress* 20: 880; 2004). The CMV promoter is one of the most widely used promoters for recombinant protein expression. For example, the glutamine synthetase (GS) system uses a murine or human CMV promoter (Kalwy, S., "Towards stronger gene expression – a promoter’s tale," 19th European Society for Animal Cell Technology (ESACT) meeting, 2005). The commercial expression plasmid pcDNA™3 (Life Technologies Corp., Carlsbad, CA) carries a CMV promoter derived from the major immediate-early (IE) gene (GenBank Accession # K03104.1) described previously (Boshart et al., *Cell* 1985; 4:521). Another commonly used CMV promoter is derived from the human CMV strain AD169 (GenBank Accession # X17403.1), also known as human herpesvirus 5.

**0007** Vectors containing CHEF1 regulatory DNA result in improved expression of recombinant proteins that is up to 280-fold greater than from CMV-controlled plasmids (Running Deer and Allison, 2004). Increased expression of a variety of proteins of interest, including secreted and membrane-bound proteins, has been achieved in different eukaryotic cell lines, including non-hamster cells, using CHEF1-driven vectors. Transfection efficiencies between CHEF1 and CMV vectors are comparable, but expression levels in clones transfected with CHEF1 vectors are generally uniformly higher.

**0008** Despite the demonstrated success of CHEF1 vectors in driving high-level expression of recombinant proteins, there exists an ongoing need to develop improved expression systems.

**SUMMARY OF THE INVENTION**
[0009] The disclosure provides an expression vector for high-level expression of recombinant proteins. In various aspects, the expression vector comprises CHEF1 transcriptional regulatory DNA elements and a CMV promoter and/or a human adenovirus tripartite leader (AdTPL) sequence.

[0010] In various aspects, an expression vector according to the disclosure comprises 5' CHEF1 transcriptional regulatory DNA. In various embodiments, the 5' CHEF1 transcriptional regulatory DNA comprises SEQ ID NO: 1. In various embodiments, the 5' CHEF1 transcriptional regulatory DNA comprises DNA located between position 1 and position 11,716 in SEQ ID NO: 1. In various aspects, the 5' CHEF1 transcriptional regulatory DNA comprises DNA located between position 10,744 and 11,716 in SEQ ID NO: 1. In various embodiments, the 5' CHEF1 transcriptional regulatory DNA comprises SEQ ID NO: 2. In various embodiments, the 5' CHEF1 transcriptional regulatory DNA comprises DNA located between position 1 and position 4057 in SEQ ID NO: 2.

[0011] In various aspects, an expression vector according to the disclosure further comprises 3' CHEF1 transcriptional regulatory DNA. In various embodiments, the 3' CHEF1 transcriptional regulatory DNA comprises SEQ ID NO: 3. In various embodiments, the 3' CHEF1 transcriptional regulatory DNA comprises DNA located between position 1 and position 4180 in SEQ ID NO: 3. In various aspects, the 3' CHEF1 transcriptional regulatory DNA comprises DNA located between position 1 and position 209 in SEQ ID NO: 3. In various embodiments, 3' CHEF1 transcriptional regulatory DNA comprises about 4.2 kilobases.

[0012] In various embodiments, an expression vector according to the disclosure comprises CHEF1 transcriptional regulatory DNA and a CMV promoter. In various embodiments, the expression vector comprises CHEF1 transcriptional regulatory DNA and an AdTPL sequence. In various aspects, the expression vector comprises CHEF1 transcriptional regulatory DNA, a CMV promoter, and an AdTPL sequence.

[0013] In various aspects, in an expression vector according to the disclosure, a CMV promoter and/or an AdTPL sequence is inserted into 5' CHEF1 transcriptional regulatory DNA. In various embodiments, in an expression vector comprising DNA set out in SEQ ID NO: 1, one or more bases between position 1 and position 11,716 in SEQ ID NO: 1 is/are deleted and replaced with a CMV promoter and/or an AdTPL sequence. In various aspects, one or more
bases between position 10,512 and position 11,716 in SEQ ID NO: 1 is/are deleted and replaced with a CMV promoter and/or AdTPL sequence. In various aspects, an expression vector according to the disclosure comprises one or more of the polynucleotides set forth in SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6 and SEQ ID NO: 7.

[0014] In various embodiments, an expression vector according to the disclosure further comprises a selectable marker gene. In various embodiments, an expression vector according to the disclosure further comprises a polynucleotide encoding a protein of interest that is operably linked to the 5' CHEF1 transcriptional regulatory DNA, the 3' CHEF1 transcriptional regulatory DNA, the CMV promoter, and/or the AdTPL sequence.

[0015] The disclosure also provides host cells transformed, transduced, or transfected with an expression vector comprising CHEF1 transcriptional regulatory DNA and a CMV promoter and/or an AdTPL sequence. In various aspects, the host cell is a prokaryotic or eukaryotic cell. In various aspects, the host cell is a hamster cell, and in various embodiments, the host cell is a Chinese hamster ovary (CHO) cell. In various aspects, the host cell is a non-hamster mammalian cell. In various embodiments, the host cell is a human cell. The expression vector of the disclosure comprising CHEF1 transcriptional regulatory DNA in combination with a CMV promoter and/or an AdTPL sequence achieves a synergistic increase in the protein expression capacity of the host cells.

BRIEF DESCRIPTION OF THE DRAWINGS

[0016] Figure 1 shows a map of the expression vector pDEF85 comprising 5' and 3' CHEF1 transcriptional regulatory DNA and a CMV promoter. The CMV promoter replaces 1217 nucleotides (from position 2866 to position 4083) of the 5' CHEF1 DNA in vector pDEF38 to create pDEF85. GP1 and MAb1 reporter genes were cloned into the XhoI – Xbal cloning sites to make the expression vectors pDEF85-GP1 and pDEF85-MAb1.

[0017] Figure 2 shows a map of the expression vector pDEF86 comprising 5' and 3' CHEF1 transcriptional regulatory DNA, a CMV promoter, and an AdTPL sequence. The CMV promoter and AdTPL sequence replace 1217 nucleotides (from position 2866 to position 4083) of the 5' CHEF1 DNA in vector pDEF38 to create pDEF86. GP1 and MAb1 reporter genes were cloned into the XhoI – Xbal cloning sites to make the expression vectors pDEF86-GP1 and pDEF86-MAb1.
Figure 3 shows viability of CHO host cells transfected with the vector pDEF38-GP1, pDEF85-GP1, or pDEF86-GP1. The cells recovered for 2-3 days in non-select CD-CIM1 plus HT media and were resuspended in selection media lacking HT (Passage 0). Cells were passaged every 2 to 3 days as cell number permitted. Passage number is provided on the x-axis, and percent cell viability is shown on the y-axis.

Figure 4 shows the viability and productivity of CHO host cells transfected with the vector pDEF38-GP1, pDEF85-GP1, or pDEF86-GP1. Replicate transfection pools were run in 12-day fed-batch production models and fed Feed C on Days 3, 5, and 7 and CB on Days 0, 3, 5, and 7 in CD-CIM1 base media. Productions were run at 37 °C and shifted to 34 °C on Day 3. Viable cell density, percent viability, and productivity were measured on Days 3, 5, 7, 10, and 12. Figure 4A shows viable cell density, with days shown on the x-axis and viable cell density, measured in 10e5 cells per milliliter, depicted on the y-axis. Figure 4B shows percent viability, with days shown on the x-axis and percent cell viability depicted on the y-axis. Figure 4C shows productivity, with days shown on the x-axis and protein titer, in micrograms per milliliter, depicted on the y-axis.

Figure 5 shows the specific productivity of CHO host cells transfected with the vector pDEF38-GP1, pDEF85-GP1, or pDEF86-GP1. The integrated cell area (ICA), measured in million cells per milliliter multiplied by the day, is shown on the x-axis and the protein titer, measured in micrograms per milliliter, is depicted on the y-axis. The specific productivity values were calculated as picograms of protein per cell per day averaged over the entire cell culture duration.

Figure 6 shows the productivity of CHO host cells transfected with the vector pDEF38-GP1, pDEF85-GP1, or pDEF86-GP1 grown in BF1-supplemented media. Replicate transfection pools were run in 12-day fed-batch production models and fed BF1 in CD-CIM1 base media plus CB on Days 4, 6, 8, 10, and 12. Productions were run at 37 °C and shifted to 34 °C on Day 3. Titer samples were measured on Days 5, 7, 10 and 12. The days are shown on the x-axis and the protein titer, measured in micrograms per milliliter, is depicted on the y-axis.

Figure 7 shows the growth and productivity of CHO host cells transfected with the vector pDEF38-MAb1, pDEF85-MAb1, or pDEF86-MAb1 grown in BF1-supplemented media. Replicate transfection pools were run in 12-day fed-batch production models and fed BF1 in CD-
CIM1 base media plus CB on Days 4, 6, 8, 10, and 12. Productions were run at 37 °C and shifted to 32.5 °C on Day 5. Antibody titer samples were measured on Days 7, 10, 12 and 14. Figure 7A shows growth, with days shown on the x-axis and viable cell density, measured in 10e5 cells per milliliter, depicted on the y-axis. Figure 7B shows productivity, with days shown on the x-axis and protein titer, measured in micrograms per milliliter, depicted on the y-axis.

[0023] Figure 8 shows the growth and productivity of CHO host cells transfected with the vector pDEF38-MAb1 or pDEF85-MAb1. Twelve randomly selected clones expressing MAb1 after transfection with pDEF38 or pDEF85 were run in 12-day fed-batch production models and fed BF1 in CD-CIM1 base media plus CB on Days 4, 6, 8, 10, and 12. Productions were run at 37 °C and shifted to 32.5 °C on Day 5. Viable cell density and antibody titer were measured on Days 4, 6, 11 and 13. Figure 8A shows the growth of each clone, with days shown on the x-axis and viable cell density, measured in 10e5 cells per milliliter, depicted on the y-axis. Figure 8B shows the productivity of each clone, with days shown on the x-axis and the antibody titer, measured in micrograms per milliliter, depicted on the y-axis.

[0024] Figure 9 shows the productivity of CHO host cells transfected with the vector pDEF38-GP1 or pDEF85-GP1. Eight clones expressing GP1 after transfection with pDEF38-GP1 or pDEF85-GP1 were selected using flow cytometry and run in 12-day fed-batch production model. The clones were fed BF1 in CD-CIM1 base media plus CB on Days 4, 6, 8, 10, and 12. Productions were run at 37 °C and shifted to 32.5 °C on Day 3. Titers were measured on Days 5, 7, 10, 12 and 14. Figure 9A shows the productivity of clones transfected with the CHEF1 vector pDEF38-GP1, with days shown on the x-axis and the protein titer, measured in micrograms per milliliter, depicted on the y-axis. Figure 9B shows the productivity of clones transfected with the CHEF1-CMV vector pDEF85-GP1, with days shown on the x-axis and the protein titer, measured in micrograms per milliliter, depicted on the y-axis.

DETAILED DESCRIPTION OF THE INVENTION

[0025] The present disclosure provides an expression vector comprising a combination of regulatory DNA elements for achieving high-level protein expression and improved productivity compared to vectors known in the field. The disclosure also provides a host cell transformed with an expression vector described herein. The expression vector of the disclosure comprises CHEF1 transcriptional regulatory DNA combined with a CMV promoter and/or an AdTPL
sequence. The use of CHEF1 transcriptional regulatory DNA elements in an expression vector to achieve high-level expression of recombinant proteins has been previously described (U.S. Patent No. 5,888,809; Running Deer and Allison, 2004). Protein expression from CHEF1-driven vectors has been shown to be significantly higher than from CMV promoter-controlled vectors for a number of different protein and host cell types, and the increase can be greater than 250-fold (Running Deer and Allison, 2004). The AdTPL sequence is a 200-nucleotide 5’ noncoding sequence found on late viral mRNAs that enhances their translation (Logan, PNAS 81: 3655; 1984).

[0026] Considering the improved protein expression obtained using CHEF1-controlled vectors, the addition of non-CHEF1 control regions, such as a CMV promoter or an AdTPL sequence, to a CHEF1 expression vector would be counterintuitive. The presence of such non-CHEF1 control regions could disrupt the cooperative action of individual CHEF1 transcriptional regulatory elements and would not be expected to act in concert with the CHEF1 regulatory DNA to yield improved protein expression. However, the expression vector of the present disclosure, which comprises CHEF1 transcriptional regulatory DNA and a CMV promoter and/or an AdTPL sequence, surprisingly yields increased protein expression compared to vectors comprising only CHEF1 control regions.

[0027] As used herein, the following definitions may be useful in aiding the skilled practitioner in understanding the disclosure:

[0028] The term "expression vector" refers to any molecule used to transfer coding information to a host cell. In various aspects, the expression vector is a nucleic acid, a plasmid, a cosmid, a virus, or an artificial chromosome.

[0029] The term "host cell" refers to a cell that has been transformed, transfected, or transduced by an expression vector bearing a gene of interest, which is then expressed by the cell. A host cell is, in various aspects, a prokaryotic or eukaryotic cell. In various aspects, the host cell is a bacteria cell, a protist cell, a fungal cell, a plant cell, or an animal cell. The term also refers to progeny of the parent host cell, regardless of whether the progeny is identical in genotype or phenotype to the parent, as long as the gene of interest is present.

[0030] The term "CMV promoter" refers to CMV promoter sequences known in the art. In various aspects, the CMV promoter is of any origin, including murine (mCMV) or human
(hCMV). In various aspects, a hCMV is derived from any CMV strain. In various aspects, the CMV strain is AD169, Davis, Toledo, or Towne. In various embodiments of the disclosure, the CMV promoter contains the polynucleotide set forth in SEQ ID NO: 4.

[0031] The term "AdTPL sequence" refers to the approximately 200 nucleotide, 5' noncoding sequence present in human adenovirus late viral mRNAs that is known in the art. In various embodiments, the AdTPL sequence contains the polynucleotide set forth in SEQ ID NO: 5.

[0032] The term "CHEF1 transcriptional regulatory DNA" refers to noncoding sequences containing cis-acting regulatory elements capable of controlling transcription of the CHEF1 gene, such as the promoter region and elements such as enhancers, insulators, and scaffold/matrix attachment regions.

[0033] The term "5' CHEF1 transcriptional regulatory DNA" refers to DNA, when in nature, located 5', i.e., upstream, of the start codon in the CHEF1 gene in the Chinese hamster genome.

[0034] The term "3' CHEF1 transcriptional regulatory DNA" refers to DNA, when in nature, located 3', i.e., downstream, of the stop codon in the CHEF1 gene in the Chinese hamster genome.

[0035] The terms “approximately” and “about” refer to quantities that are within close range of a reference amount. With respect to polynucleotides, a sequence that is approximately/about a specified length is within 5% of the recited length.

[0036] In various aspects, an expression vector according to the disclosure comprises CHEF1 transcriptional regulatory DNA and a CMV promoter and/or an AdTPL sequence. In various aspects, the CHEF1 transcriptional regulatory DNA comprises 5' CHEF1 transcriptional regulatory DNA and/or 3' CHEF1 transcriptional regulatory DNA.

[0037] In various embodiments, the 5' CHEF1 transcriptional regulatory DNA comprises the polynucleotide set forth in SEQ ID NO: 1. The disclosure also provides 5' CHEF1 transcriptional regulatory DNA that is at least 99%, at least 98%, at least 97%, at least 96%, at least 95%, at least 94%, at least 93%, at least 92%, at least 91%, at least 90%, at least 85%, at least 80%, at least 75% or at least 70% identical to the polynucleotide set out in SEQ ID NO: 1. In various embodiments, the 5' CHEF1 transcriptional regulatory DNA comprises DNA located between position 1 and position 11,716 of SEQ ID NO: 1, i.e., a portion of SEQ ID NO: 1. The
disclosure also provides 5’ CHEF1 transcriptional regulatory DNA that is at least 99%, at least 98%, at least 97%, at least 96%, at least 95%, at least 94%, at least 93%, at least 92%, at least 91%, at least 90%, at least 85%, at least 80%, at least 75% or at least 70% identical to DNA located between position 1 and position 11,716 in SEQ ID NO: 1. In various aspects, the 5’ CHEF1 transcriptional regulatory DNA comprises DNA located between position 10,744 and position 11,716 in SEQ ID NO: 1. The disclosure also provides 5’ CHEF1 transcriptional regulatory DNA that is at least 99%, at least 98%, at least 97%, at least 96%, at least 95%, at least 94%, at least 93%, at least 92%, at least 91%, at least 90%, at least 85%, at least 80%, at least 75% or at least 70% identical to DNA located between position 10,744 and position 11,716 in SEQ ID NO: 1. In various embodiments, the 5’ CHEF1 transcriptional regulatory DNA comprises the polynucleotide set forth in SEQ ID NO: 2. The disclosure also provides 5’ CHEF1 transcriptional regulatory DNA that is at least 99%, at least 98%, at least 97%, at least 96%, at least 95%, at least 94%, at least 93%, at least 92%, at least 91%, at least 90%, at least 85%, at least 80%, at least 75% or at least 70% identical to the polynucleotide set out in SEQ ID NO: 2. In various embodiments, the 5’ CHEF1 transcriptional regulatory DNA comprises DNA located between position 1 and position 4057 of SEQ ID NO: 2, i.e., a portion of SEQ ID NO: 2. The disclosure also provides 5’ CHEF1 transcriptional regulatory DNA that is at least 99%, at least 98%, at least 97%, at least 96%, at least 95%, at least 94%, at least 93%, at least 92%, at least 91%, at least 90%, at least 85%, at least 80%, at least 75% or at least 70% identical to DNA located between position 1 and position 4057 of SEQ ID NO: 2.

[0038] In various aspects, the expression vector according to the disclosure further comprises 3’ CHEF1 transcriptional regulatory DNA. In various embodiments, the 3’ CHEF1 transcriptional regulatory DNA comprises the polynucleotide set forth in SEQ ID NO: 3. The disclosure also provides 3’ CHEF1 transcriptional regulatory DNA that is at least 99%, at least 98%, at least 97%, at least 96%, at least 95%, at least 94%, at least 93%, at least 92%, at least 91%, at least 90%, at least 85%, at least 80%, at least 75% or at least 70% identical to the polynucleotide set out in SEQ ID NO: 3. In various embodiments, the 3’ CHEF1 transcriptional regulatory DNA comprises DNA located between position 1 and position 4180 in SEQ ID NO: 3, i.e., a portion of SEQ ID NO: 3. The disclosure also provides 3’ CHEF1 transcriptional regulatory DNA that is at least 99%, at least 98%, at least 97%, at least 96%, at least 95%, at least 94%, at least 93%, at least 92%, at least 91%, at least 90%, at least 85%, at least 80%, at
least 75% or at least 70% identical to DNA located between position 1 and position 4180 in SEQ ID NO: 3, i.e., a portion of SEQ ID NO: 3. In various aspects, the 3' CHEF1 transcriptional regulatory DNA comprises DNA located between position 1 and position 209 in SEQ ID NO: 3. The disclosure also provides 3' CHEF transcriptional regulatory DNA that is at least 99%, at least 98%, at least 97%, at least 96%, at least 95%, at least 94%, at least 93%, at least 92%, at least 91%, at least 90%, at least 85%, at least 80%, at least 75% or at least 70% identical to DNA located between position 1 and position 209 in SEQ ID NO: 3. In various embodiments, the 3' CHEF1 transcriptional regulatory DNA may comprise about 4.2 kilobases.

[0039] In various embodiments, the expression vector according to the disclosure comprises CHEF1 transcriptional regulatory DNA elements and a CMV promoter. In various aspects, the CMV promoter comprises the polynucleotide set forth in SEQ ID NO: 4. The disclosure also provides a CMV promoter that is at least 99%, at least 98%, at least 97%, at least 96%, at least 95%, at least 94%, at least 93%, at least 92%, at least 91%, at least 90%, at least 85%, at least 80%, at least 75% or at least 70% identical to the polynucleotide set forth in SEQ ID NO: 4. In various aspects, the expression vector comprising 5' CHEF1 transcriptional regulatory DNA and a CMV promoter comprises the polynucleotide set forth in SEQ ID NO: 6. The disclosure also provides 5' CHEF1 transcriptional regulatory DNA and a CMV promoter that is at least 99%, at least 98%, at least 97%, at least 96%, at least 95%, at least 94%, at least 93%, at least 92%, at least 91%, at least 90%, at least 85%, at least 80%, at least 75% or at least 70% identical to the polynucleotide set forth in SEQ ID NO: 6. In various embodiments, the expression vector comprises CHEF1 transcriptional regulatory DNA and an AdTPL sequence. In various aspects, the AdTPL sequence comprises the polynucleotide set forth in SEQ ID NO: 5. The disclosure also provides an AdTPL sequence that is at least 99%, at least 98%, at least 97%, at least 96%, at least 95%, at least 94%, at least 93%, at least 92%, at least 91%, at least 90%, at least 85%, at least 80%, at least 75% or at least 70% identical to the polynucleotide set forth in SEQ ID NO: 5. In various embodiments, the expression vector comprises CHEF1 transcriptional regulatory DNA, a CMV promoter and an AdTPL sequence. In various aspects, the expression vector comprising 5' CHEF1 transcriptional regulatory DNA, a CMV promoter, and an AdTPL sequence comprises the polynucleotide set forth in SEQ ID NO: 7. The disclosure also provides 5' CHEF1 transcriptional regulatory DNA, a CMV promoter, and an AdTPL sequence that is at least 99%, at least 98%, at least 97%, at least 96%, at least 95%, at least 94%, at least 93%, at least 92%, at least 91%, at least 90%, at least 85%, at least 80%, at least 75% or at least 70% identical to the polynucleotide set forth in SEQ ID NO: 7. The disclosure also provides 5' CHEF1 transcriptional regulatory DNA, a CMV promoter, and an AdTPL sequence that is at least 99%, at least 98%, at least 97%, at least 96%, at least 95%, at least 94%, at least 93%, at least 92%, at least 91%, at least 90%, at least 85%, at least 80%, at least 75% or at least 70% identical to the polynucleotide set forth in SEQ ID NO: 7.
least 92%, at least 91%, at least 90%, at least 85%, at least 80%, at least 75% or at least 70% identical to the polynucleotide set forth in SEQ ID NO: 7.

[0040] In various embodiments, a CMV promoter and/or an AdTPL sequence is inserted into the 5’ CHEF1 transcriptional regulatory DNA in an expression plasmid according to the disclosure. In various embodiments, in an expression vector comprising DNA set out in SEQ ID NO: 1, one or more bases between position 1 and position 11,716 of SEQ ID NO: 1 is/are deleted and replaced with a CMV promoter and/or AdTPL sequence. In various embodiments, the proximal 5’ CHEF1 promoter region is replaced with a CMV promoter and/or an AdTPL sequence. For example and without limitation, in various aspects, one or more bases between position 10,512 and position 11,716 of SEQ ID NO: 1 is/are deleted and replaced with a CMV promoter, an AdTPL sequence, or a CMV promoter and an AdTPL sequence.

[0041] The expression vector according to the disclosure further comprises a polynucleotide encoding a protein of interest. In various aspects, the polynucleotide is operably linked to the 5’ CHEF1 transcriptional regulatory DNA, the 3’ CHEF1 transcriptional regulatory DNA, the CMV promoter, and/or the AdTPL sequence. The expression vector is useful for any protein and is expected to provide higher protein expression than CHEF1 or CMV alone. In various aspects, the expression vector further comprises a selectable marker gene for identification of transformed cells. Examples of suitable selectable marker genes include, but are not limited to, neomycin phosphotransferase (npt II), hygromycin phosphotransferase (hpt), dihydrofolate reductase (dhfr), zeocin, phleomycin, bleomycin resistance gene (ble), gentamycin acetyltransferase, streptomycin phosphotransferase, mutant form of acetolactate synthase (als), bromoxynil nitrilase, phosphinothricin acetyl transferase (bar), enolpyruvylshikimate-3-phosphate (EPSP) synthase (aro A), muscle specific tyrosine kinase receptor molecule (MuSK-R), copper-zinc superoxide dismutase (sod1), metallothioneins (cup1, MT1), beta-lactamase (BLA), puromycin N-acetyl-transferase (pac), blasticidin acetyl transferase (bsl), blasticidin deaminase (bsr), histidinol dehydrogenase (HDH), N-succinyl-5-amiomimidazole-4-carboxamide ribotide (SAICAR) synthetase (ade1), argininosuccinate lyase (arg4), beta-isopropylmalate dehydrogenase (leu2), invertase (suc2), orotidine-5’-phosphate (OMP) decarboxylase (ura3), and orthologs of any of the foregoing.
The disclosure also provides host cells transformed, transduced, or transfected with an expression vector comprising CHEF1 transcriptional regulatory DNA and a CMV promoter and/or an AdTPL sequence. In various aspects, the host cell is a prokaryotic or eukaryotic cell. In various aspects, the host cell is a hamster cell. In various aspects, the host cell is a CHO cell. In various embodiments, the host cell is a non-hamster mammalian cell, and in various aspects, the cell is a human cell.

Embodiments contemplated in view of the foregoing description include, but are not limited to, the following numbered embodiments:

1. An expression vector comprising Chinese Hamster Elongation Factor-1α (CHEF1) transcriptional regulatory DNA and a cytomegalovirus (CMV) promoter and/or an adenovirus tripartite leader (AdTPL) sequence.

2. The expression vector of embodiment 1, wherein the CHEF1 transcriptional regulatory DNA comprises 5' CHEF transcriptional regulatory DNA.

3. The expression vector of embodiment 2, wherein the 5' CHEF1 transcriptional regulatory DNA comprises Sequence ID NO: 1 or a polynucleotide at least 95% identical to Sequence ID NO: 1.

4. The expression vector of embodiment 2, wherein the 5' CHEF1 transcriptional regulatory DNA comprises DNA located between position 1 and position 11,716 in Sequence ID NO: 1 or a polynucleotide at least 95% identical to DNA located between position 1 and position 11,716 in Sequence ID NO: 1.

5. The expression vector of embodiment 4, wherein the 5' CHEF1 transcriptional regulatory DNA comprises DNA located between position 10,774 and position 11,716 in Sequence ID NO: 1 or a polynucleotide at least 95% identical to DNA located between position 10,774 and position 11,716 in Sequence ID NO: 1.

6. The expression vector of embodiment 2, wherein the 5' CHEF1 transcriptional regulatory DNA comprises Sequence ID NO: 2 or a polynucleotide at least 95% identical to Sequence ID NO: 2.

7. The expression vector of embodiment 2, wherein the 5' CHEF1 transcriptional regulatory DNA comprises DNA located between position 1 and position 4057 in Sequence ID NO: 2.
NO: 2 or a polynucleotide at least 95% identical to DNA located between position 1 and position 4057 in Sequence ID NO: 2.

[0051] 8. The expression vector of any one of the preceding embodiments, further comprising 3' CHEF1 transcriptional regulatory DNA.

[0052] 9. The expression vector of embodiment 8, wherein the 3' CHEF1 transcriptional regulatory DNA comprises Sequence ID NO: 3 or a polynucleotide at least 95% identical to Sequence ID NO: 3.

[0053] 10. The expression vector of embodiment 8, wherein the 3' CHEF1 transcriptional regulatory DNA comprises DNA located between position 1 and position 4180 in Sequence ID NO: 3 or a polynucleotide at least 95% identical to DNA located between position 1 and position 4180 in Sequence ID NO: 2.

[0054] 11. The expression vector of embodiment 10, wherein the 3' CHEF1 transcriptional regulatory DNA comprises DNA located between position 1 to position 209 in Sequence ID NO: 3 or a polynucleotide at least 95% identical to DNA located between position 1 to position 209 in Sequence ID NO: 3.

[0055] 12. The expression vector of any one of embodiments 8-11, wherein the 3' CHEF1 transcriptional regulatory DNA comprises about 4.2 kilobases.

[0056] 13. The expression vector of any one of the preceding embodiments comprising a CMV promoter.

[0057] 14. The expression vector of any one of the preceding embodiments comprising an AdTPL sequence.

[0058] 15. The expression vector of any one of the preceding embodiments comprising a CMV promoter and an AdTPL sequence.

[0059] 16. The expression vector of embodiment 2, wherein the 5' CHEF1 transcriptional regulatory DNA comprises DNA set out in Sequence ID NO: 1, wherein one or more bases between position 1 and position 11,716 of Sequence ID NO: 1 is/are deleted and replaced with a CMV promoter and/or an AdTPL sequence.
17. The expression vector of embodiment 16, wherein one or more bases between position 10,512 and position 11,716 of Sequence ID NO: 1 is/are deleted and replaced with a CMV promoter and/or an AdTPL sequence.

18. The expression vector of embodiment 17, wherein one or more bases between position 10,512 and position 11,716 of Sequence ID NO: 1 is/are deleted and replaced with a CMV promoter.

19. The expression vector of embodiment 17, wherein one or more bases between position 10,512 and position 11,716 of Sequence ID NO: 1 is/are deleted and replaced with an AdTPL sequence.

20. The expression vector of embodiment 17, wherein one or more bases between position 10,512 and position 11,716 of Sequence ID NO: 1 is/are deleted and replaced with a CMV promoter and an AdTPL sequence.

21. The expression vector of any one of the preceding embodiments comprising Sequence ID NO: 4 or a polynucleotide at least 95% identical to Sequence ID NO: 4.

22. The expression vector of any one of the preceding embodiments comprising Sequence ID NO: 5 or a polynucleotide at least 95% identical to Sequence ID NO: 5.

23. The expression vector of any one of the preceding embodiments comprising Sequence ID NO: 6 or a polynucleotide at least 95% identical to Sequence ID NO: 6.

24. The expression vector of any one of the preceding embodiments comprising Sequence ID NO: 7 or a polynucleotide at least 95% identical to Sequence ID NO: 7.

25. The expression vector of any one of the preceding embodiments, further comprising a selectable marker gene.

26. The expression vector of any one of the preceding embodiments, further comprising a polynucleotide encoding a protein of interest operably linked to the 5’ CHEF1 transcriptional regulatory DNA, the 3’ CHEF1 transcriptional regulatory DNA, the CMV promoter and/or the AdTPL sequence.

27. A host cell transformed, transduced or transfected with an expression vector according to any one of the preceding embodiments.
28. The host cell of embodiment 27, wherein the host cell is a prokaryotic cell.

29. The host cell of embodiment 27, wherein the host cell is a eukaryotic cell.

30. The host cell of embodiment 29, wherein the host cell is a hamster cell.

31. The host cell of embodiment 30, wherein the host cell is a Chinese Hamster Ovary (CHO) cell.

32. The host cell of embodiment 29, wherein the host cell is a non-hamster mammalian cell.

33. The host cell of embodiment 32, wherein the host cell is a human cell.

An expression plasmid according to the disclosure is further described in the following Example. The Example serves only to illustrate the invention and is not intended to limit the scope of the invention in any way.

EXAMPLE

Gene Sequence and Expression Vectors - DNA fragments encoding the CMV promoter (SEQ ID NO: 4) and CMV-AdTPL promoter (SEQ ID NO: 5) were chemically synthesized and cloned into pDEF38, a CHEF1 expression vector previously described (Running Deer and Allison, 2004), creating the CHEF1-CMV-promoter vector designated pDEF85 (Figure 1) and the CHEF1-CMV-AdTPL promoter vector designated pDEF86 (Figure 2). Derivative vectors expressing a Fc-glycoprotein fusion (GP1) and an IgG1 antibody (MAb1) were created using standard molecular biology techniques (Maniatis et al., J. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory. 545, 1982) and designated pDEF38-GP1, pDEF85-GP1, pDEF86-GP1, pDEF38-MAb1, pDEF85-MAb1 and pDEF86-MAb1.

Cell Line Construction - The pDEF38-GP1, pDEF85-GP1, pDEF86-GP1, pDEF38-MAb1, pDEF85-MAb1 and pDEF86-MAb1 expression vectors were transfected individually into CHO DG44 cells by standard electroporation methods, grown for two days in non-select media containing hypoxanthine and thymidine (HT), and then selected for about two weeks in media lacking HT. The selected cell populations, or transfection pools, were expanded and split into production model cultures to assess productivity and also simultaneously split into cultures for single cell cloning.
[0080] **Production Models** - Small-scale fed-batch production models were run to assess culture productivity (titer) following standard biologics manufacturing processes. Cultures were inoculated at seed densities of 0.5 million cells per milliliter in shake flasks of chemically defined media (CD-CIM1, CMC Biologics, Bothell, WA) lacking HT. The cultures were run for 3 to 5 days at 37 °C and then shifted to lower temperatures (30 °C to 34 °C) to slow growth and promote production. Cultures were fed the supplements Balanced Feed 1 (BF1, CMC Biologies), Efficient Feed C (Feed C, Life Technologies, Grand Island, NY) or Cell Boost (CB, Thermo Fisher Scientific, Waltham, MA) to prolong culture health. Samples for titer and cell densities were collected on Days 3, 5, 7, 10, 12, 14 and 16. The study was concluded by Day 12 to 16. Harvest supernatants were filtered through 0.2 micrometer filters and then assayed for GP1 or MAb1 production by Protein A high performance liquid chromatography (HPLC).

[0081] **Cell Line Cloning** - Selected GP1- and MAb1-expressing transfection pools were diluted to seed single cells into individual wells of 96-well plates. The plates were imaged from inoculation out to two weeks to identify monoclonal cell lines originating from single cells. Wells containing monoclonal colonies were expanded and either randomly chosen or selected using flow cytometry to identify highly-expressing cells from each transfection pool. Cell lines were expanded to grow in suspension culture and split into production model cultures to assess productivity.

[0082] **Flow Cytometry** - Fluorescence activated cell sorting (FACS) analysis was performed with Day 2 normal growing cells that were harvested and stained with fluorescent anti-IgG1 Fc antibody (RPE) to detect recombinant GP1 and MAb1 expression.

Results and Discussion

[0083] Stable cell lines expressing the reporter protein GP1 or MAb1 were made using the expression vectors pDEF38, pDEF85 and pDEF86 using standard DHFR selection methods. Transfection pools were selected in media lacking hypoxanthine and thymidine (HT) without using methotrexate. Cell viabilities dropped in media lacking HT and then recovered as cells with DHFR vectors grew out in the population. The transfection cultures dropped initially to about 10% to about 30% viability and then attained greater than 90% viability by around Day 12 (Passage 6). The growth of cells transfected with a CHEF1-CMV vector (pDEF85-GP1) or CHEF1-CMV-AdTPL vector (pDEF86-GP1) compared to cells transfected with a CHEF1 only
vector (pDEF38-GP1) showed similar recovery for the GP1 expressing constructs, with consistent high viability growth after the recovery period (Figure 3). Similar results were obtained for antibody expressing cell lines (data not shown).

[0084] Transfection pools were placed directly into production models or advanced into single cell cloning and then clonal cell lines were compared in production models. Figure 4 shows that fed-batch shake flask production model growth was comparable for the GP1 expressing pools (Figure 4A); however, the protein expression (titer) at harvest, typically 12 to 16 days from inoculation, was significantly different. The Day 12 harvest titers for the CHEF1-CMV or CHEFl-CMV-AdTPL expression vectors (pDEF85 and pDEF86) were much higher than for the standard CHEF1 vector (pDEF38) (Figure 4C). The amount of recombinant GP1 protein produced from pooled transfectants in fed-batch shake flasks from the CHEF1-CMV or CHEFl-CMV-AdTPL vectors was about twice the standard CHEF1 vector. Growth of the CHEF1-CMV and CHEF1-CMV-AdTPL pools peaked slightly earlier and showed a more rapid decline in viability (Figure 4B). The viability drop was not anticipated to improve expression and could instead be detrimental. Later experiments showed that improving ending viability increased titer for the CHEF1-CMV cultures.

[0085] The increase in titer seen with the CHEF1-CMV and CHEF1-CMV-AdTPL vectors was the result of increased specific productivity, as seen in Figure 5. Specific productivity was calculated as picograms of protein per cell per day averaged over the entire culture duration. There was a slight expression difference between the CHEF1-CMV (pDEF85) and CHEF1-CMV-AdTPL (pDEF86) constructs, indicating a possible benefit of the addition of the AdTPL sequence with respect to specific productivity. The specific productivities in picograms per cell per day (PCD) are shown in Table 1. The specific productivity achieved using the CHEF1-CMV or CHEF1-CMV-AdTPL vectors was more than two-fold greater than the specific productivity of the CHEF1 vector.

<table>
<thead>
<tr>
<th>Vector</th>
<th>Promoter</th>
<th>PCD</th>
</tr>
</thead>
<tbody>
<tr>
<td>pDEF38</td>
<td>CHEF1</td>
<td>2.4</td>
</tr>
<tr>
<td>pDEF85</td>
<td>CHEF1-CMV</td>
<td>6.7</td>
</tr>
<tr>
<td>pDEF86</td>
<td>CHEF1-CMV-AdTPL</td>
<td>7.4</td>
</tr>
</tbody>
</table>
Improved expression with the CHEF1-CMV and CHEF1-CMV-AdTPL vectors was confirmed using different reporter proteins and varied fed-batch production conditions. Cells grown in CD-CIM1 base media supplemented with CB that were fed proprietary BF1 supplement on Days 4, 6, 8, 10 and 12 demonstrated similar GP1 productivity profiles compared to cells supplemented with commercial media (Feed C). Figure 6 shows the CHEF1-CMV and CHEF1-CMV-AdTPL vector cultures had increased GP1 titers over the control CHEF1 vector in the BF1 process.

Antibody (MAb1) production utilizing the CHEF1-CMV and CHEF1-CMV-AdTPL vectors was also tested in the BF1 process. The MAb1 transfection pools were created with the same methodology as the GP1 pools and once fully recovered from selection, were put into fed-batch shake flask production models. As shown in Figure 7, the CHEF1-CMV and CHEF1-CMV-AdTPL MAb1 pools produced higher titer antibody than the pDEF38 controls after Day 12. The productivity profile in transfection pools expressing antibody was novel compared to glycoprotein production because the antibody productivity increased dramatically as cells entered stationary phase. A lower initial titer (Day 7) for the CHEF1-CMV and CHEF1-CMV-AdTPL MAb1 pools compared to the CHEF1 MAb1 pool was seen as the cell cultures were actively growing, followed by rapid increases in productivity for the CHEF1-CMV and CHEF1-CMV-AdTPL cultures as the growth slowed down and finally declined after Day 10 (Figure 7B). Even as the viable cell density decreased (Figure 7A), productivity increased out to Day 14 in the CHEF1-CMV and CHEF1-CMV-AdTPL pools, whereas it started to slow down in the CHEF1 pool, although the terminal Day 14 percent viabilities were similar for all cultures (about 80% viable, data not shown).

Clonal cultures were developed from transfection pools expressing both GP1 and MAb1. Monoclonal cell lines were identified by imaging of limiting dilution plates and then expanded into suspension culture. Twelve MAb1 clonal cultures were selected randomly from each of the pDEF38-MAb1 and pDEF85-MAb1 transfection pools and were run in fed-batch shake flask production models. Clonal CHEF1-promoter (pDEF38-MAb1) antibody production matched the transfection pool profiles, showing higher expression than the pDEF85-MAb1 clones during the growth phase and then slower production as the cultures entered stationary phase (Figure 8B). Antibody production from the CHEF1-CMV vector (pDEF85-MAb1) clones looked very similar to the transfection pool, wherein the majority of antibody expression
occurred after Day 6 after exponential growth slowed down and the cells transitioned to stationary phase (Figure 8A).

[0089] Clonal cell lines expressing GP1 were selected using a FACS based assay to detect GP1 expression early in development. More than 100 clones from each of the pDEF38-GP1 and pDEF85-GP1 transfection pools were screened and ranked by FACS mean fluorescence. The top eight GP1-expressing cultures, based on FACS analysis from each set, were further examined in fed-batch production models using CD-CIM1 base media and BF1 feeds (Figure 9). The average titers and specific productivities, shown in Table 2 and Table 3, indicated that expression from the CHEF1-CMV promoter (pDEF85-GP1) was much higher than from the CHEF1 promoter alone (pDEF38-GP1) and was driven by an increase in specific productivity.

<table>
<thead>
<tr>
<th>Average Glycoprotein Titer</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pDEF38-GP1</td>
<td>pDEF85-GP1</td>
</tr>
<tr>
<td>Day 12</td>
<td>760 (n=8)</td>
<td>1199 (n=8)</td>
</tr>
<tr>
<td>Day 14</td>
<td>735 (n=8)</td>
<td>1456 (n=8)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Specific Productivity</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pDEF38-GP1</td>
<td>pDEF85-GP1</td>
</tr>
<tr>
<td>Average</td>
<td>6.4 (n=8)</td>
<td>15.0 (n=8)</td>
</tr>
<tr>
<td>Range</td>
<td>2.7 to 13.6</td>
<td>10.6 to 20.1</td>
</tr>
</tbody>
</table>

[0090] As demonstrated in the foregoing Example, novel CHEF1-CMV and CHEF1-CMV-AdTPL expression vectors increased expression of both glycoprotein and antibody in stable CHO cell transfection pools. Stable clonal cell lines derived from the CHEF1-CMV and CHEF1-CMV-AdTPL pools also showed improved protein expression compared to the CHEF1-promoter pool. Increased expression in the CHEF1-CMV and CHEF1-CMV-AdTPL clonal cell lines resulted from higher specific productivity compared to CHEF1-promoter, indicating that combining CHEF1 transcriptional regulatory DNA with a CMV promoter increased cellular expression capacity and did not just improve growth performance. The expression pattern from the CHEF1-CMV constructs differed from the CHEF1-promoter alone, with maximal expression occurring later, during the stationary phase of cell growth, indicating that regulation by the CHEF1-CMV-promoter was different than from CHEF1 alone and possessed unique
recombinant protein production characteristics. The delayed temporal expression of protein from CHEF1-CMV compared to CHEF1 alone is evidence that the combined regulatory elements alter CHEF1 growth-dependent expression, thus presenting a novel mechanism to control CHEF1 protein production. The high level expression achieved from a combination of CHEF1 and CMV is unexpected considering previous findings wherein the CHEF1 promoter outperformed the CMV promoter (Running Deer and Allison, 2004). Achieving increased specific productivity, combined with the observed temporal shift in expression, is beneficial because the culture feeding conditions can be optimized for biphasic growth and production in biologic manufacturing processes. An expression vector according to the disclosure comprising CHEF1 transcriptional regulatory DNA and a CMV promoter and/or an AdTPL sequence, therefore, provides an improved option for achieving high titer and productivity in recombinant protein expression systems.

[0091] All of the compositions disclosed and claimed herein can be made and executed without undue experimentation in light of the present disclosure. While the compositions of this disclosure have been described in terms of specific embodiments, it will be apparent to those of skill in the art that variations of the compositions can be made without departing from the concept and scope of the disclosure. More specifically, it will be apparent that certain polynucleotides which are both chemically and biologically related may be substituted for the polynucleotides described herein with the same or similar results achieved. All such similar substitutes and modifications apparent to those skilled in the art are deemed to be within the scope and concept of the invention as defined by the appended claims.

[0092] The references cited herein throughout, to the extent that they provide exemplary procedural or other details supplementary to those set forth herein, are all specifically incorporated herein by reference.
WHAT IS CLAIMED:

1. An expression vector comprising Chinese Hamster Elongation Factor-1α (CHEF1) transcriptional regulatory DNA and a cytomegalovirus (CMV) promoter and/or an adenovirus tripartite leader (AdTPL) sequence.

2. The expression vector of claim 1, wherein the CHEF1 transcriptional regulatory DNA comprises 5' CHEF transcriptional regulatory DNA.

3. The expression vector of claim 2, wherein the 5' CHEF1 transcriptional regulatory DNA comprises Sequence ID NO: 1 or a polynucleotide at least 95% identical to Sequence ID NO: 1.

4. The expression vector of claim 2, wherein the 5' CHEF1 transcriptional regulatory DNA comprises DNA located between position 1 and position 11,716 in Sequence ID NO: 1 or a polynucleotide at least 95% identical to DNA located between position 1 and position 11,716 in Sequence ID NO: 1.

5. The expression vector of claim 4, wherein the 5' CHEF1 transcriptional regulatory DNA comprises DNA located between position 10,774 and position 11,716 in Sequence ID NO: 1 or a polynucleotide at least 95% identical to DNA located between position 10,774 and position 11,716 in Sequence ID NO: 1.

6. The expression vector of claim 2, wherein the 5' CHEF1 transcriptional regulatory DNA comprises Sequence ID NO: 2 or a polynucleotide at least 95% identical to Sequence ID NO: 2.

7. The expression vector of claim 2, wherein the 5' CHEF1 transcriptional regulatory DNA comprises DNA located between position 1 and position 4057 in Sequence ID NO: 2 or a polynucleotide at least 95% identical to DNA located between position 1 and position 4057 in Sequence ID NO: 2.
8. The expression vector of any one of the preceding claims, further comprising 3' CHEF1 transcriptional regulatory DNA.

9. The expression vector of claim 8, wherein the 3' CHEF1 transcriptional regulatory DNA comprises Sequence ID NO: 3 or a polynucleotide at least 95% identical to Sequence ID NO: 3.

10. The expression vector of claim 8, wherein the 3' CHEF1 transcriptional regulatory DNA comprises DNA located between position 1 and position 4180 in Sequence ID NO: 3 or a polynucleotide at least 95% identical to DNA located between position 1 and position 4180 in Sequence ID NO: 2.

11. The expression vector of claim 10, wherein the 3' CHEF1 transcriptional regulatory DNA comprises DNA located between position 1 to position 209 in Sequence ID NO: 3 or a polynucleotide at least 95% identical to DNA located between position 1 to position 209 in Sequence ID NO: 3.

12. The expression vector of any one of claims 8-11, wherein the 3' CHEF1 transcriptional regulatory DNA comprises about 4.2 kilobases.

13. The expression vector of any one of the preceding claims comprising a CMV promoter.

14. The expression vector of any one of the preceding claims comprising an AdTPL sequence.

15. The expression vector of any one of the preceding claims comprising a CMV promoter and an AdTPL sequence.

16. The expression vector of claim 2, wherein the 5' CHEF1 transcriptional regulatory DNA comprises DNA set out in Sequence ID NO: 1, wherein one or more bases between
position 1 and position 11,716 of Sequence ID NO: 1 is/are deleted and replaced with a CMV promoter and/or an AdTPL sequence.

17. The expression vector of claim 16, wherein one or more bases between position 10,512 and position 11,716 of Sequence ID NO: 1 is/are deleted and replaced with a CMV promoter and/or an AdTPL sequence.

18. The expression vector of claim 17, wherein one or more bases between position 10,512 and position 11,716 of Sequence ID NO: 1 is/are deleted and replaced with a CMV promoter.

19. The expression vector of claim 17, wherein one or more bases between position 10,512 and position 11,716 of Sequence ID NO: 1 is/are deleted and replaced with an AdTPL sequence.

20. The expression vector of claim 17, wherein one or more bases between position 10,512 and position 11,716 of Sequence ID NO: 1 is/are deleted and replaced with a CMV promoter and an AdTPL sequence.

21. The expression vector of any one of the preceding claims comprising Sequence ID NO: 4 or a polynucleotide at least 95% identical to Sequence ID NO: 4.

22. The expression vector of any one of the preceding claims comprising Sequence ID NO: 5 or a polynucleotide at least 95% identical to Sequence ID NO: 5.

23. The expression vector of any one of the preceding claims comprising Sequence ID NO: 6 or a polynucleotide at least 95% identical to Sequence ID NO: 6.

24. The expression vector of any one of the preceding claims comprising Sequence ID NO: 7 or a polynucleotide at least 95% identical to Sequence ID NO: 7.
25. The expression vector of any one of the preceding claims, further comprising a selectable marker gene.

26. The expression vector of any one of the preceding claims, further comprising a polynucleotide encoding a protein of interest operably linked to the 5' CHEF1 transcriptional regulatory DNA, the 3' CHEF1 transcriptional regulatory DNA, the CMV promoter and/or the AdTPL sequence.

27. A host cell transformed, transduced or transfected with an expression vector according to any one of the preceding claims.

28. The host cell of claim 27, wherein the host cell is a prokaryotic cell.

29. The host cell of claim 27, wherein the host cell is a eukaryotic cell.

30. The host cell of claim 29, wherein the host cell is a hamster cell.

31. The host cell of claim 30, wherein the host cell is a Chinese Hamster Ovary (CHO) cell.

32. The host cell of claim 29, wherein the host cell is a non-hamster mammalian cell.

33. The host cell of claim 32, wherein the host cell is a human cell.
FIGURES

Figure 1

- **pDEF85**
  - GP1 or MAb1
  - XhoI
  - XbaI
  - CHEF1 5'
  - CMV Promoter
  - DHFR
  - CHEF1 3'

- **pDEF38**
  - CHEF1 5'
  - Intron
  - PolyA
  - DHFR
  - CHEF1 3'
Figure 2

Diagram showing the genetic structure of vectors pDEF86 and pDEF38:
- pDEF86 contains a CMV-AdTPL promoter and DHFR cassette flanked by CHEF1 5' and 3' regions.
- pDEF38 includes an intron followed by a polyA site, DHFR cassette, and CHEF1 3' region.
- Both vectors are cut by XhoI and XbaI restriction enzymes.
- GP1 or MAb1 is indicated as an insert in the CMV-AdTPL promoter region.
Figure 3

Transfection Recovery Viability (GP1)

Viability (%) vs Passage Number

- pDEF38-GP1
- pDEF85-GP1
- pDEF86-GP1
Figure 4

Viable Cell Density (Feed C+CB4)

Viability (Feed C+CB4)
Figure 4C

Productivity (Feed C+CB4)
Figure 5

GP1 Specific Productivity

Titer (ug/ml)

ICA ((million cells/ml)*day)

pDEF38-GP1

pDEF85-GP1

pDEF86-GP1
Figure 6

Productivity (BF1 Feed)

Titer (μg/ml)

Day

pDEF38-GP1
pDEF85-GP1
pDEF86-GP1
Figure 7

Viable Cell Density (BF1 Feed)

Figure 7A

![Graph showing VCD (10^5 c/ml) over days for different constructs.]

Figure 7B

Productivity (BF1 Feed)

![Graph showing Titer (ug/ml) over days for different constructs.]

- pDEF38-MAb1
- pDEF85-MAb1
- pDEF86-MAb1
**Figure 8A**

**Viable Cell Density (Clones: BF1)**

![Graph showing viable cell density over days for pDEF38-MAb1 and pDEF85-MAb1 clones.](image)

**Figure 8B**

**Productivity (Clones: BF1)**

![Graph showing productivity over days for pDEF38-MAb1 and pDEF85-MAb1 clones.](image)
**Figure 9**

**Figure 9A**

**pDEF38-GP1 Clone Productivity**

**Figure 9B**

**pDEF85-GP1 Clone Productivity**
44744A_SeqListing.TXT

SEQUENCE LISTING

CMC Icos Biologics
Clarke, et al.

IMPROVED RECOMBINANT PROTEIN EXPRESSION USING A HYBRID CHEF1 PROMOTER

31351/44744A
61/777,603
2013-03-12

PatentIn version 3.5

1 1
DNA
Cricetulus griseus

ctagaagata tatttgaatat ctttactatt ttgtttcttg aaatcacagg acaggggctt 60
tgcttttcag ccctgtgggc tctgatactt ttcttcacat ataaagctag ttatttttgg 120
tggatgccat tgagcattta tttttctagt ccagacacag tcagcaaaag ctgctatag 180
ttgactgaaa ttaagtcaat taaggaattg aggcaataat gtatgacagt acaggttaa 240
gatctgcagt ggcttcacaggg tcagacagac aacaattgaac gtgggaggag aacacceagg 300

61/777,603
2013-03-12
gctgaggcaagaccatctcaagctgaggcaatttagcagaacctgcttcaaaacacaaag5520
caagacaggctgcaaacacttggaaatctgtagcacttctcataagaggcttagc5580
tatgtaacattgctcaaaattcaagggactaataaatattttttaactgctcata5640
gatacaatgctgcttgaagactagctgtggactgtctcactactaataacaaca5700
acagtaataacagtaacagtggccaaatataaaggcttaaccttctttgatatcagag5760
cattttggtactaaatatttttcttgaatattcactgcttggaaattt5820
tatgattttttaggcaattcatttcattttcttggaaattt5880
tatgattttttaggcttatataatttatttagatcactgatcttgggattt5940
ttgccatagcttcatatgttgttagttctgtgctttatgatttggatg5990
agccagatgcacaccaccatgtgacagtacaacttcatcattgcttatattttgatc6050
tttgagggttggagggtatctgctcttgactattattttgttatctttgattttt6110
agacacggttagttatcattgcctgacttgcacaccttctgttttttgcctgatc6170
tttgtagttctcttaccttcttttattcatttgacgatccagtttattttttttttt6230
caggtggcattcacttgagagctatattttagtttttattattttttttttt6290
ttttgatcttggatttttcgaggtttttctttttttttttttttttggacg6350
tggccagaaaaatgctcggccctggccctttggcccttggccctggccctggcc6410
tggccagaaaaatgctcggccctggccctttggcccttggccctggccctggcc6470
tggccagaaaaatgctcggccctggccctttggcccttggccctggccctggcc6530
tggccagaaaaatgctcggccctggccctttggcccttggccctggccctggcc6590
tggccagaaaaatgctcggccctggccctttggcccttggccctggccctggcc6650
tggccagaaaaatgctcggccctggccctttggcccttggccctggccctggcc6710
tggccagaaaaatgctcggccctggccctttggcccttggccctggccctggcc6770
tggccagaaaaatgctcggccctggccctttggcccttggccctggccctggcc6830
tggccagaaaaatgctcggccctggccctttggcccttggccctggccctggcc6890
tggccagaaaaatgctcggccctggccctttggcccttggccctggccctggcc6950
tggccagaaaaatgctcggccctggccctttggcccttggccctggccctggcc7010
acaaatataggg ctggagagat ctggagagat caggtcagtt tcggagagac 7560
agttcagttc caggtcagtt caggtcagtt caggtcagtt caggtcagtt 7620
gagaccctgac ccgccctgac ccgccctgac ccgccctgac ccgccctgac 7680
gccggccaaaa cctccccaca cctccccaca cctccccaca cctccccaca 7740
ttttccttttc ttttccttttc ttttccttttc ttttccttttc ttttccttttc 7800
atggcgccctt ttttccttttc ttttccttttc ttttccttttc ttttccttttc 7860
ggacacctgc acctgcacct gcacacctgc gcacacctgc gcacacctgc 7920
agagaaacct ttaaaccatat ttaaaccatat ttaaaccatat ttaaaccatat 7980
agcagacagt gctgacactg gctgacactg gctgacactg gctgacactg 8040
actactacta tattactata tagtactata tagtactata tagtactata 8100
caggtcgtcc aacacactcc tggctgagtaa tggctgagtaa tggctgagtaa 8160
tttcgtatat ttctctatct ttttttcctt ttttttcctt ttttttcctt 8220
ttaacgcacct gctgacacct caggtcgtcc gctgacacct caggtcgtcc 8280
agggtaatgt gcggactacta gatgctgtct gatgctgtct gatgctgtct 8340
ctgtcctctt ctggcctgctt ttttcctttt ttttcctttt ttttcctttt 8400
atataatcttt ttctccccaa ttttttttttttttttttt ttttttttttttttttttt 8460
gcaaaaaag cagagacagt gctgacacct gctgacacct gctgacacct 8520
tccctgatata gatgctgtct gatgctgtct gatgctgtct gatgctgtct 8580
agcagagagc gatgctgtct gatgctgtct gatgctgtct gatgctgtct 8640
gtttgacact accttcctgc acatcaggtt atctctctct atctctctct 8700
tttttttttttttgtgtgagtt ttctttttttttttttttttttttttttttttttttttttttttttttt 8760
caggtgagcc ttcgctgctt gatgctgtct gatgctgtct gatgctgtct 8820
caggccccaa gttggggtgt cagctctcct cagctctcct cagctctcct 8880
cagttggttc cggctgctt gctgctgctt gctgctgctt gctgctgctt 8940
gtggctggcc aagagacagt cttgctgctt cttgctgctt cttgctgctt 9000
acagcagctg cattcagata cactacata cactacata cactacata 9060
caggaggtgt ctggtgggtc cagctctcct cagctctcct cagctctcct 9120
tgctttttaa cattacata cactacata cactacata cactacata 9180
tggattgctt ccctggcctt caacatactg cttggccttt cttggccttt 9240
actacgctag cttcatcttc cctgactgct cttcatcttc cttcatcttc 9300
gcgagagatcc aagttccttc gctgctgctt gctgctgctt gctgctgctt 9360
acattctaat tactacata cctgactgct cttgctgctt cttgctgctt 9420
cctccttcct cattacata cactacata cactacata cactacata 9480
cattagaaaa cattagaaaa cattagaaaa cattagaaaa cattagaaaa 9540
44744A_SeqListing.TXT

ggtgggtgga gactgtagcc aggccagcct ggccatggaa gtaattcttg gaatttgccc 11640
attttgagtt tggagcgaag ctgattgaca aagctgctta gccgttcaaa gttattcttc 11700
gaacttttt ttttaag 11716

<210>  2
<211>  4057
<212>  DNA
<213>  Cricetulus griseus
<400>  2
gtgcatagaa acagatgcag gcaaacaatc cacacatat aaaaaaaaaa ttaaaaccaaa 60
taaaacctct aaaccttttgg tttttcttga atctccaact cttcaggtta tgaataatac 120
atttatcgag tcaaaaaattt gcccattcttg ttggcaggtg tgtgtgatga tcggggaagc 180
agaagcaggc agatctctgt gaatgaggcc agcttggtct acaaaagtga tcccaggaca 240
gtcaggcctg ttacacagag aaacctgaaa aaaaaaaga taatatgtac tttgttatta 300
cccccaataa taaggtcaaa cttccagaag cacaacagtt ttaagttcag aaaaataaatat 360
ctagtggtgt acagttacta ctactataa acactaatat agctgtggga aactagttcc 420
aaagatgaat tactaaccag tgtttczaag gaaataataaag aaagcagaga gattaggttc 480
attgctagtg tttctatttt gaataattttct ctttttcaac atagggcacta 540
ggtatcaag ataattttta cgcattgtct gaagcccttag aaatctcttg tgaaaaaaggg 600
atttttgaa taagagaggc taatgtgaggc attataaaaa agggttttttt tgtgcctttgca 660
tgcacagacc ctgtgctttga ttttcttaaca cttctcttga ccagaaagaag cttttggtgga 720
tagaaaaatg taagttatat atacttttag ggggaagctg ttttggttaac tttgttttaca 780
attaacagaa ttaagttgca acaaaacccag aaacctctctg ataaatgaga aaaaaagtcttt 840
gtgaaggagtt gtaaggtcct ctataattag aatttagggg aaaaaaactt gttggtggggg 900
gcacgtctgt aatccccagc ctggggaagta gaggtgaaag attagaatac aaagggcagc 960
tcagcaaca cagtgagttt gagccacccc tgaactatac cagtttctgt tctcctttttt 1020
@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
aattgttttg	 tacattggag	 tgaattgtga	 catcctggag	 agatcctcttg	 agcccataat

1680
caggttatta
ggaataaatc
tctcaagcctt
aaaaatgtag	cttattgata	agagtgttgc

tcttacacctt
tctagtacatt	aaaagaacaat	acaacaagaagttttttcttcttcttgaga

tctgtgccccgg	agaggctttttt
aaaactgccccggggcacaatatttttttttttttctaatatatcagatctcagaagtgggaggtggttttttttattgctaatattttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
44744A_SeqListing.TXT

accagttgcg tgcgcggaaa gatgcccgcct cccgggccct gtgaagagga ggctaaatag 3720

gaggacgcgg cagcccgcccg caggcgggcc gccggtagtc cacacaaaag gagaagggcc 3780

ttgcccctcg cggcccgctg cttcctgtga cccggcggct taccggccgc acttcagtca 3840

cccgggggcc tttttcgagag ccacctgggcct cctccgctggg ggaggggatc tgtctaatgg 3900
cgctgaggtt tgcctcattc tggatggtggg agacagttgc caggccagcc tggccatagga 3960

tagtaattcct ggatatggcc cattttgagt ttggagcgaa gctgattgac aaagctgctt 4020

tagcccccaag aggtattctt cgaacttttt ttttaag 4057

<210> 3
<211> DNA
<212> Homo sapiens
<400> 3

tccatattga agaagactggt taatgataac aatgcctcgag aaacccttca 60
ggagggcgaa agaatgatatt ttggaacatt ttgctgttgt gggcatattttag 120
ttttcaaat cagattctttt taatggaac aacttcggaca aaaaactgct acagaaatatt 180
gagaccacatt aaaaatcaag tttattgaga agttccattc ttttagcttg gaagtcattta 240
ttaagttgctt agcttgacaa ggtatgtgga ggctcaagaga gggagttttag 300
ttcatcagag cccagagctg aagttcttcc agoctcaagaga gctgatttta 360
tctgaggtt tctggttggg acacagatgg aagttcttttg gcagacatgg tgactctgct 420
catctggttt agttctgttg attttgctttg agccttctttg gtttaacaca gtaaatgaca 480
tccataattt agttcattttg tagtgaatat taatatagta gcagagctttg gctctgttcag 540
ttttaaaact agttctttcgg aaagttccttctattctgc gtagttgattt attcattcatc 600

ttagagttta aacttcgaag gctcaagaggg ttttaattta gctgagctat aaataagctt 660
tctttagaat ctgataggtta tcacagatctttaa aaaaagcctt ctggggcaca aagattagc 720

agaaaaggtg agaatggtaga ttgggtggct ttgagttctt cacacacttt aatgtgtagt 780
tggttttagtg catatagttg aacttcggaca aaaaagcctt ctggggcaca aagttatggg 840
gcacacagg gttttcggga attcataagtt gggagaaaaa tgttccccaa cgtgtaagcc 900
tcttgaatact tggagttcgtttttctttgctttctcctttaa ggtttctttt cagtttaactctc 960

gttgggttagc ttggagttagt tcaattcagc aagttctctat aatgtttagt agggatcact 1020
tgaaatatcct tgggcctactttgctttacactttcagagactat gtcgagctaa 1080
ttatccattc tataagcttg ccctttgcttg gtttctttt gctttctgtt gacctaaagga 1140
gtttcttctc ataataaggt gggtagttgg cagacggtga ggaatttat tcatgccaaag 1200
gagttcattt cttgattttgt gggagttttt aatgtttttttgcagagctttt cagatcctgtggt 1260

agtatatgta aaaaattttgc aacgctagtt tcaagcttttattcttactt tggagttctg ggtttttttt 1320
gcactttttaa agctttgcat agagagctcg gtttttggg tggagttcct cagtttattag 1380
cctcaggttg tgcgatctaa taaaagcctt ttacacataa acatttactt cagcatacag 1440
cagtccttta gtgttgaaaa acagaagggga aagctaatat atatatgtcct tgcttttttt 1500
aagtctagt gattacgtgt tgggttgcca gtgtgacagt tctggagttg aatttgtcct 1560
cagacacgta aatggaatt tgggattcac aacactctag tatgagggac ctaatggcct 1620
gtaccaggca caaacgtgct tataaactac acaaaagcga ggaatttaca ggaatttagga 1680
aggtattctt acactttataa aattttttaa aagcttttgac aggattttctt 1740
tgtcacagct gcctttggagc tagttgtgta gacattcataa aatttttttt 1800
gcctgcttga gacacttttt tattatgtat aacactctag gctttcatgt atatctgac 1860
attagaagac gcacacagat ctcctaatgg atggttgtga gccacatgt ggttgtctggg 1920
aattgaacctt aggaccccttg gaagacagct gctcttaacc tctgagccat ccacacccc 1980
agctgctgca cattttaata ggtcggtgaa tcaaaaaacc taggcctttct gtcagatag 2040
aagggcttttt ggtcacccag agtaggattt aagtttattc ggagctcgag gtttgtcctca 2100
gtgcaggtttt gggaggctcag gatcttcagaa atgcagcttg agctatatttt 2160
tggtaaaaa aaaaataagt gggttaagtg ctgctgagcc tggtaaagct gctgggacac 2220
aagttggaaga acatagggca atgctttata ttaaaagcga ggtgtcatttt taatgtcctt 2280
gagaagccta tgctctacact aaccttcagcc accgcagcgt ttatttattc gtaaatattt 2340
gaaatttttt ttggggtttaa gctatttata ctccttctttt ggcgggttgga 2400
tctctcccta ttcttttttg tttttataag atttcagttca ggcgtggtcct gcccttttaac 2460
ttggtaatttc ggcctgagac agttacattg atgagcttat tagtagtaccc atcaataatgt 2520
gaattttttt gggatggggt tcagccttcc gcctgcttccca ctaaagcgtg tggtggtgga 2580
caccaccaca acctggctct tgaataatct ctgcttctttt ggcaggtata ctcctgacta 2640
gacaggggtt tctctgtatttt gagaaggctc ttttgattga ctcctttgagc tctctgtattt 2700
tgagcttcga cttgccctgt tctgctcttcct aaattttggg attaagagtg ttggtctagt 2760
ctgcctggct acaaaagcat ttttttttttt ttttatattta aaaaacaaaag tgggttttttt 2820
agaaggggtg gttggttgtgg cacactttcc acacactcag gtttttgagt tgcctcgagg 2880
atgaagcactg cattactgcc gcccctccctt ggttaaggctc acacagagaa atctttatgg 2940
gagctcttac gctggtgac ctccctgcttgg ctcctgtgac gcccttatttt 3000
tgcctgcttgc aacgcagcagg tttttttttt gttataatat tcaatttaccg ctcctcctttt 3060
agcagccaga ctttcttctct tttttttttt taataataac ctccttttttt taattctgc 3120
ttcattatc atcctgccag tagaagaggg ctacctcctcg ccagcttttaa ggttttgaga 3180
caccaagtgg tgggtggaag tcgatcactg aaccctctgga accacacca cccagcttaa acaatgtcct 3240
cctctgagcc atcctccag cccctaaaaa tggtcttttg gatagggctt caagtagtattt 3300
gagactgagtc ttgctatata aacagggctg gcacacatag ccatgtacag ctggtgagggt 3360	tttcataggc gtgtttttgt tctgctaggcc aggaggatcc ttgagcata cggagttaat 3420
agtgaggtca tttttttatc actctttctga attgagaact aagctgtatgc aaagcaagtt 3480
44744A_SeqListing.TXT

tgactgaaga agtccagttt atgagaacaa ggtgtaaac aatgtgtc a aagatg gcct 3540
tgcagtggt ttagatgatg acccagtc ac cttgggaatta cttgagttgt aagac ctata 3600
tcttgacaggt agtgaacgt gc tcttat acg tccttatag aagaaatg aacatacc cat 3660
tttgttcccc ctaagaatcc actttctctta acctggttcca tgcattatag gttaatttac 3720
ttgcacatcc taggtgctcc cttaaccag tt atgctttag tggcatca aaggactacctact 3780
ccccagttt gcaagtctct gggaacttcc atgacaacct aatagctgtt cggatgagtt a 3840
ttaatagtga ggtgatc tttctcttttc acatctgatgc cagac gat gaaat cctg cttaceca aatacataca 3880
ccatcagcaag ttaaggttgc atccatttttc ctgaagttgctt gccatcagatgc ctagatg 3940
tttgacagtt aatctttgct ag tagttgctta taatgctttgct ggtgatgtc 3990
taaatagtga ggttgatc tttctctctttc cagacaacct aatagctgtt cggatgagtt a 4040
cagcagcagc attttttatc ccagc caagtttttg aatcttttgc aagactctttc tgcattattttttc 4100
tggacaaacc ttgctgcatt gc gagcttgcag atcctttcct gggccggacct ggtcttgcac 4160
tggtgacattg attattgact agttattaat aatagcagtt tggattttact 60
gcccatatat ggagtttccc gc tacaataac ttacagttaaa tggccgctct ggtctgacgcc 120
ccaacgaccc ccgcccattg acgtcaataa tgcgttatgt tccccatgta acgccaatag 180
ggacatttct cc ttgaagctca tgggtgagtt attttgc agtagttgctt cggatgagtt a 240
atcaagtgtca tcatatgcca agtacgcccct tttggcacag cttaagctgtt tcacatctttttc 300
ccgctgctgt ttttggcacag cttaagctgtt tcacatctttttc 360
tattagctca agctatttcc atgctttc ct ggtttggtgc gaggctgtgac 420
agcgagttca ctttggggga cccttttctgcc tttttttt cccttttccc ctttggccgct 480
atatgtgctt catcattttc atgatttttc at ggttttgggca gtacatcataa gggctgtgac ct ggttttgggca gtacatcataa gggctgtgac 540
aatgcgctgct gccttttccc tttttttt cccttttccc ctttggccgct 600
agtgcgctgct gccttttccc tttttttt cccttttccc ctttggccgct 660
agtgcgctgct gccttttccc tttttttt cccttttccc ctttggccgct 660
ctgcctctcccc cttaagctgtt tcacatctttttc 676
<210> 4
<211> DNA
<212> Homo sapiens
<400> 4

gttgacattg attattgact agttattaat aatagcagtt tggattttact 60
<210> 5
<211> DNA
<212> Artificial Sequence
<220> Synthetic polynucleotide
<400> 5
actctttttcc gcatcgcgctg ctgcgcgggca cagctttttc gctgcgcttttg gaggacaaac 60
<table>
<thead>
<tr>
<th>Start Position</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1680</td>
<td>aattgtttgt tacttggcag tgatggtgtga cacttggagg agatctcttg agcccaaat</td>
</tr>
<tr>
<td>1740</td>
<td>caggatatga ggaataaaacttcatagggct gtagatgata agagtgttggag</td>
</tr>
<tr>
<td>1800</td>
<td>ctctggttgc tgaagcctctc gttgctactct ccaacacccc cattaccact aacaaatcc</td>
</tr>
<tr>
<td>1860</td>
<td>ttttcacgct ctctagcatt aagaaacaaaa acaacaaga agttttttttt cttctctcagaa</td>
</tr>
<tr>
<td>1920</td>
<td>ttctgccccg agagggcttt aaacactggc cggggcaaaaa aaaaaaaga aaaaagaaaaa</td>
</tr>
<tr>
<td>1980</td>
<td>aaaaagagggctttagggcagct tggggtggcag cgccttttaaat cccagcagc</td>
</tr>
<tr>
<td>2040</td>
<td>aggaggagccag ggcagggcgg agctctgtga gttgtgaggtc acgctgtgtct acttagtggag</td>
</tr>
<tr>
<td>2100</td>
<td>ttctcagggcag ccaggctgta aagagactgt tcctaaaaac aacagcccc acaatcagaa</td>
</tr>
<tr>
<td>2160</td>
<td>ccagacaaaa acgcagtttat gatcccttggag actgtagag taagagcttactacatataa</td>
</tr>
<tr>
<td>2220</td>
<td>gagacgcagcc tttttgagaa gctagctgat ggtgaggtc acgctgtgtct acttagtggag</td>
</tr>
<tr>
<td>2280</td>
<td>taagattgct aatgctggtgc tctaaatgag cctcaaacagcggcc</td>
</tr>
<tr>
<td>2340</td>
<td>gttctctgctgt ctctatgcctactactactactag acagagctg</td>
</tr>
<tr>
<td>2400</td>
<td>atgagtttgcag ggaataaata ctctctctct cttccgagctt cttccgagctt cttccgagctt</td>
</tr>
<tr>
<td>2460</td>
<td>ttccagagga gcagctgctgct cttggaggtga gaaacaaagc</td>
</tr>
<tr>
<td>2520</td>
<td>tcgctgacct tcctcagccc ggggagtttg ccagctgctgct cttccgagctt cttccgagctt</td>
</tr>
<tr>
<td>2580</td>
<td>gattgtgaccc cccagccccgc cgggagtttg ccagctgctgct cttccgagctt cttccgagctt</td>
</tr>
<tr>
<td>2640</td>
<td>gcagactgtcag ctttcagcccc gcagctgctgct cttccgagctt cttccgagctt cttccgagctt</td>
</tr>
<tr>
<td>2700</td>
<td>tagagctgtctg cttcgagctgc gtagctgctgct cttccgagctt cttccgagctt cttccgagctt</td>
</tr>
<tr>
<td>2760</td>
<td>taggacagatg tagagctgtctg cttcgagctgc gtagctgctgct cttccgagctt cttccgagctt</td>
</tr>
<tr>
<td>2820</td>
<td>agagagaatctgtagctgctgct cttcgagctgc gtagctgctgct cttccgagctt cttccgagctt cttccgagctt</td>
</tr>
<tr>
<td>2880</td>
<td>taaattcagtt cctccagctg gttccctggagag gagaagctgtagctgctgct cttcgagctgc gtagctgctgct cttccgagctt cttccgagctt</td>
</tr>
<tr>
<td>2940</td>
<td>taggcccagaaaaa cccagccccgc cgggagtttg ccagctgctgct cttccgagctt cttccgagctt cttccgagctt</td>
</tr>
<tr>
<td>3000</td>
<td>ctagcagcagcgtagctgctgct cttcgagctgc gtagctgctgct cttccgagctt cttccgagctt cttccgagctt</td>
</tr>
<tr>
<td>3060</td>
<td>atgagctgtctg cttcgagctgc gtagctgctgct cttccgagctt cttccgagctt cttccgagctt</td>
</tr>
<tr>
<td>3120</td>
<td>tattttcagtt cctccagctg gttccctggagag gagaagctgtagctgctgct cttcgagctgc gtagctgctgct cttccgagctt cttccgagctt</td>
</tr>
<tr>
<td>3180</td>
<td>cctagcagctg cttccagctg gttccctggagag gagaagctgtagctgctgct cttcgagctgc gtagctgctgct cttccgagctt cttccgagctt</td>
</tr>
<tr>
<td>3240</td>
<td>ctagcagcagcgtagctgctgct cttcgagctgc gtagctgctgct cttccgagctt cttccgagctt cttccgagctt</td>
</tr>
<tr>
<td>3300</td>
<td>atgagctgtctg cttcgagctgc gtagctgctgct cttccgagctt cttccgagctt cttccgagctt</td>
</tr>
<tr>
<td>3360</td>
<td>cctgctgagctg gttccctggagag gagaagctgtagctgctgct cttcgagctgc gtagctgctgct cttccgagctt cttccgagctt</td>
</tr>
<tr>
<td>3420</td>
<td>ctagcagcagcgtagctgctgct cttcgagctgc gtagctgctgct cttccgagctt cttccgagctt cttccgagctt</td>
</tr>
<tr>
<td>3480</td>
<td>ctagcagcagcgtagctgctgct cttcgagctgc gtagctgctgct cttccgagctt cttccgagctt cttccgagctt</td>
</tr>
<tr>
<td>3527</td>
<td>ctagcagcagcgtagctgctgct cttcgagctgc gtagctgctgct cttccgagctt cttccgagctt cttccgagctt</td>
</tr>
</tbody>
</table>

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

Synthetic polynucleotide

7

ttgcatagaa acagatgcag gcacaacatc cacacatatata aacaacaaaaa ttataaccaaa 60
taaacacctt caaacttttgt gctttcttga atcttcaacct cttcaggtta tgaataatac
atttatagc tccccaaatttt gcacacttttg tggcaggggt tgggtgtgat tggggaagc 120
agaagcaggc agatctctgtg gaatagggcc agctctgtct cacaagttgac tccccaggaca
gtcaggccctg ttaacacagag aacacctgaa aaaaaaaaga taatatgtac tatacttta 180
ccccaaata taagcctaaa cacaacactg ttaagtcagc aaaaataaat 240
taqtggaggt acagtacta ctacatatata acaacattatt agcgttgagg aactattttc 300
aaagatgaat taagttttgga ctaaagctcc agagcaacgatt cccaataaat 360
tgcatagacc ctgtgtcttga tctctttgta cacaagttcacc tttccgtgt 420
aaagatcaat taactacacag tttgctcagaa gaaataattg aagaccaaga gattagttct 480
attgtcagtg tttcacttttc gtatatattt tacaattcct tttggtaca aatggcacta 540
gggtacagc ataatatatta cgaagcggctc aagaacccct aaaaactttc 600
attgtgagaa taatggtggcc attatatagaa agccccctttt gttccttgca 660
tgcatagac ccgtgttttga tctctttaaca cccctctttga ccaagaaaagct cttctgtgga 720
tagaaaaagta ttagtttata actatattaag gggacacgatt ttctgatttc tttggtttca 780
attaacagaa ttaagttgca aacagccagg aacccctttcg ataataagtga aaaaacgctt 840
gtagaggggt tgaaggtctct gtaatatagaa aataagagaa aggctttttgt gttgcttgga 900
gcacgctgtc aatcccaagca tttgggaagta gaggtgaag attagaaatc aagggccagc 960
ttccagcaaca caggtgttattt cagggccacct tgaactcaca cagttctttg tctctttctt 1020
ttttttttttttttttttt tttttttttttgt tttcctgtgt agtttttggag ccctatcctgg 1080
cactagctct gaaagacgggt ttggcctctga actcagatg cagccagctt ctgctgggat 1140
taaagtattgc cccacacac ccccgagtct ttgtctcaaa caacaaaaaa taacatcagg 1200
aggtggtgag agggctcagtt gttcagctcc aacttctgca aacgcttgcaact ctgagtgtgga 1260
tcttttaagc tctagttctgc ggaggaggaac actctgaacc tggaggttct ccctttggtcc 1320
ccacacatatg ctatatacag cagttctctgct cccacactct ttaatatggtt aaaaaaataa 1380
aaatattttta caacaacagt ttgggtttgtg aaaaactgaa ctatgataata gtagaagatac 1440
aagtattcag taaattttct ctcacaactc cccaaatttt gtttatatt tcaagttttttt 1500
tccttctagt cttggacagt ggtctctttgc gccaaggtc cagatcttcttc 1560
caatgcagca cccgcagaga ggctctctct gctctcttat taggtgttgg ctataatact 1620
aatgttttgt tacagctcag ctgattccttg gacatctctgg agcctcataat 1680
caggattata ggaataaatc cttcagcgtt ccaatagttcctactagtgata agagtgtcctg 1740
tctggtgtgc tgaagcctct cggctccatt ccacacccct ataacctcaaacatgctct 1800
ttttcaccgct cccagcactc aagaacaaaaa caaacaagaag atggtttttctt tttttgtgga 1860