VECTORS FOR PROTEIN MANUFACTURE

CROSS-REFERENCE TO RELATED APPLCIATIONS

This application claims priority to U.S. Provisional Application No. 62/775,194, filed December 4, 2018, the entire contents of which are incorporated herein by reference.

FIELD OF THE INVENTION

The present invention relates to vectors and their use to develop host cell lines for production of a protein of interest, and in particular to vectors which utilize a weak promoter to drive a selectable marker.

BACKGROUND OF THE INVENTION

Therapeutic protein drugs are an important class of medicines serving patients most in need of novel therapies. Recently approved recombinant protein therapeutics have been developed to treat a wide variety of clinical indications, including cancers, autoimmunity/inflammation, exposure to infectious agents, and genetic disorders. The latest advances in protein-engineering technologies have allowed drug developers and manufacturers to fine-tune and exploit desirable functional characteristics of proteins of interest while maintaining (and in some cases enhancing) product safety or efficacy or both.

The manufacturing and production of therapeutic proteins are highly complex processes. For example, a typical protein drug may include in excess of 5,000 critical process steps, many times greater than the number required for manufacturing a small-molecule drug.

Similarly, protein therapeutics, which include monoclonal antibodies as well as large or fusion proteins, can be orders-of-magnitude larger in size than small-molecule drugs, having molecular weights exceeding 100 kDa. In addition, protein therapeutics exhibit complex secondary and tertiary structures that must be maintained. Protein therapeutics cannot be completely synthesized by chemical processes and have to be manufactured in living cells or organisms; consequently, the choices of the cell line, species origin, and culture conditions all affect the final product characteristics. Moreover, most biologically active proteins require post-translational modifications that can be compromised when heterologous expression systems are used. Additionally, as the products are synthesized by cells or organisms, complex purification processes are involved. Furthermore, viral clearance processes such as removal of virus particles by using filters or resins, as well as inactivation steps by using low pH or detergents, are implemented to prevent the serious safety issue of

viral contamination of protein drug substances. Given the complexity of therapeutic proteins with respect to their large molecular size, post-translational modifications, and the variety of biological materials involved in their manufacturing process, the ability to enhance particular functional attributes while maintaining product safety and efficacy achieved through protein-engineering strategies is highly desirable.

While the integration of novel strategies and approaches to modify protein drug products is not a trivial matter, the potential therapeutic advantages have driven the increased use of such strategies during drug development. A number of protein-engineering platform technologies are currently in use to increase the circulating half-life, targeting, and functionality of novel therapeutic protein drugs as well as to increase production yield and product purity. For example, protein conjugation and derivatization approaches, including Fc-fusion, albumin-fusion, and PEGylation, are currently being used to extend a drug's circulating half-life.

The production of protein pharmaceutical (biologics) is expensive and time consuming. What is needed in the art are more efficient tools and processes for producing this important class of drugs.

SUMMARY

The present disclosure describes vectors and their use to develop host cell lines for production of a protein of interest, and in particular to vectors which utilize a weak promoter to drive a selectable marker.

In an aspect, the present disclosure refers to a vector for expression of a protein of interest comprising a nucleic acid sequence encoding a selectable marker in operable association with a viral Self-Inactivating (SIN) Long Terminal Repeat (LTR) promoter sequence and a nucleic acid sequence encoding the protein of interest operably linked to a second promoter sequence.

In some examples, the first promoter sequence that has been altered to reduce promoter activity as compared to a non-altered or wild-type version of the first promoter sequence is a viral Self-Inactivating (SIN) Long Terminal Repeat (LTR) promoter sequence. In some preferred embodiments, the SIN LTR promoter sequence is at least 95% identical to SEQ ID NO:3. In some examples, the SIN LTR promoter sequence is SEQ ID NO:3.

In some examples, the selectable maker is Glutamine Synthetase (GS). In some examples, the selectable maker is Dihydrofolate Reductase (DHFR).

In some examples, the vector comprises a single poly A signal sequence in operable association with the selectable marker and the nucleic acid encoding a protein of interest. In other examples, the vector comprises a first poly A signal sequence in operable association with the selectable marker and a second poly A signal sequence in operable association the nucleic acid encoding a protein of interest.

In some examples, the protein of interest is selected from the group consisting of an Fc-fusion protein, an enzyme, an albumin fusion, a growth factor, a protein receptor, a single chain antibody (scFv), a single chain-Fc (scFv-Fc), a diabody, and minibody (scFv-CH3), Fab, single chain Fab (scFab), an immunoglobulin heavy chain, and an immunoglobulin light chain. In some examples, the protein of interest is an Fc-fusion protein.

In some examples, the vector is a plasmid. In some examples, the vector is a viral vector. In some examples, the vector is a retroviral vector. In some examples, the vector is a lentiviral vector.

In another aspect, the present disclosure refers to a host cell comprising the vector as disclosed herein. In some examples, the host cell line is a GS knockout cell line. In some examples, the host cell line is a DHFR knockout cell line. In some examples, the host cell line is Chinese Hamster Ovary (CHO) cell line. In some examples, the host cell line is a HEK 293 or CAP cell line. In some examples, the host cell comprises from about 1, 20, 50 to 1000 copies of the vector. In some examples, the host cell comprises from about 10 to 200 copies of the vector. In some examples, the host cell comprises from about 10 to 100 copies of the vector. In some examples, the host cell comprises from about 20 to 100 copies of the vector.

In some examples, the host cell further comprises at least a second vector that encodes and allows for expression of a second protein of interest, and wherein said second vector does not include a selectable marker. In some examples, the host cell further comprises at least a second vector that encodes and allows for expression of a second protein of interest, and wherein said second vector includes a selectable marker that is different from the selectable marker in the first vector. In some examples, the first protein of interest in the first vector is one of an immunoglobulin heavy or light chain and the second protein in the second vector is the other of an immunoglobulin heavy or light chain. In some examples, the first protein of interest is an immunoglobulin heavy chain and the second protein of interest is an immunoglobulin light chain. In some examples, the host cell comprises from about 1, 20, 50 or 100 to 1000 copies of the second vector. In some examples, the host cell comprises from about 10 to 200 copies of the second vector. In some examples, the host cell comprises from about 10 to 100 copies of the

second vector. In some examples, the host cell comprises from about 20 to 100 copies of the second vector.

In yet another aspect, the present disclosure refers to a host cell culture comprising host cells as disclosed herein. In some examples, the culture produces from 1 to 50 grams/liter/day of the protein of interest. In some examples, the culture produces from 2 to 10 grams/liter/day of the protein of interest.

In an aspect, the present disclosure refers to a process for producing a protein of interest comprising culturing host cells as disclosed herein and purifying the protein of interest from the host cell culture.

In another aspect, the present disclosure refers to an infectious retroviral particle comprising the following elements in 5' to 3' order: 1) a 5' LTR; 2) a retroviral packaging region; 3) a nucleic acid encoding a selectable marker; 4) an internal promoter; 5) a nucleic acid sequence encoding a protein of interest that is operably linked to the internal promoter; and 6) a SIN 3' LTR. In some preferred embodiments, the 5'LTR is an MoMuSV LTR or a SIN LTR. In some embodiments, the 3' LTR comprises a poly A signal sequence. In some examples, the packaging region comprises a plurality of potential translation start sites. In some examples, the selectable marker is GS. In some examples, the internal promoter is a CMV promoter. In some examples, the particle comprises a single poly A signal sequence downstream of the nucleic acid encoding the protein of interest. In some examples, the particles comprises a first poly A signal sequence in operable association with the selectable marker and a second poly A signal sequence in operable association with the nucleic acid encoding the gene of interest.

In yet another aspect, the present disclosure refers to a plasmid comprising the following elements in 5' to 3' order: 1) a 5' SIN LTR (e.g., SIN LTR); 2) a packaging region; 3) a selectable marker (e.g., GS); 4) an internal promoter (e.g., a CMV promoter); 5) a nucleic acid sequence encoding a protein of interest that is operably linked to the internal promoter; and 6) a poly A signal sequence. In some examples, the plasmid comprises a single poly A signal sequence downstream of the nucleic acid encoding the protein of interest.

In an aspect, the present disclosure refers to a system, comprising: a) a first vector comprising a nucleic acid sequence encoding a selectable marker in operable association with a viral Self-Inactivating (SIN) Long Terminal Repeat (LTR) promoter sequence and a nucleic acid sequence encoding a first protein of interest operably linked to a second promoter sequence; and b) a second vector comprising a nucleic acid sequence encoding a second protein of interest operably linked to a promoter sequence, and wherein the second vector does not include a selectable marker.

In another aspect, the present disclosure refers to a process for producing a protein of interest comprising: transducing or transfecting a host cell or cells with an infectious retroviral particle, plasmid, or vector system as disclosed herein, developing a host cell line that expresses the protein of interest from the host cell or cells; culturing the host cells under conditions such that the protein of interested is produced by the host cell line; and purifying the protein of interest from the host cell culture.

In yet another aspect, the present disclosure refers to the infectious retroviral particle or the plasmid of the vector system as disclosed herein for use in transducing a host cell or cells for production of a protein of interest.

DESCRIPTION OF THE FIGURES

FIG. 1. Map of full length MMLV construct.

FIG. 2. Map of SIN MMLV LTR construct.

FIG. 3. Sequence of full length MMLV construct (SEQ ID NO:1).

FIG. 4. Sequence of SIN MMLV LTR construct (SEQ ID NO:2).

FIG. 5. Pooled cell line titer comparison between the full length MMLV construct and SIN MMLV LTR construct.

FIG. 6. Pooled cell line titer comparison between processes using the full length MMLV construct and SIN MMLV LTR construct.

FIG. 7. SIN LTR sequence (SEQ ID NO:3).

FIG. 8. Map of proviral plasmid construct (SEQ ID NO: 4).

FIG. 9. Sequence of proviral plasmid construct (SEQ ID NO: 4).

DEFINITIONS

To facilitate understanding of the invention, a number of terms are defined below.

As used herein, the term "host cell" refers to any eukaryotic cell (e.g., mammalian cells, avian cells, amphibian cells, plant cells, fish cells, and insect cells), whether located in vitro or in vivo.

As used herein, the term "cell culture" refers to any in vitro culture of cells. Included within this term are continuous cell lines (e.g., with an immortal phenotype), primary cell cultures, finite cell lines (e.g., non-transformed cells), and any other cell population maintained in vitro, including oocytes and embryos.

As used herein, the term "vector" refers to any genetic element, such as a plasmid, phage, transposon, cosmid, chromosome, virus, virion, etc., which is capable of replication when associated with the proper control elements and which can transfer gene sequences between cells. Thus, the term includes cloning and expression vehicles, as well as viral vectors.

As used herein, the term “multiplicity of infection” or "MOI" refers to the ratio of integrating vectors:host cells used during transfection or transduction of host cells. For example, if 1,000,000 vectors are used to transduce 100,000 host cells, the multiplicity of infection is 10. The use of this term is not limited to events involving transduction, but instead encompasses introduction of a vector into a host by methods such as lipofection, microinjection, calcium phosphate precipitation, and electroporation.

As used herein, the term “genome” refers to the genetic material (*e.g.*, chomosomes) of an organism.

The term "nucleotide sequence of interest" refers to any nucleotide sequence (*e.g.*, RNA or DNA), the manipulation of which may be deemed desirable for any reason (*e.g.*, treat disease, confer improved qualities, expression of a protein of interest in a host cell, expression of a ribozyme, etc.), by one of ordinary skill in the art. Such nucleotide sequences include, but are not limited to, coding sequences of structural genes (*e.g.*, reporter genes, selection marker genes, oncogenes, drug resistance genes, growth factors, etc.), and non-coding regulatory sequences which do not encode an mRNA or protein product (*e.g.*, promoter sequence, polyadenylation sequence, termination sequence, enhancer sequence, etc.).

As used herein, the term “protein of interest” refers to a protein encoded by a nucleic acid of interest.

As used herein, the terms "nucleic acid molecule encoding," "DNA sequence encoding," "DNA encoding," "RNA sequence encoding," and "RNA encoding" refer to the order or sequence of deoxyribonucleotides or ribonucleotides along a strand of deoxyribonucleic acid or ribonucleic acid. The order of these deoxyribonucleotides or ribonucleotides determines the order of amino acids along the polypeptide (protein) chain. The DNA or RNA sequence thus codes for the amino acid sequence.

The term "promoter," "promoter element," or "promoter sequence" as used herein, refers to a DNA sequence which when ligated to a nucleotide sequence of interest is capable of controlling the transcription of the nucleotide sequence of interest into mRNA. A promoter is typically, though not necessarily, located 5' (*i.e.*, upstream) of a nucleotide sequence of interest whose transcription into mRNA it controls, and provides a site for specific binding by RNA polymerase and other transcription factors for initiation of transcription.

As used herein, the term "altered," when used in reference to a promoter, refers to promoters that have an altered nucleic acid sequence as compared to a reference wild type sequence. For example, the sequence promoter may be altered by deleting certain promoter and/or enhancer elements.

Transcriptional control signals in eukaryotes comprise "promoter" and "enhancer" elements. Promoters and enhancers consist of short arrays of DNA sequences that interact specifically with cellular proteins involved in transcription (Maniatis *et al.*, Science 236:1237 [1987]). Promoter and enhancer elements have been isolated from a variety of eukaryotic sources including genes in yeast, insect and mammalian cells, and viruses (analogous control elements, *i.e.*, promoters, are also found in prokaryotes). The selection of a particular promoter and enhancer depends on what cell type is to be used to express the protein of interest. Some eukaryotic promoters and enhancers have a broad host range while others are functional in a limited subset of cell types (for review see, Voss *et al.*, Trends Biochem. Sci., 11:287 [1986]; and Maniatis *et al.*, supra). For example, the SV40 early gene enhancer is very active in a wide variety of cell types from many mammalian species and has been widely used for the expression of proteins in mammalian cells (Dijkema *et al.*, EMBO J. 4:761 [1985]). Two other examples of promoter/enhancer elements active in a broad range of mammalian cell types are those from the human elongation factor 1α gene (Uetsuki *et al.*, J. Biol. Chem., 264:5791 [1989]; Kim *et al.*, Gene 91:217 [1990]; and Mizushima and Nagata, Nuc. Acids. Res., 18:5322 [1990]) and the long terminal repeats of the Rous sarcoma virus

(Gorman *et al.*, Proc. Natl. Acad. Sci. USA 79:6777 [1982]) and the human cytomegalovirus (Boshart *et al.*, Cell 41:521 [1985]).

As used herein, the term "promoter/enhancer" denotes a segment of DNA which contains sequences capable of providing both promoter and enhancer functions (*i.e.*, the functions provided by a promoter element and an enhancer element, see above for a discussion of these functions). For example, the long terminal repeats of retroviruses contain both promoter and enhancer functions. The enhancer/promoter may be "endogenous" or "exogenous" or "heterologous." An "endogenous" enhancer/promoter is one that is naturally linked with a given gene in the genome. An "exogenous" or "heterologous" enhancer/promoter is one that is placed in juxtaposition to a gene by means of genetic manipulation (*i.e.*, molecular biological techniques such as cloning and recombination) such that transcription of that gene is directed by the linked enhancer/promoter.

As used herein, the term "long terminal repeat" of "LTR" refers to transcriptional control elements located in or isolated from the U3 region 5' and 3' of a retroviral genome. As is known in the art, long terminal repeats may be used as control elements in retroviral vectors, or isolated from the retroviral genome and used to control expression from other types of vectors.

As used herein, the terms "complementary" or "complementarity" are used in reference to polynucleotides (i.e., a sequence of nucleotides) related by the base-pairing rules. For example, the sequence "5'-A-G-T-3'," is complementary to the sequence "3'-T-C-A-5'." Complementarity may be "partial," in which only some of the nucleic acids' bases are matched according to the base pairing rules. Or, there may be "complete" or "total" complementarity between the nucleic acids. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of hybridization between nucleic acid strands. This is of particular importance in amplification reactions, as well as detection methods that depend upon binding between nucleic acids.

The terms "homology" and "percent identity" when used in relation to nucleic acids refers to a degree of complementarity. There may be partial homology (i.e., partial identity) or complete homology (i.e., complete identity). A partially complementary sequence is one that at least partially inhibits a completely complementary sequence from hybridizing to a target nucleic acid sequence and is referred to using the functional term "substantially homologous." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or Northern blot, solution hybridization and the like) under conditions of low stringency. A substantially

homologous sequence or probe (i.e., an oligonucleotide which is capable of hybridizing to another oligonucleotide of interest) will compete for and inhibit the binding (i.e., the hybridization) of a completely homologous sequence to a target sequence under conditions of low stringency. This is not to say that conditions of low stringency are such that non-specific binding is permitted; low stringency conditions require that the binding of two sequences to one another be a specific (i.e., selective) interaction. The absence of non-specific binding may be tested by the use of a second target which lacks even a partial degree of complementarity (e.g., less than about 30% identity); in the absence of non-specific binding the probe will not hybridize to the second non-complementary target.

The terms "in operable combination," "in operable order," and "operably linked" as used herein refer to the linkage of nucleic acid sequences in such a manner that a nucleic acid molecule capable of directing the transcription of a given gene and/or the synthesis of a desired protein molecule is produced. The term also refers to the linkage of amino acid sequences in such a manner so that a functional protein is produced.

As used herein, the term “selectable marker” refers to a gene that encodes an enzymatic activity or other protein that confers the ability to grow in medium lacking what would otherwise be an essential nutrient; in addition, a selectable marker may confer resistance to an antibiotic or drug upon the cell in which the selectable marker is expressed.

As used herein, the term “retrovirus" refers to a retroviral particle which is capable of entering a cell (*i.e.*, the particle contains a membrane-associated protein such as an envelope protein or a viral G glycoprotein which can bind to the host cell surface and facilitate entry of the viral particle into the cytoplasm of the host cell) and integrating the retroviral genome (as a double-stranded provirus) into the genome of the host cell. The term "retrovirus" encompasses Oncovirinae (*e.g.*, Moloney murine leukemia virus (MoMLV), Moloney murine sarcoma virus (MoMSV), and Mouse mammary tumor virus (MMTV), Spumavirinae, amd Lentivirinae (*e.g.*, Human immunodeficiency virus, Simian immunodeficiency virus, Equine infection anemia virus, and Caprine arthritis-encephalitis virus; *See, e.g.*, U.S. Pat. Nos. 5,994,136 and 6,013,516, both of which are incorporated herein by reference).

As used herein, the term "retroviral vector" refers to a retrovirus that has been modified to express a gene of interest. Retroviral vectors can be used to transfer genes efficiently into host cells by exploiting the viral infectious process. Foreign or heterologous genes cloned (*i.e.*, inserted using molecular biological techniques) into the retroviral genome can be delivered efficiently to host cells that are susceptible to infection by the retrovirus. Through well known genetic manipulations, the replicative capacity of the retroviral genome

can be destroyed. The resulting replication-defective vectors can be used to introduce new genetic material to a cell but they are unable to replicate. A helper virus or packaging cell line can be used to permit vector particle assembly and egress from the cell. Such retroviral vectors comprise a replication-deficient retroviral genome containing a nucleic acid sequence encoding at least one gene of interest (*i.e.*, a polycistronic nucleic acid sequence can encode more than one gene of interest), a 5' retroviral long terminal repeat (5' LTR); and a 3' retroviral long terminal repeat (3' LTR).

The term “pseudotyped retroviral vector” refers to a retroviral vector containing a heterologous membrane protein. The term "membrane-associated protein" refers to a protein (*e.g.*, a viral envelope glycoprotein or the G proteins of viruses in the Rhabdoviridae family such as VSV, Piry, Chandipura and Mokola) that are associated with the membrane surrounding a viral particle; these membrane-associated proteins mediate the entry of the viral particle into the host cell. The membrane associated protein may bind to specific cell surface protein receptors, as is the case for retroviral envelope proteins or the membrane-associated protein may interact with a phospholipid component of the plasma membrane of the host cell, as is the case for the G proteins derived from members of the Rhabdoviridae family.

The term "heterologous membrane-associated protein" refers to a membrane-associated protein which is derived from a virus that is not a member of the same viral class or family as that from which the nucleocapsid protein of the vector particle is derived. "Viral class or family" refers to the taxonomic rank of class or family, as assigned by the International Committee on Taxonomy of Viruses.

As used herein, the term “lentivirus vector” refers to retroviral vectors derived from the Lentiviridae family (*e.g.*, human immunodeficiency virus, simian immunodeficiency virus, equine infectious anemia virus, and caprine arthritis-encephalitis virus) that are capable of integrating into non-dividing cells (*See, e.g.*, U.S. Pat. Nos. 5,994,136 and 6,013,516, both of which are incorporated herein by reference).

The term “pseudotyped lentivirus vector” refers to lentivirus vector containing a heterologous membrane protein (*e.g.*, a viral envelope glycoprotein or the G proteins of viruses in the Rhabdoviridae family such as VSV, Piry, Chandipura and Mokola).

As used herein, the term “transposon” refers to transposable elements (*e.g.*, Tn5, Tn7, and Tn10) that can move or transpose from one position to another in a genome. In general, the transposition is controlled by a transposase. The term "transposon vector," as used herein, refers to a vector encoding a nucleic acid of interest flanked by the terminal ends of

transposon. Examples of transposon vectors include, but are not limited to, those described in U.S. Pat. Nos. 6,027,722; 5,958,775; 5,968,785; 5,965,443; and 5,719,055, all of which are incorporated herein by reference.

As used herein, the term "adeno-associated virus (AAV) vector" refers to a vector derived from an adeno-associated virus serotype, including without limitation, AAV-1, AAV-2, AAV-3, AAV-4, AAV-5, AAVX7, etc. AAV vectors can have one or more of the AAV wild-type genes deleted in whole or part, preferably the rep and/or cap genes, but retain functional flanking ITR sequences.

AAV vectors can be constructed using recombinant techniques that are known in the art to include one or more heterologous nucleotide sequences flanked on both ends (5' and 3') with functional AAV ITRs. In the practice of the invention, an AAV vector can include at least one AAV ITR and a suitable promoter sequence positioned upstream of the heterologous nucleotide sequence and at least one AAV ITR positioned downstream of the heterologous sequence. A "recombinant AAV vector plasmid" refers to one type of recombinant AAV vector wherein the vector comprises a plasmid. As with AAV vectors in general, 5' and 3' ITRs flank the selected heterologous nucleotide sequence.

AAV vectors can also include transcription sequences such as polyadenylation sites, as well as selectable markers or reporter genes, enhancer sequences, and other control elements that allow for the induction of transcription. Such control elements are described above.

As used herein, the term "AAV virion" refers to a complete virus particle. An AAV virion may be a wild type AAV virus particle (comprising a linear, single-stranded AAV nucleic acid genome associated with an AAV capsid, *i.e.*, a protein coat), or a recombinant AAV virus particle (described below). In this regard, single-stranded AAV nucleic acid molecules (either the sense/coding strand or the antisense/anticoding strand as those terms are generally defined) can be packaged into an AAV virion; both the sense and the antisense strands are equally infectious.

As used herein, the term "recombinant AAV virion" or "rAAV" is defined as an infectious, replication-defective virus composed of an AAV protein shell encapsidating (*i.e.*, surrounding with a protein coat) a heterologous nucleotide sequence, which in turn is flanked 5' and 3' by AAV ITRs. A number of techniques for constructing recombinant AAV virions are known in the art (*See, e.g.*, U.S. Patent No. 5,173,414; WO 92/01070; WO 93/03769; Lebkowski *et al.*, Molec. Cell. Biol. 8:3988-3996 [1988]; Vincent *et al.*, Vaccines 90 [1990] (Cold Spring Harbor Laboratory Press); Carter, Current Opinion in Biotechnology 3:533-539

[1992]; Muzyczka, Current Topics in Microbiol. and Immunol. 158:97-129 [1992]; Kotin, Human Gene Therapy 5:793-801 [1994]; Shelling and Smith, Gene Therapy 1:165-169 [1994]; and Zhou *et al.*, J. Exp. Med. 179:1867-1875 [1994], all of which are incorporated herein by reference).

Suitable nucleotide sequences for use in AAV vectors (and, indeed, any of the vectors described herein) include any functionally relevant nucleotide sequence. Thus, the AAV vectors of the present invention can comprise any desired gene that encodes a protein that is defective or missing from a target cell genome or that encodes a non-native protein having a desired biological or therapeutic effect (*e.g.*, an antiviral function), or the sequence can correspond to a molecule having an antisense or ribozyme function. Suitable genes include those used for the treatment of inflammatory diseases, autoimmune, chronic and infectious diseases, including such disorders as AIDS, cancer, neurological diseases, cardiovascular disease, hypercholestemia; various blood disorders including various anemias, thalassemias and hemophilia; genetic defects such as cystic fibrosis, Gaucher's Disease, adenosine deaminase (ADA) deficiency, emphysema, etc. A number of antisense oligonucleotides (*e.g.*, short oligonucleotides complementary to sequences around the translational initiation site (AUG codon) of an mRNA) that are useful in antisense therapy for cancer and for viral diseases have been described in the art. (*See, e.g.*, Han *et al.*, Proc. Natl. Acad. Sci. USA 88:4313-4317 [1991]; Uhlmann *et al.*, Chem. Rev. 90:543-584 [1990]; Helene *et al.*, Biochim. Biophys. Acta. 1049:99-125 [1990]; Agarwal *et al.*, Proc. Natl. Acad. Sci. USA 85:7079-7083 [1989]; and Heikkila *et al.*, Nature 328:445-449 [1987]). For a discussion of suitable ribozymes, *see, e.g.*, Cech *et al.* (1992) J. Biol. Chem. 267:17479-17482 and U.S. Patent No. 5,225,347, incorporated herein by reference.

As used herein, the term "purified" refers to molecules, either nucleic or amino acid sequences, that are removed from their normal environment, isolated or separated. An "isolated nucleic acid sequence" is therefore a purified nucleic acid sequence. "Substantially purified" molecules are at least 60% free, preferably at least 75% free, and more preferably at least 90% free from other components with which they are normally associated.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to vectors and their use to develop host cell lines for production of a protein of interest, and in particular to vectors which utilize a weak promoter to drive a selectable marker.

In some preferred embodiments, the expression systems of the present invention utilize an expression vector that includes a nucleic acid sequence encoding a protein of interest (i.e., a therapeutic protein or other protein that is desired to be produced) in operable association with additional nucleic acid sequences that serve various functions. Thus, for example, the nucleic acid sequence of interest may be included in any one of a variety of expression vectors for expressing a polypeptide. In some embodiments of the present invention, vectors include, but are not limited to, retroviral vectors, chromosomal, nonchromosomal and synthetic DNA sequences (e.g., derivatives of SV40, bacterial plasmids, phage DNA; baculovirus, yeast plasmids, vectors derived from combinations of plasmids and phage DNA, and viral DNA such as vaccinia, adenovirus, fowl pox virus, and pseudorabies). It is contemplated that any vector may be used as long as it is replicable and viable in the host. In some preferred embodiments, the vectors are retroviral vectors as described in US Pat. Nos. 6,852,510 and 7,332,333 and U.S. pat. Publ. Nos. 200402335173 and 20030224415, all of which are incorporated herein by references in their entirety. In some especially preferred embodiments, the vectors are pseudotyped retroviral vectors. In some preferred embodiments of the present invention, mammalian expression vectors comprise an origin of replication, a suitable promoter and enhancer, and also any necessary ribosome binding sites, polyadenylation sites, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking non-transcribed sequences. In other embodiments, DNA sequences derived from the SV40 splice, and polyadenylation sites may be used to provide the required non-transcribed genetic elements.

In some preferred embodiments, the vectors are retroviral vectors. The organization of the genomes of numerous retroviruses is well known to the art and this has allowed the adaptation of the retroviral genome to produce retroviral vectors. The production of a recombinant retroviral vector carrying a gene of interest is typically achieved in two stages.

First, the gene of interest is inserted into a retroviral vector which contains the sequences necessary for the efficient expression of the gene of interest (including promoter and/or enhancer elements which may be provided by the viral long terminal repeats (LTRs) or by an internal promoter/enhancer and relevant splicing signals), sequences required for the efficient packaging of the viral RNA into infectious virions (e.g., the packaging signal (Psi), the tRNA primer binding site (-PBS), the 3' regulatory sequences required for reverse transcription (+PBS)) and the viral LTRs. The LTRs contain sequences required for the association of viral genomic RNA, reverse transcriptase and integrase functions, and sequences involved in directing the expression of the genomic RNA to be packaged in viral

particles. For safety reasons, many recombinant retroviral vectors lack functional copies of the genes that are essential for viral replication (these essential genes are either deleted or disabled); therefore, the resulting virus is said to be replication defective.

Second, following the construction of the recombinant vector, the vector DNA is introduced into a packaging cell line. Packaging cell lines provide proteins required in trans for the packaging of the viral genomic RNA into viral particles having the desired host range (i.e., the viral-encoded gag, pol and env proteins). The host range is controlled, in part, by the type of envelope gene product expressed on the surface of the viral particle. Packaging cell lines may express ecotropic, amphotropic or xenotropic envelope gene products. Alternatively, the packaging cell line may lack sequences encoding a viral envelope (env) protein. In this case the packaging cell line will package the viral genome into particles that lack a membrane-associated protein (e.g., an env protein). In order to produce viral particles containing a membrane associated protein that will permit entry of the virus into a cell, the packaging cell line containing the retroviral sequences is transfected with sequences encoding a membrane-associated protein (e.g., the G protein of vesicular stomatitis virus (VSV)). The transfected packaging cell will then produce viral particles, which contain the membrane-associated protein expressed by the transfected packaging cell line; these viral particles, which contain viral genomic RNA derived from one virus encapsidated by the envelope proteins of another virus are said to be pseudotyped virus particles.

The retroviral vectors of the present invention can be further modified to include additional regulatory sequences. As described above, the retroviral vectors of the present invention include the following elements in operable association: a) a 5' LTR; b) a packaging signal; c) a 3' LTR and d) a nucleic acid encoding a protein of interest located between the 5' and 3' LTRs. In some embodiments of the present invention, the nucleic acid of interest may be arranged in opposite orientation to the 5' LTR when transcription from an internal promoter is desired. Suitable internal promoters include, but are not limited to, the alpha-lactalbumin promoter, the CMV promoter (human or simian), and the thymidine kinase promoter.

In other embodiments of the present invention, where secretion of the protein of interest is desired, the vectors are modified by including a signal peptide sequence in operable association with the protein of interest. The sequences of several suitable signal peptides are known to those in the art, including, but not limited to, those derived from tissue plasminogen activator, human growth hormone, lactoferrin, alpha-casein, and alpha-lactalbumin.

In other embodiments of the present invention, the vectors are modified by incorporating an RNA export element (See, e.g., U.S. Pat. Nos. 5,914,267; 6,136,597; and 5,686,120; and WO99/14310, all of which are incorporated herein by reference) either 3' or 5' to the nucleic acid sequence encoding the protein of interest. It is contemplated that the use of RNA export elements allows high levels of expression of the protein of interest without incorporating splice signals or introns in the nucleic acid sequence encoding the protein of interest.

In still other embodiments, the vector further comprises at least one internal ribosome entry site (IRES) sequence. The sequences of several suitable IRES's are available, including, but not limited to, those derived from foot and mouth disease virus (FDV), encephalomyocarditis virus, and poliovirus. The IRES sequence can be interposed between two transcriptional units (e.g., nucleic acids encoding different proteins of interest or subunits of a multisubunit protein such as an antibody) to form a polycistronic sequence so that the two transcriptional units are transcribed from the same promoter.

The most commonly used recombinant retroviral vectors are derived from the amphotropic Moloney murine leukemia virus (MoMLV) (See e.g., Miller and Baltimore Mol. Cell. Biol. 6:2895 [1986]). The MoMLV system has several advantages: 1) this specific retrovirus can infect many different cell types, 2) established packaging cell lines are available for the production of recombinant MoMLV viral particles and 3) the transferred genes are permanently integrated into the target cell chromosome. The established MoMLV vector systems comprise a DNA vector containing a small portion of the retroviral sequence (e.g., the viral long terminal repeat or "LTR" and the packaging or "psi" signal) and a packaging cell line. The gene to be transferred is inserted into the DNA vector. The viral sequences present on the DNA vector provide the signals necessary for the insertion or packaging of the vector RNA into the viral particle and for the expression of the inserted gene. The packaging cell line provides the proteins required for particle assembly (Markowitz et al., J. Virol. 62:1120 [1988]).

In some preferred embodiments, the retroviral vectors are pseudotyped, and for example utilize the G protein of VSV as the membrane associated protein. Unlike retroviral envelope proteins that bind to a specific cell surface protein receptor to gain entry into a cell, the VSV G protein interacts with a phospholipid component of the plasma membrane (Mastromarino et al., J. Gen. Virol. 68:2359 [1977]). Because entry of VSV into a cell is not dependent upon the presence of specific protein receptors, VSV has an extremely broad host range. Pseudotyped retroviral vectors bearing the VSV G protein have an altered host range

characteristic of VSV (i.e., they can infect almost all species of vertebrate, invertebrate and insect cells). Importantly, VSV G-pseudotyped retroviral vectors can be concentrated 2000-fold or more by ultracentrifugation without significant loss of infectivity (Burns et al. Proc. Natl. Acad. Sci. USA 90:8033 [1993]).

The present invention is not limited to the use of the VSV G protein when a viral G protein is employed as the heterologous membrane-associated protein within a viral particle (See, e.g., U.S. Pat. No. 5,512,421, which is incorporated herein by reference). The G proteins of viruses in the Vesiculovirus genera other than VSV, such as the Piry and Chandipura viruses, that are highly homologous to the VSV G protein and, like the VSV G protein, contain covalently linked palmitic acid (Brun et al. Intervirol. 38:274 [1995] and Masters et al., Virol. 171:285 (1990]). Thus, the G protein of the Piry and Chandipura viruses can be used in place of the VSV G protein for the pseudotyping of viral particles. In addition, the VSV G proteins of viruses within the Lyssa virus genera such as Rabies and Mokola viruses show a high degree of conservation (amino acid sequence as well as functional conservation) with the VSV G proteins. For example, the Mokola virus G protein has been shown to function in a manner similar to the VSV G protein (i.e., to mediate membrane fusion) and therefore may be used in place of the VSV G protein for the pseudotyping of viral particles (Mebatsion et al., J. Virol. 69:1444 [1995]). Viral particles may be pseudotyped using either the Piry, Chandipura or Mokola G protein as described in Example 2, with the exception that a plasmid containing sequences encoding either the Piry, Chandipura or Mokola G protein under the transcriptional control of a suitable promoter element (e.g., the CMV intermediate-early promoter; numerous expression vectors containing the CMV IE promoter are available, such as the pcDNA3.1 vectors (Invitrogen)) is used in place of pHCMV-G. Sequences encoding other G proteins derived from other members of the Rhabdoviridae family may be used; sequences encoding numerous rhabdoviral G proteins are available from the GenBank database.

In some preferred embodiments, the vectors are lentiviral vectors. The lentiviruses (e.g., equine infectious anemia virus, caprine arthritis-encephalitis virus, human immunodeficiency virus) are a subfamily of retroviruses that are able to integrate into non-dividing cells. The lentiviral genome and the proviral DNA have the three genes found in all retroviruses: gag, pol, and env, which are flanked by two LTR sequences. The gag gene encodes the internal structural proteins (e.g., matrix, capsid, and nucleocapsid proteins); the pol gene encodes the reverse transcriptase, protease, and integrase proteins; and the pol gene encodes the viral envelope glycoproteins. The 5' and 3' LTRs control transcription and

polyadenylation of the viral RNAs. Additional genes in the lentiviral genome include the vif, vpr, tat, rev, vpu, nef, and vpx genes.

A variety of lentiviral vectors and packaging cell lines are known in the art and find use in the present invention (See, e.g., U.S. Pat. Nos. 5,994,136 and 6,013,516, both of which are herein incorporated by reference). Furthermore, the VSV G protein has also been used to pseudotype retroviral vectors based upon the human immunodeficiency virus (HIV) (Naldini et al., Science 272:263 [1996]). Thus, the VSV G protein may be used to generate a variety of pseudotyped retroviral vectors and is not limited to vectors based on MoMLV. The lentiviral vectors may also be modified as described above to contain various regulatory sequences (e.g., signal peptide sequences, RNA export elements, and IRES's). After the lentiviral vectors are produced, they may be used to transfect host cells as described above for retroviral vectors.

In some preferred embodiments, the vectors are adeno-associated virus (AAV) vectors. AAV is a human DNA parvovirus, which belongs to the genus Dependovirus. The AAV genome is composed of a linear, single-stranded DNA molecule that contains approximately 4680 bases. The genome includes inverted terminal repeats (ITRs) at each end that function in cis as origins of DNA replication and as packaging signals for the virus. The internal nonrepeated portion of the genome includes two large open reading frames, known as the AAV rep and cap regions, respectively. These regions code for the viral proteins involved in replication and packaging of the virion. A family of at least four viral proteins are synthesized from the AAV rep region, Rep 78, Rep 68, Rep 52 and Rep 40, named according to their apparent molecular weight. The AAV cap region encodes at least three proteins, VP1, VP2 and VP3 (for a detailed description of the AAV genome, see e.g., Muzyczka, Current Topics Microbiol. Immunol. 158:97-129 [1992]; Kotin, Human Gene Therapy 5:793-801 [1994]).

AAV requires coinfection with an unrelated helper virus, such as adenovirus, a herpesvirus or vaccinia, in order for a productive infection to occur. In the absence of such coinfection, AAV establishes a latent state by insertion of its genome into a host cell chromosome. Subsequent infection by a helper virus rescues the integrated copy, which can then replicate to produce infectious viral progeny. Unlike the non-pseudotyped retroviruses, AAV has a wide host range and is able to replicate in cells from any species so long as there is coinfection with a helper virus that will also multiply in that species. Thus, for example, human AAV will replicate in canine cells coinfected with a canine adenovirus. Furthermore, unlike the retroviruses, AAV is not associated with any human or animal disease, does not

appear to alter the biological properties of the host cell upon integration and is able to integrate into nondividing cells. It has also recently been found that AAV is capable of site-specific integration into a host cell genome.

In light of the above-described properties, a number of recombinant AAV vectors have been developed for gene delivery (See, e.g., U.S. Patent Nos. 5,173,414; 5,139,941; WO 92/01070 and WO 93/03769, both of which are incorporated herein by reference; Lebkowski et al., Molec. Cell. Biol. 8:3988-3996 [1988]; Carter, Current Opinion in Biotechnology 3:533-539 [1992]; Muzyczka, Current Topics in Microbiol. and Immunol. 158:97-129 [1992]; Kotin, (1994) Human Gene Therapy 5:793-801; Shelling and Smith, Gene Therapy 1:165-169 [1994]; and Zhou et al., J. Exp. Med. 179:1867-1875 [1994]).

Recombinant AAV virions can be produced in a suitable host cell that has been transfected with both an AAV helper plasmid and an AAV vector. An AAV helper plasmid generally includes AAV rep and cap coding regions, but lacks AAV ITRs. Accordingly, the helper plasmid can neither replicate nor package itself. An AAV vector generally includes a selected gene of interest bounded by AAV ITRs that provide for viral replication and packaging functions. Both the helper plasmid and the AAV vector bearing the selected gene are introduced into a suitable host cell by transient transfection. The transfected cell is then infected with a helper virus, such as an adenovirus, which transactivates the AAV promoters present on the helper plasmid that direct the transcription and translation of AAV rep and cap regions. Recombinant AAV virions harboring the selected gene are formed and can be purified from the preparation. Once the AAV vectors are produced, they may be used to transfect (See, e.g., U.S. Pat. 5,843,742, herein incorporated by reference) host cells at the desired multiplicity of infection to produce high copy number host cells. As will be understood by those skilled in the art, the AAV vectors may also be modified as described above to contain various regulatory sequences (e.g., signal peptide sequences, RNA export elements, and IRES's).

In some preferred embodiments, the vectors are transposon-based vectors. Transposons are mobile genetic elements that can move or transpose from one location another in the genome. Transposition within the genome is controlled by a transposase enzyme that is encoded by the transposon. Many examples of transposons are known in the art, including, but not limited to, Tn5 (See e.g., de la Cruz et al., J. Bact. 175: 6932-38 [1993], Tn7 (See e.g., Craig, Curr. Topics Microbiol. Immunol. 204: 27-48 [1996]), and Tn10 (See e.g., Morisato and Kleckner, Cell 51:101-111 [1987]). The ability of transposons to integrate into genomes has been utilized to create transposon vectors (See, e.g., U.S. Pat. Nos.

5,719,055; 5,968,785; 5,958,775; and 6,027,722; all of which are incorporated herein by reference.) Because transposons are not infectious, transposon vectors are introduced into host cells via methods known in the art (e.g., electroporation, lipofection, or microinjection). Therefore, the ratio of transposon vectors to host cells may be adjusted to provide the desired multiplicity of infection to produce the high copy number host cells of the present invention.

Transposon vectors suitable for use in the present invention generally comprise a nucleic acid encoding a protein of interest interposed between two transposon insertion sequences. Some vectors also comprise a nucleic acid sequence encoding a transposase enzyme. In these vectors, one of the insertion sequences is positioned between the transposase enzyme and the nucleic acid encoding the protein of interest so that it is not incorporated into the genome of the host cell during recombination. Alternatively, the transposase enzyme may be provided by a suitable method (e.g., lipofection or microinjection). As will be understood by those skilled in the art, the transposon vectors may also be modified as described above to contain various regulatory sequences (e.g., signal peptide sequences, RNA export elements, and IRES's).

In some preferred embodiments, the vectors include a selectable marker. Suitable selectable markers include but are not limited to glutamine synthetase (GS), dihydrofolate reductase (DHFR) and the like. These genes are described in U.S. Pat. Nos. 5,770,359; 5,827,739; 4,399,216; 4,634,665; 5,149,636; and 6,455,275; all of which are incorporated herein by reference. In some particularly preferred embodiments, the selectable marker is GS and has a sequence that shares at least 80%, 90%, 95%, 99% or 100% identity with the sequence from positions 1789 to 2910 of SEQ ID NO:2. In some preferred embodiments, the selectable marker that is utilized is compatible with a host cell line that is deficient in the production of the enzyme encoded by the selectable marker nucleic acid sequence. Suitable host cell lines are described in more detail below. In other embodiments, the selectable marker is an antibiotic resistance marker, i.e., a gene that produces a protein that provides cells expressing this protein with resistance to an antibiotic. Suitable antibiotic resistance markers include genes that provide resistance to neomycin (neomycin resistance gene (neo)), hygromycin (hygromycin B phosphotransferase gene), puromycin (puromycin N-acetyl-transferase), and the like.

In some embodiments, the nucleic acid sequence encoding the selectable marker is operably linked to a promoter sequence. In some particularly preferred embodiments, the promoter sequence has been altered, for example by mutation, to have reduced promoter activity as compared to an unaltered version of the promoter. These promoters may be

described as weak promoters in that the expression of the gene linked to the promoter occurs at a low level as opposed to a high level. Genes regulated by strong promoters yield more mRNA and therefore more product protein than genes regulated by weak promoters. Thus, the instant invention preferably utilizes a promoter for the selectable marker that has been altered to yield less mRNA that the comparable unaltered promoter.

In some preferred embodiments, the promoter that is operably linked to the selectable marker is a long terminal repeat (LTR) from a retrovirus or lentivirus. In some particularly preferred embodiments, the promoter is an LTR that has been altered by removing either all or a portion of the U3 region of the LTR. In some embodiments, the promoter that is operably linked to the selectable marker is a self-inactivating (SIN) LTR. Suitable SIN LTRs are known in the art. In some preferred embodiments, the SIN LTRs of the present invention preferably have at least 80%, 90%, 95%, 99% or 100% identity to SEQ ID NO:3 and the SIN LTR most preferably has therein a deletion in the U3 region that reduces promoter activity as compared to promoter activity when there is no deletion in the U3 region.

It will be understood that where the vector is a retroviral vector, the vector is constructed so that the SIN LTR is the 3' LTR of the retroviral vector sequence. Due to the fact that the U3 deletion is copied to the 5′ and 3′ LTRs during reverse transcription, integrated SIN vectors contain only LTRs with U3 deleted. A map of an exemplary retroviral vector of the instant invention is provided as FIG. 2. As can be seen, the vector contains a SIN 3' LTR. When reverse transcribed, the hCMV-MOMuSV LTR depicted in the vector map is replaced by the SIN 3' LTR depicted in the vector map so that when the vector is integrated into the host cell chromosome the SIN LTR is operably linked to and drives the expression of the depicted GS cDNA.

In certain preferred embodiments of the present invention, the nucleic acid sequence encoding the protein of interest in the expression vector is operatively linked to an appropriate expression control sequence(s) (promoter) to direct mRNA synthesis. Promoters useful in the present invention include, but are not limited to, the LTR or SV40 promoter, the E. coli lac or trp, the phage lambda P_L and P_R, T3 and T7 promoters, and the cytomegalovirus (CMV) immediate early, herpes simplex virus (HSV) thymidine kinase, and mouse metallothionein-I promoters and other promoters known to control expression of gene in prokaryotic or eukaryotic cells or their viruses.

As mentioned above, an exemplary map of a vector of the present invention is provided in FIG. 3 and corresponds to SEQ ID NO:2. In some preferred embodiments, the vectors comprise the following elements in 5' to 3' order:

- 5' LTR (exemplified by the hCMV-MoMuSV LTR)
- a retroviral packaging region
- a nucleic acid encoding a selectable marker (exemplified by a GS cDNA)
- an internal promoter (exemplified by the simian CMV (sCMV) promoter
- a nucleic acid sequence encoding the protein of interest (exemplified by the "anyway" sequence) that is operably linked to the internal promoter
- a WPRE sequence
- a SIN 3' LTR.

In some embodiments, the vector is a plasmid comprising the following elements in 5' to 3' order: 1) a 5' LTR (e.g., SIN LTR); 2) a packaging region; 3) a selectable marker (e.g., GS); 4) an internal promoter (e.g., a CMV promoter); and 5) a nucleic acid sequence encoding a protein of interest that is operably linked to the internal promoter.

In some embodiments, vectors comprise a single poly A signal sequence downstream of the nucleic acid encoding the protein of interest. For example, in some embodiments, the vector comprises the following components in 5' to 3' order: Promoter (e.g., SIN) - selectable marker - stop codon - promoter (e.g., CMV) - protein of interest - stop codon-poly A.

In some embodiments, the present invention provides host cells and host cell culture wherein the host cells express the protein of interest from the vectors described above. In preferred embodiment, the host cells a mammalian host cells. A number of mammalian host cell lines are known in the art. In general, these host cells are capable of growth and survival when placed in either monolayer culture or in suspension culture in a medium containing the appropriate nutrients and growth factors, as is described in more detail below. Typically, the cells are capable of expressing and secreting large quantities of a particular protein of interest into the culture medium. Examples of suitable mammalian host cells include, but are not limited to Chinese hamster ovary cells (CHO-K1, ATCC CCl-61); bovine mammary epithelial cells (ATCC CRL 10274; bovine mammary epithelial cells); monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture; see, e.g., Graham et al., J. Gen Virol., 36:59 [1977]); baby hamster kidney cells (BHK, ATCC CCL 10); mouse sertoli cells (TM4, Mather, Biol. Reprod. 23:243-251 [1980]); monkey kidney cells (CV1 ATCC CCL 70); African green monkey kidney cells (VERO-76, ATCC CRL-1587); human cervical carcinoma cells (HELA, ATCC CCL 2); canine kidney cells (MDCK, ATCC CCL 34); buffalo rat liver cells (BRL 3A, ATCC CRL 1442); human lung cells (W138, ATCC CCL

75); human liver cells (Hep G2, HB 8065); mouse mammary tumor (MMT 060562, ATCC CCL51); TRI cells (Mather et al., Annals N.Y. Acad. Sci., 383:44-68 [1982]); MRC 5 cells; FS4 cells; rat fibroblasts (208F cells); MDBK cells (bovine kidney cells); CAP (CEVEC's Amniocyte Production) cells; and a human hepatoma line (Hep G2).

In some particularly preferred embodiments, the host cells are modified so that they are deficient, or are naturally deficient, in an enzyme activity that is required for growth or survival of the cells in the presence of a selection agent and which is provided by the selectable marker. For example, Chinese Hamster Ovary (CHO) cells have been modified to be deficient for GS. In some preferred embodiments where vector includes a GS selectable marker, the host cell line is deficient in GS. In some particularly preferred embodiments, the GS deficient host cell line is the CHOZN® $GS^{-/-}$ cell line available from Merck KGaA. In other embodiments, where the selectable marker is, for example, DHFR, the cell line may preferably be deficient for DHFR activity (i.e., $DHFR^-$). Suitable DHFR- cell lines include but are not limited to CHO-DG44 and derivatives thereof.

The vectors may be introduced in the host cells by any suitable means. In some preferred embodiments, vectors (e.g., retroviral vectors) encoding a protein of interest have been produced, they may be used to transfect or transduce host cells. Preferably, host cells are transfected or transduced with integrating vectors at a multiplicity of infection sufficient to result in the integration of at least 1, and preferably at least 2 or more retroviral vectors. In some embodiments, multiplicities of infection of from 10 to 1,000,000 may be utilized, so that the genomes of the infected host cells contain from 2 to 1000 copies of the integrated vectors, preferably from 5 to 1000 copies of the integrated vectors, and most preferably from 20 to 500 copies of the integrated vectors. In other embodiments, a multiplicity of infection of from 10 to 10,000 is utilized. When non-pseudotyped retroviral vectors are utilized for infection, the host cells are incubated with the culture medium from the retroviral producer cells containing the desired titer (i.e., colony forming units, CFUs) of infectious vectors. When pseudotyped retroviral vectors are utilized, the vectors are concentrated to the appropriate titer by ultracentrifugation and then added to the host cell culture. Alternatively, the concentrated vectors can be diluted in a culture medium appropriate for the cell type. Additionally, when expression of more than one protein of interest by the host cell is desired, the host cells can be transfected with multiple vectors each containing a nucleic acid encoding a different protein of interest.

In each case, the host cells are exposed to medium containing the infectious retroviral vectors for a sufficient period of time to allow infection and subsequent integration of the

vectors. In general, the amount of medium used to overlay the cells should be kept to as small a volume as possible so as to encourage the maximum amount of integration events per cell. As a general guideline, the number of colony forming units (cfu) per milliliter should be about 10^5 to 10^7 cfu/ml, depending upon the number of integration events desired.

In some embodiments, after transfection or transduction, the cells are allowed to multiply, and are then trypsinized and replated. Individual colonies are then selected to provide clonally selected cell lines. In still further embodiments, the clonally selected cell lines are screened by Southern blotting or PCR assays to verify that the desired number of integration events has occurred. It is also contemplated that clonal selection allows the identification of superior protein producing cell lines. In other embodiments, the cells are not clonally selected following transfection.

In some embodiments, the host cells are transfected with vectors encoding different proteins of interest. The vectors encoding different proteins of interest can be used to transfect the cells at the same time (e.g., the host cells are exposed to a solution containing vectors encoding different proteins of interest) or the transfection can be serial (e.g., the host cells are first transfected with a vector encoding a first protein of interest, a period of time is allowed to pass, and the host cells are then transfected with a vector encoding a second protein of interest). In some preferred embodiments, the host cells are transfected with an integrating vector encoding a first protein of interest, high expressing cell lines containing multiple integrated copies of the integrating vector are selected (e.g., clonally selected), and the selected cell line is transfected with an integrating vector encoding a second protein of interest. This process may be repeated to introduce multiple proteins of interest. In some embodiments, the multiplicities of infection may be manipulated (e.g., increased or decreased) to increase or decrease the expression of the protein of interest. Likewise, the different promoters may be utilized to vary the expression of the proteins of interest. It is contemplated that these transfection methods can be used to construct host cell lines containing an entire exogenous metabolic pathway or to provide host cells with an increased capability to process proteins (e.g., the host cells can be provided with enzymes necessary for post-translational modification).

In still further embodiments, cell lines are serially transfected with vectors encoding the same gene. In some preferred embodiments, the host cells are transfected (e.g., at an MOI of about 10 to 1,000,000, preferably 100 to 10,000) with an integrating vector encoding a protein of interest, cell lines containing single or multiple integrated copies of the integrating vector or expressing high levels of the desired protein are selected (e.g., clonally

selected), and the selected cell line is retransfected with the vector (e.g., at an MOI of about 10 to 1,000,0000; preferably 100 to 10,000). In some embodiments, cell lines comprising at least two integrated copies of the vector are identified and selected. This process may be repeated multiple times until the desired level of protein expression is obtained and may also be repeated to introduce vectors encoding multiple proteins of interest. Unexpectedly, serial transfection with the same gene results in increases in protein production from the resulting cells that are not merely additive.

The present invention contemplates a variety of serial transfection procedures. In some embodiments, where retroviral vectors are utilized, serial transduction procedures are provided. In preferred embodiments, serial transduction is carried out on a pool of cells. In these embodiments, an initial pool of host cells is contacted with retroviral vectors, preferably at a multiplicity of infection ranging from about 0.5 to about 1000 vectors/host cell. The cells are then cultured for several days in an appropriate medium (e.g., with a selection agent). An aliquot of the cells in then taken to determine the number of integrated vectors and to freeze for future possible use. The remaining cells are then recontacted with retroviral vectors, again preferably at a multiplicity of infection ranging from about 0.5 to about 1000 vectors/host cell. This process is repeated until cells with a desired number of integrated vectors are obtained. For example, the process can be repeated up to 10 to 20 or more times. In some embodiments, cells can be clonally selected after any particular transduction step if so desired, however, utilizing a pool of cells in the absence of transduction results in a decreased time to the desired integrated vector copy number.

In some embodiments of the present invention, following transformation of a suitable host strain and growth of the host strain to an appropriate cell density in media, the protein of interest is secreted during culture of the host cells. In some preferred embodiments where amplifiable markers are utilized, it is contemplated that culture of transduced host cells in a medium comprising an inhibitor of the gene. Suitable inhibitors include, but are not limited to methotrexate for inhibition of DHFR and methionine sulphoximine (Msx) or phosphinothricin for inhibition of GS. It is contemplated that as concentrations of these inhibitors are increased in a cell culture system, cells with higher copy numbers of the amplifiable marker (and thus the genes or genes of interest) or which contain higher-producing insertions are selected.

Accordingly, the host cells containing vectors as described above are cultured according to methods known in the art. Suitable culture conditions for mammalian cells are well known in the art (See e.g., J. Immunol. Methods (1983) 56:221-234 [1983], Animal Cell

Culture: A Practical Approach 2nd Ed., Rickwood, D. and Hames, B. D., eds. Oxford University Press, New York [1992]).

The host cell cultures of the present invention are prepared in a media suitable for the particular cell being cultured. Commercially available media such as ActiPro media (HyClone), ExCell Advanced Fed Batch Medium (SAFC), Ham's F10 (Sigma, St. Louis, MO), Minimal Essential Medium (MEM, Sigma), RPMI-1640 (Sigma), and Dulbecco's Modified Eagle's Medium (DMEM, Sigma) are exemplary nutrient solutions. Suitable media are also described in U.S. Pat. Nos. 4,767,704; 4,657,866; 4,927,762; 5,122,469; 4,560,655; and WO 90/03430 and WO 87/00195; the disclosures of which are herein incorporated by reference. Any of these media may be supplemented as necessary with serum, hormones and/or other growth factors (such as insulin, transferrin, or epidermal growth factor), salts (such as sodium chloride, calcium, magnesium, and phosphate), buffers (such as HEPES), nucleosides (such as adenosine and thymidine), antibiotics (such as gentamycin (gentamicin), trace elements (defined as inorganic compounds usually present at final concentrations in the micromolar range) lipids (such as linoleic or other fatty acids) and their suitable carriers, and glucose or an equivalent energy source. In some preferred embodiments where selectable markers such as GS are utilized, for example, the media will lack glutamine. Any other necessary supplements may also be included at appropriate concentrations that would be known to those skilled in the art.

The present invention also contemplates the use of a variety of culture systems (e.g., petri dishes, 96 well plates, roller bottles, and bioreactors) for the transfected host cells. For example, the transfected host cells can be cultured in a perfusion system. Perfusion culture refers to providing a continuous flow of culture medium through a culture maintained at high cell density. The cells are suspended and do not require a solid support to grow on. Generally, fresh nutrients must be supplied continuously with concomitant removal of toxic metabolites and, ideally, selective removal of dead cells. Filtering, entrapment and micro-capsulation methods are all suitable for refreshing the culture environment at sufficient rates.

As another example, in some embodiments a fed batch culture procedure can be employed. In the preferred fed batch culture the mammalian host, cells and culture medium are supplied to a culturing vessel initially and additional culture nutrients are fed, continuously or in discrete increments, to the culture during culturing, with or without periodic cell and/or product harvest before termination of culture. The fed batch culture can include, for example, a semi-continuous fed batch culture, wherein periodically whole culture (including cells and medium) is removed and replaced by fresh medium. Fed batch culture is

distinguished from simple batch culture in which all components for cell culturing (including the cells and all culture nutrients) are supplied to the culturing vessel at the start of the culturing process. Fed batch culture can be further distinguished from perfusion culturing insofar as the supernatant is not removed from the culturing vessel during the process (in perfusion culturing, the cells are restrained in the culture by, e.g., filtration, encapsulation, anchoring to microcarriers etc. and the culture medium is continuously or intermittently introduced and removed from the culturing vessel). In some particularly preferred embodiments, the batch cultures are performed in roller bottles.

Further, the cells of the culture may be propagated according to any scheme or routine that may be suitable for the particular host cell and the particular production plan contemplated. Therefore, the present invention contemplates a single step or multiple step culture procedure. In a single step culture, the host cells are inoculated into a culture environment and the processes of the instant invention are employed during a single production phase of the cell culture. Alternatively, a multi-stage culture is envisioned. In the multi-stage culture cells may be cultivated in a number of steps or phases. For instance, cells may be grown in a first step or growth phase culture wherein cells, possibly removed from storage, are inoculated into a medium suitable for promoting growth and high viability. The cells may be maintained in the growth phase for a suitable period of time by the addition of fresh medium to the host cell culture.

Fed batch or continuous cell culture conditions are devised to enhance growth of the mammalian cells in the growth phase of the cell culture. In the growth phase cells are grown under conditions and for a period of time that is maximized for growth. Culture conditions, such as temperature, pH, dissolved oxygen (dO2) and the like, are those used with the particular host and will be apparent to the ordinarily skilled artisan. Generally, the pH is adjusted to a level between about 6.5 and 7.5 using either an acid (e.g., CO2) or a base (e.g., Na2CO3 or NaOH). A suitable temperature range for culturing mammalian cells such as CHO cells is between about 30o to 38o C and a suitable dO2 is between 5-90% of air saturation.

Following the polypeptide production phase, the polypeptide of interest is recovered from the culture medium using techniques that are well established in the art. The protein of interest preferably is recovered from the culture medium as a secreted polypeptide (e.g., the secretion of the protein of interest is directed by a signal peptide sequence), although it also may be recovered from host cell lysates. As a first step, the culture medium or lysate is centrifuged to remove particulate cell debris. The polypeptide thereafter is purified from

contaminant soluble proteins and polypeptides, with the following procedures being exemplary of suitable purification procedures: by fractionation on immunoaffinity or ion-exchange columns; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation-exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex G-75; and protein A Sepharose columns to remove contaminants such as IgG. A protease inhibitor such as phenyl methyl sulfonyl fluoride (PMSF) also may be useful to inhibit proteolytic degradation during purification. Additionally, the protein of interest can be fused in frame to a marker sequence that allows for purification of the protein of interest. Non-limiting examples of marker sequences include a hexahistidine tag, which may be supplied by a vector, preferably a pQE-9 vector, and a hemagglutinin (HA) tag. The HA tag corresponds to an epitope derived from the influenza hemagglutinin protein (See e.g., Wilson et al., Cell, 37:767 [1984]). One skilled in the art will appreciate that purification methods suitable for the polypeptide of interest may require modification to account for changes in the character of the polypeptide upon expression in recombinant cell culture.

EXPERIMENTAL

The instant invention provides a unique way of combining retroviral transduction (referred to herein as GPEx® technology) in combination with a Glutamine Synthase (GS) knock-out CHO cell line system that has yielded unexpected improvement in the ability of the cells to produce high titers and specific productivities of the protein of interest. The GS knock-out cell line that was used in these studies was the CHOZN cell line available from MilliporeSigma. The GPEx® technology was practiced in a similar manner to what has been done previously in normal CHO cells. However, in the viral expression vector, the GS gene that is used for selection in GS knock-out cell line, was driven off the very weak promoter from the SIN (self-inactivating) LTR present in one of the versions of the GPEx expression vector that we use at Catalent. The combination of that vector, the normal GPEx cell line production process and the GS knock-out CHO cell line gave production levels of up to 7-fold greater than the traditional GPEx process in an unmodified CHO cell line.

Experiment 1:

Two pooled cell lines were produced from two different viral vectors. Both lines were produced using normal GPEx transduction processes and the GS knock-out cell line CHOZN. Both expression vectors were designed to express a test protein “Anyway”.

Anyway is an Fc fusion protein. The only difference between the two expression vectors was that one vector after insertion into the cell line would have the full length Moloney Murine Leukemia Virus (MMLV) LTR driving expression of the GS gene and the other vector would have the MMLV SIN LTR driving expression of GS. The gene constructs used to produce the retrovector particles are provided in FIGs. 1 and 2. The sequences for the gene constructs are provided in FIGs. 3 and 4.

CHOZN Cell Line Development

Retrovector Production: The expression constructs outlined above were introduced into a HEK 293 cell line that constitutively produces the MLV *gag*, *pro*, and *pol proteins*. An envelope containing expression plasmid was also co-transfected with each of the gene constructs. The co-transfection resulted in the production of replication incompetent high titer retrovector that was concentrated by ultracentrifugation and used for cell transductions (1,2). Transduction using retrovectors produced using the above constructs into CHO cells results in the 3' LTR sequence being duplicated to the 5' end of the sequence replacing the hCMV-MoMuSV 5'LTR and subsequently controlling the expression of the GS gene in the pooled cell line.

Transduction of CHOZN Cells with Retrovector: Pooled cell lines containing retrovector insertions from each of the above constructs were made by performing transduction of the CHOZN Chinese Hamster Ovary parental cell line with retrovector made from the gene constructs developed to express the Anyway protein. A single cycle of transduction was performed to generate a pooled cell line for each of the two constructs. Upon completion of the cell transductions, the cell lines were placed in glutamine free medium to perform GS selection.

Fed Batch Production Anyway from the Two Pooled Population of Cells: Post-transduction, the pooled cell lines were scaled up for productivity in a fed batch study in duplicate 250 mL shake flasks. Each shake flask was seeded with 300,000 viable cells per mL in a 50 mL working volume of ExCell Advanced Fed Batch Medium (SAFC) and incubated in a humidified (70-80%) shaking incubator at 110 rpm with 5% CO_2 and temperature of 37°C. Cultures were fed every other day starting on day 3 during the production using one feed supplement. Glucose was monitored daily and supplemented if the level dropped below 4 g/L. Cultures were terminated when viabilities were $\leq$ 50%.

Results

The results are presented in Table 1 and FIG. 5. These results are surprising. At similar average gene copy number for the two pools, there was major titer difference between the two pooled cell lines. So even at a slightly higher average gene copy number, the wild type or full length LTR gene construct gave significantly lower titers and this translated into significantly lower cell specific productivity per gene insert. A selection or competitive advantage for high expressing cells containing the SIN version of the gene construct is occurring.

Table 1. Pooled cell line comparison between the two gene constructs.

Pooled Cell Lines	**Average Pool Gene Copy Number**	**Final Titer (mg/L)**	**PCD/Gene Copy**
wt-LTR	19	101	0.0385
SIN-LTR	16	246	0.113

Experiment 2:

Based on the previous results we designed an experiment to compare the way we were traditionally practicing the GPEx technology to practicing the technology in exactly the same way but using the SIN-GS vector described above in combination with the CHOZN GS knock-out cell line. The processes were kept as similar as possible, with only two main differences. The first being that traditional GPEx uses the GPEx® Chinese Hamster Ovary (GCHO) parental cell line and the new version uses the CHOZN cell line. The second difference is the gene constructs used to make the retrovector were different. For traditional GPEx the construct does not contain the GS gene and for the new version it does contain the GS gene. All other components of the gene construct were similar. Each of the constructs again expresses the Anyway protein. After the pooled cell lines for each of the methods were completed, they were compared for production in a fed-batch culture analysis.

GCHO and CHOZN Cell Line Development:

Retrovector Production: The expression constructs were introduced into a HEK 293 cell line that constitutively produces the MLV *gag*, *pro*, and *pol proteins*. An envelope containing expression plasmid was also co-transfected with the each of the gene constructs.

The co-transfection resulted in the production of replication incompetent high titer retrovector that was concentrated by ultracentrifugation and used for cell transductions (1,2).

Transduction of CHOZN Cells with Retrovector: Pooled cell lines containing retrovector insertions from each of the constructs were made by performing transduction of either the CHOZN Chinese Hamster Ovary parental cell line or the GCHO Chinese Hamster Ovary parental cell line with retrovector made from the either the gene construct containing GS or the gene construct without GS, both developed to express the Anyway protein. Three cycles of transduction were performed to generate a pooled cell line for each of the two constructs in the corresponding cell line. Additional cycles typically increase the number of inserts in the cell line and subsequently the gene copy number. Upon completion of each cycle of transduction the CHOZN based cell pools were placed in glutamine free medium for GS selection. The traditional GPEx cells underwent no selection as is the normal procedure.

Results

As expected the copy numbers increased for both methods with repeated cycles of transduction. However, significantly higher gene copy numbers were observed in each of the new GPEx cell pools compared to traditional GPEx as shown in Table 2.

Table 2. Pooled cell line gene copy numbers compared between the two different processes

Transduction Cycle Number	**Pooled Cell Line Copy Number**	
	Traditional GPEx	**New GPEx**
1x	15	47
2x	38	69
3x	50	79

Fed Batch Production Anyway from the Two Pooled Population of Cells (3 Cycles of Transduction): Post-transduction, the pooled cell lines were scaled up for productivity in a fed batch study in duplicate 250 mL shake flasks. Each shake flask was seeded with 300,000 viable cells per mL in a 50 mL working volume of ActiPro media (HyClone) and incubated in a humidified (70-80%) shaking incubator at 120 rpm with 5% CO_2 and temperature of

37°C (34°C starting day 6). Cultures were fed six times during the production run using two different feed supplements. Glucose was monitored daily and supplemented if the level dropped below 4 g/L. Cultures were terminated when viabilities were ≤ 50%.

The results are provided in Table 3 and FIG. 6. Again, a very unexpected result was observed. Optimized traditional GPEx clonal cell lines producing the Anyway product expressed at a maximum of 1.8 g/L and the new GPEx pool is over double that expression before any clonal selection. The new GPEx pool has both a higher copy number as well as more production per gene insert like was seen in experiment #1. The two of these together resulted in a significant difference in cellular specific productivity with close to a 3-fold higher pg/cell/day for the new GPEx pool. Overall viable cell density was also higher for the new GPEx pool compared to the traditional pool also aiding the substantial titer difference that was observed.

Table 3. Pooled cell line comparison between the two different processes for Anyway

Pooled Cell Lines	Average Pool Gene Copy Number	Final Titer (g/L)	Specific Productivity pg/cell/day (PCD)	PCD/Gene Copy
Traditional GPEx 3x	50	0.59	6.6	0.13
New GPEx 3x	79	4.25	15.2	0.19

Experiments 3 and 4:

The above data is for a Fc-Fusion produced from a single gene. Experiments 3 and 4 utilize transduction with separate heavy and light chain vectors (2 light chain transductions and 3 heavy chain transductions for both) with two different antibodies "Peelaway" and "Yourway" using both GPEx® technology and the new process using the SIN LTR to drive GS expression.

Retrovector Production: The expression constructs were introduced into a HEK 293 cell line that constitutively produces the MLV *gag*, *pro*, and *pol proteins*. An envelope containing expression plasmid was also co-transfected with the each of the gene constructs. The co-transfection resulted in the production of replication incompetent high titer retrovector that was concentrated by ultracentrifugation and used for cell transductions (1,2).

Transduction of CHOZN Cells with Retrovector: Pooled cell lines containing retrovector insertions from each of the constructs were made by performing transduction of either the CHOZN Chinese Hamster Ovary parental cell line or the GCHO Chinese Hamster Ovary parental cell line with retrovector made from the either the heavy and light chain gene constructs containing GS or the gene constructs without GS, developed to express the two different antibodies. Two cycles of light chain transduction and three cycles of heavy chain transduction were performed to generate a pooled cell line for each of the four constructs in each of the corresponding antibody cell lines. Additional cycles typically increase the number of inserts in the cell line and subsequently the gene copy number. Upon completion of each cycle of transduction the CHOZN based cell pools were placed in glutamine free medium for GS selection. The traditional GPEx cells underwent no selection as is the normal procedure.

Fed Batch Production Peelaway and Yourway from the Pooled Population Cell Lines (2 Light Chain Cycles of Transduction and 3 Heavy Chain Cycles of Transduction): Post-transduction, the pooled cell lines were scaled up for productivity in a fed batch study in duplicate 250 mL shake flasks. Each shake flask was seeded with 300,000 viable cells per mL in a 50 mL working volume of ActiPro media (HyClone) and incubated in a humidified (70-80%) shaking incubator at 120 rpm with 5% CO_2 and temperature of 37°C (34°C starting day 6). Cultures were fed six times during the production run using two different feed supplements. Glucose was monitored daily and supplemented if the level dropped below 4 g/L. Cultures were terminated when viabilities were ≤ 50%.

Results

The results are presented in Tables 4 and 5. Significant advantages were observed for titer with the new vector and process.

Table 4. Pooled cell line comparison for Peelaway antibody between the two different processes

Pooled Cell Lines	**Final Titer (g/L)**	**Specific Productivity pg/cell/day (PCD)**
Traditional GPEx 5x	0.57	5.4
New GPEx 5x	1.79	9.6

Table 5. Pooled cell line comparison for Yourway antibody between the two different processes

Pooled Cell Lines	Final Titer (g/L)	Specific Productivity pg/cell/day (PCD)
Traditional GPEx 5x	2.09	18.5
New GPEx 5x	3.66	14.0

Additional copy number analysis was performed on 4 Yourway cell pools. Cell pools generated after 1 light chain transduction and 2 heavy chain transduction cycles were compared to cell pools generated after 2 light chain transduction cycles and 3 heavy chain transduction cycles. Each of these two pooled cell lines were either set up to continue to grow in glutamine containing media or were place in media lacking glutamine. The gene copy number for the total number of genes, the number heavy chain genes and the number of light chain genes were calculated for each of the four pooled cell lines.

Table 6.

Pooled Cell Line	Glutamine +/-	Total Gene Copy #	Heavy Chain Gene Copy #	Light Chain Gene Copy #
LC1x/HC2x	+	31	15	18
LC1x/HC2x	-	63	17	48
LC2x/HC3x	+	64	28	43
LC2x/HC3x	-	94	32	65

These data indicate that even though each of the cells in these pools have numerous copies of the GS gene being driven off the SIN LTR, there is significant selection occurring for those with more copies and, based on data above, those copies that are expressing at higher levels.

Experiment 5

For this experiment, the Yourway antibody product was used. The heavy chain gene construct used to generate the cell line was the same used in Experiment 4. The light chain

gene construct was identical to the heavy chain gene construct in all aspects except that it contained the light chain coding sequence and it lacked the GS gene. A single transduction was performed for each the light chain (-GS) and heavy chain (+GS) containing retrovectors to generate a pooled cell line.

Retrovector Production: The expression constructs were introduced into a HEK 293 cell line that constitutively produces the MLV *gag*, *pro*, and *pol proteins*. An envelope containing expression plasmid was also co-transfected with the each of the gene constructs. The co-transfection resulted in the production of replication incompetent high titer retrovector that was concentrated by ultracentrifugation and used for cell transductions (1,2).

Transduction of CHOZN Cells with Retrovector: Pooled cell lines containing retrovector insertions from each of the constructs were made by performing transduction of the CHOZN Chinese Hamster Ovary parental cell line with retrovector made from light chain gene constructs without GS and heavy chain constructs containing GS. One cycle of light chain transduction and one cycle of heavy chain transduction were performed to generate a pooled cell line. Upon completion of both cycles of transduction, the CHOZN based cell pool was split, with half being placed in glutamine free media and the other half continuing to be supplemented with glutamine during cell expansion and subsequent productivity assessment.

Fed Batch Production of Yourway from the Pooled Population Cell Lines (1 Light Chain Transduction and 1 Heavy Chain Transduction): The pooled cell lines were scaled up for productivity in a fed batch study in a 250 mL shake flask. The shake flasks were seeded with 300,000 viable cells per mL in a 50 mL working volume of ActiPro media (HyClone) and incubated in a humidified (70-80%) shaking incubator at 120 rpm with 5% CO_2 and temperature of 37°C (34°C starting day 6). Cultures were fed six times during the production run using two different feed supplements. Glucose was monitored daily and supplemented if the level dropped below 4 g/L. Glutamine supplemented cultures were monitored daily for glutamine and fed if the level dropped below 2 mM. Cultures were terminated when viabilities were $\leq$ 50%.

Results

The results are presented in Table 7. Selection by removing glutamine significantly increased the heavy chain gene copy number of the cell pool. Somewhat surprisingly the light chain copy number also increased, but not to the same extent as the heavy chain. Total antibody production was also significantly higher than the culture that was unselected with

levels reaching those shown in Table 5 with only two transductions performed instead of the five that were completed for that experiment.

Table 7. Pooled cell line comparison for Yourway antibody with and without glutamine selection. The pooled cell line was produced from a single light chain transduction (-GS) and a single heavy chain transduction (+GS).

Pooled Cell Lines	Average Pool Heavy Chain Copy Number	Average Pool Light Chain Copy Number	Final Titer (mg/L)
+ Glutamine	2.1	12.5	516
- Glutamine	24.1	36.0	3,323

Experiment 6

A gene construct was designed and produced to test if the technology would behave similarly using traditional cell transfection methodologies as compared to retrovector transduction. One example of a traditional plasmid transfection construct that was generated is shown in FIG. 8 (Plasmid map) and FIG. 9 (Actual plasmid sequence; SEQ ID NO:4).

1. Bleck, G.T. 2005 An alternative method for the rapid generation of stable, high-expressing mammalian cell lines (A Technical Review). Bioprocessing J. *Sep/Oct.* pp 1-7.

2. Bleck, G.T., 2010. GPEx® A Flexible Method for the Rapid Generation of Stable, High Expressing, Antibody Producing Mammalian Cell Lines Chapter 4 In: Current Trends in Monoclonal Antibody Development and Manufacturing, Biotechnology: Pharmaceutical Aspects, Edited by: S.J. Shire et al. © 2010 American Association of Pharmaceutical Scientists, DOI 10.1007/978-0-387-76643-0_4.

All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and

spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention that are obvious to those skilled in the field of this invention are intended to be within the scope of the following claims.

CLAIMS

What is claimed is:

1. A vector for expression of a protein of interest comprising a nucleic acid sequence encoding a selectable marker in operable association with a first promoter sequence and a nucleic acid sequence encoding the protein of interest operably linked to a second promoter sequence, wherein the first promoter sequence is a viral Self-Inactivating (SIN) Long Terminal Repeat (LTR) promoter sequence that is at least 95% identical to SEQ ID NO: 3.

2. The vector of claim 1, wherein the SIN LTR promoter sequence is SEQ ID NO:3.

3. The vector of any one of claims 1 to 2, wherein the selectable maker is Glutamine Synthetase (GS).

4. The vector of any one of claims 1 to 2, wherein the selectable maker is Dihydrofolate Reductase (DHFR).

5. The vector of any one of claims 1 to 4, wherein the vector comprises a single poly A signal sequence in operable association with said selectable marker and said nucleic acid encoding a protein of interest.

6. The vector of any one of claims 1 to 4, wherein the vector comprises a first poly A signal sequence in operable association with said selectable marker and a second poly A signal sequence in operable association with said nucleic acid encoding a protein of interest.

7. The vector of any one of claims 1 to 6, wherein the protein of interest is selected from the group consisting of an Fc-fusion protein, an enzyme, an albumin fusion, a growth factor, a protein receptor, a single chain antibody (scFv), a single chain-Fc (scFv-Fc), a diabody, and minibody (scFv-CH3), Fab, single chain Fab (scFab), an immunoglobulin heavy chain, and an immunoglobulin light chain.

8. The vector of claim 7, wherein the protein of interest is an Fc-fusion protein.

9. The vector of any one of claims 1 to 8, wherein the vector is a viral vector.

10. The vector of claim 9, wherein the vector is a retroviral vector.

11. The vector of any one of claims 1 to 10, wherein said vector is a plasmid.

12. A host cell comprising the vector of any one of claims 1 to 11.

13. The host cell of claim 12, wherein the host cell line is a GS knockout cell line.

14. The host cell of claim 13, wherein the host cell line is a DHFR knockout cell line.

15. The host cell of any one of claims 12 to 14, wherein the host cell line is selected from the group consisting of a Chinese Hamster Ovary (CHO) cell, a HEK 293 cell and a CAP cell.

16. The host cell of any one of claims 12 to 15, wherein the host cell comprises from 1 to 1000 copies of the vector.

17. The host cell of any one of claims 12 to 15, wherein the host cell comprises from 10 to 200 copies of the vector.

18. The host cell of any one of claims 12 to 15, wherein the host cell comprises from 10 to 100 copies of the vector.

19. The host cell of any one of claims 12 to 15, wherein the host cell comprises from 20 to 100 copies of the vector.

20. The host cell of any one of claims 12 to 19, wherein the host cell further comprises at least a second vector that encodes and allows for expression of a second protein of interest, and wherein said second vector does not include a selectable marker.

21. The host cell of any one of claims 12 to 19, wherein the host cell further comprises at least a second vector that encodes and allows for expression of a second protein of interest, and wherein said second vector includes a selectable marker that is different from the selectable marker in the first vector.

22. The host cell of any one of claims 20 to 21, wherein the first protein of interest in the first vector is one of an immunoglobulin heavy or light chain and the second protein in the second vector is the other of an immunoglobulin heavy or light chain.

23. The host cell of claim 22, wherein the first protein of interest is an immunoglobulin heavy chain and the second protein of interest is an immunoglobulin light chain.

24. The host cell of any one of claims 20 to 23, wherein the host cell comprises from 1 to 1000 copies of the second vector.

25. The host cell of any one of claims 20 to 23, wherein the host cell comprises from 10 to 200 copies of the second vector.

26. The host cell of any one of claims 20 to 23, wherein the host cell comprises from 10 to 100 copies of the second vector.

27. The host cell of any one of claims 20 to 23, wherein the host cell comprises from 20 to 100 copies of the second vector.

28. A host cell culture comprising host cells according to any one of claims 12 to 27.

29. The host cell culture of claim 28, wherein said culture produces from 1 to 150 picograms/cell/day of the protein of interest.

30. The host cell culture of claim 28, wherein said culture produces from 2 to 50 picograms/cell/day of the protein of interest.

31. A process for producing a protein of interest comprising culturing host cells according to any one of claims 12 to 30 and purifying the protein of interest from the host cell culture.