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 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 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' 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 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 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).