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.

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.

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 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.

2. The vector of claim 1, wherein the SIN LTR promoter sequence is at least 95% identical to SEQ ID NO:3.

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

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

5. The vector of any one of claims 1 to 3, wherein the selectable maker is Dihydrofolate Reductase (DHFR).

6. The vector of any one of claims 1 to 5, 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.

7. The vector of any one of claims 1 to 5, 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.

8. The vector of any one of claims 1 to 7, 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.

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

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

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

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

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

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

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

16. The host cell of any one of claims 13 to 15, wherein the host cell line is selected from the group consisting of a Chinese Hamster Ovary (CHO) cell line, a HEK 293 cell line and a CAP cell line.

17. The host cell of any one of claims 13 to 16, wherein the host cell comprises from about 1 to 1000 copies of the vector.

18. The host cell of any one of claims 13 to 16, wherein the host cell comprises from about 10 to 200 copies of the vector.

19. The host cell of any one of claims 13 to 16, wherein the host cell comprises from about 10 to 100 copies of the vector.

20. The host cell of any one of claims 13 to 16, wherein the host cell comprises from about 20 to 100 copies of the vector.

21. The host cell of any one of claims 13 to 20, 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.

22. The host cell of any one of claims 13 to 20, 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.

23. The host cell of any one of claims 21 to 22, 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.

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

25. The host cell of any one of claims 21 to 24, wherein the host cell comprises from about 1 to 1000 copies of the second vector.

26. The host cell of any one of claims 21 to 24, wherein the host cell comprises from about 10 to 200 copies of the second vector.

27. The host cell of any one of claims 21 to 24, wherein the host cell comprises from about 10 to 100 copies of the second vector.

28. The host cell of any one of claims 21 to 24, wherein the host cell comprises from about 20 to 100 copies of the second vector.

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

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

31. The host cell culture of claim 29, wherein said culture produces from 2 to 50 picograms/cell/day of the protein of interest.

32. A process for producing a protein of interest comprising culturing host cells according to any one of claims 13 to 31 and purifying the protein of interest from the host cell culture.

33. 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;
6) a SIN 3' LTR.

34. The infectious retroviral particle of claim 33, wherein said 5'LTR is an MoMuLV LTR.

35. The infectious retroviral particle of claim 33, wherein said 5'LTR is a SIN LTR.

36. The infectious retroviral particle of claims 33 to 35, wherein said 3' LTR comprises a poly A signal sequence.

37. The infectious retroviral particle of claims 33 to 36, wherein said packaging region comprises a plurality of potential translation start sites.

38. The infectious retroviral particle of any one of claims 33 to 37, wherein the selectable marker is GS.

39. The infectious retroviral particle of any one of claims 33 to 38, wherein the internal promoter is a CMV promoter.

40. The infectious retroviral particle of any one of claims 33 to 39, wherein said particle comprises a single poly A signal sequence downstream of said nucleic acid encoding a protein of interest.

41. A plasmid comprising the following elements in 5' to 3' order:

1) a 5' SIN LTR;
2) a packaging region;
3) 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 poly A signal sequence.

42. The plasmid of claim 41, wherein said packaging region comprises a plurality of potential translation start sites.

43. The plasmid of any one of claims 41 to 42, wherein the selectable marker is GS.

44. The plasmid of any one of claims 41 to 43, wherein the internal promoter is a CMV promoter.

45. The plasmid of any one of claims 41 to 44, wherein said particle comprises a single poly A signal sequence downstream of said nucleic acid encoding a protein of interest.

46. 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 said second vector does not include a selectable marker.

47. The system of claim 46, 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.

48. A process for producing a protein of interest comprising:

Transducing or transfecting a host cell or cells with the infectious retroviral particle of any one of claims 33 to 40, the plasmid of any one of claims 41 to 45, or the vector system of claims 46 to 47;

developing a host cell line that expresses the protein of interest;

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.

49. Infectious retroviral particle of any one of claims 33 to 40, the plasmid of any one of claims 41 to 45, of the vector system of any one of claims 46 to 47 for use in transducing a host cell or cells for production of a protein of interest.