purification. Poor protein expression limits the large-scale use of certain technologies, for example, problems in expressing enough antigen from a DNA vaccine to generate an immune response to conduct a phase 3 clinical trial.

SUMMARY

[0006] There is a need in the art for improving the efficiency and stability of protein translation and improving protein yield and concentration, for example, in the industrial production of protein drugs.

[0006a] In one aspect, there is provided a method of improving full-length protein expression efficiency comprising:

a) providing a polynucleotide comprising:

i) a coding sequence for the full-length protein;

ii) a primary initiation codon that is upstream of the coding sequence of the full-length protein, said primary initiation codon encoding the first amino acid of the coding sequence of the full-length protein; and

iii) one or more secondary initiation codons located within the coding sequence of the full-length protein downstream of the primary initiation codon; and

b) mutating the one or more secondary initiation codons, located within the coding sequence of the full-length protein downstream of the primary initiation codon, wherein the mutation results in a decrease in initiation of protein synthesis at the one or more secondary initiation codons resulting in a reduction of ribosomal diversion away from the primary initiation codon,

thereby increasing expression efficiency of the full-length protein initiated at the primary initiation codon,

wherein mutating the one or more secondary initiation codons located within the coding sequence of the full-length protein downstream of the primary initiation codon comprises mutating one or more nucleotides such that the amino acid sequence of the protein remains unaltered.

[0007] Disclosed is a method of improving full-length protein expression efficiency. The method includes providing a polynucleotide having a coding sequence for the protein; a primary initiation codon that is upstream of the coding sequence; and one or more secondary initiation codons located within the coding sequence. The method also includes mutating one or more secondary initiation codons resulting in a decrease in initiation of protein synthesis at the one or more secondary initiation codons resulting in a reduction of ribosomal diversion away from the primary initiation codon, thereby increasing full-length protein expression efficiency.

[0008] The method can also include mutating one or more nucleotides such that the amino acid sequence remains unaltered. The one or more secondary initiation codons can be in the same reading frame as the coding sequence or out-of-frame with the coding sequence. The one or more secondary initiation codons can be located one or more nucleotides upstream or downstream from a ribosomal recruitment site. The ribosomal recruitment site can include a cap or an IRES. The one or more secondary initiation codons can be selected from AUG, ACG, GUG, UUG, CUG, AUA, AUC, and AUU. The method can include mutating more than one secondary initiation codon within the coding sequence. The method can include mutating all the secondary initiation codons within the coding sequence. A flanking nucleotide can be mutated to a less favorable nucleotide context. The mutation of the one or more secondary initiation codons can avoid introducing new initiation codons. The mutation of the one or more secondary initiation codons can avoid introducing miRNA seed sequences. The mutation of the one or more secondary initiation codons can avoid altering usage bias of mutated codons. The generation of truncated proteins, polypeptide, or peptides other than the full-

CLAIMS

What is claimed is:

1. A method of improving full-length protein expression efficiency comprising:

a) providing a polynucleotide comprising:

i) a coding sequence for the protein;

ii) a primary initiation codon that is upstream of the coding sequence; and

iii) one or more secondary initiation codons located within the coding sequence; and

b) mutating one or more secondary initiation codons, wherein the mutation results in a decrease in initiation of protein synthesis at the one or more secondary initiation codons resulting in a reduction of ribosomal diversion away from the primary initiation codon,

thereby increasing full-length protein expression efficiency, wherein mutating the one or more secondary initiation codons comprises mutating one or more nucleotides such that the amino acid sequence remains unaltered.

2. The method of claim 1, wherein the one or more secondary initiation codons is in the same reading frame as the coding sequence.

3. The method of claim 1, wherein the one or more secondary initiation codons is out-of-frame with the coding sequence.

4. The method of claim 1, wherein the one or more secondary initiation codons is located one or more nucleotides upstream or downstream from a ribosomal recruitment site.

5. The method of claim4 , wherein the ribosomal recruitment site comprises a cap or an IRES.

6. The method of claim 1, wherein the one or more secondary initiation codons is selected from the group consisting of AUG, ACG, GUG, UUG, CUG, AUA, AUC, and AUU.

7. The method of claim 1, wherein more than one secondary initiation codon within the coding sequence is mutated.

8. The method of claim 1, wherein all secondary initiation codons within the coding sequence are mutated.

9. The method of claim 1, wherein mutating the one or more secondary initiation codons comprises mutating one or more nucleotides flanking the one or more secondary initiation codons to diminish the efficiency of the one or more secondary initiation codons .

10. The method of claim 1, wherein mutating the one or more secondary initiation codons does not introduce new initiation codons.

11. The method of claim 1, wherein mutating the one or more secondary initiation codons does not alter usage bias of mutated codons.

12. The method of claim 1, further comprising decreasing the generation of truncated proteins, polypeptide, or peptides other than the full-length encoded protein.

13. The method of claim 1, wherein mutating one or more secondary initiation codons does not introduce miRNA seed sequences, splice donor site, splice acceptor site, or mRNA destabilization elements.

~~20. The method of claim 14, wherein the one or more miRNA binding sites is~~ located within the 3' untranslated region.

21. The method of claim 14, wherein the one or more miRNA binding sites is located within the 5' leader sequence.

Claims:

1. A method of improving full-length protein expression efficiency comprising:

a) providing a polynucleotide comprising:

i) a coding sequence for the full-length protein;

ii) a primary initiation codon that is upstream of the coding sequence of the full-length protein, said primary initiation codon encoding the first amino acid of the coding sequence of the full-length protein; and

iii) one or more secondary initiation codons located within the coding sequence of the full-length protein downstream of the primary initiation codon; and

b) mutating the one or more secondary initiation codons, located within the coding sequence of the full-length protein downstream of the primary initiation codon, wherein the mutation results in a decrease in initiation of protein synthesis at the one or more secondary initiation codons resulting in a reduction of ribosomal diversion away from the primary initiation codon,

thereby increasing expression efficiency of the full-length protein initiated at the primary initiation codon,

wherein mutating the one or more secondary initiation codons located within the coding sequence of the full-length protein downstream of the primary initiation codon comprises mutating one or more nucleotides such that the amino acid sequence of the protein remains unaltered.

2. The method of claim 1, wherein the one or more secondary initiation codons is in the same reading frame as the coding sequence.

3. The method of claim 1, wherein the one or more secondary initiation codons is out-of-frame with the coding sequence.

4. The method of claim 1, wherein the one or more secondary initiation codons is located one or more nucleotides upstream or downstream from a ribosomal recruitment site.

5. The method of claim 4, wherein the ribosomal recruitment site comprises a cap or an IRES.

6. The method of claim 1, wherein the one or more secondary initiation codons is selected from the group consisting of AUG, ACG, GUG, UUG, CUG, AUA, AUC, and AUU.

7. The method of claim 1, wherein more than one secondary initiation codon within the coding sequence is mutated.

8. The method of claim 1, wherein all secondary initiation codons within the coding sequence are mutated.

9. The method of claim 1, wherein mutating the one or more secondary initiation codons comprises mutating one or more nucleotides flanking the one or more secondary initiation codons to diminish the efficiency of the one or more secondary initiation codons.

10. The method of claim 1, wherein mutating the one or more secondary initiation codons does not introduce new initiation codons.

11. The method of claim 1, wherein mutating the one or more secondary initiation codons does not alter usage bias of mutated codons.

12. The method of claim 1, wherein the generation of truncated proteins, polypeptide, or peptides other than the full-length encoded protein is decreased.

13. The method of claim 1, wherein mutating one or more secondary initiation codons does not introduce miRNA seed sequences, splice donor site, splice acceptor site, or mRNA destabilization elements.

14. The method of claim 1, wherein the polynucleotide sequence further comprises one or more miRNA binding sites located within the coding sequence, said method further

comprises the step of mutating the one or more miRNA binding sites, wherein the mutation results in a decrease in miRNA binding at the one or more miRNA binding sites resulting in a reduction of miRNA-mediated down regulation of protein translation.

15. The method of claim 14, wherein mutating the one or more miRNA binding sites comprises mutating one or more nucleotides in a miRNA seed sequence.

16. The method of claim 14, wherein mutating the one or more miRNA binding sites comprises mutating one or more nucleotides such that initiation codons are not introduced into the polynucleotide sequence.

17. The method of claim 14, wherein mutating the one or more miRNA binding sites comprises mutating one or more nucleotides such that rare codons are not introduced into the polynucleotide sequence.

18. The method of claim 14, wherein mutating the one or more miRNA binding sites mutating one or more nucleotides such that additional miRNA seed sequences are not introduced into the polynucleotide sequence.