tion terminator database.

Unniraman et al. (Nucleic Acids Research 30 (3) (2002): 675-684) provide a review of certain terminator structures and classification.

Quiagen's QIAexpress pQE vector ("The QIAexpress System", The QIAexpressionist 06/2003, 5th ed., page 15) contains two terminators.

Unniraman et al. (Journal of Biological Chemistry, 276 (45) (2001): 41850- 41855) relates to intrinsic transcription terminators in eubacteria.

Friedman et al. (Annual Review of Genetics, 21 (1) (1987): 453-488) provide a review on type I and type II terminators.

Lesnik et al. (Nucleic Acid Research, 29(17) (2001): 3583-3594) describe the program RNAMotif for use in searching intrinsic rho-independent terminators.

**Summary**

In one aspect, the present disclosure refers to a polynucleotide comprising at least two consecutive transcription termination signals, characterized in that said consecutive transcription termination signals comprise at least a first and a second transcription termination signal that are at most 1000 nucleotides apart, and at least one of the termination signal has or encodes a RNA hairpin structure, said hairpin structure comprising a stem of at least 12 internal base pairs and wherein said hairpin structure is defined by a sequence Y-X-Z, wherein Y is a nucleotide sequence of at least 12 nucleotides in length and with a G/C content of at least 70%,
X is a nucleotide sequence of 4 nucleotides in length,
X consists of the sequence U/T-NNG, with N being any nucleotide selected from U(T), A, C, G, and
Z is a nucleotide sequence with at least 70% complementarity to Y,
the complementary nucleotides of Z being base paired with the nucleotides of Y, and X being a loop with no base pairing to Y or Z.

In another aspect, the present disclosure refers to an expression cassette comprising a polynucleotide sequence disclosed herein and a coding sequence, wherein the sequence is operatively positioned for termination of the coding sequence.

In yet another aspect, the present disclosure refers to a

vector comprising a polynucleotide disclosed herein or the expression cassette disclosed herein, the vector preferably being a plasmid.

In a further aspect, the present disclosure refers to a cell comprising a polynucleotide molecule with a gene operatively linked to the transcription termination signals of the polynucleotide disclosed herein, preferably in a vector; preferably wherein said cell comprises the T7 RNA polymerase or a polynucleotide encoding the T7 RNA polymerase.

In another aspect, the present disclosure refers to a method of producing a gene product, comprising providing a cell disclosed herein and cultivating said cell under conditions allowing expression of said gene.

In yet another aspect, the present disclosure refers to an in vitro method of producing a mRNA, comprising providing a polynucleotide disclosed herein and a coding sequence upstream of said transcription termination signal and contacting said polynucleotide with the T7 RNA polymerase.

**Description**

An objective of the present invention is to provide improved expression systems with increased efficiency for the expression of a desired product by effective transcriptional termination, especially of highly processive RNA polymerases like the T7 RNA polymerase.

This objective is solved by the subject matter of the claims and definitions described herein.

In a first aspect, the present invention provides a polynucleotide comprising at least two consecutive transcription termination signals, wherein said consecutive transcription termination signals comprise at least a first and a second transcription termination signal that are at most 1000 nucleotides apart, and at least one of the termination signal has or encodes a RNA hairpin structure. According to the present invention it has been found that consecutive termination signals can influence on another and lead to a dramatically increased termination efficiency, especially if at least one terminator comprises a hairpin structure. Said hairpin has a structure comprising a stem of at least 12 internal base pairs with at least 60% G or C ("G/C") nucleic acids. In preferred embodiments the hairpin has a special structure as defined in the second aspect. The invention

provides polynucleotides comprising combinations of 2, 3, 4 or more termination signals that are operatively linked to allow termination after transcription of a preceding coding sequence in concerted action to increase termination efficiency.

Claims:

1. A polynucleotide comprising at least two consecutive transcription termination signals, **characterized in that** said consecutive transcription termination signals comprise at least a first and a second transcription termination signal that are at most 1000 nucleotides apart, and at least one of the termination signal has or encodes a RNA hairpin structure, said hairpin structure comprising a stem of at least 12 internal base pairs and wherein said hairpin structure is defined by a sequence Y-X-Z, wherein Y is a nucleotide sequence of at least 12 nucleotides in length and with a G/C content of at least 70%,
X is a nucleotide sequence of 4 nucleotides in length,
X consists of the sequence U/T-NNG, with N being any nucleotide selected from U(T), A, C, G, and
Z is a nucleotide sequence with at least 70% complementarity to Y,
the complementary nucleotides of Z being base paired with the nucleotides of Y, and X being a loop with no base pairing to Y or Z.

2. The polynucleotide of claim 1 comprising a third transcription termination signal, optionally comprising or encoding a RNA hairpin structure, said third transcription termination signal being at most 1000 nucleotides apart from the first or second transcription termination signal.

3. The polynucleotide according to claim 1 or 2, wherein the hairpin structure with a stem of at least 12 internal base pairs of the at least one termination signal having or encoding a RNA hairpin structure has at least 70% G or C nucleic acids.

4. The polynucleotide of any one of claims 1 to 3, **characterized in that** X consists of the sequence U/T-NCG, especially U/T-U/T-CG.

5. The polynucleotide of any one of claims 1 to 4, **characterized in that** said first, second and/or third transcription termination signal comprises said hairpin structures as defined in any one of claims 1, 3 or 4, preferably at least one of the tran-

scription termination signals comprises a hairpin structure as defined in claim 4 and at least a further transcription termination signal comprises a hairpin structure of claim 1, with Y, X and Z being selected independently.

6. The polynucleotide according to any one of claims 1 to 5 wherein at least one of said terminator signals is a class II terminator.

7. The polynucleotide according to any one of claims 1 to 6, **characterized in that** at least one transcription termination signal comprises a A/T(U)-rich region of at least 4 nucleotides in length comprising at least 75% A or T(U), wherein preferably said A/T(U)-rich region is downstream of a hairpin, in particular preferred not more than 20 nucleotides apart from said hairpin.

8. The polynucleotide according to any one of claims 1 to 7, **characterized in that** Y is of 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, or 30 nucleotides in length.

9. The polynucleotide according to any one of claims 1 to 8, **characterized in that** Y and Z are of m nucleotides in length, Y consists of the nucleotides $y_1$ to $y_m$ and Z consists of the nucleotides $z_1$ to $z_m$, the hairpin consists of the sequence ymym-1$y_m$-2...$y_8y_7y_6y_5y_4y_3y_2$y1-X-$z_1z_2z_3z_4z_5z_6z_7z_8$...$z_{m_1}z_m$, wherein $z_1$ is complementary to $y_1$ and $z_m$ is complementary to $y_m$.

10. The polynucleotide according to any one of claims 1 to 9, **characterized in that** Y and Z are at least 90% complementary, in particular preferred at least 95% complementary or 100% complementary.

11. The polynucleotide according to any one of claims 1 to 10, **characterized in that** in the hairpin, between nucleotides of Y and Z, at most 3, 2 or 1 G-U complementarities are adjacent to another G-U complementarity, preferably the hairpin being without adjacent G-U complementaries.

12. The polynucleotide according to any one of claims 1 to 11 being DNA, encoding the termination signals as defined in any one of claims 1 to 10; or RNA comprising the termination signals as defined in claims 1 to 10.

13. A polynucleotide according to any one of claims 1 to 12 having a size of at most 10000 nucleotides in length.

14. A polynucleotide according to any one of claims 1 to 13, characterized in that the polynucleotide comprises one or more restriction sites flanking said transcription termination signals and/or a cloning site upstream of the transcription termination signals, or a coding sequence upstream of the transcription termination signals, preferably further operatively linked to a promoter.

15. A polynucleotide according to any one of claims 1 to 14 comprising a first terminator signal comprising a hairpin structure defined by a sequence Y-X-Z, wherein Y is a nucleotide sequence of at least 12 nucleotides in length and with a G/C content of at least 70%,
X is a nucleotide sequence of 4 nucleotides in length,
X consists of the sequence U/T-NNG, with N being any nucleotide selected from U(T), A, C, G, and
Z is a nucleotide sequence with at least 80% complementarity to Y, the complementary nucleotides of Z being base paired with the nucleotides of Y, and X being a loop with no base pairing to Y or Z;

a second terminator signal comprising a class II termination signal, preferably comprising the consensus sequence TCTGTT, especially preferred comprising a sequence of at least 20 nucleotides in length with a T content of at least 40%.

16. A polynucleotide according to any one of claims 1 to 15

i) comprising a first terminator signal comprising a hairpin structure defined by a sequence Y-X-Z, wherein Y is a nucleotide sequence of at least 14 nucleotides in length and with a G/C content of at least 70%
X is a nucleotide sequence of 3 to 9 nucleotides in length,
X comprises or consists of the sequence U/T-NNG, with N being any nucleotide selected from U(T), A, C, G, and
Z is a nucleotide sequence with at least 80% complementarity to Y, the complementary nucleotides of Z being base paired with the nucleotides of Y, and X being a loop with no base pairing to Y or Z;

a second terminator signal comprising a hairpin structure defined by a sequence Y-X-Z, wherein Y is a nucleotide sequence of at least 8 nucleotides in length and with a G/C content of at least 40%, preferably at least 50%, especially at least 60%,
X is a nucleotide sequence of 3 to 9 nucleotides in length, and
Z is a nucleotide sequence with at least 70% complementarity to Y, the complementary nucleotides of Z being base paired with the nucleotides of Y, and X being a loop with no base pairing to Y or Z; with the first and second terminator being at most 1000 nucleotides apart to provide a concerted termination by the first and second termination signal, wherein Y, X, and Z of the first and second terminator can be selected independently, in particular as defined in claims 1, 3, 4, 8 to 11; or

ii) comprising a first terminator signal comprising a hairpin structure defined by a sequence Y-X-Z, wherein Y is a nucleotide sequence of at least 12 nucleotides in length and with a G/C content of at least 70%, X is a nucleotide sequence of 3 to 9 nucleotides in length, X comprises or consists of the sequence U/T-NNG, with N being any nucleotide selected from U(T), A, C, G, and Z is a nucleotide sequence with at least 80% complementarity to Y, the complementary nucleotides of Z being base paired with the nucleotides of Y, and X being a loop with no base pairing to Y or Z;

a second terminator signal comprising a class II termination signal, preferably comprising the consensus sequence TCTGTT, es-

pecially preferred comprising a sequence of at least 20 nucleotides in length with a T content of at least 40%; or

iii) comprising a first terminator signal comprising a hairpin structure defined by a sequence Y-X-Z, wherein Y is a nucleotide sequence of at least 12 nucleotides in length and with a G/C content of at least 70%, X is a nucleotide sequence of 3 to 9 nucleotides in length, X comprises or consists of the sequence U/T-NNG, with N being any nucleotide selected from U(T), A, C, G, and Z is a nucleotide sequence with at least 80% complementarity to Y, the complementary nucleotides of Z being base paired with the nucleotides of Y, and X being a loop with no base pairing to Y or Z;

a second terminator signal comprising a hairpin structure defined by a sequence Y-X-Z, wherein Y is a nucleotide sequence of at least 8 nucleotides in length and with a G/C content of at least 40%, preferably at least 50%, especially at least 60%,
X is a nucleotide sequence of 3 to 9 nucleotides in length, and
Z is a nucleotide sequence with at least 70% complementarity to Y, the complementary nucleotides of Z being base paired with the nucleotides of Y, and X being a loop with no base pairing to Y or Z;

a third terminator signal comprising a hairpin structure defined by a sequence Y-X-Z, wherein Y is a nucleotide sequence of at least 8 nucleotides in length and with a G/C content of at least 40%, preferably at least 50%, especially at least 60%,
X is a nucleotide sequence of 3 to 9 nucleotides in length, and
Z is a nucleotide sequence with at least 70% complementarity to Y, the complementary nucleotides of Z being base paired with the nucleotides of Y, and X being a loop with no base pairing to Y or Z;

with the first, second and third terminator being at most 1000 nucleotides apart to provide a concerted termination by the first, second and third termination signal, wherein Y, X, and Z of the first, second or third terminator can be selected can be selected independently, in particular as defined in claims 1, 3, 4, 8 to 11; or

iv) comprising a first terminator signal comprising a hairpin structure defined by a sequence Y-X-Z, wherein Y is a nucleotide sequence

of at least 12 nucleotides in length and with a G/C content of at least 70%, X is a nucleotide sequence of 3 to 9 nucleotides in length, X comprises or consists of the sequence U/T-NNG, with N being any nucleotide selected from U(T), A, C, G, and Z is a nucleotide sequence with at least 80% complementarity to Y, the complementary nucleotides of Z being base paired with the nucleotides of Y, and X being a loop with no base pairing to Y or Z;

a second terminator signal comprising a class II termination signal, preferably comprising the consensus sequence TCTGTT, especially preferred comprising a sequence of at least 20 nucleotides in length with a T content of at least 40%;

a third terminator signal comprising a hairpin structure defined by a sequence Y-X-Z, wherein Y is a nucleotide sequence of at least 8 nucleotides in length and with a G/C content of at least 40%, preferably at least 50%, especially at least 60%,
X is a nucleotide sequence of 3 to 9 nucleotides in length, and
Z is a nucleotide sequence with at least 70% complementarity to Y, the complementary nucleotides of Z being base paired with the nucleotides of Y, and X being a loop with no base pairing to Y or Z; with the first, second and third terminator being at most 1000 nucleotides apart to provide a concerted termination by the first, second and third termination signal, wherein Y, X, and Z of the first or third terminator and the consensus and T-rich sequence of the class II termination signal can be selected independently, in particular as defined in claims 1, 3, 4, 8 to 11; or

v) comprising a first terminator signal comprising a hairpin structure defined by a sequence Y-X-Z, wherein Y is a nucleotide sequence of at least 12 nucleotides in length and with a G/C content of at least 70%, X is a nucleotide sequence of 3 to 9 nucleotides in length, X comprises or consists of the sequence U/T-NNG, with N being any nucleotide selected from U(T), A, C, G, and Z is a nucleotide sequence with at least 80% complementarity to Y, the complementary nucleotides of Z being base paired with the nucleotides of Y, and X being a loop with no base pairing to Y or Z;

a second terminator signal comprising a hairpin structure defined by a sequence Y-X-Z, wherein Y is a nucleotide sequence of at least 8 nucleotides in length and with a G/C content of at least 40%, preferably at least 50%, especially at least 60%,
X is a nucleotide sequence of 3 to 9 nucleotides in length, and
Z is a nucleotide sequence with at least 70% complementarity to

Y, the complementary nucleotides of Z being base paired with the nucleotides of Y, and X being a loop with no base pairing to Y or Z and a class II termination signal adjacent to the hairpin structure, preferably comprising the consensus sequence TCTGTT, especially preferred comprising a sequence of at least 20 nucleotides in length with a T content of at least 40%;

a third terminator signal comprising a hairpin structure defined by a sequence Y-X-Z, wherein Y is a nucleotide sequence of at least 8 nucleotides in length and with a G/C content of at least 40%, preferably at least 50%, especially at least 60%,
X is a nucleotide sequence of 3 to 9 nucleotides in length, and
Z is a nucleotide sequence with at least 70% complementarity to Y, the complementary nucleotides of Z being base paired with the nucleotides of Y, and X being a loop with no base pairing to Y or Z; with the first, second and third terminator being at most 1000 nucleotides apart to provide a concerted termination by the first, second and third termination signal, wherein Y, X, and Z of the first, second or third terminator and the consensus and T-rich sequence of the class II termination signal can be selected independently, in particular as defined in claims 1, 3, 4, 8 to 11.

17. A polynucleotide according to any one of claims 1 to 16, wherein at least one of the termination signals comprises a sequence selected from SEQ ID NO: 4, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18 or SEQ ID NO: 19.

18. An expression cassette comprising a polynucleotide sequence according to any one of claims 1 to 17 and a coding sequence, wherein the sequence is operatively positioned for termination of the coding sequence.

19. A vector comprising a polynucleotide as defined in any one of claims 1 to 17 or the expression cassette according to claim 18, the vector preferably being a plasmid.

20. A cell comprising a polynucleotide molecule with a gene operatively linked to the transcription termination signals of the polynucleotide according to any one of claims 1 to 17, preferably in a vector; preferably wherein said cell comprises the T7 RNA polymerase or a polynucleotide encoding the T7 RNA polymerase.

21. A method of producing a gene product, comprising providing a cell according to claim 20 and cultivating said cell under conditions allowing expression of said gene.

22. An in vitro method of producing a mRNA, comprising providing a polynucleotide according to any one of claims 1 to 17 and a coding sequence upstream of said transcription termination signal and contacting said polynucleotide with the T7 RNA polymerase.