CLAIMS

1. A nucleic acid sequence comprising a sequence of contiguous nucleotides, wherein said sequence of contiguous nucleotides has at least 90% sequence identity to the nucleic acid sequence of SEQ ID NO: 1,
and wherein said sequence of contiguous nucleotides encodes a single-chain BoNT/E1 protein.

2. The nucleic acid sequence according to Claim 1, wherein the nucleic acid has a maximum of 160 slow codons.

3. The nucleic acid sequence according to Claim 1 or Claim 2, wherein said single-chain BoNT/E1 protein comprises a sequence of contiguous amino acids, and wherein said sequence of contiguous amino acids has at least 95% sequence identity to the amino acid sequence of SEQ ID NO: 2.

4. The nucleic acid sequence according to any one of Claims 1-3, with the proviso that the single-chain BoNT/E1 according to Claim 1 or Claim 2 or said sequence of contiguous amino acids according to Claim 3 includes one or more of the following amino acids (wherein the amino acid position numbering starts with the N-terminal amino acid residue and ends with the C-terminal amino acid residue of the BoNT/E1 protein):
glycine at position 177
serine at position 198
alanine at position 340
leucine at position 773
leucine at position 963
glutamine at position 964
alanine at position 967
asparagine at position 1195.

5. The nucleic acid sequence according to any preceding claim, wherein said sequence of contiguous nucleotides has at least 785 synonymous codons when compared to the nucleic acid sequence of wild-type BoNT/E1 (SEQ ID NO: 3).

6. A method for producing soluble single-chain BoNT/E1 protein in an *E. coli* host cell, said method comprising:
expressing a nucleic acid sequence according to any preceding claim in an *E. coli* expression system.

7. The method according to Claim 6, wherein said soluble single-chain BoNT/E1 protein is expressed in the cytoplasm of said *E. coli* host cell.

8. The method according to Claim 6 or Claim 7, wherein said soluble single-chain BoNT/E1 protein is expressed at a level of at least 5 mg/L.

9. The method according to any of Claims 6-8, comprising lysis of the *E. coli* host cell to provide an *E. coli* host cell homogenate containing said soluble single-chain BoNT/E1 protein.

10. A method for producing soluble di-chain BoNT/E1 protein, said method comprising:

providing a soluble single-chain BoNT/E1 protein comprising a sequence of contiguous amino acids, wherein said sequence of contiguous amino acids has at least 95% sequence identity to the amino acid sequence of SEQ ID NO: 2,

contacting said BoNT/E1 protein with trypsin in solution, and

separating the soluble BoNT/E1 protein from the trypsin by contacting the solution containing soluble BoNT/E1 protein and trypsin with a hydrophobic surface, wherein the soluble BoNT/E1 protein preferentially binds to the hydrophobic surface.

11. The method according to Claim 10, with the proviso that said sequence of contiguous amino acids includes one or more of the following amino acids (wherein the amino acid position numbering starts with the N-terminal amino acid residue and ends with the C-terminal amino acid residue of the BoNT/E1 protein):
glycine at position 177
serine at position 198
alanine at position 340
leucine at position 773
leucine at position 963
glutamine at position 964

alanine at position 967
asparagine at position 1195.

12. The method according to Claim 10 or Claim 11, wherein the soluble single-chain BoNT/E1 protein is provided by the method of any of Claims 6-9.

13. The method according to any one of Claims 10-12, wherein the hydrophobic surface is an inert matrix to which a ligand consisting of aryl or alkyl groups is attached.

14. The method according to Claim 13, wherein the hydrophobic surface is selected from the group consisting of: butyl, phenyl or octyl ligands.

15. An active di-chain BoNT/E1 protein,
wherein the first chain comprises a sequence of contiguous amino acids, and wherein said sequence of contiguous amino acids has at least 95% sequence identity to the amino acid sequence of positions 1-419 of SEQ ID NO: 2;
wherein the second chain comprises a sequence of contiguous amino acids, and wherein said sequence of contiguous amino acids has at least 95% sequence identity to the amino acid sequence of positions 423-1252 of SEQ ID NO: 2;
wherein the first and second chains are joined together by a disulphide bond between cysteine 412 on the first chain and cysteine 426 on the second chain;
with the proviso that said sequence of contiguous amino acids includes one or more of the following amino acids (wherein the amino acid position numbering starts with the N-terminal amino acid residue and ends with the C-terminal amino acid residue of the BoNT/E1 protein):
glycine at position 177
serine at position 198
alanine at position 340
leucine at position 773
leucine at position 963
glutamine at position 964
alanine at position 967
asparagine at position 1195.

16. An active di-chain BoNT/E1 protein obtainable by the method according to any of Claims 10-14.

17. A composition comprising an active di-chain BoNT/E1 protein according to Claim 15 or Claim 16, or an active di-chain BoNT/E1 protein obtainable by proteolytic cleavage of the single-chain BoNT/E1 protein according to any of Claims 6-9; wherein said composition is substantially free from trypsin.

18. The composition according to Claim 17, where said composition contains less than 10 pg trypsin per 100 ng BoNT/E1 protein, or less than 7 pg trypsin per 100 ng BoNT/E1 protein, or less than 5 pg trypsin per 100 ng BoNT/E1 protein.

19. A liquid pharmaceutical composition comprising:

an active di-chain BoNT/E1 protein according to Claim 15 or Claim 16, or an active di-chain BoNT/E1 protein obtainable by proteolytic cleavage of the single-chain BoNT/E1 protein according to any of Claims 6-9;

a non-protein stabilising agent that is a surfactant; and

water;

wherein said liquid pharmaceutical composition does not comprise a protein stabilising agent; and

wherein said liquid pharmaceutical composition is substantially free from trypsin (e.g. said liquid pharmaceutical composition contains less than 10 pg trypsin per 100 ng BoNT/E1 protein, or less than 7 pg trypsin per 100 ng BoNT/E1 protein, or less than 5 pg trypsin per 100 ng BoNT/E1 protein).

20. The liquid pharmaceutical composition according to claim 19, wherein said liquid pharmaceutical composition further comprises:

sodium chloride;

a buffer to maintain pH between 5.5 and 7.5, and

a disaccharide;

wherein the water is sterile water.

21. An active di-chain BoNT/E1 protein according to Claim 15 or Claim 16, or an active di-chain BoNT/E1 protein obtainable by proteolytic cleavage of the single-chain BoNT/E1

protein according to any of Claims 6-9, or a composition according to Claim 17 or Claim 18, or a liquid pharmaceutical composition according to Claim 19 or Claim 20, for use in therapy.

CLAIMS

1. A method for producing soluble di-chain BoNT/E1 protein, said method comprising:

providing a soluble single-chain BoNT/E1 protein comprising a sequence of contiguous amino acids, wherein said sequence of contiguous amino acids has at least 95% sequence identity to the amino acid sequence of SEQ ID NO: 2,

contacting said BoNT/E1 protein with trypsin in solution, and

separating the soluble BoNT/E1 protein from the trypsin by contacting the solution containing soluble BoNT/E1 protein and trypsin with a hydrophobic surface, wherein the soluble BoNT/E1 protein preferentially binds to the hydrophobic surface.

2. The method according to Claim 1, with the proviso that said sequence of contiguous amino acids includes one or more of the following amino acids (wherein the amino acid position numbering starts with the N-terminal amino acid residue and ends with the C-terminal amino acid residue of the BoNT/E1 protein):
glycine at position 177
serine at position 198
alanine at position 340
leucine at position 773
leucine at position 963
glutamine at position 964
alanine at position 967
asparagine at position 1195.

3. The method according to Claim 1 or Claim 2, wherein the soluble single-chain BoNT/E1 protein is provided by a method comprising expressing in an *E. coli* expression system a nucleic acid sequence comprising a sequence of contiguous nucleotides, wherein said sequence of contiguous nucleotides has at least 90% sequence identity to the nucleic acid sequence of SEQ ID NO: 1, and wherein said sequence of contiguous nucleotides encodes a single-chain BoNT/E1 protein.

4. The method of claim 3, wherein the nucleic acid has a maximum of 160 slow codons.

5. The method of claim 3 or claim 4, wherein said single-chain BoNT/E1 protein comprises a sequence of contiguous amino acids, and wherein said sequence of contiguous amino acids has at least 95% sequence identity to the amino acid sequence of SEQ ID NO: 2.

6. The method of any one of claims 3 to 5, wherein said sequence of contiguous nucleotides has at least 785 synonymous codons when compared to the nucleic acid sequence of wild-type BoNT/E1 (SEQ ID NO: 3).

7. The method according to any one of Claims 1 to 6, wherein the hydrophobic surface is an inert matrix to which a ligand consisting of aryl or alkyl groups is attached.

8. The method according to Claim 7, wherein the hydrophobic surface is selected from the group consisting of: butyl, phenyl or octyl ligands.

9. An active di-chain BoNT/E1 protein,
wherein the first chain comprises a sequence of contiguous amino acids, and wherein said sequence of contiguous amino acids has at least 95% sequence identity to the amino acid sequence of positions 1-419 of SEQ ID NO: 2;
wherein the second chain comprises a sequence of contiguous amino acids, and wherein said sequence of contiguous amino acids has at least 95% sequence identity to the amino acid sequence of positions 423-1252 of SEQ ID NO: 2;
wherein the first and second chains are joined together by a disulphide bond between cysteine 412 on the first chain and cysteine 426 on the second chain;
with the proviso that said sequence of contiguous amino acids includes one or more of the following amino acids (wherein the amino acid position numbering starts with the N-terminal amino acid residue and ends with the C-terminal amino acid residue of the BoNT/E1 protein):
glycine at position 177
serine at position 198
alanine at position 340
leucine at position 773
leucine at position 963
glutamine at position 964
alanine at position 967
asparagine at position 1195,

and wherein said sequence of contiguous amino acids does not include residues 420, 421 and 422 of SEQ ID NO: 2.

10. A composition comprising an active di-chain BoNT/E1 protein according to Claim 9, wherein said composition contains less than 10 pg trypsin per 100 ng BoNT/E1 protein, or less than 7 pg trypsin per 100 ng BoNT/E1 protein, or less than 5 pg trypsin per 100 ng BoNT/E1 protein.

11. A liquid pharmaceutical composition comprising:

an active di-chain BoNT/E1 protein according to Claim 9;

a non-protein stabilising agent that is a surfactant; and

water;

wherein said liquid pharmaceutical composition does not comprise a protein stabilising agent; and

wherein said liquid pharmaceutical composition contains less than 10 pg trypsin per 100 ng BoNT/E1 protein, or less than 7 pg trypsin per 100 ng BoNT/E1 protein, or less than 5 pg trypsin per 100 ng BoNT/E1 protein).

12. The liquid pharmaceutical composition according to claim 11, wherein said liquid pharmaceutical composition further comprises:

sodium chloride,

a buffer to maintain pH between 5.5 and 7.5, and

a disaccharide;

wherein the water is sterile water.

13. An active di-chain BoNT/E1 protein according to Claim 9, or a composition according to Claim 10, or a liquid pharmaceutical composition according to Claim 11 or Claim 12, for use in therapy.