

A polypeptide

Field of the invention

The invention relates to a polypeptide, a corresponding nucleic acid molecule, a
5 construct and/or a vector and/or a cell comprising such nucleic acid molecule and/or a
composition comprising said polypeptide, nucleic acid molecule, construct, vector
and/or cell. The invention further relates to such polypeptide, corresponding nucleic
acid molecule, construct and/or vector and/or cell comprising such nucleic acid
10 molecule and/or composition for medical use, preferably for use in treating an
infectious disease. Furthermore, the invention relates to the use of said polypeptide,
nucleic acid molecule, construct, vector, cell and/or composition as an antimicrobial,
preferably as a food additive or disinfectant, or for detecting bacteria, preferably in a
diagnostic application.

Background of the invention

Staphylococcus aureus is a major human pathogen frequently implicated in
several serious infectious diseases and food poisoning. Its treatment becomes more and
more difficult because of emerging antibiotic resistant strains. Endolysins from phages
infecting *Staphylococcus aureus* have been shown to potentially control these
20 pathogens and can be used for their specific detection. In most cases, major obstacles in
the application of endolysins targeting *Staphylococcus* species are low enzyme activity,
difficult production in large quantities and/or protein stability.

There is always a need for new antimicrobials with improved characteristics on
for example antimicrobial activity and/or stability.

25

Summary

In a first aspect, there is provided a nucleic acid molecule comprising a first
nucleotide sequence, wherein said nucleotide sequence has at least 85% sequence
identity with SEQ ID NO: 12,
30 wherein said first nucleotide sequence encodes a cell wall-binding domain
binding the peptidoglycan cell wall of *Staphylococcus* genera,

wherein said nucleic acid molecule has at least 85% sequence identity with SEQ ID NO: 1 without consisting of SEQ ID NO: 1, or wherein said nucleic acid molecule further comprises a heterologous nucleotide sequence encoding a lytic domain.

In a second aspect, there is provided a polypeptide encoded by a nucleic acid molecule as defined above.

In a third aspect, there is provided a nucleic acid construct comprising a nucleic acid molecule as defined above.

In a fourth aspect, there is provided an expression vector comprising a nucleic acid construct as defined above operably linked to one or more control sequences, which direct the production or expression of the encoded polypeptide in a cell, a subject, or a cell-free expression system.

In a fifth aspect, there is provided a cell comprising the nucleic acid construct as defined above or the expression vector as defined above, said cell being a microbial, prokaryotic or eukaryotic cell.

In a sixth aspect, there is provided a method for producing, optionally purifying and optionally freeze-drying a polypeptide as defined above, said method comprising the steps of:

- i) producing said polypeptide in a cell as defined above comprising a nucleic acid construct as defined above, optionally
- ii) purifying said polypeptide, and optionally
- iii) freeze-drying said purified polypeptide.

In a seventh aspect, there is provided a method for producing a polypeptide with an enhanced lytic activity by treating a polypeptide as defined above and/or a polypeptide as obtainable by the method as defined above, wherein said method comprises the steps of:

- i) dialyzing said polypeptide against a buffer comprising a chelating compound,
- ii) dialyzing said polypeptide against a divalent metal ion-containing buffer, preferably said divalent metal ion being selected from the group consisting of Co^{2+} , Cu^{2+} , Mn^{2+} and Zn^{2+} .

In an eighth aspect, there is provided a composition comprising a nucleic acid molecule as defined above, a polypeptide as defined above and/or as obtainable by a method as defined above, and/or a nucleic acid construct as defined above and/or a vector as defined above and/or a cell as defined above.

In a ninth aspect, there is provided a composition as defined above for use as a medicament.

In a tenth aspect, there is provided the use of a composition as defined above in the preparation of a medicament for the treatment of an infectious disease.

5 In an eleventh aspect, there is provided the use of a nucleic acid molecule as defined above, a polypeptide as defined above, and/or as obtainable by a method as defined above, a nucleic acid construct as defined above, a vector as defined above, a cell as defined above and/or a composition as defined above in the manufacture of an antimicrobial.

10 In a twelfth aspect, there is provided the use of a nucleic acid molecule as defined above, a polypeptide as defined above, and/or as obtainable by a method as defined above, a nucleic acid construct as defined above, a vector as defined above, a cell as defined above and/or a composition as defined above as a food additive or a disinfectant.

15 In a thirteenth aspect, there is provided the use of a nucleic acid molecule as defined above, a polypeptide encoded by such nucleic acid molecule, a nucleic acid construct and/or vector and/or cell comprising such nucleic acid construct or a composition comprising any of such polypeptide, nucleic acid molecule, nucleic acid, vector and/or cell for detecting a *Staphylococcus* in food or on a surface.

20

Description of the invention

Reported here is the newly characterised Ply2638, the endolysin of *S. aureus* bacteriophage Φ 2638a. The enzyme and several engineered derivatives were expressed in a soluble way in *E. Coli* and showed surprising stability after lyophilisation as
 25 proven by their lytic activity after reconstitution. In addition to a cell wall-binding domain which binds the cell wall of *Staphylococcus* genera, we showed that two functional domains, i.e. an M23 endopeptidase domain and an amidase domain, are crucial for optimal lytic activity.

We showed that retrofitting of the enzyme with catalytic domains and/or duplication of the cell wall-binding domain originating from *S. aureus* Φ 11 endolysin, Φ Twort endolysin, and Lysostaphin resulted in a heterologous polypeptide fusion product with an enhanced lytic activity and/or a shifted pH optimum and/or an increased stability after lyophilisation and reconstitution.

Nucleic acid molecule

In a first aspect, there is provided a nucleic acid molecule comprising or consisting of a first nucleotide sequence, said nucleotide sequence having 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO: 12 (referred is to Table 1 for an overview of all SEQ ID NOs used herein). Preferably, said first nucleotide sequence has a length of at least 282, 285, 290, 300, 310, 320, 330, 340, 350, 360 or 370 nucleotides, more preferably, at least 381 nucleotides and/or a length of at most 510, 480, 450, 420 or 390 nucleotides. Preferably, said nucleic acid molecule has a length of at least 282, 285, 290, 300, 310, 320, 330, 340, 350, 360, 370 nucleotides, more preferably at least 381 nucleotides and/or a length of at most 4500, 4200, 3900 and/or 3600 nucleotides. Preferably, said nucleic acid molecule has a length of at least 1200, 1230, 1260, 1290, 1320, 1350, 1380, 1410, 1440, 1470, 1500, 1530 or 1560 nucleotides and/or a length of at most 4500, 4200, 3900 and/or 3600 nucleotides. Also preferred is a nucleic acid molecule according to the invention with a length of at least 1890, 1920, 1950, 1980, 2010, 2040, 2070, 2100, 2130 or 2160 nucleotides and/or a length of at most 4500, 4200, 3900 and/or 3600 nucleotides. Preferably, said first nucleotide sequence encodes a cell wall-binding domain which binds the peptidoglycan cell wall of *Staphylococcus* genera. Preferably, said first nucleotide sequence originates from *S. aureus* bacteriophage Φ 2638a endolysin.

As estimated from alignments with the crystal structure of the C-terminal 92 residues of ALE-1 (Lu et al., J. Biol. Chem., 2006, 281(1):549-58), it was estimated that a minimum of 94 amino acids from the cell wall-binding domain originating from *S. aureus* bacteriophage Φ 2638a endolysin may be sufficient to direct the enzyme to the cell wall of *Staphylococcus* genera.

Binding of a domain to the peptidoglycan cell wall of *Staphylococcus* genera may be assessed using assays well known to the artisan. In a preferred embodiment, an immunohistochemical technique and/or a gene fusion technique resulting in labelled

constructs are used for assessing specific binding of peptides, polypeptides or proteins to the peptidoglycan cell wall of *Staphylococcus* genera. Quantification methods of signals used in the above mentioned immunohistochemical or fusion techniques are well known in the art.

5 In one embodiment, *Staphylococcus* peptidoglycan cell wall-binding can be quantified using a fluorescent fusion construct comprising a polypeptide comprising a domain encoded by a first nucleotide sequence. Such a cell wall-binding assay is described in detail by Loessner et al (Molecular Microbiology 2002, 44(2): 335–349) and in Example 1. In this assay a solution comprising said fluorescent fusion construct
10 or a negative control, preferably Green Fluorescent Protein (GFP), is subjected to *Staphylococcus* cells, preferably *S. aureus* cells, more preferably *S. aureus* BB255 for an indicated time period where after the cells are sedimented by centrifugation together with the bound fluorescent fusion constructs. The fluorescent signal of the *Staphylococcus* cells exposed to a fluorescent fusion construct subtracted by the
15 fluorescence signal of the *Staphylococcus* cells exposed to a negative control, preferably GFP, is a measure for cell binding as meant in this disclosure.

Preferably, within the context of the invention a nucleic acid molecule will be said to encode a polypeptide domain that binds the peptidoglycan cell wall of *Staphylococcus* genera when using this assay an increase in fluorescent signal of the
20 sedimented cells above the negative control as defined herein is detected. The binding is preferably said to be specific. Preferably, the invention relates to a nucleic acid molecule encoding a polypeptide or a domain which exhibits binding as defined herein of at least 50, 60, 70, 80, 90 or 100, 150 or 200% of peptidoglycan cell wall binding of *S. aureus* bacteriophage Φ 2638a endolysin (Ply2638) encoded by SEQ ID NO: 1.

25 In an embodiment, the invention relates to a nucleic acid molecule that has 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO: 1 encoding for *S. aureus* bacteriophage Φ 2638 endolysin.

The present invention further relates to a nucleic acid molecule comprising in
30 addition to said first nucleotide sequence, a heterologous nucleotide sequence encoding a lytic domain. Preferably, said lytic domain exhibits peptidoglycan hydrolase activity as defined later herein. Said nucleic acid molecule comprising heterologous nucleotide sequences being defined herein as a “retrofitted construct”.

As used herein the term "heterologous sequence" or "heterologous nucleic acid" is one that is not naturally found operably linked as neighbouring sequence of said first nucleotide sequence. As used herein, the term "heterologous" may mean "recombinant". "Recombinant" refers to a genetic entity distinct from that generally found in nature. As applied to a nucleotide sequence or nucleic acid molecule, this means that said nucleotide sequence or nucleic acid molecule is the product of various combinations of cloning, restriction and/or ligation steps, and other procedures that result in the production of a construct that is distinct from a sequence or molecule found in nature.

10 A "peptidoglycan hydrolase activity" herein also defined as a "lytic activity" can be assessed by methods well known to the artisan. In an embodiment, lytic activity can be assessed spectrophotometrically by measuring the drop in turbidity of substrate cell suspensions. Preferably, lytic activity can be assessed spectrophotometrically measuring the drop in turbidity of a *S. aureus* suspension, wherein turbidity is
15 quantified by measuring OD₅₉₅ spectrophotometrically (Libra S22, Biochrom). More preferably, 200 nM of a polypeptide encoded by a nucleic acid molecule as identified herein is incubated together with an *S. aureus* suspension having an initial OD₆₀₀ of 1 ± 0.05 , as assessed spectrophotometrically (Libra S22, Biochrom), in PBS buffer pH 7.4, 120 mM sodium chloride for 30 min at 37°C. The drop in turbidity is calculated by
20 subtracting the OD₅₉₅ after 30 min of incubation from the OD₅₉₅ before 30 min of incubation. Within the context of the invention a nucleic acid molecule will be said to comprise a nucleic acid sequence encoding a lytic domain when using this assay a drop in turbidity of at least 10, 20, 30, 40, 50 or 60% is detected. Preferably, a drop of at least 70% is detected. Preferably, the invention relates to a nucleic acid molecule
25 encoding a polypeptide which exhibits a lytic activity of at least 50, 60, 70, 80, 90, 100, 150 or 200% or more of a lytic activity of *S. aureus* bacteriophage Φ 2638a endolysin (Ply2638) encoded by SEQ ID NO: 1.

In an embodiment a nucleic acid molecule of the invention may not comprise or consist of SEQ ID NO:1. SEQ ID NO: 1 encodes for *S. aureus* bacteriophage Φ 2638
30 endolysin.

A preferred embodiment encompasses a nucleic acid molecule comprising said first nucleotide sequence as identified herein and further comprising as a lytic domain a second and third nucleotide sequences, wherein said second sequence encodes an

endopeptidase domain and third nucleotide sequence encodes an amidase domain. Accordingly, the invention relates to a nucleic acid molecule comprising said first nucleotide sequence, wherein said nucleic acid molecule has 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO: 1, or wherein said nucleic acid molecule further comprises a heterologous nucleotide sequence encoding a lytic domain.

An endopeptidase domain as used herein preferably cleaves the pentaglycine cross-bridges (Trayer, H. R. and Buckley, C. E. (1970) *Molecular properties of lysostaphin, a bacteriolytic agent specific for Staphylococcus aureus*. *J. Biol. Chem.* 245, 4842–4846) that are found in the cell wall of *Staphylococcus* genera, preferably in the cell wall of *S. aureus*, *S. simulans* and *S. carnosus*. An amidase domain as used herein preferably hydrolyzes gamma-glutamyl-containing substrates. The functionality and activity of these domains in a polypeptide can be confirmed by characterizing the cleavage products upon incubation of said polypeptides containing any of these domains with purified peptidoglycan. Preferably, each of the nucleotide sequences encoding the second or third domain is of bacterial or bacteriophage origin. In a preferred embodiment, said second and third nucleotide sequences originate from a gene encoding for an enzyme selected from the group consisting of *S. aureus* bacteriophage Φ 2638a endolysin, *S. aureus* bacteriophage Φ 11 endolysin, *S. aureus* bacteriophage Φ Twort endolysin and *S. Simulans* lysostaphin. Preferably, said second nucleotide sequence has 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO: 14 or 15 and said third nucleotide sequence has 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO: 17 or 18.

The invention encompasses all constructs as defined herein containing the functional domains as meant in the invention at any possible location within the construct. In a preferred embodiment, a nucleic acid molecule as defined herein encodes for a polypeptide with a C-terminal domain encoded by a first nucleic acid sequence as identified herein, which is shown herein to encode for functional polypeptides able to target for *Staphylococcus* genera. Even more preferred is a nucleic acid molecule as defined herein comprising a nucleic acid molecule that has 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO: 9. SEQ ID NO: 9 comprises a first nucleotide sequence

encoding a C-terminal SH3b homologue cell wall-binding domain (both domains from Ply2638 encoded by SEQ ID NO: 1: Leu138 - Lys486), a second nucleotide sequence encoding a polypeptide comprising an N-terminal M23 glycyl-glycine endopeptidase homologue domain (mature Lysostaphin encoded by SEQ ID NO: 33: Ala1 - Gly154) and a third nucleotide sequence encoding a central amidase-2 homologue domain. It has a theoretical size of 58.266 kDa. A polypeptide encoded by SEQ ID NO: 9 differs from *S. aureus* bacteriophage Φ 2638a endolysin in that the N-terminal M23 endopeptidase domain is substituted by an M23 endopeptidase domain from *S. Simulans* lysostaphin. We showed here that a polypeptide encoded by SEQ ID NO: 9 demonstrated at least 20% increased lytic activity as compared to *S. aureus* bacteriophage Φ 2638a endolysin while the lytic activity is maintained after lyophilisation and reconstitution. In a preferred embodiment, a nucleic acid molecule comprising said first, second and third nucleotide sequences encodes for a polypeptide exhibiting a lytic activity of at least 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9 or 2 fold as compared to a lytic activity of *S. aureus* bacteriophage Φ 2638a endolysin encoded by SEQ ID NO:1. In a preferred embodiment, a nucleic acid molecule comprising said first, second and third nucleotide sequences encodes for a polypeptide exhibiting a decrease in lytic activity of at most 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10% after lyophilisation and reconstitution as defined herein as compared to freshly prepared polypeptide, wherein "freshly prepared" is preferably defined herein as at most 2 days storage at 1.63 mg/mL in lyophilisation buffer (50mM Tris, 500mM sucrose, 200mM mannitol, 0.05% polysorbate 20 + 50% glycerol) at -20 °C and thawed immediately before assessing lytic activity in an assay as identified herein.

Lyophilisation and reconstitution is defined herein as dehydration by freeze-drying and subsequent reconstitution of the sample by adding water. In an embodiment, lyophilisation and reconstitution may be done by dialyzing against 3 changes of 300 ml lyophilization buffer (50 mM phosphate or Tris, 500 mM sucrose, 200 mM mannitol, pH 7.4) aliquot and freezing in the gaseous phase of liquid nitrogen. The freeze-drying can be done under standard conditions, preferably at -40°C and vacuum at 75 mTorr for 60 minutes, followed by increasing temperature during 5 hours to -10°C and another 60 minutes at -10°C at the same vacuum levels. As final step, temperature is preferably increased to 25°C during 10 hours. Samples are reconstituted by the addition of water.

In another preferred embodiment, a nucleic acid molecule as defined herein comprises in addition to the above identified first, second and third nucleotide sequences at least one duplicate identical or heterologous first, second and/or third nucleotide sequence. Preferably, a nucleic acid molecule as defined herein comprises in addition to said first, second and third nucleotide sequences, a duplicate identical first nucleotide sequence. We showed here that duplication of the first nucleotide sequence as defined herein encoding a cell wall-binding domain results in a polypeptide preferably as encoded by SEQ ID NO: 20 which exposes at least 5, 10, 20, 30, 20 or 40% increased lytic activity as compared to a lytic activity of a reference polypeptide encoded by a nucleic acid molecule lacking such duplicate first domain or as compared to a lytic activity of *S. aureus* bacteriophage Φ 2638a endolysin encoded by SEQ ID NO:1 as assessed in an assay as identified herein, more specifically using equimolar amount of a polypeptide and a modified PBS buffer containing 200, 300, 400, or 1000nM NaCl. This embodiment also encompasses a heterologous nucleic acid molecule in which said first, second and third nucleotide sequences originate from the same source being *S. aureus* bacteriophage Φ 2638a, said nucleic acid molecule comprising a sequence that has 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO:1 and an additional duplicate identical or heterologous first, second and/or third nucleotide sequence. Also preferred is a nucleic acid molecule as defined herein comprising a nucleotide sequence that has 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO: 6 and 20.

In another preferred embodiment, a nucleic acid molecule as defined herein comprises a fourth nucleotide sequence encoding a CHAP (cysteine, histidine-dependent amidohydrolases/peptidases) domain. More preferably, said fourth nucleotide sequences originates from *S. aureus* bacteriophage Φ 11 or *S. aureus* bacteriophage Φ Twort endolysin. Even more preferably, said fourth nucleotide sequence has 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO: 18 or SEQ ID NO: 19. Preferably, a nucleic acid molecule as defined herein comprises a nucleotide sequence that has 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO: 5 or 7. We showed here that a molecule comprising said first, second, third and fourth nucleotide sequences as defined by SEQ

ID NO: 5 encodes for a polypeptide exhibiting an increased lytic activity and/or a shifted, preferably decreased pH optimum as compared to a polypeptide encoded by a construct lacking said fourth domain and/or compared to *S. aureus* bacteriophage Φ 2638a endolysin encoded by SEQ ID NO:1. Lytic activity was assessed

5 spectrophotometrically and under the conditions as earlier defined herein. Preferably, said polypeptide exhibits an increase in a lytic activity of at least 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, or 2 fold as compared to a lytic activity of a polypeptide encoded by a reference polypeptide differing from said polypeptide only in lacking said fourth domain or as compared to a lytic activity of *S. aureus* bacteriophage

10 Φ 2638a endolysin encoded by SEQ ID NO:1. Even more preferably, said polypeptide exhibits an increase in a lytic activity of at least 2.5 fold as compared to the defined reference polypeptide or a polypeptide encoded by SEQ ID NO:1.

A shifted or decreased pH optimum is defined herein as a shift or decrease in optimal lytic activity to a lower pH value, where ionic strength is kept constant. Lytic

15 activity is preferably assessed spectrophotometrically as defined herein. Preferably, the pH optimum of a lytic activity is decreased 0.5 – 1 pH unit as compared to a lytic activity of a polypeptide encoded by a reference polypeptide differing from said polypeptide only in lacking said fourth domain or as compared to a lytic activity of *S. aureus* bacteriophage Φ 2638a endolysin encoded by SEQ ID NO:1.

20 The invention preferably relates to a nucleic acid molecule comprising a first, second, third and optionally fourth nucleotide sequences as identified herein encoding a polypeptide which has the same lytic activity and/or the same pH optimum or which has an increased lytic activity and/or a decreased pH optimum as compared to a lytic activity of *S. aureus* bacteriophage Φ 2638a endolysin encoded by SEQ ID NO:1. The

25 same being identified herein as no detectable difference when using the assay as identified herein or a method well known by the artisan. The current invention also relates to a nucleic acid molecule comprising a first, second, and fourth nucleotide sequence as identified herein encoding a polypeptide which has the same lytic activity and/or the same pH optimum or which has an increased lytic activity and/or a

30 decreased pH optimum as compared to a lytic activity of *S. aureus* bacteriophage Φ 2638a endolysin encoded by SEQ ID NO:1. The current invention also relates to a nucleic acid molecule comprising a first, third, and fourth nucleotide sequence as identified herein encoding a polypeptide which has the same lytic activity and/or the

same pH optimum or which has an increased lytic activity and/or a decreased pH optimum as compared to a lytic activity of *S. aureus* bacteriophage Φ 2638a endolysin encoded by SEQ ID NO:1. Each of the nucleotide sequences identified herein, i.e. first, second, third, fourth nucleotide sequences, encoding the individual domain of a polypeptide defined herein can be assembled by any usual method known for constructing and assembling nucleic acid fragments which are well known to those skilled in the art and widely described in the literature (Sambrook, Maniatis et al. (1989) and illustrated experimental part of the disclosure. In a preferred embodiment, a first, second, third and/or fourth nucleotide sequences are operably linked together.

Accordingly, a nucleic acid molecule of the invention encodes a polypeptide, preferably a polypeptide as identified herein which is able to bind *Staphylococcus* genera via the cell wall-binding domain encoded by a first nucleotide sequence as defined herein and/or lyse said bacteria via an endopeptidase and/or amidase domain and optionally a CHAP domain encoded by a second, third and fourth nucleotide sequence, respectively, as defined herein.

In a preferred embodiment, a nucleic acid molecule of the invention as defined herein optionally comprises a sequence encoding a tag for ease of purification. Preferably, said tag is selected from, but is not limited to, the group consisting of a FLAG-tag, poly(His)-tag, HA-tag and Myc-tag. More preferably said tag is a 6xHis-tag. Even more preferably, said tag is an N-terminal 6xHis-tag identical to SEQ ID NO: 43.

Polypeptide

In a further aspect, there is provided a polypeptide encoded by a nucleic acid molecule as earlier identified herein. This polypeptide comprises a cell wall-binding domain and preferably an endopeptidase domain and/or an amidase domain as defined in the previous section.

A polypeptide domain encompassed by the current invention preferably has 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO: 35, 36, 37, 38, 39, 40, 41 and/or 42. Preferably, a polypeptide domain encompassed by the current invention preferably comprises one or more putative linkers and has 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93,

94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO: 51, 52, 53, 54, 55, 56, 57 and/or 58.

A polypeptide encompassed by the current invention preferably has 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence
5 identity with SEQ ID NO: 21, 25, 26, 27, 29 and/or 32. More preferably, a polypeptide encompassed by the current invention preferably has 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO: 25, 26, 27, 29 and/or 32 with SEQ ID NO: 25, 26, 27, 29 and/or 32.

In an embodiment of the current invention, a polypeptide preferably has at least
10 80% sequence identity with SEQ ID NO: 21, 25, 26, 27, 29, 32, 35, 36, 37, 38, 39, 40, 41, and/or 42 encoded by a nucleic acid construct with at least 80% identity with SEQ ID NO: 1, 5, 6, 7, 9, 20, 12, 13, 14, 15, 16, 17, 18 and/or 19, respectively. More preferably, a polypeptide of the current invention has at least 80% sequence identity with SEQ ID NO: 25, 26, 27, 29, 32, 35, 36, 37, 38, 39, 40, 41, and/or 42 encoded by a
15 nucleic acid construct with at least 80% identity with SEQ ID NO: 5, 6, 7, 9, 20, 12, 13, 14, 15, 16, 17, 18 and/or 19, respectively.

A polypeptide according to the invention may have a length of at least 94, 95, 96, 100, 110 or 120 amino acids, preferably 127 amino acids and/or at most 1500, 1400, 1300 or 1200 amino acids. Preferably, said polypeptide has a length of at least 400,
20 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510 or 520 amino acids and/or at most 1500, 1400, 1300 or 1200 amino acids. Also preferred is a polypeptide according to the invention with a length of at least 630, 640, 650, 660, 670, 680, 690, 700, 710 or 720 amino acids and/or at most 1500, 1400, 1300 or 1200 amino acids.

An amino acid or nucleotide sequence, encompassed by the present invention,
25 may be derived from one of the sequences as identified herein by substituting, inserting, deleting, or adding one, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20 or more nucleotides or amino acids, respectively. An amino acid sequence, encompassed by the present invention, may be derived from one of the sequences as identified herein by adding an additional N- or C- terminal amino acids or chemical moieties to increase
30 stability, solubility and activity.

An embodiment of the invention encompasses a variant polypeptide. A variant polypeptide may be a non-naturally occurring form of the polypeptide. A polypeptide variant may differ in some engineered way from the polypeptide isolated from its

native source. A variant may be made by site-directed mutagenesis starting from the nucleotide sequence of SEQ ID NO: 1, 5, 6, 7, 9, 12, 13, 14, 15, 16, 17, 18, 19 and/or 20. Preferably, a polypeptide variant contains mutations that do not alter the biological function of the encoded polypeptide. According to a preferred embodiment, a

5 polypeptide variant exhibits *Staphylococcus* peptidoglycan cell wall-binding and/or a lytic activity which is the same or enhanced as compared to the *Staphylococcus* peptidoglycan cell-wall binding and/or lytic activity of *S. aureus* bacteriophage Φ 2638a endolysin encoded by SEQ ID NO:1. A polypeptide variant of the invention preferably is a variant of SEQ ID NO: 21, 25, 26, 27, 29, 32, 35, 36, 37, 38, 39, 40, 41

10 and/or 42. A polypeptide variant with the same or an enhanced *Staphylococcus* peptidoglycan cell-wall binding and/or lytic activity is a polypeptide exhibiting a *Staphylococcus aureus* peptidoglycan cell-wall binding and/or lytic activity, which is the same or increased compared to the *Staphylococcus* peptidoglycan cell-wall binding and/or lytic activity of *S. aureus* bacteriophage Φ 2638a endolysin encoded by SEQ ID

15 NO:1 measured in an assay as earlier identified herein.

According to another preferred embodiment, a nucleotide sequence of the invention is a variant of the nucleotide sequences of SEQ ID NO: 1, 5, 6, 7, 9, 12, 13, 14, 15, 16, 17, 18, 19 and/or 20. Nucleotide sequence variants may be used for preparing a polypeptide variant as defined earlier. A nucleic acid variant may be a

20 fragment of any of the nucleotide sequences as defined above. A nucleic acid variant may also be a nucleotide sequence that differs from SEQ ID NO: 1, 5, 6, 7, 9, 12, 13, 14, 15, 16, 17, 18, 19 and/or 20 by virtue of the degeneracy of the genetic code. A nucleic acid variant may also be an allelic variant of SEQ ID NO: 1, 5, 6, 7, 9, 12, 13, 14, 15, 16, 17, 18, 19 and/or 20. An allelic variant denotes any of two or more

25 alternative forms of a gene occupying the same chromosome locus. A preferred nucleic acid variant is a nucleotide sequence, which contains silent mutation(s). Alternatively or in combination, a nucleic acid variant may also be obtained by introduction of nucleotide substitutions, which do not give rise to another amino acid sequence of the polypeptide encoded by the nucleotide sequence, but which corresponds to the codon

30 usage of the host organism intended for production of the polypeptide of the invention. According to a preferred embodiment, a nucleic acid variant encodes a polypeptide still exhibiting its biological function. More preferably, a nucleotide sequence variant encodes a polypeptide exhibiting *Staphylococcus* peptidoglycan cell wall-binding

and/or a lytic activity. Even more preferably, a nucleic acid variant encodes a polypeptide with enhanced *Staphylococcus* peptidoglycan cell wall-binding and/or lytic activity as defined earlier. Nucleic acids encoding a polypeptide exhibiting *S. Staphylococcus* peptidoglycan cell wall-binding and/or lytic activity may be isolated from any microorganism.

All these variants can be obtained using techniques known to the skilled person, such as screening of library by hybridisation (southern blotting procedures) under low to medium to high hybridisation conditions with for the nucleotide sequence SEQ ID NO: 1, 5, 6, 7, 9, 12, 13, 14, 15, 16, 17, 18, 19 and/or 20 or a variant thereof which can be used to design a probe. Low to medium to high stringency conditions means prehybridization and hybridization at 42°C in 5X SSPE, 0.3% SDS, 200pg/ml sheared and denatured salmon sperm DNA, and either 25% 35% or 50% formamide for low to medium to high stringencies respectively. Subsequently, the hybridization reaction is washed three times for 30 minutes each using 2XSSC, 0.2%SDS and either 55 °C, 65 °C, or 75 °C for low to medium to high stringencies.

The sequence information as provided herein should not be so narrowly construed as to require inclusion of erroneously identified bases. The skilled person is capable of identifying such erroneously identified bases and knows how to correct for such errors.

Nucleic acid construct

In a further aspect, there is provided a nucleic acid construct comprising a nucleic acid molecule as identified in the previous section. This nucleic acid construct may comprise a first nucleic acid sequence encoding a polypeptide comprising a cell wall-binding domain, possibly further comprising a second and third and optionally fourth nucleic acid sequence as defined in the previous section.

The invention also relates to an expression vector comprising a nucleic acid construct of the invention. Preferably, an expression vector comprises a nucleotide sequence of the invention, which is operably linked to one or more control sequences, which direct the production or expression of the encoded polypeptide in a cell, a subject, or a cell-free expression system.

An expression vector may be seen as a recombinant expression vector. This vector can be constituted by a plasmid, a cosmid, a bacteriophage or a virus which is

transformed by introducing a nucleic acid molecule according to the invention. Such transformation vectors according to the host organism to be transformed are well known to those skilled in the art and widely described in the literature.

A further subject of the invention is a process for the transformation of host organisms, by integrating a least one nucleic acid molecule of the invention, which transformation may be carried out by any suitable known means which have been widely described in the specialist literature and in particular in the references cited in the present application, more particularly by the vector according to the invention.

10 Cell

In a further aspect, the present invention relates to a cell, which comprises a nucleic acid construct or an expression vector of the invention as defined herein. A cell may be any microbial, prokaryotic or eukaryotic cell, which is suitable for expression of the polypeptide of the invention. In a preferred embodiment, said cell is an *E. Coli*.

15 In an even more preferred embodiment, said cell is *E. coli* CL1blue MRF.

Method

In a further aspect, there is provided a method for producing, optionally purifying and optionally freeze-drying a polypeptide as defined in the previous section. Said method comprising the steps of:

- i) producing said polypeptide in a cell comprising a nucleic acid construct as defined in the previous section, optionally
- ii) purifying said polypeptide, and optionally
- iii) freeze-drying said purified polypeptide.

25 In a preferred embodiment, an *E. Coli* is used in step i) for producing a polypeptide using recombinant technologies. More preferably an *E. coli* XL1blueMRF is used in step i) for producing a polypeptide using recombinant technologies. Preferably, in step ii), IMAC and Econo-Pac Chromatography columns (Biorad) packed with 5mL low density Nickel chelating agarose beads (ABT beads) in combination with gravity flow is used to purify said (6xHis-tagged recombinant) polypeptides. The eluted polypeptide can be dialyzed for 2, 4, and 12 hours against 3 x 1l lyophilization buffer, said buffer preferably comprising 50 mM phosphate, 500mM sucrose, 200mM mannitol, 0.005% polysorbate20, pH 7.4.

Method

In a further aspect, the invention also relates to a method for producing a polypeptide with an enhanced lytic activity by treating a polypeptide as defined in the previous section or as obtainable by the method described above. Said treatment
5 comprises substituting a divalent metal ion for increasing a lytic activity as compared to an untreated polypeptide, preferably said method comprising the steps of:

- i) dialyzing said polypeptide against a buffer comprising a chelating compound;
- ii) dialyzing said polypeptide against a divalent metal ion-containing buffer,
10 preferably said divalent metal ion being selected from the group consisting of Mn^{2+} , Co^{2+} , Cu^{2+} , and Zn^{2+} .

A "chelating compound" being defined herein as a compound that binds a metal ion. Well known chelating compounds are ethylene diammine tetraacetic acid (EDTA) and ethylene glycol tetraacetic acid (EGTA). Preferably EDTA is used in step i) of the
15 method of the invention.

Preferably, the divalent metal ion of step ii) is selected from the group consisting Mn^{2+} , Co^{2+} , Cu^{2+} , more preferably, said divalent metal ion is selected from the group consisting of Mn^{2+} and Co^{2+} , even more preferably said divalent metal ion is Mn^{2+} .

We showed that substituting a divalent metal ion by any of the above defined
20 resulted in an increase of a lytic activity of Ply2638 of 2 - 2.5 fold. Lytic activity was assessed spectrophotometrically as defined herein. Preferably, said method leads to an increase in a lytic activity of at least 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9 or 2 fold as compared to an untreated polypeptide. Even more preferably, the method leads to an increase in a lytic activity of at least 2.5 fold. Preferably, the treated polypeptide
25 exhibits a 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9 to 2 fold increase in lytic activity as compared to the untreated polypeptide encoded by SEQ ID NO: 1.

Composition

In a further aspect, there is provided a composition comprising a nucleic acid
30 molecule or a nucleic acid construct or a polypeptide or a vector or a cell as identified herein or obtainable by a method described herein. Preferably, the invention relates to a composition exhibiting a lytic activity as defined herein. More preferably, said composition is for use as a medicament. This medicament is preferably for treating,

preventing and/or delaying an infectious disease. The invention also relates to a pharmaceutical or medical composition. Even more preferably, the invention relates to a pharmaceutical or medical composition for the treatment of an infectious disease. Preferably, the invention relates to a pharmaceutical or medical composition for the treatment of an infectious disease caused by a bacterium, preferably a bacterium of the genus *Staphylococcus*, more preferably a bacterium of the species *S. aureus*. Preferably, said infectious disease is a skin infection, mastitis, pneumonia, meningitis, endocarditis, Toxic Shock Syndrome (TSS), sepsis, septicemia, bacteremia, or osteomyelitis. Preferably, said skin infection is selected from the group of pimples, impetigo, boils, furuncles, cellulitis folliculitis, carbuncles, scaled skin syndrome and abscesses.

A composition as defined herein may comprise a mixture of different nucleic acid molecules, and/or nucleic acid constructs and/or polypeptides and/or vectors and/or cells as identified herein or obtainable by a method described herein.

A composition as defined herein may comprise one or more additional active ingredients. Active preferably being defined herein as showing a lytic activity as defined herein. Preferably, said one or more additional active ingredients are selected from the group consisting of a bacteriophage or phage and antibiotic. A phage encompassed herein can be any phage known in literature. Preferably, a phage encompassed by the present invention belongs, but is not limited, to a family of the list consisting of *Myoviridae*, *Siphoviridae* and *Podoviridae*. A phage encompassed by the present invention may also belong to a family of the list consisting of *Tectiviridae*, *Corticoviridae*, *Lipothrixviridae*, *Plasmaviridae*, *Rudoviridae*, *Fuselloviridae*, *Inoviridae*, *Microviridae*, *Leviviridae* and *Cystoviridae*. More preferably, said one or more active ingredients comprise and/or consist of lysostaphin, preferably *S. Simulans* lysostaphin having 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or 100% sequence identity with SEQ ID NO: 34, more preferably *S. Simulans* lysostaphin of SEQ ID NO: 34.). Even more preferably, said one or more active ingredients comprise and/or consist of both one or more different bacteriophages and lysostaphin, preferably, one or more different phages and *S. Simulans* lysostaphin (SEQ ID NO: 34). Within the context of this invention, a combination of active ingredients as defined herein can be administered sequentially and/or simultaneously.

A composition as defined herein may further comprise a pharmaceutically acceptable carrier. Such composition is preferably for use as a medicine or as a medicament. Preferably the medicament is used in the treatment of infectious diseases. A composition may be in the liquid, solid or semi-liquid or semi-solid form.

5 A composition of the invention can be used to treat animals, including humans, infected with *S. aureus*. Any suitable route of administration can be used to administer said composition including but not limited to: oral, aerosol or other device for delivery to the lungs, nasal spray, intravenous, intramuscular, intraperitoneal, intrathecal, vaginal, rectal, topical, lumbar puncture, intrathecal, and direct application to the brain
10 and/or meninges.

 A composition comprising a nucleic acid molecule or a nucleic acid construct or a polypeptide or a vector or a cell as identified herein or obtainable by a method described herein is preferably said to be active, functional or therapeutically active or able to treat, prevent and/or delay an infectious disease when it decreases the amount of
15 a *Staphylococcus* genera present in a patient or in a cell of said patient or in a cell line or in a cell free in vitro system and preferably means that 99%, 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10%, 5% or less of the initial amount of a *Staphylococcus* genera, is still detectable. Preferably no *Staphylococcus* genera is detectable. In this paragraph, the expression "amount of *Staphylococcus* genera" preferably means alive
20 *Staphylococcus* genera. *Staphylococcus* genera may be detected using standard techniques known by the artisan such as immunohistochemical techniques using *Staphylococcus* specific antibodies, tube coagulase tests that detect staphylocoagulase or "free coagulase", detection of surface proteins such as clumping factor (slide coagulase test) and/or protein A (commercial latex tests). Alive *Staphylococcus* genera
25 may be detected using standard techniques known by the artisan such as microbiological bacterial culture techniques and/or real-time quantitative reverse transcription polymerase chain reaction to assay for bacterial mRNA.
 Said decrease is preferably assessed in a tissue or in a cell of an individual or a patient by comparison to the amount present in said individual or patient before treatment with
30 said composition or polypeptide of the invention. Alternatively, the comparison can be made with a tissue or cell of said individual or patient which has not yet been treated with said composition or polypeptide in case the treatment is local.

A composition comprising a nucleic acid molecule or a nucleic acid construct or a polypeptide or a vector or a cell as identified herein or obtainable by a method described herein may be administered to a patient or of a cell, tissue or organ or said patient at least one week, one month, six month, one year or more.

5 In another embodiment, the invention relates to a non-medical composition exhibiting a binding and/or lytic activity as defined herein. Preferably the invention relates to an antimicrobial. Preferably, the invention relates to an antimicrobial for lysing a bacterium, preferably a bacterium of the genus *Staphylococcus*, more preferably a bacterium of the species *S. aureus*. Preferably the invention relates to an
10 antimicrobial as food preservative or disinfectant.

Use

In a further aspect, the invention relates to the use of a polypeptide comprising domains encoded by a first, second, third and optionally fourth nucleic acid sequence as
15 defined herein, a nucleic acid molecule encoding such polypeptide, a construct comprising such nucleic acid molecule, a vector comprising such construct, a cell comprising such vector and/or a composition comprising any of the above, preferably as antimicrobial. Preferably, the invention relates the use as an antimicrobial for lysing a bacterium, preferably a bacterium of the genus *Staphylococcus*, more preferably a
20 bacterium of the species *S. aureus*. Preferably the invention relates to an antimicrobial as food preservative or disinfectant. Possibly, such food preservatives or disinfectants are used together with other antimicrobial agents. Preferably, such food preservatives or disinfectants are used in combination with one or more additional active ingredients as defined herein. Preferably, said one or more additional active ingredients are
25 selected from the group consisting of a bacteriophage or phage and antibiotic as defined herein.

The above-referenced polypeptide, nucleic acid molecule, construct, vector, cell and/or composition can be applied on or into food products, and/or into various physical sites to be disinfected, by a number of means including, but not limited to,
30 admixing said polypeptide and/or cell containing polypeptide of the invention into the food products, spraying said polypeptide and/or cell containing the polypeptide of the invention onto the foodstuffs or physical sites to be disinfected.

A polypeptide of the invention can be isolated from a cell or a cell containing said polypeptide of the invention can be directly applied or administered without isolation of said polypeptide. For example, a cell which produces a polypeptide of the invention could be administered to a subject (human or animal) or applied to a surface where the polypeptide of the invention would be secreted into food, onto a surface or into the subject's gut. The polypeptide of the invention can then bind and optionally lyse bacterial cells, preferably a bacterium of the genus *Staphylococcus*, more preferably a bacterium of the species *S. aureus*, present in this environment.

Further encompassed is the use of a polypeptide comprising a domain encoded by a first nucleic acid sequence as defined herein, a nucleic acid molecule encoding such polypeptide, a construct comprising such nucleic acid molecule, a vector comprising such construct, a cell comprising such vector and/or a composition comprising any of the above, preferably for detecting bacteria, more preferably for detecting bacteria of the genus *Staphylococcus*, more preferably a bacterium of the species *S. aureus*.

Preferably, said polypeptide, nucleic acid molecule, construct, vector, cell and/or composition is used in a diagnostic application. Possibly said polypeptide, nucleic acid molecule, a construct, a vector, cell and/or a composition is used together with other detection agents.

Method

The invention further relates in a further aspect to a method for treating, delaying and/or preventing an infectious disease by administering a composition as earlier defined herein. All features of this method have already been defined herein.

Definitions

Sequence identity

"Sequence identity" is herein defined as a relationship between two or more amino acid (peptide, polypeptide, or protein) sequences or two or more nucleic acid (nucleotide, polynucleotide) sequences, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between amino acid or nucleotide sequences, as the case may be, as determined by the match between strings of such sequences. "Similarity" between two amino acid sequences is determined by comparing the amino acid sequence and its conserved amino acid substitutes of one

peptide or polypeptide to the sequence of a second peptide or polypeptide. In a preferred embodiment, identity or similarity is calculated over the whole SEQ ID NO as identified herein. "Identity" and "similarity" can be readily calculated by known methods, including but not limited to those described in Computational Molecular Biology, Lesk, A. M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D. W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part I, Griffin, A. M., and Griffin, H. G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heine, G., Academic Press, 1987; and Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991; and Carillo, H., and Lipman, D., SIAM J. Applied Math., 48:1073 (1988).

Preferred methods to determine identity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs. Preferred computer program methods to determine identity and similarity between two sequences include e.g. the GCG program package (Devereux, J., et al., Nucleic Acids Research 12 (1): 387 (1984)), BestFit, BLASTP, BLASTN, and FASTA (Altschul, S. F. et al., J. Mol. Biol. 215:403-410 (1990)). The BLAST X program is publicly available from NCBI and other sources (BLAST Manual, Altschul, S., et al., NCBI NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215:403-410 (1990)). The well-known Smith Waterman algorithm may also be used to determine identity.

Preferred parameters for polypeptide sequence comparison include the following: Algorithm: Needleman and Wunsch, J. Mol. Biol. 48:443-453 (1970); Comparison matrix: BLOSSUM62 from Hentikoff and Hentikoff, Proc. Natl. Acad. Sci. USA. 89:10915-10919 (1992); Gap Penalty: 12; and Gap Length Penalty: 4. A program useful with these parameters is publicly available as the "Ogap" program from Genetics Computer Group, located in Madison, WI. The aforementioned parameters are the default parameters for amino acid comparisons (along with no penalty for end gaps).

Preferred parameters for nucleic acid comparison include the following: Algorithm: Needleman and Wunsch, J. Mol. Biol. 48:443-453 (1970); Comparison matrix: matches=+10, mismatch=0; Gap Penalty: 50; Gap Length Penalty: 3. Available as the Gap program from Genetics Computer Group, located in Madison, Wis. Given above are the default parameters for nucleic acid comparisons.

Optionally, in determining the degree of amino acid similarity, the skilled person may also take into account so-called "conservative" amino acid substitutions, as will be clear to the skilled person. Conservative amino acid substitutions refer to the interchangeability of residues having similar side chains. For example, a group of amino acids having aliphatic side chains is glycine, alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of amino acids having amide-containing side chains is asparagine and glutamine; a group of amino acids having aromatic side chains is phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and histidine; and a group of amino acids having sulphur-containing side chains is cysteine and methionine. Preferred conservative amino acids substitution groups are: valine-leucine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, and asparagine-glutamine. Substitutional variants of the amino acid sequence disclosed herein are those in which at least one residue in the disclosed sequences has been removed and a different residue inserted in its place. Preferably, the amino acid change is conservative. Preferred conservative substitutions for each of the naturally occurring amino acids are as follows: Ala to ser; Arg to lys; Asn to gln or his; Asp to glu; Cys to ser or ala; Gln to asn; Glu to asp; Gly to pro; His to asn or gln; Ile to leu or val; Leu to ile or val; Lys to arg; gln or glu; Met to leu or ile; Phe to met, leu or tyr; Ser to thr; Thr to ser; Trp to tyr; Tyr to trp or phe; and, Val to ile or leu.

Nucleic acid construct, transformation, expression vector, operably linked, expression, control sequences, polypeptide

Construct

A nucleic acid molecule is represented by a nucleotide sequence. A polypeptide is represented by an amino acid sequence. A nucleic acid construct is defined as a nucleic acid molecule which is isolated from a naturally occurring gene or which has been modified to contain segments of nucleic acids which are combined or juxtaposed in a manner which would not otherwise exist in nature. A nucleic acid molecule is represented by a nucleotide sequence. Optionally, a nucleotide sequence present in a nucleic acid construct is operably linked to one or more control sequences, which direct the production or expression of said peptide or polypeptide in a cell or in a subject.

“Operably linked” is defined herein as a configuration in which a control sequence is appropriately placed at a position relative to the nucleotide sequence coding for the polypeptide of the invention such that the control sequence directs the production/expression of the peptide or polypeptide of the invention in a cell and/or in
5 a subject.

“Operably linked” may also be used for defining a configuration in which a sequence is appropriately placed at a position relative to another sequence coding for a functional domain such that a chimeric polypeptide is encoded in a cell and/or in a subject.

10 Expression will be understood to include any step involved in the production of the peptide or polypeptide including, but not limited to, transcription, post-transcriptional modification, translation, post-translational modification and secretion.

Control sequence is defined herein to include all components which are necessary or advantageous for the expression of a polypeptide. At a minimum, the control sequences
15 include a promoter and transcriptional and translational stop signals. Optionally, a promoter represented by a nucleotide sequence present in a nucleic acid construct is operably linked to another nucleotide sequence encoding a peptide or polypeptide as identified herein.

The term "transformation" refers to a permanent or transient genetic change
20 induced in a cell following the incorporation of new DNA (i.e. DNA exogenous to the cell). When the cell is a bacterial cell, as is intended in the current invention, the term usually refers to an extrachromosomal, self-replicating vector which harbors a selectable antibiotic resistance.

An expression vector may be any vector which can be conveniently subjected to
25 recombinant DNA procedures and can bring about the expression of a nucleotide sequence encoding a polypeptide of the invention in a cell and/or in a subject. As used herein, the term "promoter" refers to a nucleic acid fragment that functions to control the transcription of one or more genes or nucleic acids, located upstream with respect to the direction of transcription of the transcription initiation site of the gene. It is related
30 to the binding site identified by the presence of a binding site for DNA-dependent RNA polymerase, transcription initiation sites, and any other DNA sequences, including, but not limited to, transcription factor binding sites, repressor and activator protein binding sites, and any other sequences of nucleotides known to one skilled in the art to act

directly or indirectly to regulate the amount of transcription from the promoter. Within the context of the invention, a promoter preferably ends at nucleotide -1 of the transcription start site (TSS).

“Polypeptide” as used herein refers to any peptide, oligopeptide, polypeptide, gene product, expression product, or protein. A polypeptide is comprised of consecutive amino acids. The term “polypeptide” encompasses naturally occurring or synthetic molecules.

In this document and in its claims, the verb “to comprise” and its conjugations is used in its non-limiting sense to mean that items following the word are included, but items not specifically mentioned are not excluded. In addition the verb “to consist” may be replaced by “to consist essentially of” meaning that a product or a composition or a nucleic acid molecule or a peptide or polypeptide of a nucleic acid construct or vector or cell as defined herein may comprise additional component(s) than the ones specifically identified; said additional component(s) not altering the unique characteristic of the invention. In addition, reference to an element by the indefinite article “a” or “an” does not exclude the possibility that more than one of the elements is present, unless the context clearly requires that there be one and only one of the elements. The indefinite article “a” or “an” thus usually means “at least one”. All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

Description of the figures

Fig. 1. Linear relationship of Ply2638 (SEQ ID NO: 21 encoded by SEQ ID NO: 44) activity against *S. aureus* SA2638/2854 cells in dependency of endolysin concentration. Assays were performed under standard conditions (PBS buffer pH 7.4, 120 mM sodium chloride) in photometric lysis assays. Maximum activity was determined from the first derivative of regression fits from sigmoidal lysis curves, calculated with SigmaPlot software. Error bars represent standard deviation calculated from technical triplicate experiments.

Fig. 2. Influence of divalent cations on lytic activity of Ply2638 (SEQ ID NO: 21, encoded by SEQ ID NO: 44). The enzyme was EDTA treated with subsequent substitution of metal ions by dialysis against MOPS buffer containing MgCl₂, CaCl₂,
5 ZnCl₂, CuCl₂, CoCl₂, or MnSO₄. Ply2638 dialyzed against MOPS buffer omitting EDTA treatment served as reference. Error bars represent standard deviation, calculated from technical triplicate experiments.

Fig. 3. Lytic activity of 50 nM Ply2638 (SEQ ID NO: 21, encoded by SEQ ID NO: 44)
10 on *S. aureus* SA2638/2854 cells after lyophilization and reconstitution. Activity was measured in a turbidity reduction assay under standard conditions. Lyophilization buffer was taken as a control. The triple domain enzyme recovers full lytic activity after freeze-drying.

15 Fig. 4. Lytic activity of 50 nM M23-LST_Ami2638_CBD2638 (SEQ ID NO: 29, encoded by SEQ ID NO: 48) on *S. aureus* SA2638/2854 cells after lyophilization and reconstitution (indicated as lyophilized) as compared to freshly prepared M23-LST_Ami2638_CBD2638 (SEQ ID NO: 29, encoded by SEQ ID NO: 48) and Ply2638 (SEQ ID NO: 21, encoded by SEQ ID NO: 44). Lyophilization buffer was taken as a
20 control. Activity was measured in a turbidity reduction assay under standard conditions. The triple domain enzyme recovers full lytic activity after freeze-drying.

Fig. 5. Relative activity of Lysostaphin (LST; SEQ ID NO: 34, encoded by SEQ ID NO: 33) and Ply2638 derivatives (SEQ ID NO: 31, 30, 29, 21 and 24, encoded by SEQ
25 ID NO: 11, 10, 48, 44 and 4, respectively) in dependency of the concentration. The activity of LST at 50 nM was set as reference. All assays were done under standard conditions (37°C, pH 7.4 and 120 mM sodium chloride concentration) using *S. aureus* SA2638/2854 substrate cells from frozen stock.

30 Fig. 6. Relative activities at various pH values (left diagram) and sodium chloride concentrations (right diagram) of truncated and (with PlyTw domains) retrofitted Ply2638 variants determined in turbidity reduction assays. Truncation of the enzyme by one of the two catalytic domains resulted in impaired activities. Duplication of the

CBD (M23_Ami_SH3b_SH3b; SEQ ID NO: 32, encoded by SEQ ID NO: 49) accelerates lysis at basic pH values and elevated salt concentrations. Retrofitting of Ply2638 with CHAP11 domain results in an enzyme which is presumed to attack three different bonds in the peptidoglycan layer of *Staphylococcus*. It shifted pH optima to slight acidic conditions and improved antibacterial activity. However, protein stability of the chimeric enzymes remains a challenge. Maximum lysis velocity of Ply2638 at standard conditions (pH 7.4 and 120 mM sodium chloride concentration) was set as reference. Bars represent mean of triplicate assays, standard deviation is not shown. (CHAP_M23_Ami_SH3b = SEQ ID NO: 25, encoded by SEQ ID NO: 45; Ami_M23_Ami_SH3b = SEQ ID NO: 26, encoded by SEQ ID NO: 46; M23_Ami_SH3b_SH3b = SEQ ID NO: 32, encoded by SEQ ID NO: 49; M23_Ami_SH3b = SEQ ID NO: 21, encoded by SEQ ID NO: 44; Ami_SH3b = SEQ ID NO: 22, encoded by SEQ ID NO: 2; M23_SH3b = SEQ ID NO: 23, encoded by SEQ ID NO: 3).

Fig. 7.: Activity of 50 nM quadruple domain enzymes (SEQ ID NO: 27 and 28, encoded by SEQ ID NO: 47 and 8, respectively), Ply2638 (SEQ ID NO: 21, encoded by SEQ ID NO: 44), and Lyosotaphin (SEQ ID NO: 34, encoded by SEQ ID NO: 33) against using *S. aureus* SA2638/2854 substrate cells under standard conditions. Lyophilization buffer was taken as a control. Quadruple domain enzymes were constructed by combining domains of Ply2638, PlyTw, and Lysostaphin.

Examples

Materials and Methods

Bacterial strains, culture conditions, phages and plasmids

E. coli XL1BlueMRF⁺ and *E. coli* Sure was used for the over-expression of 6x-His-tagged (SEQ ID NO: 43) fusion proteins. Both strains were cultured in LB-PE medium at 30°C with 100 µg/ml ampicillin and 30 µg/ml tetracycline for plasmid selection. Phage 2638A lysate were used as template for the amplification of Ply2638A gene or domain coding regions thereof. CHAPT_{Tw} domain (SEQ ID NO: 19) was amplified from Phage Twort lysate. Cystein- Histidine dependent amidase/ peptidase (CHAP)

domain and amidase domain from phage 11 (SEQ ID NO 18 and 17, respectively; Donovan, et al., 2006 and 2008; Navarre et al., 1999; Sass and Bierbaum 2007) were amplified from a pet21a vector containing phi11 autolysin gene, a kindly gift from Donovan, D.M.. Plasmid LT1215 containing the sequence of mature Lysostaphin (SEQ ID NO: 33) was used as template for domain amplification M23-LST (SEQ ID NO: 15) and CWT-LST (SEQ ID NO: 13).

pQE-30 vector (catalogues number: 32915, Qiagen, Hilden, Germany; SEQ ID NO: 50) was used as cloning and expression vector for the production of 6x-His tagged recombinant fusion proteins in *E. coli* XL1BlueMRF⁺ or *E. coli* Sure respectively.

10

DNA techniques and cloning procedures

Standard techniques according to Sambrook, Maniatis et al. (1989) were employed for cloning of single genes and the creation of fusion proteins. High Fidelity PCR Enzyme mix (Fermentas) was used in PCR reactions. DNA concentrations were determined with a spectrophotometer (NanoDrop ND-1000 Spectrophotometer).

15

pHPI2638 is constructed by insertion of Ply2638 (SEQ ID NO: 1) encoding sequence Met1 – Lys486 into pQE30 (SEQ ID NO: 50) sites BamHI – SalI. pHPI2638-PI2638 construct has the same sequence consecutively inserted into BamHI – SacI – SalI sites. CHAP11 (SEQ ID NO: 18), Ami11 (SEQ ID NO: 17) and CHAP_Ami11 were N-terminal introduced into BamHI digested pHPL2638A. Prior to ligation reaction, the vector was dephosphorylated using shrimp alkaline phosphatase (SAP, Fermentas).

20

pHM23_CBD2638 (SEQ ID NO: 3), and pHM23-

2638_Ami2638_CBD2638_CBD2638 (SEQ ID NO: 49) were constructed by a replacement of GFP coding region from pHGFP_CBD2638A_c vector (SEQ ID NO:

25

59) with the respective inserts using BamHI and SacI restriction sites. pHM23-LST_Ami2638_CBD2638 (SEQ ID NO: 48) has a pQE-30 backbone with mature lysostaphin (SEQ ID NO: 33) coding sequence Ala1 - Gly154 inserted into BamHI and SacI and Ply2638 partial sequence encoding Leu138 - Lys486 in SacI and SalI sites.

pHM23-LST_M23-LST_CWT-LST (SEQ ID NO: 11) has mature Lysostaphin (SEQ ID NO: 33) sequences Ala1 – Gly154 inserted into BamHI and SacI and Ala1 – Lys246

30

into SacI and SalI sites of pQE30. pHLST-LST (SEQ ID NO: 10) is constructed the same way having Ala1 – Lys246 repeatedly in BamHI – SacI and SacI – SalI sites. In plasmids encoding quadruple domain constructs pHCHAPT_w_Ami2638_M23-

LST_CBD2638 (SEQ ID NO: 47) and pHCHAPT_w_Ami2638_M23-LST_CWT-LST (SEQ ID NO: 8) the domains are directly fused via splicing overlap extension PCR (SOE) and inserted into pQE30 BamHI and SalI sites. In both constructs, boarder regions of individual domains (CHAPT_w, SEQ ID NO: 19: Met1 – Ile140, Ami2638, 5 SEQ ID NO: 16: Lys141-Gly358 of SEQ ID NO: 1, CBD2638, SEQ ID NO: 12 : Trp393- Lys486 of SEQ ID NO: 1, M23-LST, SEQ ID NO: 15: Ala1 – Gly154 of SEQ ID NO: 33, CWT-LST, SEQ ID NO: 13: Trp155 – Lys246 of SEQ ID NO: 33) were determined with bioinformatics (unpublished data). Plasmids with repetitive sequences were transferred into E. coli Sure strain, all other plasmids into E. coli XL1BlueMRF'.

10

Expression and purification of recombinant fusion proteins

Protein overexpression and partial purification was essentially done as previously described by others (Loessner et al., 1996, Schmelcher et al., 2010). In brief, plasmid bearing E. coli were grown in 250 ml modified LB medium (15 g/l tryptose, 8 g/l yeast 15 extract, 5 g/l NaCl, pH 7.8) to an optical density at 600nm (OD_{600nm}) of 0.4 to 0.6 and induced with 1 mM IPTG. Cells were further incubated for 4 hours at 30°C, or 18 hours at 20°C, cooled to 4°C, and harvested by centrifugation. Cell pellets were suspended in 5 ml immobilization buffer (50 mM NaH₂PO₄, 500 mM NaCl, 5 mM imidazole, 0.1% polysorbate20, pH 7.4). Cytosolic E. coli contents containing soluble 20 recombinant proteins were liberated by a double passage through a French Pressure Cell Press (1200 psi, SLM Aminco, Urbana, IL, U.S.) operated at 1200 psi. Other downstream processing steps included removal of insoluble cell debris by centrifugation, filter sterilization (0.2 µm PES membrane, Millipore), and Immobilized Metal Affinity Chromatography (IMAC) purification using MicroBiospin (Bio-Rad, 25 Hercules, CA, U.S.) columns packet with low density Ni-NTA Superflow resin (Chemie Brunschwig AG, Basel, Switzerland). Ni-NTA immobilized proteins were on-column gravity flow washed with 5-10 column volumes immobilization buffer. Protein fractions were then eluted with elution buffer (50 mM NaH₂PO₄, 500 mM NaCl, 125 mM imidazole, 0.1% polysorbate20, pH 7.4) and dialyzed against two changes of 30 dialysis buffer (50 mM NaH₂PO₄, 100 mM NaCl, 0.1% polysorbate20, pH 7.4). Protein concentrations were defined in a NanoDrop ND-1000 spectrophotometer, corrected for specific absorbance at 280 nm as calculated from the primary amino acid sequence with Vector NTI software (Invitrogen, Carlsbad, CA, U.S.) and estimated for purity by SDS-

PAGE. Aliquots were stored at -20°C mixed with 50% glycerol.

Lyophilization of recombinant proteins

IMAC purified proteins were dialyzed against 3 changes of 300 ml lyophilization
5 buffer (50 mM phosphate or Tris, 500 mM sucrose, 200 mM mannitol, pH 7.4) aliquot
and frozen in the gaseous phase of liquid nitrogen. The freeze-drying was done at -
40°C and vacuum at 75 mTorr for 60 minutes, followed by increasing temperature
during 5 hours to -10°C and another 60 minutes at -10°C at the same vacuum levels. As
final step, temperature was increased to 25°C during 10 hours. Samples were
10 reconstituted prior to testing in lysis assays by the addition of water.

Cell wall-binding assay

As a standard assay to determine the ability of a CBD to direct a GFP fusion to the
bacterial surface and mediate tight binding to the cell wall ligand, the following
15 conditions is used: bacteria, preferably *S. aureus* BB255, from late log phase are
harvested by centrifugation, resuspended in 1/10th volume of PBS-T (50 mM
NaH₂PO₄, 120 mM NaCl [pH 8.0], 0.01% polysorbate 20) and stored on ice. GFP-CBD
proteins, preferably SEQ ID NO: 64, encoded by SEQ ID NO: 60, are diluted in the
same buffer to a concentration of 400 nM (2x GFP-CBD) and also stored on ice. In a
20 1.5 ml microcentrifuge cup, 100 µl cells and 100 µl of 2x GFP-CBD are mixed and
incubated at room temperature for 5 min. Cells are then removed from the supernatant
by centrifugation in a microfuge (16000 g, 60 s). The supernatant was discarded and
cells were washed twice in 500 µl of PBS-T buffer. For fluorescence microscopy the
pellet was finally resuspended in 50 µl of buffer. For fluorometer assays, the pellet is
25 finally resuspended in 200 µl of PBS-T and transferred to a microplate well.
Quantitative fluorescence assays can be performed using a multi label counter device
(Victor³, Perkin Elmer, Massachusetts, U.S.) with sterile, untreated, black 96-well
polystyrene microplates (Nunc, Roskilde, Denmark). As a negative control GFP can be
used.

30 Quantitative fluorescence assays

Dependency of pH and salt on CBD2638 to *S. aureus* BB255 cell surface ligand
interaction is investigated by incubation of cells from 1 ml volume set to an OD_{600nm} 1
+/- 0.05 (~4 x 10⁹ cells) with 7.5 µg GFP-CBDS2638 fusion protein, SEQ ID NO: 64,

encoded by SEQ ID NO: 60. This cell to protein ratio is close to saturation point as determined in previous experiments and enables detection of variations in binding efficiencies. Varying pH is tested using citrate buffers pH 4.5 to 6.5 and phosphate buffers pH 6 to 9. After incubation with GFP-CBD2638 protein in pH buffer, cells are washed with respective pH buffer followed by standard PBS-T (pH 8) washing. Finally, cells are adjusted to an $OD_{595nm} = 0.3$ to detect fluorescence from 200 μ l suspensions thereof with a Victor³ multi label counter. Similar experiments are performed using buffers prepared with increasing sodium chloride concentrations (10 mM NaH_2PO_4 , 0 – 1000 mM NaCl, 0.1% polysorbate 20, pH 6).

Quantification of CBD binding capacity of *Staphylococcus* strains with altered cell surface properties is tested by recording relative fluorescence units (RFU) of washed and heat killed cells previously incubated with excessive GFP-CBD2638 protein. Fluorescence of equal volumes (200 μ l) of GFP-CBD2638 labeled cells, adjusted to an $OD_{595nm} = 0.3$, are measured using appropriate filter sets in a multi label counter device. Comparison and quantification of absorption levels of GFP-CBD2638 (SEQ ID NO: 64, encoded by SEQ ID NO: 60), GFP-CBD2638-CBD2638 (SEQ ID NO 65 and/or 66, encoded by SEQ ID NO: 61 and 62, respectively), and GFP-CBD2638-CBD2638-CBD2638 (SEQ ID NO: 67, encoded SEQ ID NO: 63) on *S. aureus* BB255 cells and SDS treated cells is done the same way.

20

Lysis assays

Substrate cell for lytic activity assays were grown to an optical density at 600nm (OD_{600}) of 0.4, washed twice with PBST pH 7.4 and re- suspended in 15% glycerol containing PBS buffer pH 7.4 concentrating it at the same time 100 fold. The cells were stored at $-20^{\circ}C$. For further use in binding or lytic activity assays the cells were thaw, washed with PBS pH 7.4 and diluted to an OD_{600} of 1 ± 0.05 . In standard lytic activity assays protein samples were diluted to equimolar amounts and distributed in transparent 96-well tissue culture test plates (SPL life sciences, Pocheon, Korea). Substrate cells were added to a final volume and drop in optical density at 595nm (OD_{595nm}) were recorded for about 1 hour at $37^{\circ}C$.

30

Lytic activity of retrofitted and deletion constructs of Ply2638 were tested against Phage 2638A propagation strain *S. aureus* SA2638/2854 from frozen stock in lysis assays. We tested the activity at various buffer conditions. pH values from 4.6 to 9 in

0.4 increments were tested using Citrate/ Phosphate buffers (25 mM Citrate, 25 mM Phosphate, 120 mM NaCl, pH 4.6 – 6.6) and Tris/ Phosphate buffers (25 mM Tris, 25 mM Phosphate, 120 mM NaCl). The activity of Ply2638A derivatives at salt concentrations ranging from 0 to 1000 mM Sodium chloride (in 10 mM phosphate buffer pH 7.4) was tested. The Ply2638A derivatives were diluted to 10 μ M final concentration with MQ prior to its application in lysis assays. Here, 4 μ l of 10 μ M Ply2638A derivatives were applied to 196 μ l substrate cell suspensions using a multichannel pipette, resulting in an assay concentration of 200 nM protein. The substrate cell suspensions were prepared from frozen stocks, diluting it with pH or salt buffers and standardizing it spectrophotometrically (Libra S22, Biochrom) to an initial OD600 of 1 ± 0.05 . Decrease in optical density at 595 nm (OD595) was measured using a Victor3 1420 Multilabel Counter instrument (Perkin Elmer) during 1 hour. Plates were shaken vigorously for 1 second (double orbit, 0.1mm diameter) after every single read out. As positive control served N-terminal 6xHis tagged Lysostaphin (HLST), commercially available Lysostaphin (recombinant, E. coli originated, Sigma). As negative control we applied MilliQ water.

Influence of divalent metal ions on the activity of Ply2638

Partially purified Ply2638 (SEQ ID NO: 21) was dialyzed for 2 hours against EDTA containing buffer (50 mM MOPS, 100 mM sodium chloride, 0.005% polysorbate20, 10 mM EDTA) followed by dialysis against buffer containing the respective divalent metal ions (50 mM MOPS, 100 mM sodium chloride, 0.005% polysorbate20, and 10 mM CaCl₂, 10 mM MgCl₂, 1 mM CoCl₂, 1 mM CuCl₂, 1 mM MnSO₄, or 1 mM ZnCl₂ respectively). Cells used as substrate were SDS treated and EDTA washed prior to its application in standard lysis assays.

Tabel 1. SEQ ID NO identification

enzyme/domain/construct/vector	nucleic acid sequence	amino acid sequence	amino acid sequence of domain with putative linker	nucleic acid sequence of His-tagged construct
Mature enzyme				
Ply2638	SEQ ID NO: 1	SEQ ID NO: 21		SEQ ID NO: 44
LST	SEQ ID NO: 33	SEQ ID NO: 34		
Domain				
CBD-2638	SEQ ID NO: 12	SEQ ID NO: 35	SEQ ID NO: 51	
CWT-LST	SEQ ID NO: 13	SEQ ID NO: 36	SEQ ID NO: 52	
M23-2638	SEQ ID NO: 14	SEQ ID NO: 37	SEQ ID NO: 53	
M23-LST	SEQ ID NO: 15	SEQ ID NO: 38	SEQ ID NO: 54	
Ami-2638	SEQ ID NO: 16	SEQ ID NO: 39	SEQ ID NO: 55	
Ami- Φ 11	SEQ ID NO: 17	SEQ ID NO: 40	SEQ ID NO: 56	
CHAP- Φ 11	SEQ ID NO: 18	SEQ ID NO: 41	SEQ ID NO: 57	
CHAP- Φ Twort	SEQ ID NO: 19	SEQ ID NO: 42	SEQ ID NO: 58	
Retrofitted construct				
CHAP11_M23-2638_Ami2638_CBD2638	SEQ ID NO: 5	SEQ ID NO: 25		SEQ ID NO: 45
Ami11_M23-2638_Ami2638_CBD2638	SEQ ID NO: 6	SEQ ID NO: 26		SEQ ID NO: 46
CHAPTw_Ami2638_M23-LST_CBD2638	SEQ ID NO: 7	SEQ ID NO: 27		SEQ ID NO: 47

M23-LST_Ami2638_CBD2638	SEQ ID NO: 9	SEQ ID NO: 29		SEQ ID NO: 48
M23-2638_Ami2638_CBD2638_CBD2638	SEQ ID NO: 20	SEQ ID NO: 32		SEQ ID NO: 49
Ami2638_CBD2638		SEQ ID NO: 22		SEQ ID NO: 2
M23-2638_CBD2638		SEQ ID NO: 23		SEQ ID NO: 3
Ply2638-Ply2638		SEQ ID NO: 24		SEQ ID NO: 4
CHAPT _w _Ami2638_M23-LST_CWT-LST		SEQ ID NO: 28		SEQ ID NO: 8
LST_LST		SEQ ID NO: 30		SEQ ID NO: 10
M23-LST_M23-LST_CWT-LST		SEQ ID NO: 31		SEQ ID NO: 11
GFP_CBD2638		SEQ ID NO: 64		SEQ ID NO: 60
GFP_CBD2638_CBD2638 var.1		SEQ ID NO: 65		SEQ ID NO: 61
GFP_CBD2638_CBD2638 var. 2		SEQ ID NO: 66		SEQ ID NO: 62
GFP_CBD2638_CBD2638_CBD2638		SEQ ID NO: 67		SEQ ID NO: 63
Tag				
6xHis-tag		SEQ ID NO: 43		
Vector				
pQE-30 vector	SEQ ID NO: 50			
pHGFP_CBD2638_c vector	SEQ ID NO: 59			

SEQ ID NO: 1 (Ply2638)

ATGCTAACTGCTATTGACTATCTTACGAAAAAAGGTTGGAAAATATCATCT
GACCCTCGCACTTACGATGGTTACCCTAAAAACTACGGCTACAGAAATTAC
CATGAAAACGGCATTAATTATGATGAGTTTTGTGGTGGTTATCATAGAGCTT
5 TTGATGTTTACAGTAACGAACTAACGACGTGCCTGCTGTTACTAGCGGAA
CAGTTATTGAAGCAAACGATTACGGTAATTTTGGTGGTACATTCGTTATTAG
AGACGCTAACGATAACGATTGGATATATGGGCATCTACAACGTGGCTCAAT
GCGATTTGTTGTAGGCGACAAAGTCAATCAAGGTGACATTATTGGTTTACA
AGGTAATAGCAACTATTACGACAATCCTATGAGTGTACATTTACATTTACA
10 ATTACGCCCTAAAGACGCAAAGAAAGATGAAAAATCACAAGTATGTAGTG
GTTTGGCTATGGAAAAATATGACATTACAAATTTAAATGCTAAACAAGATA
AATCAAAGAATGGGAGCGTGAAAGAGTTGAAACATATCTATTCAAACCAT
ATTAAAGGTAACAAGATTACAGCACCAAAACCTAGTATTCAAGGTGTGGTC
ATCCACAATGATTATGGTAGTATGACACCTAGTCAATACTTACCATGGTTAT
15 ATGCACGTGAGAATAACGGTACACACGTTAACGGTTGGGCTAGTGTTTATG
CAAATAGAAACGAAGTGCTTTGGTATCATCCGACAGACTACGTAGAGTGGC
ATTGTGGTAATCAATGGGCAAATGCTAACTTAATCGGATTTGAAGTGTGTG
AGTCGTATCCTGGTAGAATCTCGGACAAATTATTCTTAGAAAATGAAGAAG
CGACATTGAAAGTAGCTGCGGATGTGATGAAGTCGTACGGATTACCAGTTA
20 ATCGCAACACTGTACGTCTGCATAACGAATTCTTCGGAACCTTCTTGTCCACA
TCGTTCTGTTGGGACTTGTCATGTTGGCAAAGGTGAGCCTTACACAACCTACTAA
TATTAATAAAAATGAAAGACTACTTCATCAAACGCATCAAACATTATTATGA
CGGTGGAAAGCTAGAAGTAAGCAAAGCAGCAACTATCAAACAATCTGACG
TTAAGCAAGAAGTTAAAAAGCAAGAAGCAAAACAAATTGTGAAAGCAACA
25 GATTGGAAACAGAATAAAGATGGCATTGTTGGTATAAAGCTGAACATGCTTCG
TTCACAGTGACAGCACCAGAGGGAATTATCACAAGATACAAAGGTCCTTGG
ACTGGTCACCCACAAGCTGGTGTATTACAAAAAGGTCAAACGATTAAATAT
GATGAGGTTCAAAAATTTGACGGTCATGTTTGGGTATCGTGGGAAACGTTT
GAGGGCGAACTGTATACATGCCGGTACGCACATGGGACGCTAAAACTGG
30 TAAAGTTGGTAAGTTGTGGGGCGAAATTAAATAA

SEQ ID NO: 2 (Ami2638_CBD2638 + 6xHis and cloning sites)

ATGAGAGGATCGCATCACCATCACCATCACGGATCCTTACGCCCTAAAGAC
GCAAAGAAAGATGAAAAATCACAAGTATGTAGTGGTTTGGCTATGGAAAA
ATATGACATTACAAATTTAAATGCTAAACAAGATAAATCAAAGAATGGGA
5 GCGTGAAAGAGTTGAAACATATCTATTCAAACCATATTAAAGGTAACAAGA
TTACAGCACCAAAACCTAGTATTCAAGGTGTGGTCATCCACAATGATTATG
GTAGTATGACACCTAGTCAATACTTACCATGGTTATATGCACGTGAGAATA
ACGGTACACACGTTAACGGTTGGGCTAGTGTTTATGCAAATAGAAACGAAG
TGCTTTGGTATCATCCGACAGACTACGTAGAGTGGCATTGTGGTAATCAAT
10 GGGCAAATGCTAACTTAATCGGATTTGAAGTGTGTGAGTCGTATCCTGGTA
GAATCTCGGACAAATTATTCTTAGAAAATGAAGAAGCGACATTGAAAGTAG
CTGCGGATGTGATGAAGTCGTACGGATTACCAGTTAATCGCAACACTGTAC
GTCTGCATAACGAATTCTTCGGAACCTCTTGTCCACATCGTTCGTGGGACTT
GCATGTTGGCAAAGGTGAGCCTTACACAACCTACTAATATTAATAAAAATGAA
15 AGACTACTTCATCAAACGCATCAAACATTATTATGACGGTGGAAAGCTAGA
AGTAAGCAAAGCAGCAACTATCAAACAATCTGACGTTAAGCAAGAAGTTA
AAAAGCAAAGAAGCAAAACAAATTGTGAAAGCAACAGATTGGAAACAGAAT
AAAGATGGCATTTGGTATAAAGCTGAACATGCTTCGTTACAGTGACAGCA
CCAGAGGGAATTATCACAAGATACAAAGGTCCTTGGACTGGTCACCCACAA
20 GCTGGTGTATTACAAAAAGGTCAAACGATTAAATATGATGAGGTTCAAAAA
TTTGACGGTCATGTTTGGGTATCGTGGGAAACGTTTGAGGGCGAAACTGTA
TACATGCCGGTACGCACATGGGACGCTAAAACTGGTAAAGTTGGTAAGTTG
TGGGGCGAAATTAAATAA

25 SEQ ID NO: 3 (M23-2638_CBD2638 + 6xHis and cloning sites)

ATGAGAGGATCGCATCACCATCACCATCACGGATCCATGCTAACTGCTATT
GACTATCTTACGAAAAAAGGTTGGAAAATATCATCTGACCCTCGCACTTAC
GATGGTTACCCTAAAAACTACGGCTACAGAAATTACCATGAAAACGGCATT
AATTATGATGAGTTTTGTGGTGGTTATCATAGAGCTTTTGATGTTTACAGTA
30 ACGAAACTAACGACGTGCCTGCTGTTACTAGCGGAACAGTTATTGAAGCAA
ACGATTACGGTAATTTTGGTGGTACATTCGTTATTAGAGACGCTAACGATA
ACGATTGGATATATGGGCATCTACAACGTGGCTCAATGCGATTTGTTGTAG
GCGACAAAGTCAATCAAGGTGACATTATTGGTTTACAAGGTAATAGCAACT

ATTACGACAATCCTATGAGTGTACATTTACATTTACAATTACGCCCTAAAGA
CGCAAAGAAAGATGAAAAATCACAAGTATGTAGTGGTTTGGCTATGGAAA
AATATGACATTACAAATTTAAATGCTAAACAAGATAAATCAAAGAATGGG
AGCGTGAAAGAGTTGAAACATATCTATTCAAACCATATTAAAGGTAACAAG
5 ATTACAGCACCAAAACCTAGTATTCAAGGTGAGCTCGGTGGAAAGCTAGAA
GTAAGCAAAGCAGCAACTATCAAACAATCTGACGTTAAGCAAGAAGTTAA
AAAGCAAGAAGCAAAACAAATTGTGAAAGCAACAGATTGGAAACAGAATA
AAGATGGCATTTTGGTATAAAGCTGAACATGCTTCGTTACAGTGACAGCAC
CAGAGGGAATTATCACAAGATACAAAGGTCCTTGGACTGGTCACCCACAAG
10 CTGGTGTATTACAAAAAGGTCAAACGATTAAATATGATGAGGTTCAAAAAT
TTGACGGTCATGTTTGGGTATCGTGGGAAACGTTTGAGGGCGAAACTGTAT
ACATGCCGGTACGCACATGGGACGCTAAAACTGGTAAAGTTGGTAAGTTGT
GGGGCGAAATTAAATAA

15 SEQ ID NO: 4 (Ply2638-Ply2638 + 6x His and cloning sites))

ATGAGAGGATCGCATCACCATCACCATCACGGATCCATGCTAACTGCTATT
GACTATCTTACGAAAAAAGGTTGGAAAAATATCATCTGACCCTCGCACTTAC
GATGGTTACCCTAAAAACTACGGCTACAGAAATTACCATGAAAACGGCATT
AATTATGATGAGTTTTGTGGTGGTTATCATAGAGCTTTTGATGTTTACAGTA
20 ACGAAACTAACGACGTGCCTGCTGTTACTAGCGGAACAGTTATTGAAGCAA
ACGATTACGGTAATTTTGGTGGTACATTCGTTATTAGAGACGCTAACGATA
ACGATTGGATATATGGGCATCTACAACGTGGCTCAATGCGATTTGTTGTAG
GCGACAAAGTCAATCAAGGTGACATTATTGGTTTACAAGGTAATAGCAACT
ATTACGACAATCCTATGAGTGTACATTTACATTTACAATTACGCCCTAAAGA
25 CGCAAAGAAAGATGAAAAATCACAAGTATGTAGTGGTTTGGCTATGGAAA
AATATGACATTACAAATTTAAATGCTAAACAAGATAAATCAAAGAATGGG
AGCGTGAAAGAGTTGAAACATATCTATTCAAACCATATTAAAGGTAACAAG
ATTACAGCACCAAAACCTAGTATTCAAGGTGTGGTCATCCACAATGATTAT
GGTAGTATGACACCTAGTCAATACTTACCATGGTTATATGCACGTGAGAAT
30 AACGGTACACACGTTAACGGTTGGGCTAGTGTTTATGCAAATAGAAACGAA
GTGCTTTGGTATCATCCGACAGACTACGTAGAGTGGCATTGTGGTAATCAA
TGGGCAAATGCTAACTTAATCGGATTTGAAGTGTGTGAGTCGTATCCTGGT
AGAATCTCGGACAAATTATTCTTAGAAAATGAAGAAGCGACATTGAAAGTA

GCTGCGGATGTGATGAAGTCGTACGGATTACCAGTTAATCGCAACACTGTA
CGTCTGCATAACGAATTCTTCGGAACCTTCTTGTCCACATCGTTTCGTGGGACT
TGCATGTTGGCAAAGGTGAGCCTTACACAACCTACTAATATTAATAAAATGA
AAGACTACTTCATCAAACGCATCAAACATTATTATGACGGTGGAAAGCTAG
5 AAGTAAGCAAAGCAGCAACTATCAAACAATCTGACGTTAAGCAAGAAGTT
AAAAAGCAAGAAGCAAAACAAATTGTGAAAGCAACAGATTGGAAACAGA
ATAAAGATGGCATTGTTGGTATAAAGCTGAACATGCTTCGTTACAGTGACAG
CACCAGAGGGAATTATCACAAGATACAAAGGTCCTTGGACTGGTCACCCAC
AAGCTGGTGTATTACAAAAAGGTCAAACGATTAAATATGATGAGGTTCAA
10 AATTTGACGGTCATGTTTGGGTATCGTGGGAAACGTTTGAGGGCGAAACTG
TATACATGCCGGTACGCACATGGGACGCTAAAACTGGTAAAGTTGGTAAGT
TGTGGGGCGAAATTAAAGAGCTCATGCTAACTGCTATTGACTATCTTACGA
AAAAAGGTTGGAAAATATCATCTGACCCTCGCACTTACGATGGTTACCCTA
AAAACCTACGGCTACAGAAATTACCATGAAAACGGCATTAAATTATGATGAGT
15 TTTGTGGTGGTTATCATAGAGCTTTTGATGTTTACAGTAACGAACTAACGA
CGTGCCCTGCTGTTACTAGCGGAACAGTTATTGAAGCAAACGATTACGGTAA
TTTTGGTGGTACATTCGTTATTAGAGACGCTAACGATAACGATTGGATATAT
GGGCATCTACAACGTGGCTCAATGCGATTTGTTGTAGGCGACAAAGTCAAT
CAAGGTGACATTATTGGTTTACAAGGTAATAGCAACTATTACGACAATCCT
20 ATGAGTGTACATTTACATTTACAATTACGCCCTAAAGACGCAAAGAAAGAT
GAAAAATCACAAGTATGTAGTGGTTTGGCTATGGAAAAATATGACATTACA
AATTTAAATGCTAAACAAGATAAATCAAAGAATGGGAGCGTGAAAGAGTT
GAAACATATCTATTCAAACCATATTAAGGTAACAAGATTACAGCACCAAA
ACCTAGTATTCAAGGTGTGGTCATCCACAATGATTATGGTAGTATGACACC
25 TAGTCAATACTTACCATGGTTATATGCACGTGAGAATAACGGTACACACGT
TAACGGTTGGGCTAGTGTTTATGCAAATAGAAACGAAGTGCTTTGGTATCA
TCCGACAGACTACGTAGAGTGGCATTGTGGTAATCAATGGGCAAATGCTAA
CTTAATCGGATTTGAAGTGTGTGAGTCGTATCCTGGTAGAATCTCGGACAA
ATTATTCTTAGAAAAATGAAGAAGCGACATTGAAAGTAGCTGCGGATGTGAT
30 GAAGTCGTACGGATTACCAGTTAATCGCAACACTGTACGTCTGCATAACGA
ATTCTTCGGAACCTTCTTGTCCACATCGTTTCGTGGGACTTGCATGTTGGCAA
GGTGAGCCTTACACAACCTACTAATATTAATAAAATGAAAGACTACTTCATC
AAACGCATCAAACATTATTATGACGGTGGAAAGCTAGAAGTAAGCAAAGC

AGCAACTATCAAACAATCTGACGTTAAGCAAGAAGTTAAAAAGCAAGAAG
CAAAACAAATTGTGAAAGCAACAGATTGGAAACAGAATAAAGATGGCATT
TGGTATAAAGCTGAACATGCTTCGTTACAGTGACAGCACCAGAGGGAATT
ATCACAAGATACAAAGGTCCTTGGACTGGTCACCCACAAGCTGGTGTATTA
5 CAAAAAGGTCAAACGATTAAATATGATGAGGTTCAAAAATTTGACGGTCAT
GTTTGGGTATCGTGGGAAACGTTTGAGGGCGAAACTGTATACATGCCGGTA
CGCACATGGGACGCTAAAACTGGTAAAGTTGGTAAGTTGTGGGGCGAAATT
AAATAA

10 SEQ ID NO: 5 (CHAP11_M23-2638_Ami2638_CBD2638)

ATGCAAGCAAAATTAATACTAAAAATGAGTTTATAGAGTGGTTGAAAACCTCT
GAGGGAAAACAATTCAATGTGGACTTATGGTATGGATTTCATGCTTTGAT
TATGCCAATGCTGGTTGGAAAGTTTTGTTTGGATTACTTCTAAAAGGTTTAG
GTGCAAAAGATATTCCGTTTCGCTAACAACTTCGACGGATTAGCTACTGTAT
15 ACCAAAATACACCGGACTTCTTAGCACAACCTGGCGACATGGTGGTATTCG
GTAGCAACTACGGTGCTGGATATGGTCACGTTGCATGGGTAATTGAAGCAA
CTTTAGATTACATCATTGTATATGAGCAGAATTGGCTAGGCGGTGGCTGGA
CTGACGGAATCGAACAACCCGGCTGGGGTTGGGAAAAAGTTACAAGACGA
CAACATGCTTATGATTTCCCTATGTGGTTTATCCGTCCGAATTTTAAAAGTG
20 AGACAGCGCCACGATCAGTTCAATCTCCTACACAAGCACCTAAAAAAGAA
ACAGCTGGATCCATGCTAACTGCTATTGACTATCTTACGAAAAAAGGTTGG
AAAATATCATCTGACCCTCGCACTTACGATGGTTACCCTAAAAACTACGGC
TACAGAAATTACCATGAAAACGGCATTAAATTATGATGAGTTTTGTGGTGGT
TATCATAGAGCTTTTGATGTTTACAGTAACGAACTAACGACGTGCCTGCT
25 GTTACTAGCGGAACAGTTATTGAAGCAAACGATTACGGTAATTTTGGTGGT
ACATTCGTTATTAGAGACGCTAACGATAACGATTGGATATATGGGCATCTA
CAACGTGGCTCAATGCGATTTGTTGTAGGCGACAAAGTCAATCAAGGTGAC
ATTATTGGTTTACAAGGTAATAGCAACTATTACGACAATCCTATGAGTGTA
CATTTACATTTACAATTACGCCCTAAAGACGCAAAGAAAGATGAAAAATCA
30 CAAGTATGTAGTGGTTTGGCTATGGAAAAATATGACATTACAAATTTAAAT
GCTAAACAAGATAAATCAAAGAATGGGAGCGTGAAAGAGTTGAAACATAT
CTATTCAAACCATATTAAAGGTAACAAGATTACAGCACCAAAACCTAGTAT
TCAAGGTGTGGTCATCCACAATGATTATGGTAGTATGACACCTAGTCAATA

CTTACCATGGTTATATGCACGTGAGAATAACGGTACACACGTTAACGGTTG
GGCTAGTGTTTATGCAAATAGAAACGAAGTGCTTTGGTATCATCCGACAGA
CTACGTAGAGTGGCATTGTGGTAATCAATGGGCAAATGCTAACTTAATCGG
ATTTGAAGTGTGTGAGTCGTATCCTGGTAGAATCTCGGACAAATTATTCTTA
5 GAAAATGAAGAAGCGACATTGAAAGTAGCTGCGGATGTGATGAAGTCGTA
CGGATTACCAGTTAATCGCAACACTGTACGTCTGCATAACGAATTCTTCGG
AACTTCTTGTCCACATCGTTCGTGGGACTTGTCATGTTGGCAAAGGTGAGCCT
TACACAACCTACTAATATTAATAAAATGAAAGACTACTTCATCAAACGCATC
AAACATTATTATGACGGTGGAAAGCTAGAAGTAAGCAAAGCAGCAACTAT
10 CAAACAATCTGACGTTAAGCAAGAAGTTAAAAAGCAAGAAGCAAAACAAA
TTGTGAAAGCAACAGATTGGAAACAGAATAAAGATGGCATTGTGGTATAAA
GCTGAACATGCTTCGTTACAGTGACAGCACCAGAGGGAATTATCACAAGA
TACAAAGGTCCTTGGACTGGTCACCCACAAGCTGGTGTATTACAAAAAGGT
CAAACGATTAAATATGATGAGGTTCAAAAATTTGACGGTCATGTTTGGGTA
15 TCGTGGGAAACGTTTGAGGGCGAAACTGTATACATGCCGGTACGCACATGG
GACGCTAAAACTGGTAAAGTTGGTAAGTTGTGGGGCGAAATTAAATAA

SEQ ID NO: 6 (Ami11_M23-2638_Ami2638_CBD2638)

AAGCCACAACCTAAAGCAGTAGAACTTAAAATCATCAAAGATGTGGTTAA
20 AGGTTATGACCTACCTAAGCGTGGTAGTAACCCTAAAGGTATAGTTATACA
CAACGACGCAGGGAGCAAAGGGGCGACTGCTGAAGCATATCGTAACGGAT
TAGTAAATGCACCTTTATCAAGATTAGAAGCGGGCATTGCGCATAGTTACG
TATCAGGCAACACAGTTTGGCAAGCCTTAGATGAATCACAAGTAGGTTGGC
ATACCGCTAATCAAATAGGTAATAAATATTATTACGGTATTGAAGTATGTC
25 AATCAATGGGCGCAGATAACGCGACATTCTTAAAAAATGAACAGGCAACTT
TCCAAGAATGCGCTAGATTGTTGAAAAAATGGGGATTACCAGCAAACAGA
AATACAATCAGATTGCACAATGAATTTACTTCAACATCATGCCCTCATAGA
AGTTCGGTTTTACACACTGGTTTTGACCCAGTAACTCGCGGTCTATTGCCAG
AAGACAAGCGGTTGCAACTTAAAGACTACTTTATCAAGCAGATTAGGGCGT
30 ACATGGATGGTAAAATACCGGTTGCCACTGTCTCTAATGAGTCAAGCGCTT
CAAGTAATACAGTTAAACCAGTTGCAAGTGCAGGATCCATGCTAACTGCTA
TTGACTATCTTACGAAAAAAGGTTGGAAAAATATCATCTGACCCCTCGCACTT
ACGATGGTTACCCTAAAAACTACGGCTACAGAAATTACCATGAAAACGGCA

TTAATTATGATGAGTTTTGTGGTGGTTATCATAGAGCTTTTGATGTTTACAG
TAACGAACTAACGACGTGCCTGCTGTTACTAGCGGAACAGTTATTGAAGC
AAACGATTACGGTAATTTTGGTGGTACATTTCGTTATTAGAGACGCTAACGA
TAACGATTGGATATATGGGCATCTACAACGTGGCTCAATGCGATTTGTTGT
5 AGGCGACAAAGTCAATCAAGGTGACATTATTGGTTTACAAGGTAATAGCAA
CTATTACGACAATCCTATGAGTGTACATTTACATTTACAATTACGCCCTAAA
GACGCAAAGAAAGATGAAAAATCACAAGTATGTAGTGGTTTGGCTATGGA
AAAATATGACATTACAAATTTAAATGCTAAACAAGATAAATCAAAGAATG
GGAGCGTGAAAGAGTTGAAACATATCTATTCAAACCATATTAAAGGTAACA
10 AGATTACAGCACCAAAACCTAGTATTCAAGGTGTGGTCATCCACAATGATT
ATGGTAGTATGACACCTAGTCAATACTTACCATGGTTATATGCACGTGAGA
ATAACGGTACACACGTTAACGGTTGGGCTAGTGTATATGCAAATAGAAACG
AAGTGCTTTGGTATCATCCGACAGACTACGTAGAGTGGCATTGTGGTAATC
AATGGGCAAATGCTAACTTAATCGGATTTGAAGTGTGTGAGTCGTATCCTG
15 GTAGAATCTCGGACAAATTATTCTTAGAAAATGAAGAAGCGACATTGAAAG
TAGCTGCGGATGTGATGAAGTCGTACGGATTACCAGTTAATCGCAACACTG
TACGTCTGCATAACGAATTCTTCGGAACCTTCTTGTCACATCGTTCGTGGGA
CTTGCAATGTTGGCAAAGGTGAGCCTTACACAACCTACTAATATTAATAAAAT
GAAAGACTACTTCATCAAACGCATCAAACATTATTATGACGGTGGAAAGCT
20 AGAAGTAAGCAAAGCAGCAACTATCAAACAATCTGACGTTAAGCAAGAAG
TTAAAAAGCAAGAAGCAAAACAAATTGTGAAAGCAACAGATTGGAAACAG
AATAAAGATGGCATTGTTGGTATAAAGCTGAACATGCTTCGTTACAGTGACA
GCACCAGAGGGAATTATCACAAGATACAAAGGTCCTTGGACTGGTCACCCA
CAAGCTGGTGTATTACAAAAAGGTCAAACGATTAAATATGATGAGGTTCAA
25 AAATTTGACGGTCATGTTTGGGTATCGTGGGAAACGTTTGAGGGCGAAACT
GTATACATGCCGGTACGCACATGGGACGCTAAAACCTGGTAAAGTTGGTAAG
TTGTGGGGCGAAATTAAATAA

SEQ ID NO: 7 (CHAPT_w_Ami2638_M23-LST_CBD2638)

30 ATGAAAACCCTGAAACAAGCAGAGTCCTACATTAAGAGTAAAGTAAATAC
AGGAACTGATTTTGATGGTTTATATGGGTATCAGTGTATGGACTTAGCAGT
AGATTATATTTACCATGTAACAGATGGTAAAATAAGAATGTGGGGTAATGC
TAAGGATGCGATAAATAACTCTTTTGGTGGTACTGCTACGGTATATAAAAA

CTACCCTGCTTTTAGACCTAAGTACGGTGATGTAGTCGTATGGACTACTGGT
AATTTTGCAACTTATGGTCATATCGCAATAGTTACTAACCCTGACCCTTATG
GAGACCTTCAATATGTTACAGTTCTTGAACAAAACCTGGAACGGTAACGGGA
TTTATAAAACCGAGTTAGCTACAATCAGAACACACGATTACACAGGAATTA
5 CACATTTTATTAAAGACGCAAAGAAAGATGAAAAATCACAAGTATGTAGTG
GTTTGGCTATGGAAAAATATGACATTACAAATTTAAATGCTAAACAAGATA
AATCAAAGAATGGGAGCGTGAAAGAGTTGAAACATATCTATTCAAACCAT
ATTAAAGGTAACAAGATTACAGCACCAAAACCTAGTATTCAAGGTGTGGTC
ATCCACAATGATTATGGTAGTATGACACCTAGTCAATACTTACCATGGTTAT
10 ATGCACGTGAGAATAACGGTACACACGTTAACGGTTGGGCTAGTGTTTATG
CAAATAGAAACGAAGTGCTTTGGTATCATCCGACAGACTACGTAGAGTGCC
ATTGTGGTAATCAATGGGCAAATGCTAACTTAATCGGATTTGAAGTGTGTG
AGTCGTATCCTGGTAGAATCTCGGACAAATTATTCTTAGAAAATGAAGAAG
CGACATTGAAAGTAGCTGCGGATGTGATGAAGTCGTACGGATTACCAGTTA
15 ATCGCAACACTGTACGTCTGCATAACGAATTCTTCGGAACCTTCTTGTCACA
TCGTTCTGTGGGACTTGCATGTTGGCAAAGGTGAGCCTTACACAACCTACTAA
TATTAATAAAAATGAAAGACTACTTCATCAAACGCATCAAACATTATTATGA
CGGTGGAAAGCTAGAAGTAAGCAAAGCAGCAACTATCAAACAATCTGACG
TTAAGCAAGAAGTTAAAAAGCAAGAAGCAAAACAAATTGTGAAAGCAACA
20 GATGCTGCAACACATGAACATTCAGCACAATGGTTGAATAATTACAAAAAA
GGATATGGTTACGGTCCTTATCCATTAGGTATAAATGGCGGTATGCACTAC
GGAGTTGATTTTTTTTATGAATATTGGAACACCAGTAAAAGCTATTTCAAGC
GGAAAAATAGTTGAAGCTGGTTGGAGTAATTACGGAGGAGGTAATCAAAT
AGGTCTTATTGAAAATGATGGAGTGCATAGACAATGGTATATGCATCTAAG
25 TAAATATAATGTTAAAGTAGGAGATTATGTCAAAGCTGGTCAAATAATCGG
TTGGTCTGGAAGCACTGGTTATTCTACAGCACCAATTTACACTTCCAAAGA
ATGGTTAATTCATTTTCAAATTCAACTGCCCAAGATCCAATGCCTTTCTTAA
AGAGCGCAGGATATGGAAAAGCAGGTGGTACAGTAACTCCAACGCCGAAT
ACAGGTTGGAAACAGAATAAAGATGGCATTGTTGGTATAAAGCTGAACATGCT
30 TCGTTCACAGTGACAGCACCAAGAGGGAATTATCACAAGATACAAAGGTCCT
TGGACTGGTCACCCACAAGCTGGTGTATTACAAAAAGGTCAAACGATTA
TATGATGAGGTTCAAAAATTTGACGGTCATGTTTGGGTATCGTGGGAAACG

TTTGAGGGCGAAACTGTATACATGCCGGTACGCACATGGGACGCTAAAACT
GGTAAAGTTGGTAAGTTGTGGGGCGAAATTAAATAA

SEQ ID NO: 8 (CHAPTw_Ami2638_M23-LST_CWT-LST +6xHis and cloning site)

5 ATGAGAGGATCGCATCACCATCACCATCACGGATCCATGAAAACCCTGAAA
CAAGCAGAGTCCTACATTAAGAGTAAAGTAAATACAGGAACTGATTTTGAT
GGTTTATATGGGTATCAGTGTATGGACTTAGCAGTAGATTATATTTACCATG
TAACAGATGGTAAAATAAGAATGTGGGGTAATGCTAAGGATGCGATAAAT
AACTCTTTTGGTGGTACTGCTACGGTATATAAAAACTACCCTGCTTTTAGAC
10 CTAAGTACGGTGATGTAGTCGTATGGACTACTGGTAATTTTGCAACTTATGG
TCATATCGCAATAGTTACTAACCCTGACCCTTATGGAGACCTTCAATATGTT
ACAGTTCCTTGAACAAAACCTGGAACGGTAACGGGATTTATAAAACCGAGTTA
GCTACAATCAGAACACACGATTACACAGGAATTACACATTTTATTAAAGAC
GCAAAGAAAGATGAAAAATCACAAGTATGTAGTGGTTTGGCTATGGAAAA
15 ATATGACATTACAAATTTAAATGCTAAACAAGATAAATCAAAGAATGGGA
GCGTGAAAGAGTTGAAACATATCTATTCAAACCATATTAAAGGTAACAAGA
TTACAGCACCAAAACCTAGTATTCAAGGTGTGGTCATCCACAATGATTATG
GTAGTATGACACCTAGTCAATACTTACCATGGTTATATGCACGTGAGAATA
ACGGTACACACGTTAACGGTTGGGCTAGTGTTTATGCAAATAGAAACGAAG
20 TGCTTTGGTATCATCCGACAGACTACGTAGAGTGGCATTGTGGTAATCAAT
GGGCAAATGCTAACTTAATCGGATTTGAAGTGTGTGAGTCGTATCCTGGTA
GAATCTCGGACAAATTATTCTTAGAAAATGAAGAAGCGACATTGAAAGTAG
CTGCGGATGTGATGAAGTCGTACGGATTACCAGTTAATCGCAACACTGTAC
GTCTGCATAACGAATTCTTCGGAACCTTCTTGTCACATCGTTCGTGGGACTT
25 GCATGTTGGCAAAGGTGAGCCTTACACAACCTACTAATATTAATAAAATGAA
AGACTACTTCATCAAACGCATCAAACATTATTATGACGGTGGAAAGCTAGA
AGTAAGCAAAGCAGCAACTATCAAACAATCTGACGTTAAGCAAGAAGTTA
AAAAGCAAGAAGCAAAACAAATTGTGAAAGCAACAGATGCTGCAACACAT
GAACATTACGCACAATGGTTGAATAATTACAAAAAAGGATATGGTTACGGT
30 CCTTATCCATTAGGTATAAATGGCGGTATGCACTACGGAGTTGATTTTTTTA
TGAATATTGGAACACCAGTAAAAGCTATTTCAAGCGGAAAAATAGTTGAAG
CTGGTTGGAGTAATTACGGAGGAGGTAATCAAATAGGTCTTATTGAAAATG
ATGGAGTGCATAGACAATGGTATATGCATCTAAGTAAATATAATGTTAAAG

TAGGAGATTATGTCAAAGCTGGTCAAATAATCGGTTGGTCTGGAAGCACTG
GTTATTCTACAGCACCACATTTACACTTCCAAAGAATGGTTAATTCATTTTC
AAATTCAAAGTCCCCAAGATCCAATGCCTTTCTTAAAGAGCGCAGGATATGG
AAAAGCAGGTGGTACAGTAACTCCAACGCCGAATACAGGTGGAAAAACAA
5 ACAAATATGGCACACTATATAAATCAGAGTCAGCTAGCTTCACACCTAATA
CAGATATAATAACAAGAACGACTGGTCCATTTAGAAGCATGCCGCAGTCAG
GAGTCTTAAAAGCAGGTCAAACAATTCATTATGATGAAGTGATGAAACAAG
ACGGTCATGTTTGGGTAGGTTATACAGGTAACAGTGGCCAACGTATTTACT
TGCCTGTAAGAACATGGAATAAATCTACTAATACTTTAGGTGTTCTTTGGGG
10 AACTATAAAGTAA

SEQ ID NO: 9 (M23-LST_Ami2638_CBD2638)

GCTGCAACACATGAACATTCAGCACAATGGTTGAATAATTACAAAAAAGG
ATATGGTTACGGTCCTTATCCATTAGGTATAAATGGCGGTATGCACTACGG
15 AGTTGATTTTTTTATGAATATTGGAACACCAGTAAAAGCTATTTCAAGCGG
AAAAATAGTTGAAGCTGGTTGGAGTAATTACGGAGGAGGTAATCAAATAG
GTCTTATTGAAAATGATGGAGTGCATAGACAATGGTATATGCATCTAAGTA
AATATAATGTTAAAGTAGGAGATTATGTCAAAGCTGGTCAAATAATCGGTT
GGTCTGGAAGCACTGGTTATTCTACAGCACCACATTTACACTTCCAAAGAA
20 TGGTTAATTCATTTTCAAATTCAAAGTCCCCAAGATCCAATGCCTTTCTTAAA
GAGCGCAGGATATGGAAAAGCAGGTGGTACAGTAACTCCAACGCCGAATA
CAGGTGAGCTCTTACGCCCTAAAGACGCAAAGAAAGATGAAAAATCACAA
GTATGTAGTGGTTTGGCTATGGAAAAATATGACATTACAAATTTAAATGCT
AAACAAGATAAATCAAAGAATGGGAGCGTGAAAGAGTTGAAACATATCTA
25 TTCAAACCATATTAAGGTAACAAGATTACAGCACCAAAACCTAGTATTCA
AGGTGTGGTCATCCACAATGATTATGGTAGTATGACACCTAGTCAATACTT
ACCATGGTTATATGCACGTGAGAATAACGGTACACACGTTAACGGTTGGGC
TAGTGTTTATGCAAATAGAAACGAAGTGCTTTGGTATCATCCGACAGACTA
CGTAGAGTGGCATTGTGGTAATCAATGGGCAAATGCTAACTTAATCGGATT
30 TGAAGTGTGTGAGTCGTATCCTGGTAGAATCTCGGACAAATTATTCTTAGA
AAATGAAGAAGCGACATTGAAAGTAGCTGCGGATGTGATGAAGTCGTACG
GATTACCAGTTAATCGCAACACTGTACGTCTGCATAACGAATTCTTCGGAA
CTTCTTGTCCACATCGTTCGTGGGACTTGCATGTTGGCAAAGGTGAGCCTTA

CACAACTACTAATATTAATAAAAATGAAAGACTACTTCATCAAACGCATCAA
ACATTATTATGACGGTGGAAAGCTAGAAGTAAGCAAAGCAGCAACTATCA
AACAACTCTGACGTTAAGCAAGAAGTTAAAAAGCAAGAAGCAAAACAAATT
GTGAAAGCAACAGATTGGAAACAGAATAAAGATGGCATTGTTGGTATAAAGC
5 TGAACATGCTTCGTTACAGTGACAGCACCAGAGGGAATTATCACAAGATA
CAAAGGTCCTTGGACTGGTCACCCACAAGCTGGTGTATTACAAAAAGGTCA
AACGATTAAATATGATGAGGTTCAAAAATTTGACGGTCATGTTTGGGTATC
GTGGGAAACGTTTGGAGGGCGAACTGTATACATGCCGGTACGCACATGGG
ACGCTAAAACCTGGTAAAGTTGGTAAGTTGTGGGGCGAAATTAAATAA

10

SEQ ID NO: 10 (LST_LST + 6xHis and cloning sites)

ATGAGAGGATCGCATCACCATCACCATCACGGATCCGCTGCAACACATGAA
CATTCAGCACAATGGTTGAATAATTACAAAAAAGGATATGGTTACGGTCCT
TATCCATTAGGTATAAATGGCGGTATGCACTACGGAGTTGATTTTTTTATGA
15 ATATTGGAACACCAGTAAAAGCTATTTCAAGCGGAAAAATAGTTGAAGCTG
GTTGGAGTAATTACGGAGGAGGTAATCAAATAGGTCTTATTGAAAATGATG
GAGTGCATAGACAATGGTATATGCATCTAAGTAAATATAATGTTAAAGTAG
GAGATTATGTCAAAGCTGGTCAAATAATCGGTTGGTCTGGAAGCACTGGTT
ATTCTACAGCACCACATTTACACTTCCAAAGAATGGTTAATTCATTTTCAAA
20 TTCAACTGCCCAAGATCCAATGCCTTTCTTAAAGAGCGCAGGATATGGAAA
AGCAGGTGGTACAGTAACTCCAACGCCGAATACAGGTTGGAAAACAAACA
AATATGGCACACTATATAAATCAGAGTCAGCTAGCTTCACACCTAATACAG
ATATAATAACAAGAACGACTGGTCCATTTAGAAGCATGCCGCAGTCAGGAG
TCTTAAAAGCAGGTCAAACAATTCATTATGATGAAGTGATGAAACAAGACG
25 GTCATGTTTGGGTAGGTTATACAGGTAACAGTGGCCAACGTATTTACTTGCC
TGTAAGAACATGGAATAAATCTACTAATACTTTAGGTGTTCTTTGGGGAAC
TATAAAGGAGCTCGCTGCAACACATGAACATTCAGCACAATGGTTGAATAA
TTACAAAAAAGGATATGGTTACGGTCCTTATCCATTAGGTATAAATGGCGG
TATGCACTACGGAGTTGATTTTTTTTATGAATATTGGAACACCAGTAAAAGCT
30 ATTTCAAGCGGAAAAATAGTTGAAGCTGGTTGGAGTAATTACGGAGGAGGT
AATCAAATAGGTCTTATTGAAAATGATGGAGTGCATAGACAATGGTATATG
CATCTAAGTAAATATAATGTTAAAGTAGGAGATTATGTCAAAGCTGGTCAA
ATAATCGGTTGGTCTGGAAGCACTGGTTATTCTACAGCACCACATTTACACT

TCCAAAGAATGGTTAATTCATTTTCAAATTCAACTGCCCAAGATCCAATGCC
TTTCTTAAAGAGCGCAGGATATGGAAAAGCAGGTGGTACAGTAACTCCAAC
GCCGAATACAGGTTGGAAAACAAACAAATATGGCACACTATATAAATCAG
AGTCAGCTAGCTTCACACCTAATACAGATATAATAACAAGAACGACTGGTC
5 CATTTAGAAGCATGCCGCAGTCAGGAGTCTTAAAAGCAGGTCAAACAATTC
ATTATGATGAAGTGATGAAACAAGACGGTCATGTTTGGGTAGGTTATACAG
GTAACAGTGGCCAACGTATTTACTTGCCTGTAAGAACATGGAATAAATCTA
CTAATACTTTAGGTGTTCTTTGGGGAAGTATAAAAGTGA

10 SEQ ID NO: 11 (M23-LST_M23-LST_CWT-LST +6xHis and cloning sites)
ATGAGAGGATCGCATCACCATCACCATCACGGATCCGCTGCAACACATGAA
CATTTCAGCACAATGGTTGAATAATTACAAAAAAGGATATGGTTACGGTCCT
TATCCATTAGGTATAAATGGCGGTATGCACTACGGAGTTGATTTTTTTATGA
ATATTGGAACACCAGTAAAAGCTATTTCAAGCGGAAAAATAGTTGAAGCTG
15 GTTGGAGTAATTACGGAGGAGGTAATCAAATAGGTCTTATTGAAAATGATG
GAGTGCATAGACAATGGTATATGCATCTAAGTAAATATAATGTTAAAGTAG
GAGATTATGTCAAAGCTGGTCAAATAATCGGTTGGTCTGGAAGCACTGGTT
ATTCTACAGCACCACATTTACACTTCCAAAGAATGGTTAATTCATTTTCAA
TTCAACTGCCCAAGATCCAATGCCTTTCTTAAAGAGCGCAGGATATGGAAA
20 AGCAGGTGGTACAGTAACTCCAACGCCGAATACAGGTGAGCTCGCTGCAAC
ACATGAACATTTCAGCACAATGGTTGAATAATTACAAAAAAGGATATGGTTA
CGGTCCTTATCCATTAGGTATAAATGGCGGTATGCACTACGGAGTTGATTTT
TTTATGAATATTGGAACACCAGTAAAAGCTATTTCAAGCGGAAAAATAGTT
GAAGCTGGTTGGAGTAATTACGGAGGAGGTAATCAAATAGGTCTTATTGAA
25 AATGATGGAGTGCATAGACAATGGTATATGCATCTAAGTAAATATAATGTT
AAAGTAGGAGATTATGTCAAAGCTGGTCAAATAATCGGTTGGTCTGGAAGC
ACTGGTTATTCTACAGCACCACATTTACACTTCCAAAGAATGGTTAATTCAT
TTTCAAATTCAACTGCCCAAGATCCAATGCCTTTCTTAAAGAGCGCAGGAT
ATGGAAAAGCAGGTGGTACAGTAACTCCAACGCCGAATACAGGTTGGAAA
30 ACAAACAAATATGGCACACTATATAAATCAGAGTCAGCTAGCTTCACACCT
AATACAGATATAATAACAAGAACGACTGGTCCATTTAGAAGCATGCCGCAG
TCAGGAGTCTTAAAAGCAGGTCAAACAATTCATTATGATGAAGTGATGAAA
CAAGACGGTCATGTTTGGGTAGGTTATACAGGTAACAGTGGCCAACGTATT

TACTTGCCTGTAAGAACATGGAATAAATCTACTAATACTTTAGGTGTTCTTT
GGGGAAC TATAAAGTGA

SEQ ID NO: 12 (CBD-2638)

5 TGGAAACAGAATAAAGATGGCATT TGGTATAAAGCTGAACATGCTTCGTTT
ACAGTGACAGCACCAGAGGGAATTATCACAAGATACAAAGGTCCTTGGAC
TGGTCACCCACAAGCTGGTGTATTACAAAAAGGTCAAACGATTAAATATGA
TGAGGTTCAAAAATTTGACGGTCATGTTTGGGTATCGTGGGAAACGTTTGA
GGGCGAAACTGTATACATGCCGGTACGCACATGGGACGCTAAAACTGGTA
10 AAGTTGGTAAGTTGTGGGGCGAAATTAAATAA

SEQ ID NO: 13 (CWT-LST)

TGGAAAACAAACAAATATGGCACACTATATAAATCAGAGTCAGCTAGCTTC
ACACCTAATACAGATATAATAACAAGAACGACTGGTCCATTTAGAAGCATG
15 CCGCAGTCAGGAGTCTTAAAAGCAGGTCAAACAATTCATTATGATGAAGTG
ATGAAACAAGACGGTCATGTTTGGGTAGGTTATACAGGTAACAGTGGCCAA
CGTATTTACTTGCCTGTAAGAACATGGAATAAATCTACTAATACTTTAGGTG
TTCTTTGGGGAAC TATAAAGTGA

20 SEQ ID NO: 14 (M23-2638)

ATGCTAACTGCTATTGACTATCTTACGAAAAAAGGTTGGAAAATATCATCT
GACCCTCGCACTTACGATGGTTACCCTAAAAACTACGGCTACAGAAATTAC
CATGAAAACGGCATTAATTATGATGAGTTTTGTGGTGGTTATCATAGAGCTT
TTGATGTTTACAGTAACGAACTAACGACGTGCCTGCTGTTACTAGCGGAA
25 CAGTTATTGAAGCAAACGATTACGGTAATTTTGGTGGTACATTCGTTATTAG
AGACGCTAACGATAACGATTGGATATATGGGCATCTACAACGTGGCTCAAT
GCGATTTGTTGTAGGCGACAAAGTCAATCAAGGTGACATTATTGGTTTACA
AGGTAATAGCAACTATTACGACAATCCTATGAGTGTACATTTACATTTACA
ATTACGCCCTAAAGACGCAAAGAAAGAT

30

SEQ ID NO: 15 (M23-LST)

GCTGCAACACATGAACATT CAGCACAATGGTTGAATAATTACAAAAAAGG
ATATGGTTACGGTCCTTATCCATTAGGTATAAATGGCGGTATGCACTACGG

AGTTGATTTTTTTATGAATATTGGAACACCAGTAAAAGCTATTTCAAGCGG
AAAAATAGTTGAAGCTGGTTGGAGTAATTACGGAGGAGGTAATCAAATAG
GTCTTATTGAAAATGATGGAGTGCATAGACAATGGTATATGCATCTAAGTA
AATATAATGTTAAAGTAGGAGATTATGTCAAAGCTGGTCAAATAATCGGTT
5 GGTCTGGAAGCACTGGTTATTCTACAGCACCACATTTACACTTCCAAAGAA
TGGTTAATTCATTTTCAAATTCAACTGCCCAAGATCCAATGCCTTTCTTAAA
GAGCGCAGGATAT

SEQ ID NO: 16 (Ami-2638)

10 GGTAACAAGATTACAGCACCAAAACCTAGTATTCAAGGTGTGGTCATCCAC
AATGATTATGGTAGTATGACACCTAGTCAATACTTACCATGGTTATATGCAC
GTGAGAATAACGGTACACACGTTAACGGTTGGGCTAGTGTTTATGCAAATA
GAAACGAAGTGCTTTGGTATCATCCGACAGACTACGTAGAGTGGCATTGTG
GTAATCAATGGGCAAATGCTAACTTAATCGGATTTGAAGTGTGTGAGTCGT
15 ATCCTGGTAGAATCTCGGACAAATTATTCTTAGAAAATGAAGAAGCGACAT
TGAAAGTAGCTGCGGATGTGATGAAGTCGTACGGATTACCAGTTAATCGCA
ACACTGTACGTCTGCATAACGAATTCTTCGGAACCTTCTTGTCACATCGTTC
GTGGGACTTGCATGTTGGCAAAGGTGAGCCTTACACAACCTACTAATATTAA
TAAAATGAAAGACTACTTCATCAAACGCATCAAACATTATTATGACGGT

20

SEQ ID NO: 17 (Ami-□11)

AAGCCACAACCTAAAGCAGTAGAACTTAAAATCATCAAAGATGTGGTTAA
AGGTTATGACCTACCTAAGCGTGGTAGTAACCCTAAAGGTATAGTTATACA
CAACGACGCAGGGAGCAAAGGGGGCGACTGCTGAAGCATATCGTAACGGAT
25 TAGTAAATGCACCTTTATCAAGATTAGAAGCGGGCATTGCGCATAGTTACG
TATCAGGCAACACAGTTTGGCAAGCCTTAGATGAATCACAAGTAGGTTGGC
ATACCGCTAATCAAATAGGTAATAAATATTATTACGGTATTGAAGTATGTC
AATCAATGGGCGCAGATAACGCGACATTCTTAAAAAATGAACAGGCAACTT
TCCAAGAATGCGCTAGATTGTTGAAAAAATGGGGATTACCAGCAAACAGA
30 AATACAATCAGATTGCACAATGAATTTACTTCAACATCATGCCCTCATAGA
AGTTCGGTTTTACACACTGGTTTTGACCCAGTAACTCGCGGTCTATTGCCAG
AAGACAAGCGGTTGCAACTTAAAGACTACTTTATCAAGCAGATTAGGGCGT
ACATGGATGGTAAAATACCGGTTGCCACTGTCTCTAATGAGTCAAGCGCTT

CAAGTAATACAGTTAAACCAGTTGCAAGTGCA

SEQ ID NO:18 (CHAP-□11)

ATGCAAGCAAAATTAACATAAAATGAGTTTATAGAGTGGTTGAAAACCTTCT
5 GAGGGAAAACAATTCAATGTGGACTTATGGTATGGATTTCATGCTTTGAT
TATGCCAATGCTGGTTGGAAAGTTTTGTTTGGATTACTTCTAAAAGGTTTAG
GTGCAAAAGATATTCCGTTTCGCTAACAACTTCGACGGATTAGCTACTGTAT
ACCAAAATACACCGGACTTCTTAGCACAACTGGCGACATGGTGGTATTTCG
GTAGCAACTACGGTGCTGGATATGGTCACGTTGCATGGGTAATTGAAGCAA
10 CTTTAGATTACATCATTGTATATGAGCAGAATTGGCTAGGCGGTGGCTGGA
CTGACGGAATCGAACAACCCGGCTGGGGTTGGGAAAAAGTTACAAGACGA
CAACATGCTTATGATTTCCCTATGTGGTTTATCCGTCCGAATTTTAAAAGTG
AGACAGCGCCACGATCAGTTCAATCTCCTACACAAGCACCTAAAAAAGAA
ACAGCT

15

SEQ ID NO:19 (CHAP-□Twort)

ATGAAAACCCTGAAACAAGCAGAGTCCTACATTAAGAGTAAAGTAAATAC
AGGAACTGATTTTGATGGTTTATATGGGTATCAGTGTATGGACTTAGCAGT
AGATTATATTTACCATGTAACAGATGGTAAAATAAGAATGTGGGGTAATGC
20 TAAGGATGCGATAAATAACTCTTTTGGTGGTACTGCTACGGTATATAAAAA
CTACCCTGCTTTTAGACCTAAGTACGGTGATGTAGTCGTATGGACTACTGGT
AATTTTGCAACTTATGGTCATATCGCAATAGTTACTAACCCTGACCCTTATG
GAGACCTTCAATATGTTACAGTTCTTGAACAAAACCTGGAACGGTAACGGGA
TTTATAAAACCGAGTTAGCTACAATCAGAACACACGATTACACAGGAATTA
25 CACATTTTATT

SEQ ID NO:20 (M23-2638_Ami2638_CBD2638_CBD2638)

ATGCTAACTGCTATTGACTATCTTACGAAAAAAGGTTGGAAAATATCATCT
GACCCTCGCACTTACGATGGTTACCCTAAAAACTACGGCTACAGAAATTAC
30 CATGAAAACGGCATTAAATTATGATGAGTTTTGTGGTGGTTATCATAGAGCTT
TTGATGTTTACAGTAACGAACTAACGACGTGCCTGCTGTTACTAGCGGAA
CAGTTATTGAAGCAAACGATTACGGTAATTTTGGTGGTACATTCGTTATTAG
AGACGCTAACGATAACGATTGGATATATGGGCATCTACAACGTGGCTCAAT

GCGATTTGTTGTAGGCGACAAAGTCAATCAAGGTGACATTATTGGTTTACA
AGGTAATAGCAACTATTACGACAATCCTATGAGTGTACATTTACATTTACA
ATTACGCCCTAAAGACGCAAAGAAAGATGAAAAATCACAAGTATGTAGTG
GTTTGGCTATGGAAAAATATGACATTACAAATTTAAATGCTAAACAAGATA
5 AATCAAAGAATGGGAGCGTGAAAGAGTTGAAACATATCTATTCAAACCAT
ATTAAAGGTAACAAGATTACAGCACCAAAACCTAGTATTCAAGGTGTGGTC
ATCCACAATGATTATGGTAGTATGACACCTAGTCAATACTTACCATGGTTAT
ATGCACGTGAGAATAACGGTACACACGTTAACGGTTGGGCTAGTGTTTATG
CAAATAGAAACGAAGTGCTTTGGTATCATCCGACAGACTACGTAGAGTGGC
10 ATTGTGGTAATCAATGGGCAAATGCTAACTTAATCGGATTTGAAGTGTGTG
AGTCGTATCCTGGTAGAATCTCGGACAAATTATTCTTAGAAAATGAAGAAG
CGACATTGAAAGTAGCTGCGGATGTGATGAAGTCGTACGGATTACCAGTTA
ATCGCAACACTGTACGTCTGCATAACGAATTCTTCGGAACCTTCTTGTCACA
TCGTTTCGTGGGACTTGTCATGTTGGCAAAGGTGAGCCTTACACAACACTACTAA
15 TATTAATAAAAATGAAAGACTACTTCATCAAACGCATCAAACATTATTATGA
CGGTGGAAAGCTAGAAGTAAGCAAAGCAGCAACTATCAAACAATCTGACG
TTAAGCAAGAAGTTAAAAAGCAAGAAGCAAAACAAATTGTGAAAGCAACA
GATTGGAAACAGAATAAAGATGGCATTGTTGGTATAAAGCTGAACATGCTTCG
TTCACAGTGACAGCACCAGAGGGAATTATCACAAGATACAAAGGTCCTTGG
20 ACTGGTCACCCACAAGCTGGTGTATTACAAAAAGGTCAAACGATTAAATAT
GATGAGGTTCAAAAATTTGACGGTCATGTTTGGGTATCGTGGGAAACGTTT
GAGGGCGAAACTGTATACATGCCGGTACGCACATGGGACGCTAAACTGG
TAAAGTTGGTAAGTTGTGGGGCGAAATTAAAGAGCTCGGTGGAAAGCTAG
AAGTAAGCAAAGCAGCAACTATCAAACAATCTGACGTTAAGCAAGAAGTT
25 AAAAAGCAAGAAGCAAAACAAATTGTGAAAGCAACAGATTGGAAACAGA
ATAAAGATGGCATTGTTGGTATAAAGCTGAACATGCTTCGTTACAGTGACAG
CACCAGAGGGAATTATCACAAGATACAAAGGTCCTTGGACTGGTCACCCAC
AAGCTGGTGTATTACAAAAAGGTCAAACGATTAAATATGATGAGGTTCAAA
AATTTGACGGTCATGTTTGGGTATCGTGGGAAACGTTTGAGGGCGAAACTG
30 TATACATGCCGGTACGCACATGGGACGCTAAACTGGTAAAGTTGGTAAGT
TGTGGGGCGAAATTAAATAA

MRGSHHHHHHGSMLTAIDYLTKKGWKISSDPRTYDGYPKNYGYRNYHENG
NYDEFCGGYHRAFDVYSNETNDVPAVTS GTVIEANDYGNFGGTFVIRDANDN
DWIYGHLQRGSMRFVVGDKVNQGDII GLQGNSNYYDNPM SVHLHLQLRPKD
AKKDEKSQVCSGLAMEKYDITNLNAKQDKSKNGSVKELKHIYSNHIKGNKITA
 5 **PKPSIQGVVIHNDYGSMTPSQYLPWLYARENNGTHVNGWASVYANRNEVLW**
YHPTDYVEWHCGNQWANANLIGFEVCESYPGRISDKLFLENEEATLKVAADV
MKSYGLPVNRNTVRLHNEFFGTSCPHRSWDLHVGKGEPYTTTNINKMKDYFI
KRIKHYYDGGKLEVSKAATIKQSDVKQEVKKQEAKQIVKATDWKQNKDGIW
YKAEHASFTVTAPEGIITRYKGPWTGHPQAGVLQKGQTIKYDEVQKFDGHVW
 10 **VSWETFEGETVYMPVRTWDAKTGKVGKLWGEIK**

SEQ ID NO: 22 (Ami2638_CBD2638)

MRGSHHHHHHGSRLPKDAKKDEKSQVCSGLAMEKYDITNLNAKQDKSKNG
SVKELKHIYSNHIKGNKITAPKPSIQGVVIHNDYGSMTPSQYLPWLYARENNGT
 15 **HVNGWASVYANRNEVLWYHPTDYVEWHCGNQWANANLIGFEVCESYPGRIS**
DKLFLENEEATLKVAADVMKSYGLPVNRNTVRLHNEFFGTSCPHRSWDLHVG
KGEPYTTTNINKMKDYFIKRIKHYYDGGKLEVSKAATIKQSDVKQEVKKQEAK
QIVKATDWKQNKDGIWYKAEHASFTVTAPEGIITRYKGPWTGHPQAGVLQKG
QTIKYDEVQKFDGHVWVSWETFEGETVYMPVRTWDAKTGKVGKLWGEIK

20

SEQ ID NO: 23 (M23-2638_CBD2638)

MRGSHHHHHHGSMLTAIDYLTKKGWKISSDPRTYDGYPKNYGYRNYHENG
NYDEFCGGYHRAFDVYSNETNDVPAVTS GTVIEANDYGNFGGTFVIRDANDN
DWIYGHLQRGSMRFVVGDKVNQGDII GLQGNSNYYDNPM SVHLHLQLRPKD
 25 **AKKDEKSQVCSGLAMEKYDITNLNAKQDKSKNGSVKELKHIYSNHIKGNKITA**
PKPSIQGELGGKLEVSKAATIKQSDVKQEVKKQEAKQIVKATDWKQNKDGIW
YKAEHASFTVTAPEGIITRYKGPWTGHPQAGVLQKGQTIKYDEVQKFDGHVW
VSWETFEGETVYMPVRTWDAKTGKVGKLWGEIK

30 SEQ ID NO: 24 (Ply2638-Ply2638)

MRGSHHHHHHGSMLTAIDYLTKKGWKISSDPRTYDGYPKNYGYRNYHENG
NYDEFCGGYHRAFDVYSNETNDVPAVTS GTVIEANDYGNFGGTFVIRDANDN
DWIYGHLQRGSMRFVVGDKVNQGDII GLQGNSNYYDNPM SVHLHLQLRPKD

AKKDEKSQVCSGLAMEKYDITNLNAKQDKSKNGSVKELKHIYSNHIKGNKITA
PKPSIQGVVIHNDYGSMTPSQYLPWLYARENNGTHVNGWASVYANRNEVLW
YHPTDYVEWHCGNQWANANLIGFEVCESYPGRISDKLFLNEEEATLKVAADV
MKSYPGLPVNRNTVRLHNEFFGTSCPHRSWDLHVGKGEPYTTTNINKMKDYFI
5 KRIKHYYDGGKLEVSKAATIKQSDVKQEVKKQEAQIVKATDWKQNKDGIW
YKAEHASFTVTAPEGIITRYKGPWTGHPQAGVLQKGQTIKYDEVQKFDGHVW
VSWETFEGETVYMPVRTWDAKTGKVGKLWGEIKELMLTAIDYLTKKGWKISS
DPRTYDGYPKNYGYRNYHENGINYDEFCCGGYHRAFDVYSNETNDVPAVTS
VIEANDYGNFGGTFVIRDANDNDWIYGHLQRGSMRFFVVGDKVNQGGDIIGLQ
10 NSNYNDNPMSVHLHLQLRPKDAKKDEKSQVCSGLAMEKYDITNLNAKQDKS
KNGSVKELKHIYSNHIKGNKITAPKPSIQGVVIHNDYGSMTPSQYLPWLYARE
NNGTHVNGWASVYANRNEVLWYHPTDYVEWHCGNQWANANLIGFEVCESYP
GRISDKLFLNEEEATLKVAADV MKSYGLPVNRNTVRLHNEFFGTSCPHRSWDL
HVGKGEPYTTTNINKMKDYFIKRIKHYYDGGKLEVSKAATIKQSDVKQEVKK
15 QEAQIVKATDWKQNKDGIWYKAEHASFTVTAPEGIITRYKGPWTGHPQAGV
LQKGQTIKYDEVQKFDGHVWVSWETFEGETVYMPVRTWDAKTGKVGKLWG
EIK

SEQ ID NO: 25 (CHAP11_M23-2638_Ami2638_CBD2638)

20 **MRGSHHHHHH**GSMQAKLTKNEFIEWLKTSEGKQFNVDLWYGFQCFDYANA
GWKVLFGLLLKGLGAKDIPFANNFDGLATVYQNTPDFLAQPGDMVVFGSNYG
AGYGHVAWVIEATLDYIIVYEQNWLGGGWTDGIEQPGWGWEKVTRRQHAYD
FPMWFIRPNFKSETAPRSVQSPTQAPKKETAGSMLTAIDYLTKKGWKISSDPT
YDGYPKNYGYRNYHENGINYDEFCCGGYHRAFDVYSNETNDVPAVTS
25 NDYGNFGGTFVIRDANDNDWIYGHLQRGSMRFFVVGDKVNQGGDIIGLQ
YDNPMMSVHLHLQLRPKDAKKDEKSQVCSGLAMEKYDITNLNAKQDKSKNGS
VKELKHIYSNHIKGNKITAPKPSIQGVVIHNDYGSMTPSQYLPWLYARENNGTH
VNGWASVYANRNEVLWYHPTDYVEWHCGNQWANANLIGFEVCESYPGRISD
KLFLNEEEATLKVAADV MKSYGLPVNRNTVRLHNEFFGTSCPHRSWDLHVGK
30 GEPYTTTNINKMKDYFIKRIKHYYDGGKLEVSKAATIKQSDVKQEVKKQEAQ
IVKATDWKQNKDGIWYKAEHASFTVTAPEGIITRYKGPWTGHPQAGVLQKGQ
TIKYDEVQKFDGHVWVSWETFEGETVYMPVRTWDAKTGKVGKLWGEIK

SEQ ID NO: 26 (Ami11_M23-2638_Ami2638_CBD2638)

MRGSHHHHHHGSKPQPKAVELKIIKD VVKGYDLPKRGSNPKGIVIHNDAGSK
GATAEAYRNGLVNAPLSRLEAGIAHSYVSGNTVWQALDESQVGWHTANQIGN
KYYYGIEVCQSMGADNATFLKNEQATFQECARLLKKWGLPANRNTIRLHNEF
5 TSTSCPHRSSVLHTGFDPVTRGLLPEDKRLQLKD YFIKQIRAYMDGKIPVATVS
NESSASSNTVKPVASAGSMLTAIDYLT KKGWKISSDPRTYDGYPKNYGYRNYH
ENGINYDEFCCGYHRAFDVYSNETNDVPAVTSGTVIEANDYGNFGGTFVIRDA
NDNDWTYGHLQRGSMRFVVGDKVNQGDII GLQGNSNYNDNPM SVHLHLQLR
PKDAKKDEKSQVCSGLAMEKYDITNLNAKQDKSKNGSVKELKHIYSNHIKGN
10 KITAPKPSIQGVVIHNDYGSMTPSQYLPWLYARENNGTHVNGWASVYANRNE
VLWYHPTDYVEWHCGNQWANANLIGFEVCESYPGRISDKLFLNEEEATLKVA
ADVMKSYGLPVNRNTVRLHNEFFGTSCPHRSWDLHVGKGEPYTTTNINKMKD
YFIKRIKHYYDGGKLEVSKAATIKQSDVKQEVKKQEAKQIVKATDWKQNKDG
IWKAEHASFTVTAPEGIITRYKGPWTGHPQAGVLQKGQTIKYDEVQKFDGHV
15 WWSWETFEGETVYMPVRTWDAKTGKVGKLWGEIK

SEQ ID NO: 27 (CHAPT_w_Ami2638_M23-LST_CBD2638)

MRGSHHHHHHGSMKTLKQAESYIKSVNTGTDFDGLYGYQCMDLAVDYIY
HVTDGKIRMWGNKDAINNSFGGTATVYKNYPAFRPKYGDVVVWTTGNFAT
20 YGHIAIVTNPDYQDLQYVTVLEQNWNGNGIYKTELATIRTHDYTGITHFIKDA
KKDEKSQVCSGLAMEKYDITNLNAKQDKSKNGSVKELKHIYSNHIKGNKITAP
KPSIQGVVIHNDYGSMTPSQYLPWLYARENNGTHVNGWASVYANRNEVLWY
HPTDYVEWHCGNQWANANLIGFEVCESYPGRISDKLFLNEEEATLKVAADVM
KSYGLPVNRNTVRLHNEFFGTSCPHRSWDLHVGKGEPYTTTNINKMKDYFIKR
25 IKHYYDGGKLEVSKAATIKQSDVKQEVKKQEAKQIVKATDAATHEHSAQWLN
NYKKGYGYGPYPLGINGGMHYGVDFFMNIGTPVKAISSGKIVEAGWSNYGGG
NQIGLIENDGVHRQWYMHLSKYNVKGVDYVKAGQIIGWSGSTGYSTAPHLHF
QRMVNSFSNSTAQDPMPLKSAGYGKAGGTVTPTNTGWKQNKDGIWKAE
HASFTVTAPEGIITRYKGPWTGHPQAGVLQKGQTIKYDEVQKFDGHVWVSWE
30 TFEGETVYMPVRTWDAKTGKVGKLWGEIK

SEQ ID NO: 28 (CHAPT_w_Ami2638_M23-LST_CWT-LST)

MRGSHHHHHHGSMKTLKQAESYIKSVNTGTDFDGLYGYQCMDLAVDYIY

HVTDGKIRMWGNKDAINNSFGGTATVYKNYPAPFRPKYGDVVVWTTGNFAT
 YGHIAIVTNPDYPYGDLQYVTVLEQNWNGNGIYKTELATIRTHDYTGITHFIKDA
 KKDEKSQVCSGLAMEKYDITNLNAKQDKSKNGSVKELKHIYSNHIKGNKITAP
 KPSIQGVVIHNDYGSMTSPSYLPWLYARENNGTHVNGWASVYANRNEVLWY
 5 HPTDYVEWHCGNQWANANLIGFEVCESYPGRISDKLFLNEEATLKVAADVM
 KSYGLPVNRNTVRLHNEFFGTSCPHRSWDLHVGKGEPYTTTNINKMKDYFIKR
 IKHYDGGKLEVSKAATIKQSDVKQEVKKQEAQIVKATDAATHEHSAQWLN
 NYKKGYGYGPYPLGINGGMHYGVDFFMNIGTPVKAISSGKIVEAGWSNYGGG
 NQIGLIENDGVHRQWYMHLSKYNVKGVDYVKAGQIIGWSGSTGYSTAPHLHF
 10 QRMVNSFSNSTAQDPMPFLKSAGYGKAGGTVTPTPNTGWKTNKYGTLYKSES
 ASFTPNTDIITRTTGPFRSMPQSGVLKAGQTIHYDEVMMKQDGHVWVGTYTGNNSG
 QRIYLPVRTWNKSTNTLGVLWGTHK

SEQ ID NO: 29 (M23-LST_Ami2638_CBD2638)

15 **MRGSHHHHHH**GSAATHEHSAQWLNKYKKGYGYGPYPLGINGGMHYGVDF
 MNIGTPVKAISSGKIVEAGWSNYGGGNQIGLIENDGVHRQWYMHLSKYNV
 KVDYVKAGQIIGWSGSTGYSTAPHLHFQRMVNSFSNSTAQDPMPFLKSAGY
 GKAGGTVTPTPNTGELLRPKDAKKDEKSQVCSGLAMEKYDITNLNAKQDK
 SKNGSVKELKHIYSNHIKGNKITAPKPSIQGVVIHNDYGSMTSPSYLPWLY
 20 HVNGWASVYANRNEVLWYHPTDYVEWHCGNQWANANLIGFEVCESYPGRIS
 DKLFLNEEATLKVAADVMKSYGLPVNRNTVRLHNEFFGTSCPHRSWDLHVG
 KGEPYTTTNINKMKDYFIKRIKHYYDGGKLEVSKAATIKQSDVKQEVKKQEA
 QIVKATDWKQNKDGIWYKAHASFTVTAPEGITRYKGPWTGHPQAGVLQKG
 QTIKYDEVQKFDGHVWVSWETFEGETVYMPVRTWDAKTGKVGKLWGEIK

25

SEQ ID NO: 30 (LST_LST)

MRGSHHHHHHGSAATHEHSAQWLNKYKKGYGYGPYPLGINGGMHYGVDF
 MNIGTPVKAISSGKIVEAGWSNYGGGNQIGLIENDGVHRQWYMHLSKYNV
 KVDYVKAGQIIGWSGSTGYSTAPHLHFQRMVNSFSNSTAQDPMPFLKSAGY
 30 AGGTVTPTPNTGWKTNKYGTLYKSESASFTPNTDIITRTTGPFRSMPQSGVL
 KAQTIHYDEVMMKQDGHVWVGTYTGNNSGQRIYLPVRTWNKSTNTLGVLWG
 THKE
 LAATHEHSAQWLNKYKKGYGYGPYPLGINGGMHYGVDFFMNIGTPVKAISSG
 KIVEAGWSNYGGGNQIGLIENDGVHRQWYMHLSKYNVKGVDYVKAGQIIGW

SGSTGYSTAPHLHFQRMVNSFSNSTAQDPMPFLKSAGYGKAGGTVTPTPNTG
WKTNKYGTLYKSESASFTPNTDIITRTTGPFERSMPQSGVLKAGQTIHYDEVMK
QDGHVWVG YTGNSGQRIYLPVRTWNKSTNTLGVLWG TIK

5 SEQ ID NO: 31 (M23-LST_M23-LST_CWT-LST)

MRGSHHHHHHGSAATHEHSAQWLNNYKKGYGYGPYPLGINGGMHYGVDF
MNIGTPVKAISSGKIVEAGWSNYGGGNQIGLIENDGVHRQWYMHLSKYNVKV
GDYVKAGQIIGWSGSTGYSTAPHLHFQRMVNSFSNSTAQDPMPFLKSAGYGK
AGGTVTPTPNTGELAATHEHSAQWLNNYKKGYGYGPYPLGINGGMHYGVDF
10 FMNIGTPVKAISSGKIVEAGWSNYGGGNQIGLIENDGVHRQWYMHLSKYNVK
VG DYVKAGQIIGWSGSTGYSTAPHLHFQRMVNSFSNSTAQDPMPFLKSAGYG
KAGGTVTPTPNTGWKTNKYGTLYKSESASFTPNTDIITRTTGPFERSMPQSGVLK
AGQTIHYDEVMKQDGHVWVG YTGNSGQRIYLPVRTWNKSTNTLGVLWG TIK

15 SEQ ID NO: 32 (M23-2638_Ami2638_CBD2638_CBD2638)

MRGSHHHHHHGSM LTAIDYLTKKGWKISSDPRTYDGYPKNYGYRNYHENG
NYDEFCGGYHRAFDVYSNETNDVPAVTS GTVIEANDYGNFGGT FVIRDANDN
DWIYGHLQRGSMRFVVGDKVNQGD IIGLQGNSNY YDNPM SVHLHLQLRPKD
AKKDEKSQVCSGLAMEKYDITNLNAKQDKSKNGSVKELKHIYSNHIKGNKITA
20 PKPSIQGVVIHNDYGSMTPSQYLPWLYARENNGTHVNGWASVYANRNEVLW
YHPTDYVEWHCGNQWANANLIGFEVCE SYPGRISDKLFLENEEATLKVAADV
MKSYGLPVNRNTVRLHNEFFGTSCPHRSWDLHVGKGEPYTTTNINKMKDYFI
KRIKHYYDGGKLEVSKAATIKQSDVKQEVKKQEAKQIVKATDWKQNKDGIW
YKAEHASFTVTAPEGIITRYKGPWTGHPQAGVLQKGQTIKYDEVQKFDGHVW
25 VSWETFEGETVYMPVRTWDAKTGKVGKLWGEIKELGGKLEVSKAATIKQSDV
KQEVKKQEAKQIVKATDWKQNKDGIWYKAEHASFTVTAPEGIITRYKGPWTG
HPQAGVLQKGQTIKYDEVQKFDGHVWVSWETFEGETVYMPVRTWDAKTGK
VGKLWGEIK

30 SEQ ID NO: 33 (LST)

GCTGCAACACATGAACATTCAGCACAATGGTTGAATAATTACAAAAAAGG
ATATGGTTACGGTCCTTATCCATTAGGTATAAATGGCGGTATGCACTACGG
AGTTGATTTTTTTATGAATATTGGAACACCAGTAAAAGCTATTTCAAGCGG

AAAAATAGTTGAAGCTGGTTGGAGTAATTACGGAGGAGGTAATCAAATAG
GTCTTATTGAAAATGATGGAGTGCATAGACAATGGTATATGCATCTAAGTA
AATATAATGTTAAAGTAGGAGATTATGTCAAAGCTGGTCAAATAATCGGTT
GGTCTGGAAGCACTGGTTATTCTACAGCACCACATTTACACTTCCAAAGAA
5 TGGTTAATTCATTTTCAAATTCAACTGCCCAAGATCCAATGCCTTTCTTAAA
GAGCGCAGGATATGGAAAAGCAGGTGGTACAGTAACTCCAACGCCGAATA
CAGGTTGGAAAACAAACAAATATGGCACACTATATAAATCAGAGTCAGCT
AGCTTCACACCTAATACAGATATAATAACAAGAACGACTGGTCCATTTAGA
AGCATGCCGCAGTCAGGAGTCTTAAAAGCAGGTCAAACAATTCATTATGAT
10 GAAGTGATGAAACAAGACGGTCATGTTTGGGTAGGTTATACAGGTAACAGT
GGCCAACGTATTTACTTGCCTGTAAGAACATGGAATAAATCTACTAATACT
TTAGGTGTTCTTTGGGGAACATAAAGTGA

SEQ ID NO: 34 (LST aa)

15 AATHEHSAQWLNNYKKGYGYGPYPLGINGGMHYGVDFFMNIGTPVKAISSGK
IVEAGWSNYGGGNQIGLIENDGVHRQWYMHLSKYNVKVGDYVKAGQIIGWS
GSTGYSTAPHLHFQRMVNSFSNSTAQDPMFPLKSAGYGKAGGTVTPTPNTGW
KTNKYGTLYKSESASFTPNTDIITRTTGPFERSMPQSGVLKAGQTIHYDEVKQD
GHVWVG YTGNSGQRIYLPVRTWNKSTNTLGVLWGTIK

20

SEQ ID NO: 35 (CBD-2638)

WKQNKDGIWYKAEHASFTVTAPEGIITRYKGPWTGHPQAGVLQKGQTIKYDE
VQKFDGHVWVSWETFEGETVYMPVRTWDAKTGKVGLWGEIK

25 SEQ ID NO: 36 (CWT-LST)

WKTNKYGTLYKSESASFTPNTDIITRTTGPFERSMPQSGVLKAGQTIHYDEVK
QDGHVWVG YTGNSGQRIYLPVRTWNKSTNTLGVLWGTIK

SEQ ID NO: 37 (M23-2638)

30 MLTAIDYLTKKGWKISSDPRTYDGYPKNYGYRNYHENGINYDEFCCGYHRAF
DVYSNETNDVPAVTSGTVIEANDYGNFGGTFVIRDANDNDWIYGHLQRGSMR
FVVGDKNVQGDIIQLQGNSNYDNPMSVHLHLQLRPKDAKKD

SEQ ID NO: 38 (M23-LST)

AATHEHSAQWLNNYKKGYGYGPYPLGINGGMHYGVDFFMNIGTPVKAISSGK
IVEAGWSNYGGGNQIGLIENDGVHRQWYMHLSKYNVKVGDYVKAGQIIGWS
GSTGYSTAPHLHFQRMVNSFSNSTAQDPMPLKSAGY

5

SEQ ID NO: 39 (Ami-2638)

NKITAPKPSIQGVVIHNDYGSMTSPSYLPWLYARENNGTHVNGWASVYANRN
EVLWYHPTDYVEWHCGNQWANANLIGFEVCESYPGRISDKLFLNEEEATLKV
AADVMKSYGLPVNRNTVRLHNEFFGTSCPHRSWDLHVGKGEPYTTTNINKMK
10 DYFIKRIKHYYDG

SEQ ID NO: 40 (Ami-□11)

NPKGIVIHNDAGSKGATAEAYRNGLVNAPLSRLEAGIAHSYVSGNTVWQALD
ESQVGWHTANQIGNKYYYGIEVCQSMGADNATFLKNEQATFQECARLLKKW
15 GLPANRNTIRLHNEFTSTSCPHRSSVLHTGFDPVTRGLLPEDKRLQLKDYFIKQI
RAYMD

SEQ ID NO: 41 (CHAP-□11)

MSIIMEVATMQAKLTKNEFIEWLKTSEGKQFNVDLWYGFQCFDYANAGWKV
20 LFGLLLKGLGAKDIPFANNFDGLATVYQNTPDFLAQPGDMVVFGSNYGAGYG
HVAWVIEATLDYIIVYEQNWLGGGWTDGIEQPGWGWEKVTRRQHAYDFPMW
FIRP

SEQ ID NO: 42(CHAP-□Twort)

MKTLKQAESYIKSKVNTGTDFDGLYGYQCMDLAVDYIYHVTDGKIRMWGNA
25 KDAINNSFGGTATVYKNYPAPFRPKYGDVVVWTTGNFATYGHIAIVTNPDYPG
DLQYVTVLEQNWNGNGIYKTELATIRTHDYTGITHFI

SEQ ID NO 43 (6xHis tag N terminal)

30 MRGSHHHHHHGS

SEQ ID NO: 44 (Ply2638 + 6xHis and cloning sites)

ATGAGAGGATCGCATCACCATCACCATCACGGATCCATGCTAACTGCTATT

GACTATCTTACGAAAAAAGGTTGGAAAAATATCATCTGACCCTCGCACTTAC
GATGGTTACCCTAAAAACTACGGCTACAGAAATTACCATGAAAACGGCATT
AATTATGATGAGTTTTGTGGTGGTTATCATAGAGCTTTTGATGTTTACAGTA
ACGAAACTAACGACGTGCCTGCTGTTACTAGCGGAACAGTTATTGAAGCAA
5 ACGATTACGGTAATTTTGGTGGTACATTCGTTATTAGAGACGCTAACGATA
ACGATTGGATATATGGGCATCTACAACGTGGCTCAATGCGATTTGTTGTAG
GCGACAAAGTCAATCAAGGTGACATTATTGGTTTACAAGGTAATAGCAACT
ATTACGACAATCCTATGAGTGTACATTTACATTTACAATTACGCCCTAAAGA
CGCAAAGAAAGATGAAAAATCACAAGTATGTAGTGGTTTGGCTATGGAAA
10 AATATGACATTACAAATTTAAATGCTAAACAAGATAAATCAAAGAATGGG
AGCGTGAAAGAGTTGAAACATATCTATTCAAACCATATTAAAGGTAACAAG
ATTACAGCACCAAAACCTAGTATTCAAGGTGTGGTCATCCACAATGATTAT
GGTAGTATGACACCTAGTCAATACTTACCATGGTTATATGCACGTGAGAAT
AACGGTACACACGTTAACGGTTGGGCTAGTGTATGCAAATAGAAACGAA
15 GTGCTTTGGTATCATCCGACAGACTACGTAGAGTGGCATTGTGGTAATCAA
TGGGCAAATGCTAACTTAATCGGATTTGAAGTGTGTGAGTCGTATCCTGGT
AGAATCTCGGACAAATTATTCTTAGAAAATGAAGAAGCGACATTGAAAGTA
GCTGCGGATGTGATGAAGTCGTACGGATTACCAGTTAATCGCAACACTGTA
CGTCTGCATAACGAATTCTTCGGAACCTTCTTGTCCACATCGTTCGTGGGACT
20 TGCATGTTGGCAAAGGTGAGCCTTACACAACCTACTAATATTAATAAAATGA
AAGACTACTTCATCAAACGCATCAAACATTATTATGACGGTGGAAAGCTAG
AAGTAAGCAAAGCAGCAACTATCAAACAATCTGACGTTAAGCAAGAAGTT
AAAAAGCAAGAAGCAAAACAAATTGTGAAAGCAACAGATTGGAAACAGA
ATAAAGATGGCATTGTTGGTATAAAGCTGAACATGCTTCGTTACAGTGACAG
25 CACCAGAGGGAATTATCACAAGATACAAAGGTCCTTGGACTGGTCACCCAC
AAGCTGGTGTATTACAAAAAGGTCAAACGATTAAATATGATGAGGTTCAAA
AATTTGACGGTCATGTTTGGGTATCGTGGGAAACGTTTGAGGGCGAAACTG
TATACATGCCGGTACGCACATGGGACGCTAAAACTGGTAAAGTTGGTAAGT
TGTGGGGCGAAATTAAATAA

30

SEQ ID NO: 45 (CHAP11_M23-2638_Ami2638_CBD2638+6xHis and cloning sites)
ATGAGAGGATCGCATCACCATCACCATCACGGATCCATGCAAGCAAAATTA
ACTAAAAATGAGTTTATAGAGTGGTTGAAAACCTTCTGAGGGGAAAACAATTC

AATGTGGACTTATGGTATGGATTTCAATGCTTTGATTATGCCAATGCTGGTT
GGAAAGTTTTGTTTGGATTACTTCTAAAAGGTTTAGGTGCAAAAGATATTCC
GTTTCGCTAACAACCTTCGACGGATTAGCTACTGTATACCAAATACACCGGA
CTTCTTAGCACAACCTGGCGACATGGTGGTATTCGGTAGCAACTACGGTGC
5 TGGATATGGTCACGTTGCATGGGTAATTGAAGCAACTTTAGATTACATCATT
GTATATGAGCAGAATTGGCTAGGCGGTGGCTGGACTGACGGAATCGAACA
ACCCGGCTGGGGTTGGGAAAAAGTTACAAGACGACAACATGCTTATGATTT
CCCTATGTGGTTTATCCGTCCGAATTTTAAAAGTGAGACAGCGCCACGATC
AGTTCAATCTCCTACACAAGCACCTAAAAAAGAAACAGCTGGATCCATGCT
10 AACTGCTATTGACTATCTTACGAAAAAAGGTTGGAAAATATCATCTGACCC
TCGCACTTACGATGGTTACCCTAAAAACTACGGCTACAGAAATTACCATGA
AAACGGCATTAAATTATGATGAGTTTTGTGGTGGTTATCATAGAGCTTTTGAT
GTTTACAGTAACGAACTAACGACGTGCCTGCTGTTACTAGCGGAACAGTT
ATTGAAGCAAACGATTACGGTAATTTTGGTGGTACATTCGTTATTAGAGAC
15 GCTAACGATAACGATTGGATATATGGGCATCTACAACGTGGCTCAATGCGA
TTTGTTGTAGGCGACAAAGTCAATCAAGGTGACATTATTGGTTTACAAGGT
AATAGCAACTATTACGACAATCCTATGAGTGTACATTTACATTTACAATTAC
GCCCTAAAGACGCAAAGAAAGATGAAAAATCACAAGTATGTAGTGGTTTG
GCTATGGAAAAATATGACATTACAAATTTAAATGCTAAACAAGATAAATCA
20 AAGAATGGGAGCGTGAAAGAGTTGAAACATATCTATTCAAACCATATTTAA
GGTAACAAGATTACAGCACCAAAACCTAGTATTCAAGGTGTGGTCATCCAC
AATGATTATGGTAGTATGACACCTAGTCAATACTTACCATGGTTATATGCAC
GTGAGAATAACGGTACACACGTTAACGGTTGGGCTAGTGTTTATGCAAATA
GAAACGAAGTGCTTTGGTATCATCCGACAGACTACGTAGAGTGGCATTGTG
25 GTAATCAATGGGCAAATGCTAACTTAATCGGATTTGAAGTGTGTGAGTCGT
ATCCTGGTAGAATCTCGGACAAATTATTCTTAGAAAATGAAGAAGCGACAT
TGAAAGTAGCTGCGGATGTGATGAAGTCGTACGGATTACCAGTTAATCGCA
ACACTGTACGTCTGCATAACGAATTCTTCGGAACCTTCTTGTCCACATCGTTC
GTGGGACTTGCATGTTGGCAAAGGTGAGCCTTACACAACCTACTAATATTAA
30 TAAAATGAAAGACTACTTCATCAAACGCATCAAACATTATTATGACGGTGG
AAAGCTAGAAGTAAGCAAAGCAGCAACTATCAAACAATCTGACGTTAAGC
AAGAAGTTAAAAAGCAAGAAGCAAAACAAATTGTGAAAGCAACAGATTGG
AAACAGAATAAAGATGGCATTGTTGGTATAAAGCTGAACATGCTTCGTTTACA

GTGACAGCACCAGAGGGAATTATCACAAGATACAAAGGTCCTTGGACTGGT
CACCCACAAGCTGGTGTATTACAAAAAGGTCAAACGATTAAATATGATGAG
GTTCAAAAATTTGACGGTCATGTTTGGGTATCGTGGGAAACGTTTGAGGGC
GAAACTGTATACATGCCGGTACGCACATGGGACGCTAAAACTGGTAAAGTT
5 GGTAAGTTGTGGGGCGAAATTAAATAA

SEQ ID NO: 46 (Ami11_M23-2638_Ami2638_CBD2638+6xHis and cloning sites)

ATGAGAGGATCGCATCACCATCACCATCACGGATCCAAGCCACAACCTAAA
GCAGTAGAACTTAAAATCATCAAAGATGTGGTTAAAGGTTATGACCTACCT
10 AAGCGTGGTAGTAACCTAAAGGTATAGTTATACACAACGACGCAGGGAG
CAAAGGGGCGACTGCTGAAGCATATCGTAACGGATTAGTAAATGCACCTTT
ATCAAGATTAGAAGCGGGCATTGCGCATAGTTACGTATCAGGCAACACAGT
TTGGCAAGCCTTAGATGAATCACAAGTAGGTTGGCATAACCGCTAATCAAAT
AGGTAATAAATATTATTACGGTATTGAAGTATGTCAATCAATGGGCGCAGA
15 TAACGCGACATTCTTAAAAAATGAACAGGCAACTTTCCAAGAATGCGCTAG
ATTGTTGAAAAAATGGGGATTACCAGCAAACAGAAATACAATCAGATTGC
ACAATGAATTTACTTCAACATCATGCCCTCATAGAAGTTTCGGTTTTACACAC
TGGTTTTGACCCAGTAACTCGCGGTCTATTGCCAGAAGACAAGCGGTTGCA
ACTTAAAGACTACTTTATCAAGCAGATTAGGGCGTACATGGATGGTAAAAT
20 ACCGGTTGCCACTGTCTCTAATGAGTCAAGCGCTTCAAGTAATACAGTTAA
ACCAGTTGCAAGTGCAGGATCCATGCTAACTGCTATTGACTATCTTACGAA
AAAAGGTTGGAAAATATCATCTGACCCTCGCACTTACGATGGTTACCCTAA
AAACTACGGCTACAGAAATTACCATGAAAACGGCATTAAATTATGATGAGTT
TTGTGGTGGTTATCATAGAGCTTTTGATGTTTACAGTAACGAAACTAACGAC
25 GTGCCTGCTGTTACTAGCGGAACAGTTATTGAAGCAAACGATTACGGTAAT
TTTGGTGGTACATTCGTTATTAGAGACGCTAACGATAACGATTGGATATAT
GGGCATCTACAACGTGGCTCAATGCGATTTGTTGTAGGCGACAAAGTCAAT
CAAGGTGACATTATTGGTTTACAAGGTAATAGCAACTATTACGACAATCCT
ATGAGTGTACATTTACATTTACAATTACGCCCTAAAGACGCAAAGAAAGAT
30 GAAAAATCACAAGTATGTAGTGGTTTGGCTATGGAAAAATATGACATTACA
AATTTAAATGCTAAACAAGATAAATCAAAGAATGGGAGCGTGAAAGAGTT
GAAACATATCTATTCAAACCATATTAAAGGTAACAAGATTACAGCACCAAA
ACCTAGTATTCAAGGTGTGGTCATCCACAATGATTATGGTAGTATGACACC

TAGTCAATACTTACCATGGTTATATGCACGTGAGAATAACGGTACACACGT
TAACGGTTGGGCTAGTGTTTATGCAAATAGAAACGAAGTGCTTTGGTATCA
TCCGACAGACTACGTAGAGTGGCATTGTGGTAATCAATGGGCAAATGCTAA
CTTAATCGGATTTGAAGTGTGTGAGTCGTATCCTGGTAGAATCTCGGACAA
5 ATTATTCTTAGAAAAATGAAGAAGCGACATTGAAAGTAGCTGCGGATGTGAT
GAAGTCGTACGGATTACCAGTTAATCGCAACACTGTACGTCTGCATAACGA
ATTCTTCGGAACCTTCTTGTCCACATCGTTCGTGGGACTTGCATGTTGGCAAA
GGTGAGCCTTACACAACCTACTAATATTAATAAAAATGAAAGACTACTTCATC
AAACGCATCAAACATTATTATGACGGTGGAAAGCTAGAAGTAAGCAAAGC
10 AGCAACTATCAAACAATCTGACGTTAAGCAAGAAGTTAAAAAGCAAGAAG
CAAAACAAATTGTGAAAGCAACAGATTGGAAACAGAATAAAGATGGCATT
TGGTATAAAGCTGAACATGCTTCGTTACAGTGACAGCACCAGAGGGAATT
ATCACAAGATACAAAGGTCCTTGGACTGGTCACCCACAAGCTGGTGTATTA
CAAAAAGGTCAAACGATTAAATATGATGAGGTTCAAAAATTTGACGGTCAT
15 GTTTGGGTATCGTGGGAAACGTTTGAGGGCGAAACTGTATACATGCCGGTA
CGCACATGGGACGCTAAAACTGGTAAAGTTGGTAAGTTGTGGGGCGAAATT
AAATAA

SEQ ID NO: 47 (CHAPTw Ami2638 M23-LST CBD2638 +6xHis and cloning sites)

20 ATGAGAGGATCGCATCACCATCACCATCACGGATCCATGAAAACCCTGAAA
CAAGCAGAGTCCTACATTAAGAGTAAAGTAAATACAGGAACTGATTTTGAT
GGTTTATATGGGTATCAGTGTATGGACTTAGCAGTAGATTATATTTACCATG
TAACAGATGGTAAAATAAGAATGTGGGGTAATGCTAAGGATGCGATAAAT
AACTCTTTTGGTGGTACTGCTACGGTATATAAAAACTACCCTGCTTTTAGAC
25 CTAAGTACGGTGATGTAGTCGTATGGACTACTGGTAATTTTGCAACTTATGG
TCATATCGCAATAGTTACTAACCCTGACCCTTATGGAGACCTTCAATATGTT
ACAGTTCTTGAACAAAACCTGGAACGGTAACGGGATTTATAAAACCGAGTTA
GCTACAATCAGAACACACGATTACACAGGAATTACACATTTTATTAAAGAC
GCAAAGAAAGATGAAAAATCACAAGTATGTAGTGGTTTGGCTATGGAAAA
30 ATATGACATTACAAATTTAAATGCTAAACAAGATAAATCAAAGAATGGGA
GCGTGAAAGAGTTGAAACATATCTATTCAAACCATATTAAAGGTAACAAGA
TTACAGCACCAAAACCTAGTATTCAAGGTGTGGTCATCCACAATGATTATG
GTAGTATGACACCTAGTCAATACTTACCATGGTTATATGCACGTGAGAATA

ACGGTACACACGTTAACGGTTGGGCTAGTGTTTATGCAAATAGAAACGAAG
TGCTTTGGTATCATCCGACAGACTACGTAGAGTGGCATTGTGGTAATCAAT
GGGCAAATGCTAACTTAATCGGATTTGAAGTGTGTGAGTCGTATCCTGGTA
GAATCTCGGACAAATTATTCTTAGAAAAATGAAGAAGCGACATTGAAAGTAG
5 CTGCGGATGTGATGAAGTCGTACGGATTACCAGTTAATCGCAACACTGTAC
GTCTGCATAACGAATTCTTCGGAACCTTCTTGTCCACATCGTTCGTGGGACTT
GCATGTTGGCAAAGGTGAGCCTTACACAACCTACTAATATTAATAAAATGAA
AGACTACTTCATCAAACGCATCAAACATTATTATGACGGTGGAAAGCTAGA
AGTAAGCAAAGCAGCAACTATCAAACAATCTGACGTTAAGCAAGAAGTTA
10 AAAAGCAAGAAGCAAAACAAATTGTGAAAGCAACAGATGCTGCAACACAT
GAACATTCAGCACAATGGTTGAATAATTACAAAAAAGGATATGGTTACGGT
CCTTATCCATTAGGTATAAATGGCGGTATGCACTACGGAGTTGATTTTTTTTA
TGAATATTGGAACACCAGTAAAAGCTATTTCAAGCGGAAAAATAGTTGAAG
CTGGTTGGAGTAATTACGGAGGAGGTAATCAAATAGGTCTTATTGAAAATG
15 ATGGAGTGCATAGACAATGGTATATGCATCTAAGTAAATATAATGTTAAAG
TAGGAGATTATGTCAAAGCTGGTCAAATAATCGGTTGGTCTGGAAGCACTG
GTTATTCTACAGCACCACATTTACACTTCCAAAGAATGGTTAATTCATTTTC
AAATTCAACTGCCCAAGATCCAATGCCTTTCTTAAAGAGCGCAGGATATGG
AAAAGCAGGTGGTACAGTAACTCCAACGCCGAATACAGGTGGAAACAGA
20 ATAAAGATGGCATTGTTGGTATAAAGCTGAACATGCTTCGTTACAGTGACAG
CACCAGAGGGAATTATCACAAGATACAAAGGTCCTTGGACTGGTCACCCAC
AAGCTGGTGTATTACAAAAAGGTCAAACGATTAAATATGATGAGGTTCAAA
AATTTGACGGTCATGTTTGGGTATCGTGGGAAACGTTTGAGGGCGAAACTG
TATACATGCCGGTACGCACATGGGACGCTAAACTGGTAAAGTTGGTAAGT
25 TGTGGGGCGAAATTAAATAA

SEQ ID NO: 48 (M23-LST_Ami2638_CBD2638 + 6xHis and cloning sites)

ATGAGAGGATCGCATCACCATCACCATCACGGATCCGCTGCAACACATGAA
CATTACAGCACAATGGTTGAATAATTACAAAAAAGGATATGGTTACGGTCCT
30 TATCCATTAGGTATAAATGGCGGTATGCACTACGGAGTTGATTTTTTTTATGA
ATATTGGAACACCAGTAAAAGCTATTTCAAGCGGAAAAATAGTTGAAGCTG
GTTGGAGTAATTACGGAGGAGGTAATCAAATAGGTCTTATTGAAAATGATG
GAGTGCATAGACAATGGTATATGCATCTAAGTAAATATAATGTTAAAGTAG

GAGATTATGTCAAAGCTGGTCAAATAATCGGTTGGTCTGGAAGCACTGGTT
ATTCTACAGCACCACATTTACACTTCCAAAGAATGGTTAATTCATTTTCAAA
TTCAACTGCCCAAGATCCAATGCCTTTCTTAAAGAGCGCAGGATATGGAAA
AGCAGGTGGTACAGTAACTCCAACGCCGAATACAGGTGAGCTCTTACGCCC
5 TAAAGACGCAAAGAAAGATGAAAAATCACAAGTATGTAGTGGTTTGGCTA
TGGAAAAATATGACATTACAAATTTAAATGCTAAACAAGATAAATCAAAG
AATGGGAGCGTGAAAGAGTTGAAACATATCTATTCAAACCATATTAAAGGT
AACAAGATTACAGCACCAAAACCTAGTATTCAAGGTGTGGTCATCCACAAT
GATTATGGTAGTATGACACCTAGTCAATACTTACCATGGTTATATGCACGTG
10 AGAATAACGGTACACACGTTAACGGTTGGGCTAGTGTTTATGCAAATAGAA
ACGAAGTGCTTTGGTATCATCCGACAGACTACGTAGAGTGGCATTGTGGTA
ATCAATGGGCAAATGCTAACTTAATCGGATTTGAAGTGTGTGAGTCGTATC
CTGGTAGAATCTCGGACAAATTATTCTTAGAAAATGAAGAAGCGACATTGA
AAGTAGCTGCGGATGTGATGAAGTCGTACGGATTACCAGTTAATCGCAACA
15 CTGTACGTCTGCATAACGAATTCTTCGGAACCTTCTTGTCCACATCGTTCGTG
GGACTTGCATGTTGGCAAAGGTGAGCCTTACACAACCTACTAATATTAATAA
AATGAAAGACTACTTCATCAAACGCATCAAACATTATTATGACGGTGGAAA
GCTAGAAGTAAGCAAAGCAGCAACTATCAAACAATCTGACGTTAAGCAAG
AAGTTAAAAAGCAAGAAGCAAAACAAATTGTGAAAGCAACAGATTGGAAA
20 CAGAATAAAGATGGCATTGTGGTATAAAGCTGAACATGCTTCGTTACAGTG
ACAGCACCAGAGGGAATTATCACAAGATACAAAGGTCCTTGGACTGGTCAC
CCACAAGCTGGTGTATTACAAAAAGGTCAAACGATTAAATATGATGAGGTT
CAAAAATTTGACGGTCATGTTTGGGTATCGTGGGAAACGTTTGAGGGCGAA
ACTGTATACATGCCGGTACGCACATGGGACGCTAAAACTGGTAAAGTTGGT
25 AAGTTGTGGGGCGAAATTAAATAA

SEQ ID NO: 49 (M23-2638_Ami2638_CBD2638_CBD2638 + 6xHis and cloning sites)

ATGAGAGGATCGCATCACCATCACCATCACGGATCCATGCTAACTGCTATT
30 GACTATCTTACGAAAAAAGGTTGGAAAAATATCATCTGACCCTCGCACTTAC
GATGGTTACCCTAAAAACTACGGCTACAGAAATTACCATGAAAACGGCATT
AATTATGATGAGTTTTGTGGTGGTTATCATAGAGCTTTTGATGTTTACAGTA
ACGAAACTAACGACGTGCCTGCTGTTACTAGCGGAACAGTTATTGAAGCAA

ACGATTACGGTAATTTTGGTGGTACATTTCGTTATTAGAGACGCTAACGATA
ACGATTGGATATATGGGCATCTACAACGTGGCTCAATGCGATTTGTTGTAG
GCGACAAAGTCAATCAAGGTGACATTATTGGTTTACAAGGTAATAGCAACT
ATTACGACAATCCTATGAGTGTACATTTACATTTACAATTACGCCCTAAAGA
5 CGCAAAGAAAGATGAAAAATCACAAGTATGTAGTGGTTTGGCTATGGAAA
AATATGACATTACAAATTTAAATGCTAAACAAGATAAATCAAAGAATGGG
AGCGTGAAAGAGTTGAAACATATCTATTCAAACCATATTAAAGGTAACAAG
ATTACAGCACCAAAACCTAGTATTCAAGGTGTGGTCATCCACAATGATTAT
GGTAGTATGACACCTAGTCAATACTTACCATGGTTATATGCACGTGAGAAT
10 AACGGTACACACGTTAACGGTTGGGCTAGTGTTTATGCAAATAGAAACGAA
GTGCTTTGGTATCATCCGACAGACTACGTAGAGTGGCATTGTGGTAATCAA
TGGGCAAATGCTAACTTAATCGGATTTGAAGTGTGTGAGTCGTATCCTGGT
AGAATCTCGGACAAATTATTCTTAGAAAATGAAGAAGCGACATTGAAAGTA
GCTGCGGATGTGATGAAGTCGTACGGATTACCAGTTAATCGCAACACTGTA
15 CGTCTGCATAACGAATTCTTCGGAACCTTCTTGTCACATCGTTCGTGGGACT
TGCATGTTGGCAAAGGTGAGCCTTACACAACCTACTAATATTAATAAAATGA
AAGACTACTTCATCAAACGCATCAAACATTATTATGACGGTGGAAAGCTAG
AAGTAAGCAAAGCAGCAACTATCAAACAATCTGACGTTAAGCAAGAAGTT
AAAAAGCAAGAAGCAAAACAAATTGTGAAAGCAACAGATTGGAAACAGA
20 ATAAAGATGGCATTGTTGGTATAAAGCTGAACATGCTTCGTTACAGTGACAG
CACCAGAGGGAATTATCACAAGATACAAAGGTCCTTGGACTGGTCACCCAC
AAGCTGGTGTATTACAAAAAGGTCAAACGATTAAATATGATGAGGTTCAAA
AATTTGACGGTCATGTTTGGGTATCGTGGGAAACGTTTGAGGGCGAAACTG
TATACATGCCGGTACGCACATGGGACGCTAAAACTGGTAAAGTTGGTAAGT
25 TGTGGGGCGAAATTAAAGAGCTCGGTGGAAAGCTAGAAGTAAGCAAAGCA
GCAACTATCAAACAATCTGACGTTAAGCAAGAAGTTAAAAAGCAAGAAGC
AAAACAAATTGTGAAAGCAACAGATTGGAAACAGAATAAAGATGGCATT
GGTATAAAGCTGAACATGCTTCGTTACAGTGACAGCACCAGAGGGAATTA
TCACAAGATACAAAGGTCCTTGGACTGGTCACCCACAAGCTGGTGTATTAC
30 AAAAAGGTCAAACGATTAAATATGATGAGGTTCAAAAATTTGACGGTCATG
TTTGGGTATCGTGGGAAACGTTTGAGGGCGAAACTGTATACATGCCGGTAC
GCACATGGGACGCTAAAACTGGTAAAGTTGGTAAGTTGTGGGGCGAAATTA
AATAA

SEQ ID NO: 50 (pQE-30 vector, available under cat. No.: 32915 at Qiagen - Hilden, Germany)

CTCGAGAAATCATAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATA
ATAGATTCAATTGTGAGCGGATAACAATTTACACAGAATTCATTAAAGAG
5 GAGAAATTAACCTATGAGAGGATCGCATCACCATCACCATCACGGATCCGCA
TGCGAGCTCGGTACCCCGGGTCGACCTGCAGCCAAGCTTAATTAGCTGAGC
TTGGACTCCTGTTGATAGATCCAGTAATGACCTCAGAACTCCATCTGGATTT
G TTCAGAACGCTCGGTGCGCGCGGGCGTTTTTTATTGGTGAGAATCCAAG
CTAGCTTGGCGAGATTTTCAGGAGCTAAGGAAGCTAAAATGGAGAAAAAA
10 ATCACTGGATATAACCACCGTTGATATATCCCAATGGCATCGTAAAGAACAT
TTTGAGGCATTTTCAGTCAGTTGCTCAATGTACCTATAACCAGACCGTTCAGC
TGGATATTACGGCCTTTTTAAAGACCGTAAAGAAAAATAAGCACAAGTTTT
ATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAATT
TCGTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTACCCC
15 TTGTTACACCGTTTTCCATGAGCAAACCTGAAACGTTTTTCATCGCTCTGGAGT
GAATACCACGACGATTTCCGGCAGTTTCTACACATATATTCGCAAGATGTG
GCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAAT
ATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCACCAGTTTTGATTTAA
ACGTGGCCAATATGGACAACTTCTTCGCCCCCGTTTTTCACCATGGGCAAAT
20 ATTATACGCAAGGCGACAAGGTGCTGATGCCGCTGGCGATTTCAGGTTTCATC
ATGCCGTTTGTGATGGCTTCCATGTCCGGCAGAATGCTTAATGAATTACAAC
AGTACTGCGATGAGTGGCAGGGCGGGGCGTAATTTTTTTAAGGCAGTTATT
GGTGCCCTTAAACGCCTGGGGTAATGACTCTCTAGCTTGAGGCATCAAATA
AAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTTATCTGTTGTTTGT
25 CGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCCTCTAGAGCTGCCTCGC
GCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGAC
GGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGG
GCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGCCATGACCCAGTCAC
GTAGCGATAGCGGAGTGATACTGGCTTAACCTATGCGGCATCAGAGCAGAT
30 TGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAG
GAGAAAATACCGCATCAGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTG
CGCTCGGTGCTTCGGCTGCGGGCAGCGGTATCAGCTCACTCAAAGGCGGTA
ATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGC

AAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGT
TTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATCGACGCTCAA
GTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCC
CTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATA
5 CCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGC
TGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGC
ACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCT
TGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGG
TAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAA
10 GTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGC
TCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGG
CAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGAT
TACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGG
GTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGAG
15 ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTT
AAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTACCAATGC
TTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAG
TTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCAT
CTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAG
20 ATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGT
CCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTA
GAGTAAGTAGTTCGCCAGTTAATAGTTTTCGCAACGTTGTTGCCATTGCTAC
AGGCATCGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATTTCAGCTCCGGT
TCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCG
25 GTTAGCTCCTTCGGTTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCGAGTGT
TATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATC
CGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGA
ATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA
TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCT
30 TTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGAT
GTAACCCACTCGTGCAACCCAACTGATCTTCAGCATCTTTTACTTTTACCAGC
GTTTCTGGGTGAGCAAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAAT
AAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTAT

TGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTA
TTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGC
CACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAAATA
GGCGTATCACGAGGCCCTTTCGTCTTCAC

5

SEQ ID NO: 51 (CBD-2638 with putative linker indicated in bold)

GKLEVSKAATIKQSDVKQEVKKQEAKQIVKATDWKQNKDGIWYKAEHASF
TVTAPEGIITRYKGPWTGHPQAGVLQKGQTIKYDEVQKFDGHVWVSWETFEG
ETVYMPVRTWDAKTGKVGKLWGEIK

10

SEQ ID NO: 52 (CWT-LST with putative linker indicated in bold)

GKAGGTVTPTPNTGWKTNKYGTLYKSESASFTPNTDIITRTTGPFERSMPQSGV
LKAGQTIHYDEVKQDGHVWVGTYTGNQRIYLPVRTWNKSTNTLGVLWGT
IK

15

SEQ ID NO: 53 (M23-2638 with putative linker indicated in bold)

MLTAIDYLTKKGWKISSDPRTYDGYPKNYGYRNYHENGINYDEFCCGGYHRAF
DVYSNETNDVPAVTSGTVIEANDYGNFGGTFVIRDANDNDWIYGHQLQRGSMR
FVVGDKVNQGDIIQLQGNSNYNDPMSVHLHLQLRPKDAKKDEKSQVCSGL

20

AMEKYDITNLNAKQDKSKNGSVKELKHIYSNHIKG

SEQ ID NO: 54 (M23-LST with putative linker indicated in bold)

AATHEHSAQWLNNYKKGYGYGPYPLGINGGMHYGVDFFMNIGTPVKAISSGK
IVEAGWSNYGGGNQIGLIENDGVHRQWYMHLSKYNVKVGDYVKAGQIIGWS

25

GSTGYSTAPHLHFQRMVNSFSNSTAQDMPFLKSAGYGKAGGTVTPTPNTG

SEQ ID NO: 55 (Ami-2638 with putative linker indicated in bold)

EKSQVCSGLAMEKYDITNLNAKQDKSKNGSVKELKHIYSNHIKGNKITAPK
PSIQGVVIHNDYGSMTPSQYLPWLYARENNGTHVNGWASVYANRNEVLWYH
PTDYVEWHCGNQWANANLIGFEVCESYPGRISDKLFLENEEATLKVAADVMK
SYGLPVNRNTVRLHNEFFGTSCPHRSWDLHVGKGEPYTTTNINKMKDYFIKRI
KHYDGGKLEVSKAATIKQSDVKQEVKKQEAKQIVKATD

30

SEQ ID NO: 56 (Ami-□11 with putative linker indicated in bold)

PKKETAKPQPKAVELKIIKDVVKGYDLPKRGSNPKGIVIHNDAGSKGATAE
AYRNGLVNAPLSRLEAGIAHSYVSGNTVWQALDESQVGWHTANQIGNKYYY
GIEVCQSMGADNATFLKNEQATFQECARLLKKWGLPANRNTIRLHNEFTSTSC
5 **PHRSSVLHTGFDPVTRGLLPEDKRLQLKDYFIKQIRAYMDGKIPVATVSNESSA**
SSNTVKPVASA

SEQ ID NO: 57 (CHAP-□11 with putative linker indicated in bold)

MSIIMEVATMQAKLTKNEFIEWLKTSEGKQFNVDLWYGFQCFDYANAGWKV
10 **LFGLLLKGLGAKDIPFANNFDGLATVYQNTPDFLAQPGDMVVFGSNYGAGYG**
HVAWVIEATLDYIIVYEQNWLGSGWTDGIEQPGWGWEKVTRRQHAYDFPMW
FIRPNFKSETAPRSVQSPTQAPKKETAKPQPKAVELKIIKDVVKGYDLPKRG

SEQ ID NO: 58 (CHAP-□Twort with putative linker indicated in bold)

15 **MKTLKQAESYIKSKVNTGTDFDGLYGYQCMDLAVDYTYHVTGDKIRMWGNA**
KDAINNSFGGTATVYKNYPAFRPKYGDVVVWTTGNFATYGHIAIVTNPDYTG
DLQYVTVLEQNWNGNGIYKTELATIRTHDYTGITHFIRPNFATESSVKKKDTK
KKPKPSNRDGINKDKIVYDRTNINYNMVKRG

20 SEQ ID NO: 59 (pHGFP_CBD2638_c vector)

CTCGAGAAATCATAAAAAATTTATTTGCTTTGTGAGCGGATAACAATTATA
ATAGATTCAATTGTGAGCGGATAACAATTTACACAGAATTCATTAAAGAG
GAGAAATTAACATGAGAGGATCGCATCACCATCACCATCACGGATCCATG
AGTAAAGGAGAAGAAGCTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTA
25 **GATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGT**
GATGCAACATACGGAAAACCTTACCCTTAAATTTATTTGCACTACTGGAAAA
CTACCTGTTCCATGGCCAACACTTGTCACTACTTTTCGCGTATGGTCTTCAAT
GCTTTGCGAGATACCCAGATCATATGAAACGGCATGACTTTTTCAAGAGTG
CCATGCCCGAAGGTTATGTACAGGAAAGAACTATATTTTCAAAGATGACG
30 **GGAAC TACAAGACACGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTA**
ATAGAATCGAGTTAAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTTG
GACACAAATTGGAATACAACTATAACTCACACAATGTATACATCATGGCAG
ACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAATTAGACACAACATT

GAAGATGGAAGCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATT
GGCGATGGCCCTGTCCTTTTACCAGACAACCATTACCTGTCCACACAATCTG
CCCTTTCGAAAGATCCCAACGAAAAGAGAGACCACATGGTCCTTCTTGAGT
TTGTAACAGCTGCTGGGATTACACATGGCATGGATGAACTATACAAAGAGC
5 TCGGTGGAAAGCTAGAAGTAAGCAAAGCAGCAACTATCAAACAATCTGAC
GTTAAGCAAGAAGTTAAAAAGCAAGAAGCAAAACAAATTGTGAAAGCAAC
AGATTGGAAACAGAATAAAGATGGCATTGTTGGTATAAAGCTGAACATGCTTC
GTTACACAGTGACAGCACCAGAGGGAATTATCACAAGATACAAAGGTCCTTG
GACTGGTCACCCACAAGCTGGTGTATTACAAAAAGGTCAAACGATTAAATA
10 TGATGAGGTTCAAAAATTTGACGGTCATGTTTGGGTATCGTGGGAAACGTT
TGAGGGCGAAACTGTATACATGCCGGTACGCACATGGGACGCTAAAACTG
GTAAAGTTGGTAAGTTGTGGGGCGAAATTAAATAAGTCGACCTGCAGCCAA
GCTTAATTAGCTGAGCTTGGACTCCTGTTGATAGATCCAGTAATGACCTCAG
AACTCCATCTGGATTTGTTCAGAACGCTCGGTTGCCGCCGGGCGTTTTTTAT
15 TGGTGAGAATCCAAGCTAGCTTGGCGAGATTTTCAGGAGCTAAGGAAGCTA
AAATGGAGAAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGC
ATCGTAAAGAACATTTTGAGGCATTTTCAGTCAGTTGCTCAATGTACCTATAA
CCAGACCGTTCAGCTGGATATTACGGCCTTTTTTAAAGACCGTAAAGAAAAA
TAAGCACAAGTTTTATCCGGCCTTTATTCACATTCTTGCCCCGCTGATGAAT
20 GCTCATCCGGAATTTTCGTATGGCAATGAAAGACGGTGAGCTGGTGATATGG
GATAGTGTTACACCCTTGTTACACCGTTTTCCATGAGCAAACCTGAAACGTTTT
CATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATAT
ATTCGCAAGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAG
GGTTTATTGAGAATATGTTTTTCGTCTCAGCCAATCCCTGGGTGAGTTTCAC
25 CAGTTTTGATTTAAACGTGGCCAATATGGACAACCTTCTTCGCCCCCGTTTTTC
ACCATGGGCAAATATTATACGCAAGGCCGACAAGGTGCTGATGCCGCTGGCG
ATTCAGGTTTCATCATGCCGTTTGTGATGGCTTCCATGTCGGCAGAATGCTTA
ATGAATTACAACAGTACTGCGATGAGTGGCAGGGCGGGGCGTAATTTTTTT
AAGGCAGTTATTGGTGCCCTTAAACGCCTGGGGTAATGACTCTCTAGCTTG
30 AGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTTCGTTTT
ATCTGTTGTTTGTGCGGTGAACGCTCTCCTGAGTAGGACAAATCCGCCCTCTA
GAGCTGCCTCGCGCGTTTTCGGTGATGACGGTGAAAACCTCTGACACATGCA
GCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACA

AGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGGTGTGCGGGCGCAGCCAT
GACCCAGTCACGTAGCGATAGCGGAGTGTATACTGGCTTA ACTATGCGGCA
TCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCAC
AGATGCGTAAGGAGAAAATACCGCATCAGGCGCTCTTCCGCTTCCTCGCTC
5 ACTGACTCGCTGCGCTCGGTCGTTTCGGCTGCGGCGAGCGGTATCAGCTCAC
TCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAA
GAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCG
CGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAA
ATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATAC
10 CAGGCGTTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGC
CGCTTACCGGATACTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTC
TCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTTCGCTCCAAG
CTGGGCTGTGTGCACGAACCCCCCGTTTCAGCCCGACCGCTGCGCCTTATCC
GGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTG
15 GCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGC
TACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGT
ATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGG
TAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTT
TGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTT
20 GATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACCTCACGTTAAGG
GATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAAT
TAAAAATGAAGTTTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTCT
GACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTAT
TTCGTTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACG
25 GGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACG
CTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGA
GCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGT
TGCCGGGAAGCTAGAGTAAGTAGTTTCGCCAGTTAATAGTTTGCGCAACGTT
GTTGCCATTGCTACAGGCATCGTGGTGTACGCTCGTCGTTTGGTATGGCTT
30 CATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTT
GTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAAGTAA
GTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTT
ACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCA

AGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCCGGCGT
CAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCA
TTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGA
GATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTT
5 TACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCCG
AAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCT
TTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATAC
ATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTT
CCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACATTA
10 ACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCAC

SEQ ID NO: 60 (GFP_CBD2638)

ATGAGAGGATCGCATCACCATCACCATCACGGATCCATGAGTAAAGGAGA
AGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTT
15 AATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGCAACATAC
GGAAAACCTTACCCTTAAATTTATTTGCACTACTGGAAAACCTGTTCCAT
GGCCAACACTTGTCACTACTTTTCGCGTATGGTCTTCAATGCTTTGCGAGATA
CCCAGATCATATGAAACGGCATGACTTTTTTCAAGAGTGCCATGCCCGAAGG
TTATGTACAGGAAAGAACTATATTTTTTCAAAGATGACGGGAACTACAAGAC
20 ACGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTT
AAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTTGGACACAAATTGGA
ATACAACTATAACTCACACAATGTATACATCATGGCAGACAAACAAAAGA
ATGGAATCAAAGTTAACTTCAAAATTAGACACAACATTGAAGATGGAAGC
GTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCT
25 GTCCTTTTACCAGACAACCATTACCTGTCCACACAATCTGCCCTTTCGAAAG
ATCCCAACGAAAAGAGAGACCACATGGTCCCTTCTTGAGTTTGTAACAGCTG
CTGGGATTACACATGGCATGGATGAACTATACAAAGAGCTCGGTGGAAAG
CTAGAAGTAAGCAAAGCAGCAACTATCAAACAATCTGACGTTAAGCAAGA
AGTTAAAAAGCAAGAAGCAAAACAAATTGTGAAAGCAACAGATTGGAAAC
30 AGAATAAAGATGGCATTGTTGTTATAAAGCTGAACATGCTTCGTTACAGTGA
CAGCACCAGAGGGAATTATCACAAGATACAAAGGTCCTTGGACTGGTCACC
CACAAGCTGGTGTATTACAAAAAGGTCAAACGATTAAATATGATGAGGTTTC
AAAAATTTGACGGTCATGTTTGGGTATCGTGGGAAACGTTTGAGGGCGAAA

CTGTATACATGCCGGTACGCACATGGGACGCTAAAACTGGTAAAGTTGGTA
AGTTGTGGGGCGAAATTAAATAA

SEQ ID NO: 61 (GFP_CBD2638_CBD2638 Variant 1. Restriction sites used for

5 construction: BamHI - SacI - SacI- Sall. TAA Stop codon derived from 2nd CBD2638)
ATGAGAGGATCGCATCACCATCACCATCACGGATCCATGAGTAAAGGAGA
AGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTT
AATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGCAACATAC
GGAAAACCTTACCCTTAAATTTATTTGCACTACTGGAAAACCTACCTGTTCCAT
10 GGCCAACACTTGTCACTACTTTTCGCGTATGGTCTTCAATGCTTTGCGAGATA
CCCAGATCATATGAAACGGCATGACTTTTTCAAGAGTGCCATGCCCCGAAGG
TTATGTACAGGAAAGAACTATATTTTTCAAAGATGACGGGAACTACAAGAC
ACGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTT
AAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTTGGACACAAATTGGA
15 ATACAACTATAACTCACACAATGTATACATCATGGCAGACAAACAAAAGA
ATGGAATCAAAGTTAACTTCAAAATTAGACACAACATTGAAGATGGAAGC
GTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCT
GTCCTTTTACCAGACAACCATTACCTGTCCACACAATCTGCCCTTTTCGAAAG
ATCCCAACGAAAAGAGAGACCACATGGTCCTTCTTGAGTTTGTAAACAGCTG
20 CTGGGATTACACATGGCATGGATGAACTATACAAAGAGCTCGGTGGAAAG
CTAGAAGTAAGCAAAGCAGCAACTATCAAACAATCTGACGTTAAGCAAGA
AGTTAAAAAGCAAGAAGCAAAACAAATTGTGAAAGCAACAGATTGGAAAC
AGAATAAAGATGGCATTTGGTATAAAGCTGAACATGCTTCGTTACAGTGA
CAGCACCAGAGGGAATTATCACAAGATACAAAGGTCCTTGGACTGGTCACC
25 CACAAGCTGGTGTATTACAAAAAGGTCAAACGATTAAATATGATGAGGTTC
AAAAATTTGACGGTCATGTTTGGGTATCGTGGGAAACGTTTGAGGGGCGAAA
CTGTATACATGCCGGTACGCACATGGGACGCTAAAACTGGTAAAGTTGGTA
AGTTGTGGGGCGAAATTAAAGAGCTCGGTGGAAAGCTAGAAGTAAGCAAA
GCAGCAACTATCAAACAATCTGACGTTAAGCAAGAAGTTAAAAAGCAAGA
30 AGCAAAACAAATTGTGAAAGCAACAGATTGGAAACAGAATAAAGATGGCA
TTTGGTATAAAGCTGAACATGCTTCGTTACAGTGACAGCACCAGAGGGAA
TTATCACAAGATACAAAGGTCCTTGGACTGGTCACCCACAAGCTGGTGTAT
TACAAAAAGGTCAAACGATTAAATATGATGAGGTTCAAAAATTTGACGGTC

ATGTTTGGGTATCGTGGGAAACGTTTGAGGGCGAAACTGTATACATGCCGG
TACGCACATGGGACGCTAAAACTGGTAAAGTTGGTAAGTTGTGGGGCGAA
ATTAAATAAGTCGAC

- 5 SEQ ID NO: 62 (GFP_CBD2638_CBD2638 Variant 2. Restriction sites used for
construction: BamHI - SacI - KpnI - Sall. Stop codon is encoded by vector)
ATGAGAGGATCGCATCACCATCACCATCACGGATCCATGAGTAAAGGAGA
AGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTT
AATGGGCACAAATTTTCTGTCACTGGAGAGGGTGAAGGTGATGCAACATAC
10 GGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAACTACCTGTTCCAT
GGCCAACACTTGTCACTACTTTTCGCGTATGGTCTTCAATGCTTTGCGAGATA
CCCAGATCATATGAAACAGCATGACTTTTTCAAGAGTGCCATGCCCCGAAGG
TTATGTACAGGAAAGAACTATATTTTTCAAAGATGACGGGAACACAAAGAC
ACGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTT
15 AAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTTGGACACAAATTGGA
ATACAACTATAACTCACACAATGTATACATCATGGCAGACAAACAAAAGA
ATGGAATCAAAGTTAACTTCAAAATTAGACACAACATTGAAGATGGAAGC
GTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCT
GTCCTTTTACCAGACAACCATTACCTGTCCACACAATCTGCCCTTTCGAAAG
20 ATCCCAACGAAAAGAGAGACCACATGGTCCTTCTTGAGTTTGTAACAGCTG
CTGGGATTACACATGGCATGGATGAACTATACAAAGAGCTCGGTGGAAAG
CTAGAAGTAAGCAAAGCAGCAACTATCAAACAATCTGACGTTAAGCAAGA
AGTTAAAAAGCAAGAAGCAAAACAAATTGTGAAAGCAACAGATTGGAAAC
AGAATAAAGATGGCATTGTTGTATAAAGCTGAACATGCTTCGTTACAGTGA
25 CAGCACCAGAGGGAATTATCACAAGATACAAAGGTCCTTGGACTGGTCACC
CACAAGCTGGTGTATTACAAAAAGGTCAAACGATTAAATATGATGAGGTTC
AAAAATTTGACGGTCATGTTTGGGTATCGTGGGAAACGTTTGAGGGCGAAA
CTGTATACATGCCGGTACGCACATGGGACGCTAAAACTGGTAAAGTTGGTA
AGTTGTGGGGCGAAATTAAAGGTACCGGTGGAAAGCTAGAAGTAAGCAAA
30 GCAGCAACTATCAAACAATCTGACGTTAAGCAAGAAGTTAAAAAGCAAGA
AGCAAAACAAATTGTGAAAGCAACAGATTGGAAACAGAATAAAGATGGCA
TTTGGTATAAAGCTGAACATGCTTCGTTACAGTGACAGCACCAGAGGGAA
TTATCACAAGATACAAAGGTCCTTGGACTGGTCACCCACAAGCTGGTGTAT

TACAAAAAGGTCAAACGATTAAATATGATGAGGTTCAAAAATTTGACGGTC
ATGTTTGGGTATCGTGGGAAACGTTTGAGGGCGAAACTGTATACATGCCGG
TACGCACATGGGACGCTAAAACTGGTAAAGTTGGTAAGTTGTGGGGCGAA
ATTAAAGTCGACCTGCAGCCAAGCTTAATTAGCTGA

5

SEQ ID NO: 63 (GFP CBD2638 CBD2638 CBD2638. Restriction sites used for
construction: BamHI - SacI - KpnI- Sall - PstI. Stop codon is encoded by vector)

ATGAGAGGATCGCATCACCATCACCATCACGGATCCATGAGTAAAGGAGA
AGAACTTTTCACTGGAGTTGTCCCAATTCTTGTTGAATTAGATGGTGATGTT
10 AATGGGCACAAATTTTCTGTCAGTGGAGAGGGTGAAGGTGATGCAACATAC
GGAAAACCTTACCCTTAAATTTATTTGCACTACTGGAAAACCTACCTGTTCCAT
GGCCAACACTTGTCACTACTTTTCGCGTATGGTCTTCAATGCTTTGCGAGATA
CCCAGATCATATGAAACAGCATGACTTTTTCAAGAGTGCCATGCCCCGAAGG
TTATGTACAGGAAAGAACTATATTTTCAAAGATGACGGGAACTACAAGAC
15 ACGTGCTGAAGTCAAGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTT
AAAAGGTATTGATTTTAAAGAAGATGGAAACATTCTTGGACACAAATTGGA
ATACAACATAACTCACACAATGTATACATCATGGCAGACAAACAAAAGA
ATGGAATCAAAGTTAACTTCAAATTAGACACAACATTGAAGATGGAAGC
GTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCT
20 GTCCTTTTACCAGACAACCATTACCTGTCCACACAATCTGCCCTTTCGAAAG
ATCCCAACGAAAAGAGAGACCACATGGTCCTTCTTGAGTTTGTAACAGCTG
CTGGGATTACACATGGCATGGATGAACTATACAAAGAGCTCGGTGGAAAG
CTAGAAGTAAGCAAAGCAGCAACTATCAAACAATCTGACGTTAAGCAAGA
AGTTAAAAAGCAAGAAGCAAAACAAATTGTGAAAGCAACAGATTGGAAAC
25 AGAATAAAGATGGCATTGTGTATAAAGCTGAACATGCTTCGTTACAGTGA
CAGCACCAGAGGGAATTATCACAAGATACAAAGGTCCTTGGACTGGTCACC
CACAAGCTGGTGTATTACAAAAAGGTCAAACGATTAAATATGATGAGGTTT
AAAAATTTGACGGTCATGTTTGGGTATCGTGGGAAACGTTTGAGGGCGAAA
CTGTATACATGCCGGTACGCACATGGGACGCTAAAACTGGTAAAGTTGGTA
30 AGTTGTGGGGCGAAATTAAAGGTACCGGTGGAAAGCTAGAAGTAAGCAAA
GCAGCAACTATCAAACAATCTGACGTTAAGCAAGAAGTTAAAAAGCAAGA
AGCAAAACAAATTGTGAAAGCAACAGATTGGAAACAGAATAAAGATGGCA
TTTGGTATAAAGCTGAACATGCTTCGTTACAGTGACAGCACCAGAGGGAA

TTATCACAAGATACAAAGGTCCTTGGACTGGTCACCCACAAGCTGGTGTAT
TACAAAAAGGTCAAACGATTAAATATGATGAGGTTCAAAAATTTGACGGTC
ATGTTTGGGTATCGTGGGAAACGTTTGAGGGCGAAACTGTATACATGCCGG
TACGCACATGGGACGCTAAAACTGGTAAAGTTGGTAAGTTGTGGGGCGAA
5 ATTAAAGTCGACGGTGGAAAGCTAGAAGTAAGCAAAGCAGCAACTATCAA
ACAATCTGACGTTAAGCAAGAAGTTAAAAAGCAAGAAGCAAAACAAATTG
TGAAAGCAACAGATTGGAAACAGAATAAAGATGGCATTGTTGTATAAAGCT
GAACATGCTTCGTTACAGTGACAGCACCAGAGGGAATTATCACAAGATAC
AAAGGTCCTTGGACTGGTCACCCACAAGCTGGTGTATTACAAAAAGGTCAA
10 ACGATTAAATATGATGAGGTTCAAAAATTTGACGGTCATGTTTGGGTATCG
TGGGAAACGTTTGAGGGCGAAACTGTATACATGCCGGTACGCACATGGGAC
GCTAAAACTGGTAAAGTTGGTAAGTTGTGGGGCGAAATTAAACTGCAGCCA
AGCTTAATTAGCTGA

15 SEQ ID NO: 64 (GFP_CBD2638 aa)

MRGSHHHHHHGSMSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEDATY
GKLTCLKFICTTGKLPVPWPTLVTTTFAYGLQCFARYPDHMKRHDFFKSAMPEGY
VQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNIN
SHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNH
20 YLSTQSALS KDPNEKRDHMLLEFVTAAGITHGMDELYKELGGKLEVSKAATI
KQSDVKQEVKKQEAKQIVKATDWKQNKDGIWYKAEHASFTVTAPEGIITRYK
GPWTGHPQAGVLQKGQTIKYDEVQKFDGHVWVSWETFEGETVYMPVRTWD
AKTGKVGKLWGEIK

25 SEQ ID NO: 65 (GFP_CBD2638_CBD2638 Variant 1 aa)

MRGSHHHHHHGSMSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEDATY
GKLTCLKFICTTGKLPVPWPTLVTTTFAYGLQCFARYPDHMKRHDFFKSAMPEGY
VQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNIN
SHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNH
30 YLSTQSALS KDPNEKRDHMLLEFVTAAGITHGMDELYKELGGKLEVSKAATI
KQSDVKQEVKKQEAKQIVKATDWKQNKDGIWYKAEHASFTVTAPEGIITRYK
GPWTGHPQAGVLQKGQTIKYDEVQKFDGHVWVSWETFEGETVYMPVRTWD
AKTGKVGKLWGEIKELGGKLEVSKAATIKQSDVKQEVKKQEAKQIVKATDW

KQNKDGIWYKAEHASFTVTAPEGIITRYKGPWTGHPQAGVLQKGQTIKYDEV
QKFDGHVWVSWETFEGETVYMPVRTWDAKTGKVGKLWGEIK

SEQ ID NO: 66 (GFP_CBD2638_CBD2638 Variant 2 aa)

5 **MRGSHHHHHH**GSMSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATY
GKLTCLKFICTTGKLPVPWPTLVTTTFAYGLQCFARYPDHMKQHDFFKSAMPEGY
VQERTIFFKDDGNYKTRAEVKFEEDTLVNRIELKGIDFKEDGNILGHKLEYNIN
SHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNH
YLSTQSALS KDPNEKRDHMLLEFVTAAGITHGMDELYKELGGKLEVSKAATI
10 KQSDVKQEVKKQEAQKIVKATDWKQNKDGIWYKAEHASFTVTAPEGIITRYK
GPWTGHPQAGVLQKGQTIKYDEVQKFDGHVWVSWETFEGETVYMPVRTWD
AKTGKVGKLWGEIKGTGGKLEVSKAATIKQSDVKQEVKKQEAQKIVKATDW
KQNKDGIWYKAEHASFTVTAPEGIITRYKGPWTGHPQAGVLQKGQTIKYDEV
QKFDGHVWVSWETFEGETVYMPVRTWDAKTGKVGKLWGEIKVDLQPSLIS

15

SEQ ID NO: 67 (GFP_CBD2638_CBD2638_CBD2638 aa)

MRGSHHHHHHGSMSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATY
GKLTCLKFICTTGKLPVPWPTLVTTTFAYGLQCFARYPDHMKQHDFFKSAMPEGY
VQERTIFFKDDGNYKTRAEVKFEEDTLVNRIELKGIDFKEDGNILGHKLEYNIN
20 SHNVYIMADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNH
YLSTQSALS KDPNEKRDHMLLEFVTAAGITHGMDELYKELGGKLEVSKAATI
KQSDVKQEVKKQEAQKIVKATDWKQNKDGIWYKAEHASFTVTAPEGIITRYK
GPWTGHPQAGVLQKGQTIKYDEVQKFDGHVWVSWETFEGETVYMPVRTWD
AKTGKVGKLWGEIKGTGGKLEVSKAATIKQSDVKQEVKKQEAQKIVKATDW
25 KQNKDGIWYKAEHASFTVTAPEGIITRYKGPWTGHPQAGVLQKGQTIKYDEV
QKFDGHVWVSWETFEGETVYMPVRTWDAKTGKVGKLWGEIKVDGGKLEVS
KAATIKQSDVKQEVKKQEAQKIVKATDWKQNKDGIWYKAEHASFTVTAPEGI
ITRYKGPWTGHPQAGVLQKGQTIKYDEVQKFDGHVWVSWETFEGETVYMPV
RTWDAKTGKVGKLWGEIKLQPSLIS

30

Claims

1. A nucleic acid molecule comprising a first nucleotide sequence, wherein said nucleotide sequence has at least 85% sequence identity with SEQ ID NO: 12,

5 wherein said first nucleotide sequence encodes a cell wall-binding domain binding the peptidoglycan cell wall of *Staphylococcus* genera,

 wherein said nucleic acid molecule has at least 85% sequence identity with SEQ ID NO: 1 without consisting of SEQ ID NO: 1, or wherein said nucleic acid molecule further comprises a heterologous nucleotide sequence encoding a lytic domain.

10

2. A nucleic acid molecule according to claim 1, wherein said first nucleotide sequence originates from *S. aureus* bacteriophage Φ 2638a endolysin.

3. A nucleic acid molecule according to any of the claims 1 or 2, wherein said nucleic acid molecule further comprises a heterologous nucleotide sequence encoding a lytic domain, wherein said lytic domain is a second and third nucleotide sequence, and wherein said second nucleotide sequence encodes an M23 endopeptidase domain and said third nucleotide sequences encodes an amidase domain.

15

20 4. A nucleic acid molecule according to claim 3, wherein said second and third nucleotide sequences originate from a gene encoding for an enzyme selected from the group consisting of *S. aureus* bacteriophage Φ 2638a endolysin, *S. aureus* bacteriophage Φ 11 endolysin, *S. aureus* bacteriophage Φ Twort endolysin and *S. Simulans* lysostaphin.

25 5. A nucleic acid molecule according to claim 4, wherein said second nucleotide sequence has at least 85% sequence identity with SEQ ID:14 or 15 and said third nucleotide sequences has at least 85% sequence identity with SEQ ID NO: 16 or 17.

30

6. A nucleic acid molecule according to claim 5, wherein said nucleotide molecule has at least 85% sequence identity with SEQ ID NO: 9.

7. A nucleic acid molecule according to any one of the claims 1 - 6, further comprising a fourth nucleotide sequence encoding a CHAP (cysteine, histidine-dependent amidohydrolases/peptidases) domain.

8. A nucleic acid molecule according to claim 7, wherein said fourth nucleotide sequences originates from *S. aureus* bacteriophage Φ 11 or *S. aureus* bacteriophage Φ Twort endolysin.

9. A nucleic acid molecule according to claim 8, wherein said fourth nucleotide sequences has at least 85% sequence identity with SEQ ID NO: 18 or SEQ ID NO: 19.

10. A nucleic acid molecule according to any one of the claims 1 – 9 encoding a polypeptide which has the same or an increased lytic activity and/or the same or a decreased pH optimum as compared to *S. aureus* bacteriophage Φ 2638a endolysin encoded by SEQ ID NO: 1.

11. A polypeptide encoded by a nucleic acid molecule as defined in any one of claims 1 - 10.

12. A nucleic acid construct comprising a nucleic acid molecule as defined in any of the claims 1 - 10.

13. An expression vector comprising a nucleic acid construct as defined in claim 12 operably linked to one or more control sequences, which direct the production or expression of the encoded polypeptide in a cell, a subject, or a cell-free expression system.

14. A cell comprising the nucleic acid construct as defined in claim 12 or the expression vector as defined in claim 13, said cell being a microbial, prokaryotic or eukaryotic cell.

15. A method for producing, optionally purifying and optionally freeze-drying a polypeptide as defined in claim 11, said method comprising the steps of:

i) producing said polypeptide in a cell as defined in claim 14 comprising a nucleic acid construct as defined in claim 12, optionally

5 ii) purifying said polypeptide, and optionally

iii) freeze-drying said purified polypeptide.

16. A method for producing a polypeptide with an enhanced lytic activity by treating a polypeptide as defined in claim 11 and/or a polypeptide as obtainable by the method of claim

10 15, wherein said method comprises the steps of:

i) dialyzing said polypeptide against a buffer comprising a chelating compound,

ii) dialyzing said polypeptide against a divalent metal ion-containing buffer, preferably said divalent metal ion being selected from the group consisting of Co^{2+} , Cu^{2+} , Mn^{2+} and Zn^{2+} .

15

17. A composition comprising a nucleic acid molecule as defined in any one of the claims 1 – 10, a polypeptide as defined in claim 11 and/or as obtainable by a method as defined in claim 15 and/or 16, and/or a nucleic acid construct as defined in claim 12 and/or a vector as defined in claim 13 and/or a cell as defined in claim 14.

20

18. A composition according to claim 17, comprising one or more additional active ingredients, preferably selected from the group consisting of a bacteriophage and an antibiotic.

25 19. A composition according to claim 17 and/or 18 for use as a medicament.

20. Use of a composition according to claim 19 in the preparation of a medicament for the treatment of an infectious disease.

30 21. Use of a nucleic acid molecule according to any of the claims 1 - 10, a polypeptide according to claim 11, and/or as obtainable by a method as defined in claim 15 and/or 16, a nucleic acid construct according to claim 12, a vector as defined in claim 13, a

cell as defined in claim 14 and/or a composition according to claim 17 and/or 18 in the manufacture of an antimicrobial.

22. Use of a nucleic acid molecule according to any of the claims 1 - 10, a polypeptide according to claim 11, and/or as obtainable by a method as defined in claim 15 and/or 16, a nucleic acid construct according to claim 12, a vector as defined in claim 13, a cell as defined in claim 14 and/or a composition according to claim 17 and/or 18 as a food additive or a disinfectant.
23. Use of a nucleic acid molecule according to any of the claims 1 or 2, a polypeptide encoded by such nucleic acid molecule, a nucleic acid construct and/or vector and/or cell comprising such nucleic acid construct or a composition comprising any of such polypeptide, nucleic acid molecule, nucleic acid, vector and/or cell for detecting a *Staphylococcus* in food or on a surface.