

## **DESCRIPTION**

### **Engineered Gram-negative endolysins**

## **BACKGROUND OF THE INVENTION**

### **I. Field of the Invention**

[0001] The present invention relates to the field of antimicrobial enzymes. In particular, the present invention relates to a polypeptide comprising the amino acid sequence of a globular Gram-negative endolysin and the amino acid sequence of a cell wall binding domain of i) a modular Gram-negative endolysin or ii) a bacteriophage tail/baseplate protein. The present invention relates also to corresponding nucleic acids, vectors, bacteriophages, host cells, and compositions. The present inventions also relates to the use of said polypeptides, nucleic acids, vectors, bacteriophages, host cells, and compositions in methods for treatment of the human or animal body by surgery or therapy or in diagnostic methods practiced on the human or animal body. The polypeptides, nucleic acids, vectors, bacteriophages, host cells, and compositions according to the invention may also be used as an antimicrobial in, e.g., food or feed, in cosmetics, or as disinfecting agent.

### **II. Description of Related Art**

[0002] Endolysins are peptidoglycan hydrolases encoded by bacteriophages (i.e. bacterial viruses). They are synthesized during late gene expression in the lytic cycle of phage multiplication and mediate the release of progeny virions from infected cells through degradation of the bacterial peptidoglycan. In terms of enzymatic activity they are usually either  $\beta(1,4)$ -glycosylases (lysozymes), transglycosylases, amidases or endopeptidases. Antimicrobial application of endolysins was already suggested in 1991 by Gasson (GB2243611). Although the killing capacity of endolysins has been known for a long time, the use of these enzymes as antibacterials was ignored due to the success and dominance of antibiotics. Only after the appearance of multiple antibiotic resistant bacteria this simple concept of combating human pathogens with endolysins received interest. A compelling need to develop totally new classes of antibacterial agents emerged and endolysins used as

'enzymiotics' - a hybrid term of 'enzymes' and 'antibiotics' - perfectly met this need. In 2001, Fischetti and coworkers demonstrated for the first time the therapeutic potential of bacteriophage C1 endolysin towards group A streptococci (Nelson et al., 2001). Since then many publications have established endolysins as an attractive and complementary alternative to control bacterial infections, particularly by Gram positive bacteria. Subsequently different endolysins against other Gram positive pathogens such as *Streptococcus pneumoniae* (Loeffler et al., 2001), *Bacillus anthracis* (Schuch et al., 2002), *S. agalactiae* (Cheng et al., 2005) and *Staphylococcus aureus* (Rashel et al., 2007) have proven their efficacy as enzymiotics. Nowadays, the most important challenge of endolysin therapy lies in the insensitivity of Gram-negative bacteria towards the exogenous action of endolysins, since the outer membrane shields the access of endolysins from the peptidoglycan.

**[0003]** Gram-negative bacteria possess an outer membrane, with its characteristic asymmetric bilayer as a hallmark. The outer membrane bilayer consists of an inner monolayer containing phospholipids (primarily phosphatidyl ethanolamine) and an outer monolayer that is mainly composed of a single glycolipid, lipopolysaccharide (LPS). There is an immense diversity of LPS structures in the bacterial kingdom and the LPS structure may be modified in response to prevailing environmental conditions. The stability of the LPS layer and interaction between different LPS molecules is mainly achieved by the electrostatic interaction of divalent ions ( $Mg^{2+}$ ,  $Ca^{2+}$ ) with the anionic components of the LPS molecule (phosphate groups in the lipid A and the inner core and carboxyl groups of KDO). Furthermore, the dense and ordered packing of the hydrophobic moiety of lipid A, favored by the absence of unsaturated fatty acids, forms a rigid structure with high viscosity. This makes it less permeable for lipophilic molecules and confers additional stability to the outer membrane (OM).

**[0004]** In contrast to Gram-negative bacteria, Gram-positive bacteria do not possess an outer membrane. The cytoplasmic membrane is surrounded by an up to 25 nm thick layer of peptidoglycan (which is only up to 5 nm for Gram-negative bacteria) which forms the cell wall. Main purpose of the cell wall of Gram-positives is to maintain bacterial shape and to counteract the internal bacterial cell pressure. Peptidoglycan, or murein, is a polymer consisting of sugars and amino acids. The sugar component consists of alternating residues of  $\beta$ -(1,4) linked N-acetylglucosamine and N-acetylmuramic acid residues compose the sugar components. A peptide chain of three to five amino acids is attached to the N-acetylmuramic acid. The peptide chain can be cross-linked to the peptide chain of another strand forming a

3D mesh-like layer. The peptide chain may contain D- and L- amino acid residues and the composition may vary for different bacteria.

**[0005]** When comparing endolysins originating from phages infecting Gram-positive bacterial species versus those originating from phages infecting Gram-negative bacterial species, a striking difference can be observed, namely the general structure of the endolysins themselves. Phage endolysins of phages infecting Gram-positive bacteria are modular and comprise different individual functional domains or modules. The most common architecture is an N-terminal catalytic domain and a C-terminal cell wall-binding domain (Loessner, 2005). Some Gram-positive endolysins consist of three modules: an N-terminal and central catalytic domain with different specificity and a C-terminal substrate-binding module (Navarre et al., 1999; Pritchard et al., 2004; Yokoi et al., 2005). In contrast, endolysins encoded by phages associated with Gram-negative host cells are typically non-modular but single-module, globular proteins, with only a small number of exceptions. For more information on endolysins as antimicrobials see Nelson et al. ("Endolysins as Antimicrobials", *Advances in Virus Research*, Volume 83 (2012), p. 299-365, Eds. M. Lobocka and W. Szybalski, Elsevier).

**[0006]** Meanwhile, new strategies have emerged to utilize also endolysins originating from phages infecting Gram-negative bacterial species to control infections caused by Gram-negative bacteria. For this purpose, endolysins of Gram negative bacteria are fused with, e.g. cationic, amphipathic, hydrophobic or antimicrobial peptides. This type of fusion protein allows overcoming previous problems with the outer membrane of Gram-negative bacteria.

**[0007]** However, despite the advances in the art regarding antibacterial agents, there is still a need in the art for further improvement in the design of such antibacterial agents, in particular due to the increasing resistance to conventional antibiotics.

**[0008]** This problem is solved by the subject-matter as set forth below and in the appended claims.

### **SUMMARY OF THE INVENTION**

**[0009]** The inventors of the present invention have surprisingly found that fusing globular Gram-negative endolysins with cell wall binding domains of Gram-negative endolysins or of bacteriophage tail/baseplate proteins yields more effective Gram-negative muralytic enzymes.

**[0010]** Thus, in a first aspect the present invention relates to a polypeptide comprising the amino acid sequence of a Gram-negative globular endolysin and the amino acid sequence

of a cell wall binding domain of i) a Gram-negative modular endolysin or ii) a bacteriophage tail/baseplate protein.

**[0011]** Such inventive polypeptide may degrade, usually depending on the components chosen, the cell wall of bacteria selected from the group consisting of *Acinetobacter*, *Aeromonas*, *Aggregatibacter*, *Azospirillum*, *Bacteroides*, *Burkholderia*, *Campylobacter*, *Candidatus*, *Caulobacter*, *Clavibacter*, *Cronobacter*, *Delftia*, *Enterobacter*, *Erwinia*, *Escherichia*, *Flavobacterium*, *Haemophilus*, *Iodobacteria*, *Klebsiella*, *Kluyvera*, *Mannheimia*, *Morganella*, *Neisseria*, *Pantoea*, *Pasteurella*, *Planktothrix*, *Pseudoalteromonas*, *Pseudomonas*, *Ralstonia*, *Salmonella*, *Shigella*, *Sinorhizobium*, *Sodalis*, *Synechococcus*, *Thalassomonas*, *Thermus*, *Vibrio*, *Xanthomonas*, *Xylella*, and *Yersinia*.

**[0012]** The inventive polypeptide may additionally comprise at least one amino acid sequence selected from the group consisting of amphiphatic peptides, cationic peptides, hydrophobic peptides, naturally occurring antimicrobial peptides, sushi peptides and defensins. Such further peptide can enhance the antibacterial activity of the inventive polypeptide.

**[0013]** In further aspects, the present invention relates to nucleic acids encoding an inventive polypeptide, vectors or bacteriophages comprising an inventive nucleic acid as well as host cells comprising an inventive polypeptide, nucleic acid, vector, and/or bacteriophage.

**[0014]** The present invention relates in a further aspect also to compositions comprising a polypeptide, nucleic acid, vector, bacteriophage, and/or host cell according to the present invention. Such compositions are preferably pharmaceutical compositions comprising a pharmaceutically acceptable carrier, diluent, or excipient.

**[0015]** Finally, the present invention relates to polypeptides, nucleic acids, vectors, bacteriophages, host cells, compositions and/or kits of the present invention for use in methods of treatment, in particular for the treatment or prevention of bacterial infections.

## **DESCRIPTION OF ILLUSTRATIVE EMBODIMENTS**

### **I. Definitions**

**[0016]** The term “polypeptide” as used herein refers in particular to a polymer of amino acid residues linked by peptide bonds in a specific sequence. The amino acid residues of a polypeptide may be modified by e.g. covalent attachments of various groups such as

carbohydrates and phosphate. Other substances may be more loosely associated with the polypeptide, such as heme or lipid, giving rise to conjugated polypeptides which are also comprised by the term "polypeptide" as used herein. The term as used herein is intended to encompass also proteins. Thus, the term "polypeptide" also encompasses for example complexes of two or more amino acid polymer chains. The term "polypeptide" does encompass embodiments of polypeptides which exhibit optionally modifications typically used in the art, e.g. biotinylation, acetylation, pegylation, chemical changes of the amino-, SH- or carboxyl-groups (e.g. protecting groups) etc. As will become apparent from the description below, the polypeptide according to the present invention are fusion proteins, i.e. represent the linkage of at least two amino acid sequences which do not occur in this combination in nature. The term "polypeptide" as used herein is not limited to a specific length of the amino acid polymer chain, but typically the polypeptide will exhibit a length of more than about 150 amino acids. Usually, but not necessarily, a typical polypeptide of the present invention will not exceed about 1000 amino acids in length. The inventive polypeptide may for instance be at most about 750 amino acids long, at most about 500 amino acids long or at most about 300 amino acids long. A possible length range for the inventive polypeptide, without being limited thereto, may thus for example be about 200 to about 750 amino acids, or about 250 to about 600 amino acids. A particularly preferred range is about 250 to about 300 amino acids.

**[0017]** The term "fragment" as used herein refers to an amino acid sequence which is N-terminally, C-terminally, and/or on both termini truncated with respect to the respective reference sequence, for example a given endolysin or SEQ ID NO. Thus, a fragment of an amino acid sequence as used herein is an amino acid sequence which is at least one amino acid shorter than the respective reference sequence. A fragment of an amino acid sequence as used herein is preferably an amino acid sequence which is at most 20, more preferably at most 19, more preferably at most 18, more preferably at most 17, more preferably at most 16, more preferably at most 15, more preferably at most 14, more preferably at most 13, more preferably at most 12, more preferably at most 11, more preferably at most 10, more preferably at most 9, more preferably at most 8, more preferably at most 7, more preferably at most 6, more preferably at most 5, more preferably at most 4, more preferably at most 3, more preferably at most 2, most preferably 1 amino acid residue shorter than the respective reference amino acid sequence. The fragment may for example exhibit vis-à-vis the reference sequence a truncation of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or 20 amino acids at the N-terminus, the C-terminus or both. It is understood that a polypeptide

comprising a fragment of a given amino acid sequence does not comprise the full length of said reference amino acid sequence.

**[0018]** The term “endolysin” is generally understood to be a person skilled in the art. As already mentioned previously, the term refers to a particular type of enzymes encoded in nature by bacteriophages. The bacteriophage utilizes such enzyme to release viral progeny from the inner of the infected bacterial cell. The enzyme hydrolyzes the peptidoglycan in the cell wall, leading to rupture of the bacterial cell. While all endolysins are peptidoglycan degrading enzymes, the actual reaction catalyzed, i.e. the actual bond cleaved in the peptidoglycan of bacteria, may be different. In terms of reaction catalyzed, endolysins may be for example glycosidases, amidases, endopeptidases, or lytic transglycosylases. As will be known by a person skilled in the art, the term “endolysin” does not encompass any enzymes catalyzing the same reactions, but which are not derived from bacteriophages. While such enzymes (e.g., hen egg-white lysozyme) catalyze formally the same reaction, they differ significantly from endolysins, e.g. in terms of biological function, evolutionary background and structure. The term “endolysin”, as used herein encompasses naturally occurring endolysins, enzymatically active truncated versions thereof as well as technically modified endolysins deriving from these (i.e. naturally occurring endolysins and their enzymatically active fragments), e.g. with increased heat stability, reduced aggregation etc. Such modified endolysins will usually exhibit a sequence identity of at least 80%, preferably at least 82,5%, more preferably at least 85%, more preferably at least 87,5%, more preferably at least 90%, more preferably at least 92,5%, more preferably at least 95%, more preferably at least 97,5%, or most preferably at least 99% or more with the respective naturally occurring endolysin or its enzymatically active fragment. The term “Gram negative endolysin” refers to endolysins deriving from bacteriophages targeting Gram negative bacteria.

**[0019]** A “modular” endolysin, as used herein, is an endolysin which exhibits at least two distinct functional domains, namely at least one “enzymatically active domain” (EAD) and at least one “cell-wall-binding domain” (CBD). While the former provides the actual enzymatic activity, the latter may provide for target binding. Due to their domain character, these two activities can be separated from each other. Endolysins lacking a distinct CBD do not fall under the term “modular endolysin”.

**[0020]** A “cell wall binding domain”, or CBD, is an amino acid sequence within an endolysin sequence or phage tail/baseplate protein which folds into a structurally discreet module. The role of a CBD is to bind to the peptidoglycan and direct the catalytic machinery of the full length endolysin or phage tail/baseplate protein onto its substrate, thus enhancing

the catalytic efficiency of the multimodular peptidoglycan-degrading enzyme. CBDs are themselves devoid of any catalytic activity.

**[0021]** An “enzymatically active domain” (EAD), as used herein, refers to an amino acid sequence within a modular endolysin sequence which folds into a structurally discreet module. An EAD exerts a catalytic, enzymatic function, i.e. may act for example as endopeptidase, chitinase, T4 like muraminidase, lambda like muraminidase, N-acetyl-muramoyl-L-alanine-amidase (amidase), muramoyl-L-alanine-amidase, muramidase, lytic transglycosylase (C), lytic transglycosylase (M), N-acetyl-muramidase (lysozyme), N-acetyl-glucosaminidase or transglycosylase. Since EADs are derived from modular endolysins, an EAD will exhibit a high degree of sequence identity with a modular endolysin, i.e. will exhibit a sequence identity of at least 95%, more preferably at least 97,5%, or most preferably even 100% with a naturally occurring modular endolysin or its enzymatically active fragment.

**[0022]** As used herein, “globular endolysins” are those endolysins lacking the modular organization and structure of an EAD and a CBD. The term is not intended to encompass fragments of modular endolysins which only retain the enzymatic active domain of a modular endolysin, i.e. an EAD is not a globular endolysin. Hence, a globular endolysin will not exhibit any significant sequence identity with a naturally occurring modular endolysin or its enzymatically active fragment. Usually, a globular endolysin will exhibit less than 90%, more preferably less than 85%, more preferably less than 80%, more preferably less than 75%, more preferably less than 70%, more preferably less than 60%, and most preferably less than 50% or less sequence identity with modular endolysins.

**[0023]** The term “bacteriophage tail/baseplate protein” is generally understood to be a protein skilled in the art. Tail proteins and baseplate proteins are proteins of bacteriophages. Binding structures located in the tail fiber and/or baseplate of bacteriophages play an important role in mediating injection of the phage genome into the host cell. Tail fiber proteins are positioned at the tip of the tail and are responsible for binding to the cell surface by recognizing a potential host bacterium and attaching to its outer surface. Baseplate proteins control the transfer of the genetic material and can have also cell binding properties. Especially for Myoviruses of Gram negative bacteria (e.g. T4 or P2 phages) different motifs are described which show homology to peptidoglycan binding domains like LysM. Another example is the gp5 of the ICP1 vibrio phage and related proteins encoded in the genome of phages infecting different species like e.g. *Methylobacter* sp.. These consist of a peptidoglycan binding domain and an enzymatic active domain, able to degrade the murein layer of the host bacteria.

**[0024]** The term "% sequence identity" is generally understood in the art. Two sequences to be compared are aligned to give a maximum correlation between the sequences. This may include inserting "gaps" in either one or both sequences, to enhance the degree of alignment. A % identity may then be determined over the whole length of each of the sequences being compared (so-called global alignment), that is particularly suitable for sequences of the same or similar length, or over shorter, defined lengths (so-called local alignment), that is more suitable for sequences of unequal length. In the above context, an amino acid sequence having a "sequence identity" of at least, for example, 95% to a query amino acid sequence, is intended to mean that the sequence of the subject amino acid sequence is identical to the query sequence except that the subject amino acid sequence may include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain an amino acid sequence having a sequence of at least 95% identity to a query amino acid sequence, up to 5% (5 of 100) of the amino acid residues in the subject sequence may be inserted or substituted with another amino acid or deleted. Methods for comparing the identity and homology of two or more sequences are well known in the art. The percentage to which two sequences are identical can for example be determined by using a mathematical algorithm. A preferred, but not limiting, example of a mathematical algorithm which can be used is the algorithm of Karlin et al. (1993), PNAS USA, 90:5873-5877. Such an algorithm is integrated in the BLAST family of programs, e.g. BLAST or NBLAST program (see also Altschul et al., 1990, J. Mol. Biol. 215, 403-410 or Altschul et al. (1997), Nucleic Acids Res, 25:3389-3402), accessible through the home page of the NCBI at world wide web site [ncbi.nlm.nih.gov](http://ncbi.nlm.nih.gov) and FASTA (Pearson (1990), Methods Enzymol. 83, 63-98; Pearson and Lipman (1988), Proc. Natl. Acad. Sci. U. S. A 85, 2444-2448.). Sequences which are identical to other sequences to a certain extent can be identified by these programs. Furthermore, programs available in the Wisconsin Sequence Analysis Package, version 9.1 (Devereux et al, 1984, Nucleic Acids Res., 387-395), for example the programs BESTFIT and GAP, may be used to determine the % identity between two polypeptide sequences. BESTFIT uses the "local homology" algorithm of (Smith and Waterman (1981), J. Mol. Biol. 147, 195-197.) and finds the best single region of similarity between two sequences. If herein reference is made to an amino acid sequence sharing a particular extent of sequence identity to a reference sequence, then said difference in sequence is preferably due to conservative amino acid substitutions. Preferably, such sequence retains the activity of the reference sequence, e.g. retains the activity of degrading the peptidoglycan layer of Gram-negative bacteria, albeit maybe at a slower rate. In addition, if reference is made herein to a sequence sharing "at least"

at certain percentage of sequence identity, then 100% sequence identity are preferably not encompassed.

**[0025]** As used herein, the term "cationic peptide" refers preferably to a peptide having positively charged amino acid residues. Preferably a cationic peptide has a pKa-value of 9.0 or greater. Typically, at least four of the amino acid residues of the cationic peptide can be positively charged, for example, lysine or arginine. "Positively charged" refers to the side chains of the amino acid residues which have a net positive charge at about physiological conditions. The term "cationic peptide" as used herein refers also to polycationic peptides, but also includes cationic peptides which comprise for example less than 20%, preferably less than 10% positively charged amino acid residues.

**[0026]** The term "polycationic peptide", as used herein, refers preferably to a peptide composed of mostly positively charged amino acid residues, in particular lysine and/or arginine residues. A peptide is composed of mostly positively charged amino acid residues if at least about 20, 30, 40, 50, 60, 70, 75, 80, 85, 90, 95 or about 100 % of the amino acid residues are positively charged amino acid residues, in particular lysine and/or arginine residues. The amino acid residues being not positively charged amino acid residues can be neutrally charged amino acid residues and/or negatively charged amino acid residues and/or hydrophobic amino acid residues. Preferably the amino acid residues being not positively charged amino acid residues are neutrally charged amino acid residues, in particular serine and/or glycine.

**[0027]** The term, "antimicrobial peptide" (AMP), as used herein, refers preferably to any naturally occurring peptide that has microbicidal and/or microbistatic activity on for example bacteria, viruses, fungi, yeasts, mycoplasma and protozoa. Thus, the term "antimicrobial peptide" as used herein refers in particular to any peptide having anti-bacterial, anti-fungal, anti-mycotic, anti-parasitic, anti-protozoal, anti-viral, anti-infectious, anti-infective and/or germicidal, algicidal, amoebicidal, microbicidal, bactericidal, fungicidal, parasiticidal, protozoacidal, protozoicidal properties. Preferred are anti-bacterial peptides. The antimicrobial peptide may be a member of the RNase A super family, a defensin, cathelicidin, granulysin, histatin, psoriasin, dermicidine or hepcidin. The antimicrobial peptide may be naturally occurring in insects, fish, plants, arachnids, vertebrates or mammals. Preferably the antimicrobial peptide may be naturally occurring in insects, fish, plants, arachnids, vertebrates or mammals. Preferably the antimicrobial peptide may be naturally occurring in radish, silk moth, wolf spider, frog, preferably in *Xenopus laevis*, *Rana* frogs, more preferably in *Rana catesbeiana*, toad, preferably Asian toad *Bufo bufo gargarizans*, fly, preferably in *Drosophila*,

more preferably in *Drosophila melanogaster*, in *Aedes aegypti*, in honey bee, bumblebee, preferably in *Bombus pascuorum*, flesh fly, preferably in *Sarcophaga peregrine*, scorpion, horseshoe crab, catfish, preferably in *Parasilurus asotus*, cow, pig, sheep, porcine, bovine, monkey and human. As used herein, an "antimicrobial peptide" (AMP) may in particular be a peptide which is not a cationic peptide, polycationic peptide, amphipathic peptide, sushi peptide, defensins, and hydrophobic peptide, but nevertheless exhibits antimicrobial activity.

**[0028]** The term "sushi peptide", as used herein, refers to complement control proteins (CCP) having short consensus repeats. The sushi module of sushi peptides functions as a protein-protein interaction domain in many different proteins. Peptides containing a Sushi domain have been shown to have antimicrobial activities. Preferably, sushi peptides are naturally occurring peptides.

**[0029]** The term "amphipathic peptide", as used herein, refers to synthetic peptides having both hydrophilic and hydrophobic functional groups. Preferably, the term "amphipathic peptide" as used herein refers to a peptide having a defined arrangement of hydrophilic and hydrophobic groups e.g. amphipathic peptides may be e.g. alpha helical, having predominantly non polar side chains along one side of the helix and polar residues along the rest of its surface.

**[0030]** The term "hydrophobic group", as used herein, refers preferably to chemical groups such as amino acid side chains which are substantially water insoluble, but soluble in an oil phase, with the solubility in the oil phase being higher than that in water or in an aqueous phase. In water, amino acid residues having a hydrophobic side chain interact with one another to generate a non-aqueous environment. Examples of amino acid residues with hydrophobic side chains are valine, isoleucine, leucine, methionine, phenylalanine, tryptophan, cysteine, alanine, tyrosine, and proline residues.

**[0031]** The term "hydrophobic peptide", as used herein, refers to a hydrophobic peptide, which is preferably composed of mostly amino acid residues with hydrophobic groups. Such peptide is preferably composed of mostly hydrophobic amino acid residues, i.e. at least about 20, 30, 40, 50, 60, 70, 75, 80, 85, 90, 95 or at least about 100 % of the amino acid residues are hydrophobic amino acid residues. The amino acid residues being not hydrophobic are preferably neutral and preferably not hydrophilic.

**[0032]** The term "comprising", as used herein, shall not be construed as being limited to the meaning "consisting of" (i.e. excluding the presence of additional other matter). Rather, "comprising" implies that optionally additional matter may be present. The term "comprising" encompasses as particularly envisioned embodiments falling within its scope "consisting of"

(i.e. excluding the presence of additional other matter) and "comprising but not consisting of" (i.e. requiring the presence of additional other matter), with the former being more preferred.

**[0033]** The use of the word "a" or "an", when used herein, may mean "one," but it is also consistent with the meaning of "one or more," "at least one," and "one or more than one."

## **II. Polypeptides**

**[0034]** As already mentioned, the present invention relates in a first aspect to a polypeptide comprising the amino acid sequence of a Gram-negative globular endolysin and the amino acid sequence of a cell wall binding domain of i) a Gram-negative modular endolysin or ii) a bacteriophage tail/baseplate protein.

**[0035]** Endolysins are well known to a person skilled in the art of antimicrobial agents. Many of these proteins have been sequenced and their structure and domain architecture (including CBDs and EADs) analyzed. Usually it is thus very simple to elucidate the presence of a cell wall binding domain in a given endolysin sequence, for example on basis of homology analyses. In case of doubt, whether a given sequence of an endolysin acts as cell wall binding domain or not, said property can also be analyzed by routine test known in the art. Exemplary tests are provided for instance in Mol Microbiol. 2002 Apr;44(2):335-49 and Briers et al. (Mol Microbiol. 2007 Sep;65(5):1334-44). Briefly, the candidate cell wall-binding domain is fused (e.g. N-terminally) to green fluorescent protein (GFP). Subsequently, the GFP-fusion protein is incubated with the target bacteria of the parent endolysin. If these are gram-negative bacteria, then the outer membrane is permeabilized in advance by treatment with a chloroform-saturated buffer (chloroform-saturated 0.05 M Tris-buffer (pH 7.7), 45 min (Lavigne *et al.*, Cell Mol Life Sci. 2004 Nov;61(21):2753-9). The candidate CBD-GFP fusion is then added to the permeabilized cells (e.g. final concentration 5  $\mu$ M). Purified recombinant GFP is used in the negative control. This mixture is then incubated, e.g. for 5 min at 25°C, subsequently spun down and the supernatant discarded. The cell pellet may then be washed (e.g. twice in the corresponding buffer) and analyzed for GFP-binding via suitable means such as epifluorescence microscopy, flow cytometry or confocal fluorescence microscopy. In addition, a specific peptidoglycan binding test can be performed Briers et al. (Mol Microbiol. 2007 Sep;65(5):1334-44). For this purpose murein of the target bacteria is isolated and contacted with the candidate CBD-GFP fusion protein and binding is once again analyzed, e.g. via epifluorescence microscopy.

**[0036]** Exemplary cell wall binding domains of Gram-negative endolysins, which may be used in carrying out the present invention, are cell wall binding domains deriving from the endolysins of bacteriophages  $\Phi$ KZ and EL, or from the endolysins OBPgpLYS, PVPSE1gp146, and 201 $\phi$ 2-1.

**[0037]** Phage baseplate/tail proteins are also known in the art and information on sequence, structure and domain architecture is frequently available. Otherwise, similar tests as mentioned above for endolysins are suited to identify cell wall binding domains of bacteriophage tail/baseplate proteins. An exemplary cell wall binding domain of bacteriophage tail/baseplate protein, which may be used in carrying out the present invention, is the cell wall binding domains deriving from the baseplate protein of Vibrio phage ICP1 (see YP\_004251150.1) or Vibrio phage RYC (BAV80844.1).

**[0038]** Examples for specific sequences comprising cell wall binding domains (deriving from Gram-negative endolysins or bacteriophage tail/baseplate proteins) are provided herein in SEQ ID Nos: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, and 17, and sequences having at least 80% sequence identity to any of these while retaining the property of cell wall binding. Particularly preferred sequences are SEQ ID Nos: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 and 14, and sequences having at least 80% sequence identity to any of these while retaining the property of cell wall binding. Most preferred are sequences deriving from the cell wall binding domain of KZ144 endolysin, such as SEQ ID Nos: 1, 2, 3, 4, 5, 6, 7, and 8, and sequences having at least 80% sequence identity to any of these while retaining the property of cell wall binding.

**[0039]** As previously mentioned, the inventive polypeptide comprises aside of the amino acid sequence of a cell wall binding domain of a Gram-negative modular endolysin or a bacteriophage baseplate tail protein also the amino acid sequence of a Gram-negative globular endolysin. A person skilled in the art will be readily able to ascertain whether a given polypeptide derives from a Gram-negative endolysin exhibiting a cell wall binding domain (i.e. is a modular endolysin) or derives from a Gram-negative endolysin which does not exhibit such domain structure (globular endolysin). Examples for globular endolysins are for example provided in Table 1 below.

**[0040]** Table 1: Examples for globular endolysins

| Host          | Phage name                | Protein ID     |
|---------------|---------------------------|----------------|
| Acinetobacter | Acinetobacter phage Ac42  | YP_004009379.1 |
| Acinetobacter | Acinetobacter phage Acj61 | YP_004009630.1 |

| Host            | Phage name                              | Protein ID     |
|-----------------|-----------------------------------------|----------------|
| Acinetobacter   | Acinetobacter phage Acj9                | YP_004010153.1 |
| Aeromonas       | Aeromonas phage PX29                    | ADQ53036.1     |
| Aeromonas       | Aeromonas phage PAS-1                   | AEM36042.1     |
| Aeromonas       | Aeromonas phage phiAS7                  | AEZ65050.1     |
| Aeromonas       | Aeromonas phage phiAS4                  | YP_003969055.1 |
| Aeromonas       | Aeromonas phage 44RR2.8t                | NP_932578.1    |
| Aeromonas       | Aeromonas phage 25                      | YP_656449.1    |
| Aeromonas       | Aeromonas phage 31                      | YP_238949.1    |
| Aeromonas       | Aeromonas phage Aeh1                    | NP_944217.1    |
| Aeromonas       | Aeromonas phage phiO18P                 | YP_001285657.1 |
| Aeromonas       | Aeromonas phage 65                      | YP_004300997.1 |
| Aeromonas       | Aeromonas phage phiAS5                  | YP_003969406.1 |
| Azospirillum    | Azospirillum phage Cd                   | YP_001686894.1 |
| Aggregatibacter | Aggregatibacter phage S1249             | YP_003344813.1 |
| Bacteroides     | Bacteroides phage B40-8                 | YP_002221548.1 |
| Bacteroides     | Bacteroides phage B124-14               | YP_005102482.1 |
| Burkholderia    | Burkholderia phage BcepF1               | YP_001039778.1 |
| Burkholderia    | Burkholderia phage BcepB1A              | YP_024909.1    |
| Burkholderia    | Burkholderia phage Bcep22               | YP_001531197.1 |
| Burkholderia    | Burkholderia phage BcepNazgul           | NP_918971.2    |
| Burkholderia    | Burkholderia phage phiE255              | YP_001111252.1 |
| Burkholderia    | Burkholderia phage KS9                  | YP_003090199.1 |
| Burkholderia    | Burkholderia phage phi1026b             | NP_945054.1    |
| Burkholderia    | Burkholderia phage phiE125              | NP_536381.1    |
| Burkholderia    | Burkholderia phage BcepC6B              | YP_024942.1    |
| Burkholderia    | Burkholderia phage Bcep176              | YP_355393.1    |
| Burkholderia    | Burkholderia phage BcepL02              | YP_002922746.1 |
| Burkholderia    | Burkholderia phage BcepMu               | YP_024695.1    |
| Burkholderia    | Burkholderia phage KS10                 | YP_002221425.1 |
| Burkholderia    | Burkholderia phage phi644-2             | YP_001111104.1 |
| Campylobacter   | Campylobacter phage CP220               | CBJ93929.1     |
| Campylobacter   | Campylobacter phage CPT10               | CBJ94327.1     |
| Candidatus      | Acyrtosiphon pisum bacteriophage APSE-1 | NP_050974.1    |
| Candidatus      | Bacteriophage APSE-2                    | ACJ10174.1     |
| Candidatus      | Bacteriophage APSE-4                    | ACJ10096.1     |
| Candidatus      | Bacteriophage APSE-7                    | ACJ10111.1     |
| Candidatus      | Bacteriophage APSE-5                    | ACJ10082.1     |
| Candidatus      | Bacteriophage APSE-3                    | ACJ10123.1     |
| Candidatus      | Bacteriophage APSE-6                    | ACJ10136.1     |
| Caulobacter     | Caulobacter phage Cd1                   | ADD21680.1     |
| Cronobacter     | Cronobacter phage ENT47670              | ADZ13641.1     |
| Cronobacter     | Cronobacter phage ESP2949-1             | AEM24793.1     |
| Cronobacter     | Cronobacter phage ES2                   | AEM24706.1     |
| Cronobacter     | Cronobacter phage ENT39118              | ADZ13601.1     |

| Host         | Phage name                           | Protein ID     |
|--------------|--------------------------------------|----------------|
| Delftia      | Deftia phage phiW-14                 | YP_003358866.1 |
| Enterobacter | Enterobacteria phage F20             | AEQ39188.1     |
| Escherichia  | Escherichia phage TL-2011b           | AEW24559.1     |
| Escherichia  | Escherichia phage vB_EcoM_ECO1230-10 | ADE87938.1     |
| Escherichia  | Escherichia phage HK639              | YP_004934099.1 |
| Escherichia  | Escherichia phage phiV10             | YP_512283.1    |
| Escherichia  | Escherichia phage rv5                | YP_002003587.1 |
| Escherichia  | Escherichia phage K1H                | ADA82342.1     |
| Escherichia  | Escherichia phage K1ind3             | ADA82488.1     |
| Escherichia  | Escherichia phage K1G                | ADA82292.1     |
| Escherichia  | Escherichia phage wV7                | AEM00790.1     |
| Escherichia  | Escherichia phage HK75               | YP_004934160.1 |
| Escherichia  | Stx2-converting phage 1717           | YP_002274257.1 |
| Escherichia  | Stx2-converting phage 86             | YP_794054.1    |
| Escherichia  | Stx2 converting phage II             | YP_003828995.1 |
| Escherichia  | Escherichia phage phiEB49            | AEI91208.1     |
| Escherichia  | Enterobacteria phage T1              | YP_003933.1    |
| Escherichia  | Enterobacteria phage T3              | NP_523313.1    |
| Escherichia  | Enterobacteria phage T4              | NP_049736.1    |
| Escherichia  | Enterobacteria phage T5              | YP_006868.1    |
| Escherichia  | Enterobacteria phage TLS             | YP_001285558.1 |
| Escherichia  | Enterobacteria phage vB_EcoM-VR7     | YP_004063811.1 |
| Escherichia  | Enterobacteria phage WV8             | YP_002922821.1 |
| Escherichia  | Enterobacteria phage Min27           | ABY49900.1     |
| Escherichia  | Enterobacteria phage lambda          | NP_040645.1    |
| Escherichia  | Enterobacteria phage K1F             | YP_338105.1    |
| Escherichia  | Enterobacteria phage IME10           | AER08021.1     |
| Escherichia  | Enterobacteria phage vB_EcoM-FV3     | AEZ65218.1     |
| Escherichia  | Escherichia phage TL-2011c           | AEW24625.1     |
| Escherichia  | Enterobacteria phage Bp7             | AEN93735.1     |
| Escherichia  | Enterobacteria phage RB49            | NP_891673.1    |
| Escherichia  | Enterobacteria phage RTP             | YP_399008.1    |
| Escherichia  | Enterobacteria phage BP-4795         | YP_001449285.1 |
| Escherichia  | Enterobacteria phage BA14            | YP_002003466.1 |
| Escherichia  | Enterobacteria phage 285P            | YP_004300550.1 |
| Escherichia  | Enterobacteria phage P1              | YP_006484.1    |
| Escherichia  | Enterobacteria phage JS98            | YP_001595245.1 |
| Escherichia  | Enterobacteria phage AR1             | BAI83135.1     |
| Escherichia  | Enterobacteria phage SfV             | NP_599082.1    |
| Escherichia  | Enterobacteria phage T7              | NP_041973.1    |
| Escherichia  | Enterobacteria phage phiEco32        | YP_001671762.1 |
| Escherichia  | Enterobacteria phage A5              | ABF71471.1     |
| Escherichia  | Enterobacteria phage 186             | AAC34155.1     |
| Escherichia  | Enterobacteria phage HK022           | AAF30387.1     |

| Host           | Phage name                             | Protein ID     |
|----------------|----------------------------------------|----------------|
| Escherichia    | Enterobacteria phage JS10              | YP_002922463.1 |
| Escherichia    | Enterobacteria phage P2                | NP_046765.1    |
| Escherichia    | Enterobacteria phage 933W (sensu lato) | NP_049505.1    |
| Escherichia    | Enterobacteria phage JK06              | YP_277498.1    |
| Escherichia    | Enterobacteria phage N15               | NP_046950.1    |
| Escherichia    | Enterobacteria phage K1E               | YP_425023.1    |
| Escherichia    | Enterobacteria phage K1-5              | YP_654144.1    |
| Escherichia    | Enterobacteria phage JSE               | YP_002922178.1 |
| Escherichia    | Enterobacteria phage IME08             | YP_003734260.1 |
| Escherichia    | Enterobacteria phage HK97              | NP_037753.1    |
| Escherichia    | Enterobacteria phage RB43              | YP_239135.1    |
| Escherichia    | Enterobacteria phage RB16              | YP_003858447.1 |
| Escherichia    | Enterobacteria phage phiP27            | NP_543082.1    |
| Escherichia    | Enterobacteria phage phiEcoM-GJ1       | YP_001595416.1 |
| Escherichia    | Enterobacteria phage Phi1              | YP_001469446.1 |
| Escherichia    | Enterobacteria phage EcoDS1            | YP_002003756.1 |
| Escherichia    | Enterobacteria phage cdtI              | YP_001272571.1 |
| Escherichia    | Enterobacteria phage CC31              | YP_004009990.1 |
| Escherichia    | Enterobacteria phage K30               | YP_004678738.1 |
| Escherichia    | Enterobacteria phage CUS-3             | ABQ88407.1     |
| Escherichia    | Stx1 converting phage                  | BAC77971.1     |
| Escherichia    | Stx2 converting phage I                | BAB88004.1     |
| Escherichia    | Stx1-converting phage phi-O153         | AAW21764.1     |
| Escherichia    | Escherichia phage D108                 | YP_003335769.1 |
| Escherichia    | Enterobacteria phage SPC35             | YP_004306522.1 |
| Escherichia    | Enterobacteria phage Mu                | NP_050626.1    |
| Escherichia    | Enterobacteria phage RB69              | NP_861818.1    |
| Escherichia    | Enterobacteria phage RB14              | YP_002854463.1 |
| Escherichia    | Enterobacteria phage RB32              | ABI94948.1     |
| Escherichia    | Enterobacteria phage RB51              | YP_002854084.1 |
| Erwinia        | Erwinia phage phiEa116                 | CCA66256.1     |
| Erwinia        | Erwinia phage vB_EamM-M7               | AEJ81266.1     |
| Erwinia        | Erwinia phage vB_EamM-Y2               | AEJ81402.1     |
| Erwinia        | Erwinia phage vB_EamP-L1               | AEJ81484.1     |
| Erwinia        | Erwinia phage phiEa21-4                | YP_002456060.1 |
| Erwinia        | Erwinia phage phiEa104                 | YP_004327012.1 |
| Erwinia        | Erwinia phage Era103                   | YP_001039680.1 |
| Flavobacterium | Flavobacterium phage 11b               | YP_112524.1    |
| Haemophilus    | Haemophilus phage SuMu                 | AEG42272.1     |
| Haemophilus    | Haemophilus phage HP1                  | NP_043495.1    |
| Haemophilus    | Haemophilus phage HP2                  | NP_536831.1    |
| Haemophilus    | Haemophilus phage Aaphi23              | NP_852750.1    |
| Iodobacteria   | Iodobacteriophage phiPLPE              | YP_002128449.1 |
| Klebsiella     | Klebsiella phage K11                   | YP_002003804.1 |

| Host                | Phage name                          | Protein ID     |
|---------------------|-------------------------------------|----------------|
| Klebsiella          | Klebsiella phage KP32               | YP_003347533.1 |
| Klebsiella          | Klebsiella phage KP15               | YP_003580002.1 |
| Klebsiella          | Enterobacteria phage vB_KleM-RaK2   | AFA44346.1     |
| Kluyvera            | Kluyvera phage Kvp1                 | YP_002308397.1 |
| Mannheimia          | Mannheimia phage phiMHaA1           | YP_655477.1    |
| Mannheimia          | Mannheimia phage phiMhaA1-PHL101    | ABD90561.1     |
| Morganella          | Morganella phage MmP1               | YP_002048642.1 |
| Alphaproteobacteria | Phage phiJL001                      | YP_224014.1    |
| Pseudoalteromonas   | Pseudoalteromonas phage H105/1      | YP_004327143.1 |
| Pseudomonas         | Pseudomonas phage LKA1              | YP_001522894.1 |
| Pseudomonas         | Pseudomonas phage D3                | NP_061527.1    |
| Pseudomonas         | Pseudomonas phage F10               | YP_001293405.1 |
| Pseudomonas         | Pseudomonas phage F116              | YP_164326.1    |
| Pseudomonas         | Pseudomonas phage F8                | YP_001294463.1 |
| Pseudomonas         | Pseudomonas phage gh-1              | NP_813758.1    |
| Pseudomonas         | Pseudomonas phage LBL3              | YP_002154189.1 |
| Pseudomonas         | Pseudomonas phage LKD16             | YP_001522837.1 |
| Pseudomonas         | Pseudomonas phage LMA2              | YP_002154280.1 |
| Pseudomonas         | Pseudomonas phage LUZ19             | YP_001671990.1 |
| Pseudomonas         | Pseudomonas phage PA11              | YP_001294626.1 |
| Pseudomonas         | Pseudomonas phage PAJU2             | YP_002284361.1 |
| Pseudomonas         | Pseudomonas phage PaP3              | NP_775256.1    |
| Pseudomonas         | Pseudomonas phage PB1               | YP_002455978.1 |
| Pseudomonas         | Pseudomonas phage phi13 (S-segment) | NP_690810.1    |
| Pseudomonas         | Pseudomonas phage phi15             | YP_004286199.1 |
| Pseudomonas         | Pseudomonas phage phi-2             | YP_003345505.1 |
| Pseudomonas         | Pseudomonas phage phiIBB-PF7A       | YP_004306332.1 |
| Pseudomonas         | Pseudomonas phage phikF77           | YP_002727868.1 |
| Pseudomonas         | Pseudomonas phage PT2               | YP_002117830.1 |
| Pseudomonas         | Pseudomonas phage PT5               | YP_002117771.1 |
| Pseudomonas         | Pseudomonas phage SN                | YP_002418854.1 |
| Pseudomonas         | Pseudomonas phage phi297            | YP_005098091.1 |
| Pseudomonas         | Pseudomonas phage Bf7               | YP_005098158.1 |
| Pseudomonas         | Pseudomonas phage PaP1              | AEK21612.1     |
| Pseudomonas         | PseudomonasLUZ24                    | YP_001671940.1 |
| Pseudomonas         | Pseudomonas phage phi-6 segment S   | NP_620343.1    |
| Pseudomonas         | Pseudomonas phage vB_PaeS_PMG1      | YP_005098234.1 |
| Pseudomonas         | Enterobacteria phage phiKMV         | NP_877484.1    |
| Pasteurella         | Pasteurella phage F108              | YP_654740.1    |
| Ralstonia           | Ralstonia phage RSB2                | BAJ51815.1     |
| Salmonella          | Salmonella phage SETP3              | YP_001110823.1 |
| Salmonella          | Bacteriophage PS3                   | CAA09701.1     |
| Salmonella          | Enterobacteria phage ST104          | YP_006397.1    |
| Salmonella          | Enterobacteria phage SP6            | AAR90036.1     |

| Host          | Phage name                          | Protein ID     |
|---------------|-------------------------------------|----------------|
| Salmonella    | Enterobacteria phage Felix 01       | NP_944846.1    |
| Salmonella    | Enterobacteria phage P22            | DAA01040.1     |
| Salmonella    | Bacteriophage Wphi                  | AAN28227.1     |
| Salmonella    | Enterobacteria phage EPS7           | YP_001836966.1 |
| Salmonella    | Enterobacteria phage Fels-2         | YP_001718740.1 |
| Salmonella    | Salmonella phage ES18               | YP_224214.1    |
| Salmonella    | Salmonella phage PsP3               | NP_958065.1    |
| Salmonella    | Salmonella phage E1                 | YP_001742044.1 |
| Salmonella    | Salmonella phage SETP12             | ABN70688.1     |
| Salmonella    | Salmonella phage SETP5              | ABN70687.1     |
| Salmonella    | Salmonella phage HK620              | NP_112069.1    |
| Salmonella    | Salmonella phage RE-2010            | ADQ92398.1     |
| Salmonella    | Phage Gifsy-1                       | YP_001700616.1 |
| Salmonella    | Phage Gifsy-2                       | YP_001700672.1 |
| Salmonella    | Salmonella phage c341               | YP_003090277.1 |
| Salmonella    | Salmonella phage epsilon15          | NP_848233.1    |
| Salmonella    | Salmonella phage epsilon34          | YP_002533525.1 |
| Salmonella    | Salmonella phage Fels-1             | YP_001700560.1 |
| Salmonella    | Salmonella phage phiSG-JL2          | YP_001949762.1 |
| Salmonella    | Salmonella phage SE1                | YP_002455881.1 |
| Salmonella    | Salmonella phage ST64B              | NP_700425.1    |
| Salmonella    | Salmonella phage ST64T              | NP_720320.1    |
| Salmonella    | Salmonella phage Vi06               | YP_004306666.1 |
| Salmonella    | Salmonella phage SPN1S              | YP_005098003.1 |
| Salmonella    | Salmonella phage SE2                | YP_005098118.1 |
| Salmonella    | Salmonella phage SS3e               | YP_005097816.1 |
| Salmonella    | Salmonella Phage PS34               | O80288.1       |
| Salmonella    | Enterobacteria phage 13a            | YP_002003950.1 |
| Shigella      | Shigella phage Sf6                  | NP_958236.1    |
| Shigella      | Shigella phage Shfl1                | YP_004414884.1 |
| Shigella      | Shigella phage Shfl2                | YP_004415022.1 |
| Shigella      | Shigella phage EP23                 | YP_004957490.1 |
| Sinorhizobium | Sinorhizobium phage PBC5            | NP_542265.1    |
| Sodalis       | Sodalis phage phiSG1                | YP_516184.1    |
| Sodalis       | Sodalis phage SO-1                  | YP_003344991.1 |
| Synechococcus | Synechococcus phage S-CBS2          | YP_004421540.1 |
| Synechococcus | Synechococcus phage S-CRM01         | YP_004508546.1 |
| Synechococcus | Synechococcus phage S-PM2           | YP_195189.2    |
| Synechococcus | Synechococcus phage S-PM2 (another) | YP_195188.1    |
| Thalassomonas | Thalassomonas phage BA3             | YP_001552293.1 |
| Thermus       | Thermus phage P23-45                | YP_001467961.1 |
| Thermus       | Thermus phage P23-77                | YP_003169716.1 |
| Thermus       | Thermus phage P23-77 (another)      | YP_003169717.1 |
| Thermus       | Thermus phage P74-26                | YP_001468077.1 |

| Host          | Phage name                            | Protein ID     |
|---------------|---------------------------------------|----------------|
| Vibrio        | Vibrio phage ICP3                     | YP_004251275.1 |
| Vibrio        | Vibrio phage K139                     | NP_536660.1    |
| Vibrio        | Vibrio phage kappa                    | YP_001650899.1 |
| Vibrio        | Vibrio phage N4                       | YP_003347926.1 |
| Vibrio        | Vibrio phage VP93                     | YP_002875665.1 |
| Vibrio        | Vibrio phage ICP3_2009_B              | ADX87518.1     |
| Vibrio        | Vibriophage VP4                       | YP_249586.1    |
| Vibrio        | Vibrio phage ICP3_2007_A              | ADX87661.1     |
| Xanthomonas   | Xanthomonas phage OP1                 | YP_453585.1    |
| Xanthomonas   | Xanthomonas phage phiL7               | YP_002922642.1 |
| Xanthomonas   | Xanthomonas phage Xop411              | YP_001285697.1 |
| Xanthomonas   | Xanthomonas phage Xp10                | NP_858975.1    |
| Xylella       | Xylella phage Xfas53                  | YP_003344916.1 |
| Yersinia      | Yersinia pestis phage phiA1122        | NP_848277.1    |
| Yersinia      | Yersinia phage Berlin                 | YP_918995.1    |
| Yersinia      | Yersinia phage L-413C                 | NP_839858.1    |
| Yersinia      | Yersinia phage phiR1-37               | YP_004934318.1 |
| Yersinia      | Yersinia phage phiYeO3-12             | NP_052084.1    |
| Yersinia      | Yersinia phage PY54                   | NP_892107.1    |
| Yersinia      | Yersinia phage Yepe2                  | YP_002003326.1 |
| Yersinia      | Yersinia phage PY100                  | CAJ28446.1     |
| Yersinia      | Yersinia phage Yep-phi                | ADQ83168.1     |
| Acinetobacter | Acinetobacter phage vB_AbaP_CEB1      | ALC76575.1     |
| Salmonella    | Salmonella phage phi68                | AHY18890.1     |
| Salmonella    | Salmonella phage SPN9CC               | YP_006383882   |
| Acinetobacter | Acinetobacter phage vB_AbaP_Acibel007 | YP_009103259.1 |
| Acinetobacter | Acinetobacter phage Abp1              | AFV51025.1     |
| Pseudomonas   | Pseudomonas aeruginosa phage LUZ7     | YP_003358335.1 |
| Pseudomonas   | Pseudomonas aeruginosa phage LIT1     | YP_003358446.1 |
| Pseudomonas   | Pseudomonas phage vB_PaeP_C2-10_Ab09  | YP_009031822.1 |
| Pseudomonas   | Pseudomonas phage 2b.1b               | SEQ ID NO: 18  |
| Salmonella    | Salmonella phage SBA-1781             | AFU63467.1     |
| Salmonella    | Salmonella phage Shivani              | AJA73488.1     |

**[0041]** The sequences of the globular endolysins of table 1 may be accessed for instance via the protein database of NCBI. It is understood that the sequences of the endolysins listed in table 1 may also be modified, e.g. may lack the N-terminal methionine to avoid a further start codon in the corresponding nucleic acid sequence. Using such marginally amended sequences is also within the scope of the present invention and it is understood, that when reference herein is made to endolysins of table 1, that also such modified endolysins are encompassed by said definition.

**[0042]** Particularly preferred sequences of globular endolysins, which may be used in carrying out the present invention, are sequences deriving from globular endolysins Abgp46, Lys68 of Salmonella phage phi68 and Lys394 endolysin. Exemplary sequences are provided as SEQ ID Nos. 18, 19, 20, 21, 22, 23, 24, 25, 26, 27 and 28.

**[0043]** The sequence of the cell wall binding domain (e.g. of the Gram-negative modular endolysin) and the sequence of the Gram-negative globular endolysin may be linked directly to each other or via an intermediate linker sequence, the linker sequence preferably not exceeding more than 50 amino acids in length, more preferably not more than 30 amino acids in length, even more preferably not more than 20 amino acids in length. Preferably, the cell wall binding domain is situated N-terminal of the sequence of the globular endolysin, i.e. of the domain with the enzymatic activity.

**[0044]** Non-limiting examples of polypeptides according to the present invention are provided in SEQ ID Nos. 29, 30, 31, 32, 33 and 34. If such sequence is to be combined with a further amino acid sequence stretch as defined further below, and said further amino acid sequence stretch is positioned N-terminal of the unit formed by the cell wall binding domain and globular endolysin sequences, said sequences are preferably used without methionine start codon (see SEQ ID Nos. 35, 36, 37, 38, 39 and 40)

**[0045]** A polypeptide according to the present invention exhibits preferably the activity of a peptidoglycan degrading enzyme, i.e. is capable of degrading bacterial peptidoglycan. Typically a polypeptide of the present invention will be capable of degrading the peptidoglycan of at least one type of Gram-negative bacteria, such as *K. pneumoniae*, *E. coli* or *P. aeruginosa*. The peptidoglycan degrading activity on gram negative bacteria can be measured by assays well known in the art, e.g. by muralytic assays in which the outer membrane of gram negative bacteria is permeabilized or removed (e.g. with chloroform) to allow the putative enzyme access to the peptidoglycan layer. If the enzyme is active, degradation of the peptidoglycan layer will lead to a drop of turbidity, which can be measured photometrically (see for example *Briers et al., J. Biochem. Biophys Methods* 70: 531-533, (2007)).

**[0046]** A polypeptide according to the present invention may comprise additionally at least one further amino acid sequence stretch selected from the group consisting of amphipathic peptide, cationic peptide, polycationic peptide, hydrophobic peptide, or naturally occurring antimicrobial peptide, like sushi peptide and defensin. This additional at least one amino acid sequence stretch may in principle be present at any position in the inventive polypeptide, but is preferably present at the termini, i.e. in the N- or C-terminal region of the

inventive polypeptide. Thus, this additional amino acid sequence stretch is preferably not positioned between the sequence of the cell wall binding domain (e.g. of the Gram-negative modular endolysin) and the sequence of the Gram-negative globular endolysin. Such additional amino acid sequence stretch may be fused directly, or via a peptide linker, to the rest of the polypeptide. It is understood that if one (or more) such additional amino acid sequence stretches according to the present invention are present in the N-terminal region of the inventive polypeptide, then there may be further additional amino acids on the N-terminus of the additional amino acid sequence stretch. Preferably these comprise the amino acid methionine (Met), or the sequence methionine, glycine and serine (Met-Gly-Ser).

**[0047]** This at least one additional amino acid sequence stretch preferably has the function to lead the inventive polypeptide through the outer membrane of bacteria and may have activity or may have no or only low activity when administered without being fused to the polypeptide of the invention. The function to guide the polypeptide through the outer membrane of Gram-negative bacteria is caused by the outer membrane or LPS disrupting, permeabilising or destabilizing activity of said amino acid sequence stretches.

**[0048]** Such outer membrane or LPS disrupting or permeabilising or destabilizing activity of these amino acid sequence stretches may be preferably determined in a method as follows: Exponentially growing Gram-negative cells are incubated at room temperature with protein (candidate polypeptide of the present invention exhibiting at least one additional amino acid sequence stretch) at a final concentration of 100 µg/ml in buffer (20 mM NaH<sub>2</sub>PO<sub>4</sub>-NaOH pH7.4; 0.5 M NaCl; 0.5 M imidazole) and a cell density of ~10<sup>6</sup>/ml. After 1 hour the cell suspension is diluted in PBS buffer (10<sup>-5</sup>, 10<sup>-4</sup> and 10<sup>-3</sup>), plated (standard LB-medium) and incubated overnight at 37°C. Additionally, negative controls containing cells in PBS buffer or cells incubated with the matching polypeptide without the additional amino acid sequence stretch) are plated. The residual colonies are counted after the overnight incubation for each plate. If the protein exhibits such outer membrane or LPS disrupting or permeabilising or destabilizing activity, the bacteria cells are lysed due to the treatment with the polypeptide and thus, the number of the bacteria colonies on the agar plate is reduced. Thus, the reduction in the number of bacteria colonies after treatment with the protein is indicative for an outer membrane or LPS disrupting or permeabilising or destabilizing activity of the polypeptide.

**[0049]** Especially preferred are cationic and/or polycationic amino acid sequence stretches comprising at least one motive according to SEQ ID NO:41 (KRKKRK). In particular cationic amino acid sequence stretches comprising at least 2, 3, 4, 5, 6, 7, 8, 9, 10,

11, 12, 13, 14, 15, 16 or 17 motives according to SEQ ID NO: 41 (KRKKRK) are preferred. More preferred are cationic peptide stretches comprising at least one KRK motive (lys-arg-lys), preferable at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32 or 33 KRK motives.

**[0050]** In another preferred embodiment of the present invention the cationic amino acid sequence stretch comprises beside the positively charged amino acid residues, in particular lysine and/or arginine residues, neutrally charged amino acid residues, in particular glycine and/or serine residues. Preferred are cationic amino acid sequence stretches consisting of about 70 % to about 100 %, or about 80 % to about 95 %, or about 85 % to about 90 % positively charged amino acid residues, in particular lysine, arginine and/or histidine residues, more preferably lysine and/or arginine residues and of about 0 % to about 30 %, or about 5 % to about 20 %, or about 10 % to about 20 % neutrally charged amino acid residues, in particular glycine and/or serine residues. Preferred are amino acid sequence stretches consisting of about 4 % to about 8 % serine residues, of about 33 % to about 36 % arginine residues and of about 56 % to about 63 % lysine residues. Especially preferred are amino acid sequence stretches comprising at least one motive according to SEQ ID NO: 42 (KRXKR), wherein X is any other amino acid than lysine, arginine and histidine. Especially preferred are polypeptide stretches comprising at least one motive according to SEQ ID NO: 43 (KRSKR). More preferred are cationic stretches comprising at least about 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or at least about 20 motives according to SEQ ID NO: 42 (KRXKR) or SEQ ID NO: 43 (KRSKR).

**[0051]** Also preferred are amino acid sequence stretches consisting of about 9 to about 16 % glycine residues, of about 4 to about 11 % serine residues, of about 26 to about 32 % arginine residues and of about 47 to about 55 % lysine residues. Especially preferred are amino acid sequence stretches comprising at least one motive according to SEQ ID NO: 44 (KRGSG). More preferred are cationic stretches comprising at least about 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 or at least about 20 motives according to SEQ ID NO: 44 (KRGSG).

**[0052]** In another preferred embodiment of the present invention such cationic amino acid sequence stretch comprises beside the positively charged amino acid residues, in particular lysine and/or arginine residues, hydrophobic amino acid residues, in particular valine, isoleucine, leucine, methionine, phenylalanine, tryptophan, cysteine, alanine, tyrosine, proline and glycine residues, more preferably alanine, valine, leucine, isoleucine, phenylalanine, and/or tryptophan residues. Preferred are cationic amino acid sequence

stretches consisting of about 70 % to about 100 %, or about 80 % to about 95 %, or about 85 % to about 90 % positively charged amino acid residues, in particular lysine and/or arginine residues and of about 0 % to about 30 %, or about 5 % to about 20 %, or about 10 % to about 20 % hydrophobic amino acid residues, valine, isoleucine, leucine, methionine, phenylalanine, tryptophan, cysteine, alanine, tyrosine, proline and glycine residues, more preferably alanine, valine, leucine, isoleucine, phenylalanine, and/or tryptophan residues. Examples for cationic and polycationic amino acid sequence stretches are listed in the following table:

[0053] Table 2:

| amino acid sequence stretch              | length | SEQ ID NO: |
|------------------------------------------|--------|------------|
| KRKKRK                                   | 6      | 41         |
| KRKKRKKRK                                | 9      | 45         |
| RRRRRRRRR                                | 9      | 46         |
| KKKKKKKK                                 | 8      | 47         |
| KRKKRKKRKK                               | 10     | 48         |
| KRKKRKKRKKRK                             | 12     | 49         |
| KRKKRKKRKKRKKR                           | 14     | 50         |
| KKKKKKKKKKKKKKKK                         | 16     | 51         |
| KRKKRKKRKKRKKRKKRK                       | 18     | 52         |
| KRKKRKKRKKRKKRKKRKK                      | 19     | 53         |
| RRRRRRRRRRRRRRRRRR                       | 19     | 54         |
| KKKKKKKKKKKKKKKKKK                       | 19     | 55         |
| KRKKRKKRKRKRKKRKKRK                      | 20     | 56         |
| KRKKRKKRKRKRKKRKKRKK                     | 21     | 57         |
| KRKKRKKRKKRKKRKKRKKRK                    | 21     | 58         |
| KRKKRKKRKRKGSGKRKKRKKRK                  | 22     | 59         |
| KRKKRKKRKRKGSGKRKKRKKRK                  | 24     | 60         |
| KRKKRKKRKKRKKRKKRKKRKK                   | 25     | 61         |
| KRKKRKKRKRKRKKRKKRKRKRKRKKRKK            | 31     | 62         |
| KRKKRKKRKRKGSGKRKKRKKRKGSGKRKKRKKRK      | 38     | 63         |
| KRKKRKKRKKRKKRKKRKKRKKRKKRKKRKKRKKRKK    | 39     | 64         |
| KRKKRKKRKRKRKKRKKRKRKRKRKKRKRKRKRKRKKRKK | 42     | 65         |

[0054] In a further aspect of the present invention at least one of the additional amino acid sequence stretches is an antimicrobial peptide, which comprises a positive net charge and around 50 % hydrophobic amino acids. The antimicrobial peptides are amphipathic with a length of about 12 to about 50 amino acid residues. The antimicrobial peptides are naturally occurring in insects, fish, plants, arachnids, vertebrates or mammals. Preferably the antimicrobial peptide may be naturally occurring in radish, silk moth, wolf spider, frog,

preferably in *Xenopus laevis*, *Rana* frogs, more preferably in *Rana catesbeiana*, toad, preferably Asian toad *Bufo bufo gargarizans*, fly, preferably in *Drosophila*, more preferably in *Drosophila melanogaster*, in *Aedes aegypti*, in honey bee, bumblebee, preferably in *Bombus pascuorum*, flesh fly, preferably in *Sarcophaga peregrine*, scorpion, horseshoe crab, catfish, preferably in *Parasilurus asotus*, cow, pig, sheep, porcine, bovine, monkey and human.

**[0055]** In another preferred embodiment of the present invention the antimicrobial amino acid sequence stretches consist of about 0 % to about 5 %, or about 0 % to about 35 %, or about 10 % to about 35 % or about 15 % to about 45 %, or about 20 % to about 45 % positively charged amino acid residues, in particular lysine and/or arginine residues and of about 50 % to about 80 %, or about 60 % to about 80 %, or about 55 % to about 75 %, or about 70 % to about 90 % hydrophobic amino acid residues, valine, isoleucine, leucine, methionine, phenylalanine, tryptophan, cysteine, alanine, tyrosine, proline and glycine residues, more preferably alanine, valine, leucine, isoleucine, phenylalanine, and/or tryptophan residues.

**[0056]** In another preferred embodiment of the present invention the antimicrobial amino acid sequence stretches consist of about 4 % to about 58 % positively charged amino acid residues, in particular lysine and/or arginine residues and of about 33 % to about 89 % hydrophobic amino acid residues, valine, isoleucine, leucine, methionine, phenylalanine, tryptophan, cysteine, alanine, tyrosine, proline and glycine residues, more preferably alanine, valine, leucine, isoleucine, phenylalanine, and/or tryptophan residues.

**[0057]** Examples for antimicrobial amino acid sequences which may be used in carrying out the present invention are listed in the following table.

**[0058]** Table 3:

| Peptide                                  | Sequence                                | SEQ ID NO |
|------------------------------------------|-----------------------------------------|-----------|
| LL-37                                    | LLGDFFRKSKEKIGKEFKRIVQRIKDFLRNLPRTES    | 66        |
| SMAP-29                                  | RGLRRLGRKIAHG VKKYGPTVLR IIRIAG         | 67        |
| Indolicidin                              | ILPWKWPWWPWR                            | 68        |
| Protegrin                                | RGGRLCYCRRRFCVCVGR                      | 69        |
| Cecropin P1                              | SWLSKTAKKLENSAKKR ISEGI AIAIQGGPR       | 70        |
| Magainin                                 | GIGKFLHSAKKFGKAFVGEIMNS                 | 71        |
| Pleurocidin                              | GWGSFFKKAHVGHVGAALHTYL                  | 72        |
| Cecropin A<br>( <i>A. aegypti</i> )      | GGLKKLGKKLEGAGKRVFNAAEKALP VVAGAKALRK   | 73        |
| Cecropin A<br>( <i>D. melanogaster</i> ) | GWLKKIGKKIERVGHTRDATIQGLGIPQQAANVAATARG | 74        |

| Peptide              | Sequence                                              | SEQ ID NO |
|----------------------|-------------------------------------------------------|-----------|
| Buforin II           | TRSSRAGLQFPVGRVHRLLRK                                 | 75        |
| Sarcotoxin IA        | GWLKKIGKKIERVGQHTRDATIQGLGIAQQAANV<br>AATAR           | 76        |
| Apidaecin            | ANRPVYIPPPRPPHRL                                      | 77        |
| Ascaphine 5          | GIKDWIKGAACKLIKTVASHIANQ                              | 78        |
| Nigrocine 2          | GLLSKVLGVGKKVLCGVSGLVC                                | 79        |
| Pseudin 1            | GLNTLKKVFQGLHEAIKLINNHVQ                              | 80        |
| Ranalexin            | FLGGLIVPAMICAVTKKC                                    | 81        |
| Melittin             | GIGAVLKVLTTGLPALISWIKRKRQQ                            | 82        |
| Lycotoxin 1          | IWLTALKFLGKHAACKLAKQQLSKL                             | 83        |
| Parasin 1            | KGRGKQGGKVRAKAKTRSS                                   | 84        |
| Buforin I            | AGRKGQGGKVRAKAKTRSSRAGLQFPVGRVHRLLRK<br>RKNY          | 85        |
| Dermaseptin 1        | ALWKTMLKKLGTMALHAGKAALGAAADTISQGTQ                    | 86        |
| Bactenecin 1         | RLCRIVVIRVCR                                          | 87        |
| Thanatin             | GSKKPVP I IYCNRRTGKQCQM                               | 88        |
| Brevinin 1T          | VNPIILGVLPKVCLITKKC                                   | 89        |
| Ranateurin 1         | SMLSVLKNLKGVLGFGVACKINIKQC                            | 90        |
| Esculentin 1         | GIFSKLGRKKIKNLLISGLKNVGKEVGMDVVRTG<br>IKIAGCKIKGEC    | 91        |
| Tachyplesin          | RWCFRVCYRGICYRKCR                                     | 92        |
| Androctonin          | RSVCRQIKICRRRGGCYYKCTNRPY                             | 93        |
| alpha-defensin       | DCYCRIPACIAGERRYGTCTIYQGRWAFCC                        | 94        |
| beta-defensin        | NPVSCVRNKGICVPIRCPGSMKQIGTCVGRAVKC<br>CRKK            | 95        |
| theta-defensin       | GFCRCLCRRGVCRCICTR                                    | 96        |
| defensin (sapecin A) | ATCDLLSGTGINHSACAAHCLLRGNRGGYCNGKA<br>VCVCRN          | 97        |
| Thionin (crambin)    | TTCCPSIVARSNFNVCRIPGTPEAICATYTGCI I<br>IPGATCPGDYAN   | 98        |
| defensin from radish | QKLCQRPSTWSTGVCNNACKNQICIRLEKARHG<br>SCNYVFPAHCICYFPC | 99        |
| Drosomycin           | DCLSGRYKGPCAVWDNETCRRVCKEEGRSSGHCS<br>PSLKCWCEGC      | 100       |
| Hepcidin             | DTHFPICIFCCGCCRSKCGMCCKT                              | 101       |
| Bac 5                | RFRPP IRRPP IRRPPFYPPFRPP IRRPP IRRPP<br>FRPPLGRFPF   | 102       |
| PR-39                | RRRPRPPYLPRPRPPFFFPRLPPRIPPGFPPRF<br>PPRFP            | 103       |
| Pyrrhocoricin        | VDKGSYLPRPTPPRPIYNRN                                  | 104       |
| Histatin 5           | DSHAKRHHGYKRKFHEKHSHRGY                               | 105       |
| ECP19                | RPPQFTRAQWFAIQHISLN                                   | 106       |
| MSI-594              | GIGKFLKAKKGIGAVLKVLTTG                                | 107       |
| TL-ColM              | METLTVHAPSPSTNLPSYNGAFSLSAPHVPGAG<br>P                | 108       |

| Peptide | Sequence           | SEQ ID NO |
|---------|--------------------|-----------|
| SBO     | KLKKIAQKIKNFFAKLVA | 109       |

**[0059]** In a further aspect of the present invention at least one of the additional amino acid sequence stretches may be a sushi peptide which is described by Ding JL, Li P, Ho B Cell Mol Life Sci. 2008 Apr;65(7-8):1202-19. The Sushi peptides: structural characterization and mode of action against Gram-negative bacteria. Especially preferred is the sushi 1 peptide according to SEQ ID NO: 110.

**[0060]** Preferred sushi peptides are sushi peptides S1 and S3 and multiples thereof; FASEB J. 2000 Sep;14(12):1801-13.

**[0061]** In a further aspect of the present invention at least one of the additional amino acid sequence stretches is a hydrophobic peptide, which comprises at least 90 % of the hydrophobic amino acid residues of valine, isoleucine, leucine, methionine, phenylalanine, tryptophan, cysteine, alanine, tyrosine, proline and/or glycine. In another preferred embodiment the hydrophobic peptide fused to the protein of the invention consists of about 90 % to about 95 %, or of about 90 % to about 100%, or of about 95 % to about 100 % of the hydrophobic amino acid residues of valine, isoleucine, leucine, methionine, phenylalanine, tryptophan, cysteine, alanine, tyrosine, proline and/or glycine.

**[0062]** Preferred hydrophobic peptides are Walmagh1 having the amino acid sequence according to SEQ ID NO: 111 and the hydrophobic peptide having the amino acid sequence Phe-Phe-Val-Ala-Pro (SEQ ID NO: 112).

**[0063]** In a further aspect of the present invention at least one of the additional amino acid sequence stretches is an amphipathic peptide, which comprises one or more of the positively charged amino acid residues of lysine, arginine and/or histidine, combined to one or more of the hydrophobic amino acid residues of valine, isoleucine, leucine, methionine, phenylalanine, tryptophan, cysteine, alanine, tyrosine, proline and/or glycine. Side chains of the amino acid residues are oriented in order that cationic and hydrophobic surfaces are clustered at opposite sides of the peptide. Preferably, more than about 30, 40, 50, 60 or 70 % of the amino acids in said peptide are positively charged amino acids. Preferably, more than about 30, 40, 50, 60 or 70 %, of the amino acid residues in said peptide are hydrophobic amino acid residues. Advantageously, the amphipathic peptide is present at the N-terminal or the C-terminal end of the polypeptide according to the present invention.

**[0064]** In another embodiment of the invention, the amphipathic peptide consists of at least 5, more preferably at least of 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47,

48, 49 or at least 50 amino acid residues. In a preferred embodiment at least about 30, 40, 50, 60 or 70 % of said amino acid residues of the amphipathic peptide are either arginine or lysine residues and/or at least about 30, 40, 50, 60 or 70 % of said amino acid residues of the amphipathic peptide are of the hydrophobic amino acids valine, isoleucine, leucine, methionine, phenylalanine, tryptophan, cysteine, alanine, tyrosine, proline and/or glycine.

**[0065]** In another preferred embodiment of the present invention the amphipathic peptide stretch comprises beside the positively charged amino acid residues, in particular lysine and/or arginine residues, hydrophobic amino acid residues, in particular valine, isoleucine, leucine, methionine, phenylalanine, tryptophan, cysteine, alanine, tyrosine, proline and glycine residues, more preferably alanine, valine, leucine, isoleucine, phenylalanine, and/or tryptophan residues. Preferred are amphipathic peptide stretches consisting of about 10 % to about 50 %, or about 20 % to about 50 %, or about 30 % to about 45 % or about 5 % to about 30 % positively charged amino acid residues, in particular lysine and/or arginine residues and of about 50 % to about 85 %, or about 50 % to about 90 %, or about 55 % to about 90 %, or about 60 % to about 90 %, or about 65 % to about 90 % hydrophobic amino acid residues, valine, isoleucine, leucine, methionine, phenylalanine, tryptophan, cysteine, alanine, tyrosine, proline and glycine residues, more preferably alanine, valine, leucine, isoleucine, phenylalanine, and/or tryptophan residues. In another preferred embodiment amphipathic peptide stretches consisting of 12 % to about 50 % positively charged amino acid residues, in particular lysine and/or arginine residues and of about 50 % to about 85 % hydrophobic amino acid residues, valine, isoleucine, leucine, methionine, phenylalanine, tryptophan, cysteine, alanine, tyrosine, proline and glycine residues, more preferably alanine, valine, leucine, isoleucine, phenylalanine, and/or tryptophan residues.

**[0066]** Preferred amphipathic peptides are  $\alpha$ 4-helix of T4 lysozyme according to SEQ ID NO: 113 and WLBU2-Variant having the amino acid sequence according to SEQ ID NO: 114 and Walmagh 2 according to SEQ ID NO: 115.

**[0067]** The optional additional amino acid sequence stretches as specified above consist preferably of at least 5, more preferably at least of 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 or at least 100 amino acid residues. Especially preferred are those additional amino acid sequence stretches consisting of about 5 to about 100 amino acid residues, about 5 to about 50 or about 5 to about 30 amino acid residues. More preferred are

peptide stretches consisting of about 6 to about 42 amino acid residues, about 6 to about 39 amino acid residues, about 6 to about 38 amino acid residues, about 6 to about 31 amino acid residues, about 6 to about 25 amino acid residues, about 6 to about 24 amino acid residues, about 6 to about 22 amino acid residues, about 6 to about 21 amino acid residues, about 6 to about 20 amino acid residues, about 6 to about 19 amino acid residues, about 6 to about 16 amino acid residues, about 6 to about 14 amino acid residues, about 6 to about 12 amino acid residues, about 6 to about 10 amino acid residues or about 6 to about 9 amino acid residues.

**[0068]** In a preferred embodiment the inventive polypeptide comprises at least one amino acid sequence stretch selected from the group consisting of KRK and SEQ ID NOs: 41 - 115.

**[0069]** An non-limiting example for a polypeptide according to the present invention comprising an additional amino acid sequence stretch is SEQ ID NO: 116, which comprises an N-terminal SMAP-29 peptide (SEQ ID NO: 67), a CBD deriving from KZ144 endolysin exhibiting 4 mutations (SEQ ID NO: 8) and a Lys68 endolysin sequence with a P78S mutation (SEQ ID NO: 26).

**[0070]** The additional amino acid sequence stretch of the polypeptide according to the present invention may be linked to the rest of the enzyme by intervening additional amino acid residues e.g. due to cloning reasons. Alternatively, the additional amino acid sequence stretches may be directly linked to the rest of the enzyme sequence without intervening linker sequences. The additional amino acid sequences, if more than one present in the inventive polypeptide and positioned on the same terminus of the enzyme, may likewise be linked to each other by additional intervening amino acid residues or may be directly joined to each other.

**[0071]** Preferably, said intervening additional amino acid residues may not be recognized and/or cleaved by proteases. Preferably said additional amino acid sequences are linked to each other and/or to the enzyme by at least 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 additional intervening amino acid residues.

**[0072]** In a preferred embodiment the at least one additional amino acid sequence stretch is linked to the rest of the inventive polypeptide, preferably at the N- or C-terminus of the polypeptide according to the present invention, by the additional intervening amino acid residues glycine, serine and serine (Gly-Ser-Ser), glycine, alanine, glycine and alanine (Gly-Ala-Gly-Ala; SEQ ID NO:117), glycine, alanine, glycine, alanine, glycine, alanine, glycine and alanine (Gly-Ala-Gly-Ala-Gly-Ala-Gly-Ala; SEQ ID NO:118) or glycine, alanine,

glycine, alanine, glycine, alanine, glycine, alanine, glycine, alanine, glycine and alanine (Gly-Ala-Gly-Ala-Gly-Ala-Gly-Ala-Gly-Ala; SEQ ID NO:119).

**[0073]** Aside of the enzymatic domain (i.e. a domain having the activity of degrading the peptidoglycan of Gram-negative bacteria), the cell wall binding domain, and the optional additional amino acid sequence stretches, as defined herein, the inventive polypeptide may of course also comprise other amino acid sequence elements, e.g. one or more tags, e.g. a His-tag, Strep-tag, Avi-tag, Myc-tag, Gst-tag, JS-tag, cystein-tag, FLAG-tag or other tags known in the art, thioredoxin, maltose binding proteins (MBP) etc.

**[0074]** In this context, the inventive polypeptide, preferably having the ability of degrading the peptidoglycan layer of Gram negative bacteria, may additional comprise a tag e.g. for purification. Preferred is a His<sub>6</sub>-tag (SEQ ID NO: 120), preferably at the C-terminus and/or the N-terminus of the polypeptide according to the present invention. Said tag can be linked to the polypeptide by additional amino acid residues e.g. due to cloning reasons. Preferably said tag can be linked to the protein by at least 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 additional amino acid residues. Preferably said additional amino acid residues may not be recognized and/or cleaved by proteases. In a preferred embodiment the inventive polypeptide comprises a His<sub>6</sub>-tag at its C-terminus linked to the polypeptide by the additional amino acid residues lysine and glycine (Lys-Gly) or leucine and glutamic acid (Leu-Glu). Preferably, said additional amino acid residues may be not recognized or cleaved by proteases. In another preferred embodiment the inventive polypeptide comprises a His<sub>6</sub>-tag at its N-terminus linked to the polypeptide by the additional amino acid residues lysine and glycine (Lys-Gly) or leucine and glutamic acid (Leu-Glu). In another preferred embodiment the polypeptide comprises a His<sub>6</sub>-tag at its N- and C-terminus linked to the polypeptide by the additional amino acid residues lysine and glycine (Lys-Gly) or leucine and glutamic acid (Leu-Glu).

**[0075]** A polypeptide according to the present invention can be produced by standard means known in the art, e.g. by recombinant expression of nucleic acids encoding the respective polypeptide in appropriate host cells. If the inventive polypeptide comprises for example additionally amino acid sequence stretches or tags etc., such fusion proteins may be produced by linking the required individual nucleic acid sequences using standard cloning techniques as described e.g. by Sambrook et al. 2001, Molecular Cloning: A Laboratory Manual. Such a polypeptide may be produced likewise with methods known in the art, e.g., in recombinant DNA expression systems.

### **III. Nucleic acids, vectors, bacteriophages and host cells**

[0076] The present invention does also relate to nucleic acids encoding one or more inventive polypeptides of the present invention. The inventive nucleic acid may take all forms conceivable for a nucleic acid. In particular the nucleic acids according to the present invention may be RNA, DNA or hybrids thereof. They may be single-stranded or double-stranded. They may have the size of small transcripts or of entire genomes, such as a bacteriophage genome. As used herein, a nucleic acid encoding one or more inventive polypeptides of the present invention may be a nucleic acid reflecting the sense strand. Likewise, the antisense strand is also encompassed. The nucleic acid may encompass a heterologous promotor for expression of the inventive polypeptide.

[0077] In a further aspect the present invention relates to a vector comprising a nucleic acid according to the present invention. Such vector may for example be an expression vector allowing for expression of an inventive polypeptide. Said expression may be constitutive or inducible. The vector may also be a cloning vector comprising the nucleic acid sequence of the current invention for cloning purposes.

[0078] The present invention does also relate to a bacteriophage comprising an inventive nucleic acid, in particular comprising an inventive nucleic acid encoding a polypeptide according to the present invention.

[0079] The present invention does also relate to (isolated) host cells comprising a polypeptide, nucleic acid, vector, or bacteriophage according to the present invention. The host cells may be selected in particular from the group consisting of bacterial cells and yeast cells. Where appropriate, other suitable host cells may be immortalized cell lines, e.g. of mammalian (in particular human) origin. Particularly preferred host cells comprise a polypeptide according to the present invention.

### **IV. Compositions**

[0080] In a further aspect the present invention relates to a composition comprising a polypeptide according to the present invention, a nucleic acid according to the present invention, a vector according to the present invention, a bacteriophage according to the present invention and/or a host cell according to the present invention.

[0081] A composition according to the present invention may be a pharmaceutical composition comprising a pharmaceutical acceptable diluent, excipient or carrier.

[0082] In an even further aspect the composition according to the present invention is a cosmetic composition. Several bacterial species can cause irritations on environmentally exposed surfaces of the patient's body such as the skin. In order to prevent such irritations or in order to eliminate minor manifestations of said bacterial pathogens, special cosmetic preparations may be employed, which comprise sufficient amounts of the inventive polypeptide, nucleic acid, vector, host cell and/or composition in order to achieve a comedolytic effect.

## V. Uses

[0083] In a further aspect the present invention relates to a polypeptide according to the present invention, a nucleic acid according to the present invention, a vector according to the present invention, a bacteriophage according to the present invention, a host cell according to the present invention, and/or a composition according to the present invention for use in a method of treatment of the human or animal body by surgery or therapy or in diagnostic methods practiced on the human or animal body. In such scenarios the antibacterial activity of polypeptide of the present invention can be exploited, in particular if the polypeptide comprises the additional amino acid sequence stretch specified above.

[0084] Such method typically comprises administering to a subject an effective amount of an inventive polypeptide, nucleic acid, vector, bacteriophage, host cell or a composition. The subject may for example be a human or an animal, with human subjects being more preferred. In particular, the inventive polypeptide, the inventive nucleic acid, the inventive vector, the inventive bacteriophage, the inventive host cell, and/or the inventive composition may be used in methods for the treatment or prevention of bacterial infections, such Gram-negative bacterial infections.

[0085] The dosage and route of administration used in a method of treatment (or prophylaxis) according to the present invention depends on the specific disease/site of infection to be treated. The route of administration may be for example oral, topical, nasopharyngeal, parenteral, intravenous, rectal or any other route of administration.

[0086] For application of an inventive polypeptide, nucleic acid, vector, bacteriophage, host cell or composition to a site of infection (or site endangered to be infected) a formulation may be used that protects the active compounds from environmental influences such as proteases, oxidation, immune response etc., until it reaches the site of infection. Therefore, the formulation may be capsule, dragee, pill, suppository, injectable solution or any other medical reasonable galenic formulation. Preferably, the galenic

formulation may comprise suitable carriers, stabilizers, flavourings, buffers or other suitable reagents. For example, for topical application the formulation may be a lotion or plaster, for nasopharyngeal application the formulation may be saline solution to be applied via a spray to the nose.

**[0087]** Preferably, an inventive polypeptide, nucleic acid, vector, bacteriophage, host cell or composition is used in combination with other conventional antibacterial agents, such as antibiotics, lantibiotics, bacteriocins or endolysins, etc. The administration of the conventional antibacterial agent can occur prior to, concurrent with or subsequent to administration of the inventive polypeptide, nucleic acid, vector, bacteriophage, host cell or composition.

**[0088]** In a further aspect the present invention relates to the inventive polypeptide, nucleic acid, vector, bacteriophage, host cell or composition for use as diagnostic means in medical diagnostics, food diagnostics, feed diagnostics, or environmental diagnostics, in particular as a diagnostic means for the diagnostic of bacterial infection, in particular those caused by Gram-negative bacteria. In this respect the inventive polypeptide, nucleic acid, vector, host cell or composition may be used as a tool to specifically degrade the peptidoglycan of pathogenic bacteria, in particular of Gram-negative pathogenic bacteria. The degradation of the bacterial cells by the inventive polypeptide, nucleic acid, vector, host cell or composition can be supported by the addition of detergents like Triton X-100 or other additives which weaken the bacterial cell envelope like polymyxin B. Specific cell degradation is needed as an initial step for subsequent specific detection of bacteria using nucleic acid based methods like PCR, nucleic acid hybridization or NASBA (Nucleic Acid Sequence Based Amplification), immunological methods like IMS, immunofluorescence or ELISA techniques, or other methods relying on the cellular content of the bacterial cells like enzymatic assays using proteins specific for distinct bacterial groups or species (e.g.  $\beta$ -galactosidase for enterobacteria, coagulase for coagulase positive strains).

**[0089]** In a further aspect the present invention relates to the use of the inventive polypeptide, the inventive nucleic acid, the inventive vector, the inventive bacteriophage, the inventive host cell, and/or the inventive composition, as an antimicrobial in food, feed, or cosmetics, or use as disinfecting agent. They can be used in particular for the treatment or prevention of Gram-negative bacterial contamination of foodstuff, of food processing equipment, of food processing plants, of (inanimate) surfaces coming into contact with foodstuff (such as shelves and food deposit areas), of feedstuff, of feed processing equipment, of feed processing plants, of (inanimate) surfaces coming into contact with feedstuff (such as

shelves and feed deposit areas), of medical devices, or of (inanimate) surfaces in hospitals, doctor's offices and other medical facilities.

### **BRIEF DESCRIPTION OF THE FIGURES**

**[0090]** In the following a brief description of the appended figures will be given. The figures are intended to illustrate the present invention in more detail. However, they are not intended to limit the scope of the invention to these specific examples.

**[0091]** Fig. 1: illustrates an archetypical structure of a Gram-negative modular endolysin. Depicted is the 3D structure of KZ 144 endolysin (Fokine et al. J. Biol. Chem. 2008, 283:7242-7250). The cell wall binding domain (CBD) is a domain (see top of Fig) clearly separate from the enzymatically active domain (bottom of structure).

**[0092]** Fig. 2: illustrates the structure of a Gram-negative globular endolysin. Depicted is the modelled 3D structure of Lys394 endolysin. The structure has been generated by way of homology modeling on basis of the structure of T5 lysozyme, which shares a 97% sequence identity with Lys394. A domain architecture, as can be seen for, e.g. KZ 144 endolysin, can not be found in the structure of Lys394 endolysin.

**[0093]** Fig. 3: illustrates muralytic activity for Salmonella phage endolysin Lys68 and two chimeric variants thereof (cEL188-Lys68; cKZ144-Lys68) on chloroform treated *P. aeruginosa* cells. The fusion proteins exhibiting the additional heterologous CBD show increased activity.

**[0094]** Fig. 4: illustrates muralytic activity for ABgp46 endolysin and two chimeric variants thereof (cKZ144-ABgp46; cICP-ABgp46) on chloroform treated *P. aeruginosa* cells. The fusion proteins exhibiting the additional heterologous CBD show increased activity.

**[0095]** Fig. 5: illustrates muralytic activity for Lys394 endolysin and a chimeric variant thereof (ICP-S394) on chloroform treated *P. aeruginosa* cells. The fusion protein exhibiting the additional heterologous CBD shows increased activity.

## **VI. Examples**

**[0096]** In the following, specific examples illustrating various embodiments and aspects of the invention are presented. However, the present invention shall not be limited in scope by the specific embodiments described herein. Indeed, various modifications of the invention in addition to those described herein will become readily apparent to those skilled in

the art from the foregoing description, accompanying figures and the examples below. All such modifications fall within the scope of the appended claims.

Example 1: Variants of Lys68 endolysin having an additional Gram-negative CBD exhibit increased muralytic activity

**[0097]** Two variants of Salmonella phage endolysin Lys68 were generated. The first variant is a fusion with a sequence comprising the CBD of EL188 endolysin (SEQ ID NO:12). The resulting chimeric variant cEL188-Lys68 comprises SEQ ID NO:29. The second variant is a fusion with the CBD of KZ144 endolysin (SEQ ID NO:7). The CBD used exhibits three point mutations in the sequence of the KZ144 CBD in which cysteine residues were replaced by serine residues. The resulting chimeric variant cKZ144-Lys68 comprises SEQ ID NO:30. The wildtype endolysin and its chimeric variants were expressed in *E. coli*. Subsequently, the proteins were purified. In order to test the muralytic activity of the enzymes, *Pseudomonas aeruginosa* PAO1 cells were treated with chloroform to remove the outer membrane. Therefore, 20 mM HEPES pH 7.4, 150 mM NaCl buffer was saturated with chloroform. Exponentially growing *P. aeruginosa* cells were harvested and resuspended in chloroform buffer and incubated for 45 minutes. Afterwards, the cells were washed two times in 20 mM HEPES pH 7.4 and 150 mM NaCl and subsequently diluted with the same buffer to a final OD600 of about 0.8. Subsequently, each protein was added at a final concentration of 0.005  $\mu$ M to an aliquot of 1 ml cell solution and the reduction of the OD600 was recorded over a period of 1800 seconds.

**[0098]** The two chimeric variants of Lys68 endolysin having an additional Gram-negative CBD exhibit significantly increased muralytic activity in comparison to the wildtype endolysin (see Fig.3).

Example 2: Variants of ABgp46 endolysin having an additional Gram-negative CBD exhibit increased muralytic activity

**[0099]** Two variants of ABgp46 endolysin (deriving from Acinetobacter phage vB\_AbaP\_CEB1) were generated. The first variant is a fusion with the CBD of KZ144 endolysin (SEQ ID NO: 7). The resulting chimeric variant cKZ144-ABgp46 comprises SEQ ID NO:31. The second variant is a fusion with the CBD of the baseplate tail protein of Vibrio phage ICP1 (SEQ ID NO:14). The resulting chimeric variant cICP-ABgp46 comprises SEQ ID NO:32. The wildtype endolysin and its chimeric variants were expressed in *E. coli*. Subsequently, the proteins were purified. In order to test the muralytic activity of the

enzymes, *Pseudomonas aeruginosa* PAO1 cells were treated with chloroform to remove the outer membrane. Therefore, 20 mM HEPES pH 7.4, 150 mM NaCl buffer was saturated with chloroform. Exponentially growing *P. aeruginosa* cells were harvested and resuspended in chloroform buffer and incubated for 45 minutes. Afterwards, the cells were washed two times in 20 mM HEPES pH 7.4 and 150 mM NaCl and subsequently diluted with the same buffer to a final OD600 of about 0.8. Subsequently, each protein was added at a final concentration of 0.005  $\mu$ M to an aliquot of 1 ml cell solution and the reduction of the OD600 was recorded over a period of 1800 seconds.

**[0100]** The two chimeric variants of ABgp46 endolysin having an additional Gram-negative CBD exhibit significantly increased muralytic activity in comparison to the wildtype endolysin (see Fig.5).

Example 3: Variant of Lys394 endolysin having an additional CBD exhibits de novo muralytic activity on P- aeruginosa

**[0101]** A variant of Salmonella phage endolysin Lys394 was generated. The variant is a fusion with the CBD of ICP tail/baseplate protein (SEQ ID NO: 14). The resulting chimeric variant comprises SEQ ID NO:33. The wildtype endolysin and its chimeric variant were expressed in *E. coli*. Subsequently, the proteins were purified. In order to test the muralytic activity of the enzymes, *Pseudomonas aeruginosa* PAO1 cells were treated with chloroform to remove the outer membrane. Therefore, 20 mM HEPES pH 7.4, 150 mM NaCl buffer was saturated with chloroform. Exponentially growing *P. aeruginosa* cells were harvested and resuspended in chloroform buffer and incubated for 45 minutes. Afterwards, the cells were washed two times in 20 mM HEPES pH 7.4 and 150 mM NaCl and subsequently diluted with the same buffer to a final OD600 of about 0.8. Subsequently, each protein was added at a final concentration of 0.005  $\mu$ M to an aliquot of 1 ml cell solution and the reduction of the OD600 was recorded over a period of 1800 seconds.

**[0102]** In contrast to the wildtype endolysin, the chimeric variant of Lys394 endolysin exhibits significant muralytic activity on *P. aeruginosa* cells (see Fig.6) already at a concentration of 0.005  $\mu$ M.

Example 4: Fusion protein exhibiting Lys68 endolysin variant, a CBD variant of the KZ144 endolysin CBD and an additional antimicrobial peptide SMAP-29 exhibits antibacterial activity against a broad range of Gram-negative bacteria

**[0103]** In a further experiment compatibility of a chimeric endolysin according to the present invention with an antimicrobial peptide to effectively kill Gram-negative bacteria was tested. For this purpose, a fusion protein was generated which comprises the components SMAP-29 (SEQ ID NO: 67), a derivative of the KZ144 CBD exhibiting four mutations (SEQ ID NO: 8) and a derivative of Lys68 endolysin (SEQ ID NO: 26). The resulting polypeptide comprises the consecutive sequence of SEQ ID NO: 116. Due to the mutations in SEQ ID NO: 8 and 26, the resulting fusion protein exhibited improved thermal stability. The chimeric variant was expressed in *E. coli*.

**[0104]** Briefly, bacteria were grown in (Luria-Bertani) medium and diluted 1:10 in Mueller-Hinton medium. At optical density OD<sub>600</sub> of about 0.6 bacteria were diluted in the same medium 1:10 followed by a 1:500 dilution. Protein buffer (20 mM HEPES, 500 mM NaCl, pH 7.4) and proteins were pipetted into a 96 well plate, using different concentrations of proteins and an end volume of 20 µl including 500 µM EDTA final concentration (if required). 180 µl of bacterial cells or a medium (Mueller-Hinton) control were given to the 96 well plate and mixed. The plate was incubated for 18-22 hours at 37°C and the bacterial growth was determined measuring the OD<sub>600</sub> values of the wells. The MIC which is the protein concentration of the well which showed the same OD<sub>600</sub> value as the no-bacteria control was determined.

**[0105]** The fusion protein showed antibacterial activity against a surprisingly broad range of Gram-negative bacteria including *Pseudomonas*, *Klebsiella*, *Escherichia*, *Acinetobacter* and *Salmonella* species. The corresponding results are described in table 4.

**[0106]** Table 4: MIC values

| Bacterial strain                               | EDTA   | MIC (µg/ml) |
|------------------------------------------------|--------|-------------|
| S2 <i>Pseudomonas aeruginosa</i> Br667         | 0.5 mM | 10          |
| S84 <i>Pseudomonas aeruginosa</i> Aa249        | 0.5 mM | 10          |
| S53 <i>Klebsiella pneumoniae</i> B10-03.05.700 | 0.5 mM | 20          |
| S516 <i>Klebsiella pneumoniae</i> va32842      | 0.5 mM | 17.5        |
| S441 <i>Escherichia coli</i> B12-11.20.0192    | 0.5 mM | 5           |
| S458 <i>Escherichia coli</i> B12-11.30.0188    | 0.5 mM | 5           |
| S45 <i>Acinetobacter baumannii</i> 2671        | 0.0 mM | 10          |
| S138 <i>Acinetobacter baumannii</i> NRZ-00066  | 0.0 mM | 10          |
| S795 <i>Salmonella Enteritidis</i> LGL-238     | 0.5 mM | 7.5         |
| S25 <i>Salmonella</i> Thyphimurium DSM 17058   | 0.5 mM | 15          |

**[0107]** In addition, the fusion protein comprising SEQ ID NO: 116 exhibited significant thermal stability due to the mutations K59M and P78S in the sequence of Lys68 endolysin.