

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
4 June 2009 (04.06.2009)

PCT

(10) International Publication Number
WO 2009/068858 A1

(51) International Patent Classification:
C07K 14/01 (2006.01) *C12N 9/36* (2006.01)

(74) Agent: SMITH, Stephen; Potter Clarkson LLP, Park View House, 58 The Ropewalk, Nottingham NG1 5DD (GB).

(21) International Application Number:
PCT/GB2008/003923

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(22) International Filing Date:
24 November 2008 (24.11.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/996,563 26 November 2007 (26.11.2007) US

(71) Applicant (*for all designated States except US*): PLANT BIOSCIENCE LIMITED [GB/GB]; Norwich Research Park, Colney Lane, Norwich, Norfolk NR4 7UH (GB).

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(72) Inventors; and

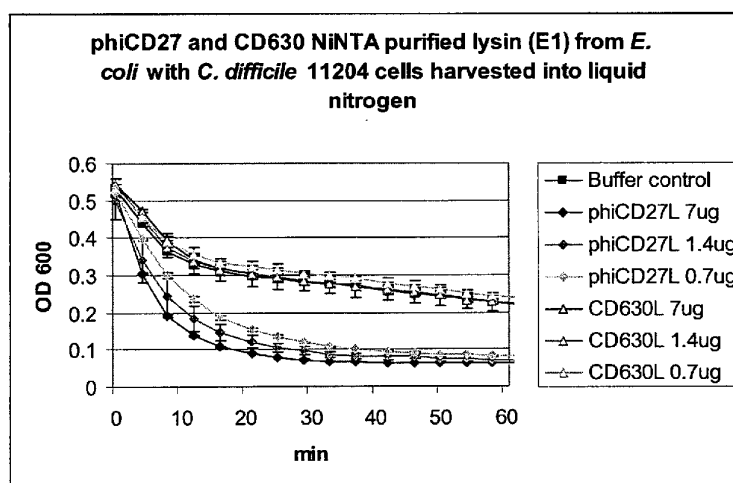
(75) Inventors/Applicants (*for US only*): GASSON, Michael [GB/GB]; Moor End Cottage, Moor End, Stibbard, Norfolk NR21 0EJ (GB). MAYER, Melinda [GB/GB]; The Laurels, Church Road, Wrenningham, Norfolk NR16 1BA (GB). NARBAD, Arjan [GB/GB]; 61 The Avenues, Norwich, Norfolk NR2 3QR (GB).

Published:

— with international search report

(54) Title: NOVEL POLYPEPTIDES HAVING ENDOLYSIN ACTIVITY AND USES THEREOF

Figure 8



(57) Abstract: The present invention provides isolated polypeptides comprising the amino acid sequence of SEQ ID NO:1, or a fragment, variant, derivative or fusion thereof which is capable of binding specifically to and/or lysing cells of *Clostridium difficile*, and means for producing the same, with the proviso that the fragment, variant, derivative or fusion is not a naturally occurring lysin of a bacteriophage of *Clostridium difficile*. The invention further provides methods for killing bacterial cells, such as cells of *Clostridium difficile*, and for diagnosing, treating and preventing diseases and conditions associated with infection of the same. The invention also provides diagnostic kits for use in such methods.

NOVEL POLYPEPTIDES HAVING ENDOLYSIN ACTIVITY AND USES THEREOF

Field of Invention

The present invention relates to novel polypeptides derived from endolysins from a bacteriophage of *Clostridium difficile* and nucleic acid molecules encoding the same, as well as compositions thereof. The invention also provides uses of such polypeptides and nucleic acid molecules in the diagnosis and treatment of conditions and diseases associated with microbial cells such as *Clostridium difficile*. In particular, the invention provides a polypeptide having endolysin activity derived from bacteriophage Φ CD27 of *Clostridium difficile* and uses thereof.

Introduction

The growing problems associated with *Clostridium difficile* are well documented, in particular its role in nosocomial infections often associated with antibiotic use (1). *C. difficile* is an anaerobic Gram positive bacterium that has the capacity to form spores that resist heating, drying and disinfectants. There is some evidence that exposure to non-chlorine based cleaning agents actually increases sporulation. These characteristics contribute the organism's capacity to persist in the hospital environment, thereby maintaining a reservoir of pathogens with the potential to infect patients. *C. difficile*-associated disease (CDAD) is a growing problem both in the UK and worldwide, with both rates and severity increasing. In England and Wales, deaths associated with *C. difficile* infection rose from 975 in 1999 to 2,247 in 2004. CDAD notifications rose from 1000 in 1999 to 15,000 in 2000 and 35,500 in 2003 (2). It should be noted that, in addition to threats to human health mentioned above, *C. difficile* is also a significant cause of morbidity and mortality in animals, particularly in farm animals such as calves and sheep. Accordingly, disclosure herein as to methods for addressing this problem in humans should likewise be read to apply to veterinary targets as well.

A particularly serious development is the emergence of a highly virulent strain of *C. difficile*, initially in Canada and the USA, but now significant in the UK and several other European countries. This new strain, defined as *C. difficile* ribotype 027, was detected in the UK in 2003 in an outbreak involving 174 cases and 19 deaths. By April 2006 there have been 450 separate UK isolates of *C. difficile* ribotype 027 from 75 hospitals (1).

C. difficile is widely distributed in soil and in the intestinal tracts of animals. It can be cultured from the stools of 3% of healthy human adults and 80% of healthy newborns and infants (1). Pathogenic potential is associated with the ability of *C. difficile* to produce potent toxins; the two major characterised toxins are a 308 kDa exotoxin, toxin A (TcdA) and a 270 kDa cytotoxin, toxin B (TcdB), which share 63% homology at the amino acid level (3). Genes encoding these toxins are associated with a pathogenicity island PaLoc (4) and strains vary in their ability to produce these two major toxins. Other virulence factors are likely to be involved, and a separate binary toxin CDT has been defined (5, 6).

The pathogenic potential of virulent *C. difficile* strains is realised when the gastro-intestinal tract (GIT) microflora becomes impaired or unbalanced, and this is a common consequence of antibiotic therapy. Thus the hospital environment is an ideal one for *C. difficile* to thrive and cause human disease (1).

CDAD occurs when pathogenic strains of *C. difficile* gain a sufficiently strong position within the GIT microflora and produce toxin(s) that damage the host epithelium. The GIT microflora is an important barrier to pathogenic microbes, representing a complex community of some 500 to 1000 different species that are maintained in a homeostatic equilibrium interacting in beneficial ways with the host. Classical antibiotic therapy is variably non-discriminatory and it can damage the fine balance of the GIT microbial community. The disruption of the normal microflora is a major factor in the manifestation of CDAD, either as consequence of prior antibiotic therapy or another factor.

Hence, there exists a growing need for new treatments and approaches for the control of *C. difficile* without damaging the protective capacity of the complex GIT microflora.

Summary of Invention

A first aspect of the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1, or a fragment, variant, derivative or fusion thereof which is capable of binding specifically to and/or lysing cells of *Clostridium difficile*

The amino acid sequence depicted below is that of the wildtype (*i.e.* naturally occurring) endolysin of bacteriophage Φ CD27 of *Clostridium difficile*.

MKICITVGHSILKSGACTSADGVVNEYQYNKSLAPVLADTFRKEGHKVDVIICPEKQFKT
KNEEKSYKIPRVNSGGYDLLIELHLNASNGQGKGSEVLYYSNKGLEYATRICDKLGTVFK
NRGAKLDKRLYLNSSKPTAVLIESFFCDNKEDYDKAKKLGHEGIAKLIVEGVLNKNINNE
GVKQMYKHTIVYDGEVDKISATVVGWGYNDGKILICDIKDYVPGQTQNLVVGGGACEK
ISSITKEKFIMIKGNDRFDTLYKALDFINR

[SEQ ID NO: 1]

See also NCBI Accession Nos. YP_002290910 and ACH91325.

In one embodiment, the polypeptide is not a naturally occurring lysin of a bacteriophage of *Clostridium difficile* (other than Φ CD27). Thus, the first aspect of the invention provides isolated polypeptides comprising or consisting of the amino acid sequence of SEQ ID NO:1 and non-naturally occurring fragments, variants, derivatives or fusions thereof.

The term 'amino acid' as used herein includes the standard twenty genetically-encoded amino acids and their corresponding stereoisomers in the 'D' form (as compared to the natural 'L' form), omega-amino acids and other naturally-occurring amino acids, unconventional amino acids (*e.g.* α,α -disubstituted amino acids, N-alkyl amino acids, *etc.*) and chemically derivatised amino acids (see below).

Thus, when an amino acid is being specifically enumerated, such as 'alanine' or 'Ala' or 'A', the term refers to both L-alanine and D-alanine unless explicitly stated otherwise. Other unconventional amino acids may also be suitable components for polypeptides of the present invention, as long as the desired functional property is retained by the polypeptide. For the peptides shown, each encoded amino acid residue, where appropriate, is represented by a single letter designation, corresponding to the trivial name of the conventional amino acid.

Preferably, the polypeptide, or fragment, variant, fusion or derivative thereof, comprises or consists of L-amino acids.

By "isolated" we mean that the polypeptide of the invention, specifically the wildtype endolysin of bacteriophage Φ CD27, is provided in a form other than that in which it may be found naturally. Preferably, the polypeptide is provided free from intact bacteriophage.

In one embodiment, the polypeptide of the invention is the naturally occurring endolysin of bacteriophage Φ CD27 [SEQ ID NO: 1], provided in an isolated form.

Other naturally occurring lysins of a bacteriophage of *Clostridium difficile* known in the prior art are not encompassed by the first aspect of the invention. In particular, the following lysins of a bacteriophage of *Clostridium difficile* are explicitly excluded from the scope of the first aspect of the invention:

- (a) the lysin of bacteriophage Φ CD119;
- (b) the lysin of bacteriophage Φ C2; and
- (c) the lysin of prophages 1 and 2 of *Clostridium difficile* strain 630 (CD630).

For example, the following known proteins (defined by reference to their NCBI accession numbers) are explicitly *excluded* from the scope of the first aspect of the invention:

PhiC2 putative endolysin	YP_001110754
CD630 phage endolysin (prophage 1)	YP_001087453
phiCD119 putative lysin	YP_529586
QCD-32g58 hypothetical protein	ZP_01803398
QCD-32g58 hypothetical protein	ZP_01803228

In one embodiment, the polypeptide of the first aspect of the invention comprises the amino acid sequence of SEQ ID NO:1. For example, the polypeptide may consist of the amino acid sequence of SEQ ID NO:1.

However, the first aspect of the invention also extends to fragments, variants, derivatives and fusions of the amino acid sequence of SEQ ID NO:1 which are capable of binding specifically to and/or lysing cells of *Clostridium difficile*.

By "capable of binding specifically to cells of *Clostridium difficile*" we mean that the polypeptide is capable of binding preferentially to cells of *Clostridium difficile*. However, it will be appreciated that such polypeptides may also bind preferentially to one or more additional types of cell. Preferably, the polypeptide binds exclusively to cells of *Clostridium* sp. Such cell binding activity may be determined using methods well known in the art.

By "capable of lysing cells of *Clostridium difficile*" we mean that the polypeptide, or fragment, variant, derivative or fusion, retains (at least in part) the ability of the wildtype endolysin of bacteriophage Φ CD27 to lyse bacterial cells. It will be appreciated that such lytic activity should be cell-specific (e.g. to cells of *Clostridium difficile*) rather than a non-specific cytotoxic activity on all cell types. Such cell lysis activity may be determined using methods well known in the art, such as those described in detail in the Examples below (see also Loessner *et al.* [37], the disclosures of which are incorporated herein by reference). Preferably, the ability of polypeptides to lyse cells of *Clostridium difficile* is determined using fresh cells.

In a preferred embodiment, the ability of polypeptides to lyse cells of *Clostridium difficile* is determined using cells of strain 11204.

It will be appreciated by persons skilled in the art that the polypeptide, or fragment, variant, derivative or fusion, need not retain all of the ability of the wildtype endolysin of bacteriophage Φ CD27 to lyse bacterial cells. Rather, it is simply necessary for said polypeptide, fragment, variant, derivative or fusion to retain at least 10% of the ability of the wildtype endolysin of bacteriophage Φ CD27 to lyse bacterial cells. Preferably, however, the polypeptide, fragment, variant, derivative or fusion exhibits at least 20%, for example at least 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 150%, 200% or more, of the ability of the wildtype endolysin of bacteriophage Φ CD27 to lyse bacterial cells.

Thus, in one embodiment of the first aspect of the invention, the polypeptide comprises or consists of a fragment of the amino acid sequence of SEQ ID NO:1, which is capable of lysing cells of *Clostridium difficile*.

It is well established that many bacteriophage endolysins consist of two distinct domains (for example, see Sheehan *et al.*, 1996, *FEMS Microbiology Letters* **140**:23-28, the

disclosures of which are incorporated herein by reference). One is a catalytic domain that is responsible for cell wall degradation and these are known to exist in several different forms. The other domain is a cell wall binding domain that recognises a cell surface motif and permits attachment of the endolysin to that target cell. The precise
5 pattern recognition involved in the latter is what provides the specificity.

The enzymatic domain can be identified by its amino acid homology to other similar regions of lytic enzymes that share the same type of lytic activity. In the case of the endolysin of bacteriophage Φ CD27, the enzymatic domain has been identified as an *N*-
10 acetylmuramoyl-L-alanine amidase and it occupies the amino-terminal region of the endolysin (this can be confirmed by alignment analysis of SEQ ID NO: 1 with known enzymatic domains, for example using the NCBI CDD search tool; see Marchler-Bauer & Bryant, 2004, *Nuc. Acids Res.* **32**[W]:327-331, the disclosures of which are incorporated herein by reference). The cell wall binding domain is believed to occupy the carboxy-
15 terminal region of the endolysin.

In one embodiment, the enzymatic domain is contained within amino acids 1 to 175 of SEQ ID NO:1. Thus, the fragment comprising the enzymatic domain may consist of the sequence of SEQ ID NO: 1 starting from any of amino acids 1, 5, 10, 15, 20, 25, 30, 35,
20 40, 45, 50, 60, 70, 80, 90 or 100 and ending at any of amino acids 175, 170, 165, 160, 155, 150, 145, 140, 135, 130, 125, 120, 115, 110 or 105. For example, the fragment comprising the enzymatic domain may consist of amino acids 10 to 140 of SEQ ID NO: 1, or amino acids 25 to 155 of SEQ ID NO: 1, or any of the other possible permutations of the above start and end points.

25 In one embodiment, the cell wall binding domain is contained within amino acids 175 to 270 of SEQ ID NO:1. Thus, the fragment comprising the cell wall binding domain may consist of the sequence of SEQ ID NO: 1 starting from any of amino acids 175, 180, 185, 190, 195, 200, 205, 210, 215, 220 and ending at any of amino acids 270, 265, 260, 255,
30 250, 245, 240, 235, 230 or 225. For example, the fragment comprising the cell wall binding domain may consist of amino acids 195 to 265 of SEQ ID NO: 1, or amino acids 180 to 240 of SEQ ID NO: 1, or any of the other possible permutations of the above start and end points.

The polypeptide of the first aspect of the invention preferably comprises or consists of one or more fragments of the amino acid sequence of SEQ ID NO:1 corresponding to both the enzymatic domain and the cell wall binding domain.

5 However, it will be appreciated by persons skilled in the art that the cell wall binding domain of SEQ ID NO:1 may alternatively be fused or otherwise coupled to an enzymatic (lytic) domain from another source capable of lysing cells of *Clostridium difficile*. The production of chimeric lysins is described in Sheehan *et al.*, 1996, *FEMS Microbiology Letters* **140**:23-28, the disclosures of which are incorporated herein by reference). Thus,
10 in an alternative embodiment, the polypeptide of the first aspect of the invention may comprise or consist of one or more fragments of the amino acid sequence of SEQ ID NO:1 corresponding to the cell wall binding domain.

The fragment may comprise or consist of at least 50 contiguous amino acids of SEQ ID
15 NO: 1, for example at least 60, 70, 80, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 175, 180, 190, 200, 210, 220, 230, 240, 250, 260 or 265 contiguous amino acids of SEQ ID NO: 1.

In an alternative embodiment, the polypeptide of the first aspect of the invention may
20 comprise or consist of a variant of the amino acid sequence of SEQ ID NO:1, or of a fragment thereof, which is capable of lysing cells of *Clostridium difficile*.

By 'variant' of the polypeptide we include insertions, deletions and/or substitutions, either conservative or non-conservative, relative to the amino acid sequence of SEQ ID NO:1.
25 In particular, the variant polypeptide may be a non-naturally occurring variant.

For example, the polypeptide may comprise an amino acid sequence with at least 60% identity to the amino acid sequence of SEQ ID NO: 1, more preferably at least 70% or 80% or 85% or 90% identity to said sequence, and most preferably at least 95%, 96%,
30 97%, 98% or 99% identity to said amino acid sequence.

It will be appreciated that the above sequence identity may be over the full length of the amino acid sequence of SEQ ID NO: 1 or over a portion thereof. Preferably, however, the sequence identity is over at least 50 amino acids of the amino acid sequence of SEQ
35 ID NO: 1, for example at least 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260 or more amino acids therein.

Percent identity can be determined by methods well known in the art, for example using the LALIGN program (Huang and Miller, *Adv. Appl. Math.* (1991) 12:337-357, the disclosures of which are incorporated herein by reference) at the ExPASy facility website:

www.ch.embnet.org/software/LALIGN_form.html

using as parameters the global alignment option, scoring matrix BLOSUM62, opening gap penalty -14, extending gap penalty -4.

Alternatively, the percent sequence identity between two polypeptides may be determined using suitable computer programs, for example AlignX, Vector NTI Advance 10 (from Invitrogen Corporation) or the GAP program (from the University of Wisconsin Genetic Computing Group).

It will be appreciated that percent identity is calculated in relation to polypeptides whose sequence has been aligned optimally.

Fragments and variants of the amino acid sequence of SEQ ID NO: 1 may be made using the methods of protein engineering and site-directed mutagenesis well known in the art (for example, see *Molecular Cloning: a Laboratory Manual*, 3rd edition, Sambrook & Russell, 2001, Cold Spring Harbor Laboratory Press, the disclosures of which are incorporated herein by reference).

It will be appreciated by skilled persons that the polypeptide of the invention, or fragment, variant or fusion thereof, may comprise one or more amino acids that are modified or derivatised. Thus, the polypeptide may comprise or consist of a derivative of the amino acid sequence of SEQ ID NO:1, or of a fragment or variant thereof.

Chemical derivatives of one or more amino acids may be achieved by reaction with a functional side group. Such derivatised molecules include, for example, those molecules in which free amino groups have been derivatised to form amine hydrochlorides, *p*-toluene sulphonyl groups, carboxybenzoxy groups, *t*-butyloxycarbonyl groups, chloroacetyl groups or formyl groups. Free carboxyl groups may be derivatised to form salts, methyl and ethyl esters or other types of esters and hydrazides. Free hydroxyl

groups may be derivatised to form O-acyl or O-alkyl derivatives. Also included as chemical derivatives are those peptides which contain naturally occurring amino acid derivatives of the twenty standard amino acids. For example: 4-hydroxyproline may be substituted for proline; 5-hydroxylysine may be substituted for lysine; 3-methylhistidine
5 may be substituted for histidine; homoserine may be substituted for serine and ornithine for lysine. Derivatives also include peptides containing one or more additions or deletions as long as the requisite activity is maintained. Other included modifications are amidation, amino terminal acylation (e.g. acetylation or thioglycolic acid amidation), terminal carboxylamidation (e.g. with ammonia or methylamine), and the like terminal
10 modifications.

It will be further appreciated by persons skilled in the art that peptidomimetic compounds may also be useful. Thus, by 'polypeptide' we include peptidomimetic compounds which exhibit endolysin activity. The term 'peptidomimetic' refers to a compound that mimics
15 the conformation and desirable features of a particular polypeptide as a therapeutic agent.

For example, the polypeptides described herein include not only molecules in which amino acid residues are joined by peptide (-CO-NH-) linkages but also molecules in
20 which the peptide bond is reversed. Such retro-inverso peptidomimetics may be made using methods known in the art, for example such as those described in Meziere *et al.* (1997) *J. Immunol.* **159**, 3230-3237, the disclosures of which are incorporated herein by reference. Such retro-inverse peptides, which contain NH-CO bonds instead of CO-NH peptide bonds, are much more resistant to proteolysis. Alternatively, the polypeptide of
25 the invention may be a peptidomimetic compound wherein one or more of the amino acid residues are linked by a - γ (CH₂NH)- bond in place of the conventional amide linkage.

It will be appreciated that the polypeptide may conveniently be blocked at its N- or C-terminus so as to help reduce susceptibility to exoproteolytic digestion, e.g. by amidation.

30

As discussed above, a variety of uncoded or modified amino acids such as D-amino acids and N-methyl amino acids may be used to modify polypeptides of the invention. In addition, a presumed bioactive conformation may be stabilised by a covalent modification, such as cyclisation or by incorporation of lactam or other types of bridges.
35 Methods of synthesis of cyclic homodetic peptides and cyclic heterodetic peptides, including disulphide, sulphide and alkylene bridges, are disclosed in US 5,643,872. Other

examples of cyclisation methods are discussed and disclosed in US 6,008,058, the relevant disclosures in which documents are hereby incorporated by reference. A further approach to the synthesis of cyclic stabilised peptidomimetic compounds is ring-closing metathesis (RCM).

5

In summary, terminal modifications are useful, as is well known, to reduce susceptibility by proteinase digestion and therefore to prolong the half-life of the peptides in solutions, particularly in biological fluids where proteases may be present. Polypeptide cyclisation is also a useful modification and is preferred because of the stable structures formed by cyclisation and in view of the biological activities observed for cyclic peptides.

10

Thus, in one embodiment the polypeptide, or fragment, variant, fusion or derivative thereof, is cyclic. However, in a preferred embodiment, the polypeptide, or fragment, variant, fusion or derivative thereof, is linear.

15

In a further embodiment of the first aspect of the invention, the polypeptide comprises or consists of a fusion of the amino acid sequence of SEQ ID NO:1, or of a fragment, variant or derivative thereof.

By 'fusion' of a polypeptide we include a polypeptide which is fused to any other polypeptide. For example, the polypeptide may comprise one or more additional amino acids, inserted internally and/or at the N- and/or C-termini of the amino acid sequence of SEQ ID NO:1, or of a fragment, variant or derivative thereof.

20

Thus, as described above, in one embodiment the polypeptide of the first aspect of the invention comprises a fragment of SEQ ID NO: 1 consisting of the cell wall binding domain (or a variant of such a domain sequence which retains the cell wall binding activity thereof), to which is fused an enzymatic domain from a different source.

25

Examples of other suitable enzymatic domains include:

30

L-alanoyl-D-glutamate endopeptidase; D-glutamyl-m-DAP endopeptidase; interpeptide bridge-specific endopeptidase; *N*-acetyl- β -D-glucosaminidase (=muramoylhydrolase); *N*-acetyl- β -D-muramidase (=lysozyme); lytic transglycosylase.

35

Also *N*-acetylmuramoyl-L-alanine amidase from other sources could be utilised (see Loessner, 2005, *Current Opinion in Microbiology* 8: 480-487, the disclosures of which are incorporated herein by reference).

5 For example, the said polypeptide may be fused to a polypeptide such as glutathione-S-transferase (GST) or protein A in order to facilitate purification of said polypeptide. Examples of such fusions are well known to those skilled in the art. Similarly, the said polypeptide may be fused to an oligo-histidine tag such as His6 or to an epitope recognised by an antibody such as the well-known Myc tag epitope. Fusions to any
10 fragment, variant or derivative of said polypeptide are also included in the scope of the invention. It will be appreciated that fusions (or variants or derivatives thereof) which retain desirable properties, namely endolysin activity are preferred. It is also particularly preferred if the fusions are ones which are suitable for use in the methods described herein.

15 For example, the fusion may comprise a further portion which confers a desirable feature on the said polypeptide of the invention; for example, the portion may be useful in detecting or isolating the polypeptide, promoting cellular uptake of the polypeptide, or directing secretion of the protein from a cell. The portion may be, for example, a biotin
20 moiety, a radioactive moiety, a fluorescent moiety, for example a small fluorophore or a green fluorescent protein (GFP) fluorophore, as well known to those skilled in the art. The moiety may be an immunogenic tag, for example a Myc tag, as known to those skilled in the art or may be a lipophilic molecule or polypeptide domain that is capable of promoting cellular uptake of the polypeptide, as known to those skilled in the art.

25 It will be appreciated by persons skilled in the art that the polypeptides of the invention also include pharmaceutically acceptable acid or base addition salts of the above described polypeptides. The acids which are used to prepare the pharmaceutically acceptable acid addition salts of the aforementioned base compounds useful in this
30 invention are those which form non-toxic acid addition salts, *i.e.* salts containing pharmacologically acceptable anions, such as the hydrochloride, hydrobromide, hydroiodide, nitrate, sulphate, bisulphate, phosphate, acid phosphate, acetate, lactate, citrate, acid citrate, tartrate, bitartrate, succinate, maleate, fumarate, gluconate, saccharate, benzoate, methanesulphonate, ethanesulphonate, benzenesulphonate, *p*-toluenesulphonate and pamoate [*i.e.* 1,1'-methylene-bis-(2-hydroxy-3 naphthoate)] salts,
35 among others.

Pharmaceutically acceptable base addition salts may also be used to produce pharmaceutically acceptable salt forms of the polypeptides. The chemical bases that may be used as reagents to prepare pharmaceutically acceptable base salts of the present compounds that are acidic in nature are those that form non-toxic base salts with such compounds. Such non-toxic base salts include, but are not limited to those derived from such pharmacologically acceptable cations such as alkali metal cations (e.g. potassium and sodium) and alkaline earth metal cations (e.g. calcium and magnesium), ammonium or water-soluble amine addition salts such as N-methylglucamine- (meglumine), and the lower alkanolammonium and other base salts of pharmaceutically acceptable organic amines, among others.

The polypeptide, or fragment, variant, fusion or derivative thereof, may also be lyophilised for storage and reconstituted in a suitable carrier prior to use. Any suitable lyophilisation method (e.g. spray drying, cake drying) and/or reconstitution techniques can be employed. It will be appreciated by those skilled in the art that lyophilisation and reconstitution can lead to varying degrees of activity loss and that use levels may have to be adjusted upward to compensate. Preferably, the lyophilised (freeze dried) polypeptide loses no more than about 20%, or no more than about 25%, or no more than about 30%, or no more than about 35%, or no more than about 40%, or no more than about 45%, or no more than about 50% of its activity (prior to lyophilisation) when rehydrated.

An essential feature of the polypeptides of the invention is the ability to lyse cells of *Clostridium difficile*. Preferably, the polypeptide is capable of lysing cells of multiple strains of *Clostridium difficile*. For example, the polypeptide may be capable of lysing one or more of the strains of *Clostridium difficile* lysed by the Φ CD27 lysin of SEQ ID NO: 1 (see Table 1 below).

It will be appreciated that the polypeptides of the invention may also be capable of lysing cells of other bacterial species, such as *Bacillus* sp. (e.g. *Bacillus cereus*, *Bacillus subtilis* and/or *Bacillus anthracis*), other *Clostridium* sp. (e.g. *Clostridium bifermentans*) and/or *Listeria* sp. (e.g. *Listeria ivanovii*).

In one embodiment, the polypeptides of the invention are substantially incapable of lysing bacteria which are useful for maintaining a healthy gut physiology. For example, it

is advantageous if the polypeptide does not lyse cells of *Clostridium leptum*, *Clostridium nexile*, *Clostridium coccoides*, *Clostridium innocuum*, *Clostridium ramosum*, and/or *Anaerococcus hydrogenalis*.

- 5 Most preferably, the polypeptide of the invention is capable of lysing cells of *Clostridium difficile* strain ribotype 027, a highly virulent strain of *Clostridium difficile* which has emerged in Canada, the US and now throughout Europe. For example, the polypeptide may exhibit at least 10% of the lysis activity of the polypeptide of SEQ ID NO: 1 on cells of *Clostridium difficile* ribotype 027, for example at least 20%, 30%, 40%, 50%, 60%,
10 70%, 80%, 90%, 100% or more. The polypeptide may even exhibit a greater lysis activity than the polypeptide of SEQ ID NO: 1 on cells of *Clostridium difficile* ribotype 027, for example at least 110%, 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, 250%, 300%, 500% or more.
- 15 Advantageously, the polypeptide is capable of lysing cells of pathogenic bacteria selectively, *i.e.* to a greater extent than cells of non-pathogenic bacteria.

- Methods for the production of polypeptides, or a fragment, variant, fusion or derivative thereof, for use in the first aspect of the invention are well known in the art.
- 20 Conveniently, the polypeptide, or fragment, variant, fusion or derivative thereof, is or comprises a recombinant polypeptide.

- Thus, a nucleic acid molecule (or polynucleotide) encoding the polypeptide, or fragment, variant, fusion or derivative thereof, may be expressed in a suitable host and the
25 polypeptide obtained therefrom. Suitable methods for the production of such recombinant polypeptides are well known in the art (for example, see Sambrook & Russell, 2000, *Molecular Cloning, A Laboratory Manual*, Third Edition, Cold Spring Harbor, New York, the relevant disclosures in which document are hereby incorporated by reference).

30

In brief, expression vectors may be constructed comprising a nucleic acid molecule which is capable, in an appropriate host, of expressing the polypeptide encoded by the nucleic acid molecule.

- 35 A variety of methods have been developed to operably link nucleic acid molecules, especially DNA, to vectors, for example, via complementary cohesive termini. For

instance, complementary homopolymer tracts can be added to the DNA segment to be inserted into the vector DNA. The vector and DNA segment are then joined by hydrogen bonding between the complementary homopolymeric tails to form recombinant DNA molecules.

5

Synthetic linkers containing one or more restriction sites provide an alternative method of joining the DNA segment to vectors. The DNA segment, e.g. generated by endonuclease restriction digestion, is treated with bacteriophage T4 DNA polymerase or *E. coli* DNA polymerase I, enzymes that remove protruding, 3'-single-stranded termini with their 3'-5'-exonucleolytic activities, and fill in recessed 3'-ends with their polymerising activities.

The combination of these activities therefore generates blunt-ended DNA segments. The blunt-ended segments are then incubated with a larger molar excess of linker molecules in the presence of an enzyme that is able to catalyse the ligation of blunt-ended DNA molecules, such as bacteriophage T4 DNA ligase. Thus, the products of the reaction are DNA segments carrying polymeric linker sequences at their ends. These DNA segments are then cleaved with the appropriate restriction enzyme and ligated to an expression vector that has been cleaved with an enzyme that produces termini compatible with those of the DNA segment.

The DNA (or in the case of retroviral vectors, RNA) is then expressed in a suitable host to produce a polypeptide. Thus, the DNA encoding the polypeptide may be used in accordance with known techniques, appropriately modified in view of the teachings contained herein, to construct an expression vector, which is then used to transform an appropriate host cell for the expression and production of the compound of the invention or binding moiety thereof. Such techniques are well known in the art.

The DNA (or in the case of retroviral vectors, RNA) encoding the polypeptide may be joined to a wide variety of other DNA sequences for introduction into an appropriate host. The companion DNA will depend upon the nature of the host, the manner of the introduction of the DNA into the host, and whether episomal maintenance or integration is desired.

Generally, the DNA is inserted into an expression vector, such as a plasmid, in proper orientation and correct reading frame for expression. If necessary, the DNA may be

linked to the appropriate transcriptional and translational regulatory control nucleotide sequences recognised by the desired host, although such controls are generally available in the expression vector. The vector is then introduced into the host through standard techniques. Generally, not all of the hosts will be transformed by the vector.

5 Therefore, it will be necessary to select for transformed host cells. One selection technique involves incorporating into the expression vector a DNA sequence, with any necessary control elements, that codes for a selectable trait in the transformed cell, such as antibiotic resistance. Alternatively, the gene for such selectable trait can be on another vector, which is used to co-transform the desired host cell.

10

Host cells that have been transformed by the expression vector are then cultured for a sufficient time and under appropriate conditions known to those skilled in the art in view of the teachings disclosed herein to permit the expression of the polypeptide, which can then be recovered.

15

Many expression systems are known, including bacteria (for example, *E. coli* and *Bacillus subtilis*), yeasts (for example *Saccharomyces cerevisiae*), filamentous fungi (for example *Aspergillus*), plant cells, animal cells and insect cells.

20 The vectors typically include a prokaryotic replicon, such as the ColE1 *ori*, for propagation in a prokaryote, even if the vector is to be used for expression in other, non-prokaryotic, cell types. The vectors can also include an appropriate promoter such as a prokaryotic promoter capable of directing the expression (transcription and translation) of the genes in a bacterial host cell, such as *E. coli*, transformed therewith.

25

Typical prokaryotic vector plasmids are pUC18, pUC19, pBR322 and pBR329 available from Biorad Laboratories, (Richmond, CA, USA) and pTrc99A and pKK223-3 available from Pharmacia, Piscataway, NJ, USA.

30 A typical mammalian cell vector plasmid is pSVL available from Pharmacia, Piscataway, NJ, USA. This vector uses the SV40 late promoter to drive expression of cloned genes, the highest level of expression being found in T antigen-producing cells, such as COS-1 cells.

An example of an inducible mammalian expression vector is pMSG, also available from Pharmacia. This vector uses the glucocorticoid-inducible promoter of the mouse mammary tumour virus long terminal repeat to drive expression of the cloned gene.

- 5 Other vectors and expression systems are well known in the art for use with a variety of host cells.

The host cell may be either prokaryotic or eukaryotic. Bacterial cells are preferred prokaryotic host cells and typically are a strain of *E. coli* such as, for example, the *E. coli* strains DH5 available from Bethesda Research Laboratories Inc., Bethesda, MD, USA, and RR1 available from the American Type Culture Collection (ATCC) of Rockville, MD, USA (No. ATCC 31343). Preferred eukaryotic host cells include yeast, insect and mammalian cells, preferably vertebrate cells such as those from a mouse, rat, monkey or human fibroblastic and kidney cell lines. Yeast host cells include YPH499, YPH500 and YPH501 which are generally available from Stratagene Cloning Systems, La Jolla, CA 92037, USA. Preferred mammalian host cells include Chinese hamster ovary (CHO) cells available from the ATCC as CRL 1658 and 293 cells which are human embryonic kidney cells. Preferred insect cells are Sf9 cells which can be transfected with baculovirus expression vectors.

20

Methods of cultivating host cells and isolating recombinant proteins are well known in the art. It will be appreciated that, depending on the host cell, the polypeptides of the invention produced may differ. For example, certain host cells, such as yeast or bacterial cells, either do not have, or have different, post-translational modification systems which may result in the production of forms of compounds of the invention which may be post-translationally modified in a different way.

25

Polypeptides of the invention may also be produced *in vitro* using a commercially available *in vitro* translation system, such as rabbit reticulocyte lysate or wheatgerm lysate (available from Promega). Preferably, the translation system is rabbit reticulocyte lysate. Conveniently, the translation system may be coupled to a transcription system, such as the TNT transcription-translation system (Promega). This system has the advantage of producing suitable mRNA transcript from an encoding DNA polynucleotide in the same reaction as the translation.

35

Automated polypeptide synthesisers may also be used, such as those available from CS Bio Company Inc, Menlo Park, USA.

Thus, a second aspect of the present invention provides an isolated nucleic acid molecule encoding a polypeptide according to the first aspect of the invention.

The nucleic acid molecule may be DNA (e.g. cDNA) or RNA.

In a preferred embodiment, the nucleic acid molecule comprises or consists of the nucleotide sequence as shown in figure 3 [SEQ ID NO:2].

A third aspect of the invention provides a vector comprising a nucleic acid molecule according to the second aspect of the invention. In one embodiment, the vector is an expression vector. Preferably, the vector is selected from the group consisting of pET15b and pACYC184.

It will be appreciated by persons skilled in the art that the choice of expression vector may be determined by the choice of host cell. Thus, for expression of the polypeptides of the invention in *Lactococcus lactis*, the nisin expression system could be used in which the polypeptide of the invention is expressed under the control of the promoter of the *nisA* operon using a background strain of *Lactococcus lactis* which also expresses the *nisR* and *nisK* genes encoding a two component regulatory system. Under this system expression is positively regulated and induced by the provision of exogenous nisin (see de Ruyter *et al.*, 1996, *Applied and Environmental Microbiology* **62**:3662-3667, the disclosures of which are incorporated herein by reference).

In an alternative embodiment, the entire nisin biosynthesis gene cluster is provided within the same host cell, in which case the inducer is synthesised by that cell.

In a further alternative embodiment, the polypeptides of the invention may be expressed in *Lactococcus lactis* under the control of the lactose catabolic operon, using either a plasmid-based or chromasomally integrated system (for example, see Payne *et al.*, 1996, *FEMS Microbiology Letters* **136**: 19-24 and van Rooijen *et al.*, 1992, *Journal of Bacteriology* **174**: 2273-2280, the disclosures of which are incorporated herein by reference).

A fourth aspect of the invention provides a host cell comprising a nucleic acid molecule according to the second aspect of the invention or a vector according to the third aspect of the invention. In one embodiment, the host cell is a microbial cell, for example a bacterial cell. Preferably, the host cell is non-pathogenic.

For example, the host cell may be selected from the group consisting of cells of *Escherichia coli*, *Lactococcus* sp., *Bacteroides* sp., *Lactobacillus* sp., *Enterococcus* sp. and *Bacillus* sp.

In a preferred embodiment, the host cell is a cell of *Lactococcus lactis*.

Alternatively, the host cell may be a yeast cell, for example *Saccharomyces* sp.

A fifth aspect of the invention provides a method for producing a polypeptide of the invention comprising culturing a population of host cells comprising a nucleic acid molecule according to the second aspect of the invention or a vector according to the third aspect of the invention under conditions in which the polypeptide is expressed, and isolating the polypeptide therefrom.

A sixth aspect of the invention provides a pharmacological composition comprising:

- (a) a polypeptide according to the first aspect of the invention;
- (b) a nucleic acid molecule according to the second aspect of the invention;
- (c) a vector according to the third aspect of the invention;
- (d) a host according to the fourth aspect of the invention; and/or
- (e) a bacteriophage capable of expressing a polypeptide according to the first aspect of the invention

and a pharmaceutically acceptable carrier, diluent or excipient.

As used herein, 'pharmaceutical composition' means a therapeutically effective formulation for use in the methods of the invention.

A 'therapeutically effective amount', or 'effective amount', or 'therapeutically effective', as used herein, refers to that amount which provides a therapeutic effect for a given

condition and administration regimen. This is a predetermined quantity of active material calculated to produce a desired therapeutic effect in association with the required additive and diluent, *i.e.* a carrier or administration vehicle. Further, it is intended to mean an amount sufficient to reduce, and most preferably prevent, a clinically significant deficit in the activity, function and response of the host. Alternatively, a therapeutically effective amount is sufficient to cause an improvement in a clinically significant condition in a host. As is appreciated by those skilled in the art, the amount of a compound may vary depending on its specific activity. Suitable dosage amounts may contain a predetermined quantity of active composition calculated to produce the desired therapeutic effect in association with the required diluent. In the methods and use for manufacture of compositions of the invention, a therapeutically effective amount of the active component is provided. A therapeutically effective amount can be determined by the ordinary skilled medical or veterinary worker based on patient characteristics, such as age, weight, sex, condition, complications, other diseases, *etc.*, as is well known in the art.

In one embodiment of the invention, the pharmacological composition comprises a polypeptide according to the first aspect of the invention.

The polypeptides can be formulated at various concentrations, depending on the efficacy/toxicity of the polypeptide being used. Preferably, the formulation comprises the polypeptide at a concentration of between 0.1 μM and 1 mM, more preferably between 1 μM and 100 μM , between 5 μM and 50 μM , between 10 μM and 50 μM , between 20 μM and 40 μM and most preferably about 30 μM . For *in vitro* applications, formulations may comprise similar concentrations of a polypeptide (however, it will be appreciated that higher concentrations may also be used).

Thus, the pharmaceutical formulation may comprise an amount of a polypeptide, or fragment, variant, fusion or derivative thereof, sufficient to inhibit at least in part the growth of cells of *Clostridium difficile* in a patient who is infected or susceptible to infection with such cells. Preferably, the pharmaceutical formulation comprises an amount of a polypeptide, or fragment, variant, fusion or derivative thereof, sufficient to kill cells of *Clostridium difficile* in the patient.

It will be appreciated by persons skilled in the art that the polypeptides of the invention are generally administered in admixture with a suitable pharmaceutical excipient, diluent

or carrier selected with regard to the intended route of administration and standard pharmaceutical practice (for example, see *Remington: The Science and Practice of Pharmacy*, 19th edition, 1995, Ed. Alfonso Gennaro, Mack Publishing Company, Pennsylvania, USA, the relevant disclosures in which document are hereby incorporated
5 by reference).

For example, the polypeptides can be administered orally, buccally or sublingually in the form of tablets, capsules, ovules, elixirs, solutions or suspensions, which may contain flavouring or colouring agents, for immediate-, delayed- or controlled-release
10 applications. The polypeptides may also be administered via direct injection (for example, into the GI tract).

Preferably, however, the polypeptides and pharmaceutical compositions thereof are for oral administration.

15 Suitable tablet formulations may contain excipients such as microcrystalline cellulose, lactose, sodium citrate, calcium carbonate, dibasic calcium phosphate and glycine, disintegrants such as starch (preferably corn, potato or tapioca starch), sodium starch glycollate, croscarmellose sodium and certain complex silicates, and granulation binders
20 such as polyvinylpyrrolidone, hydroxyl-propylmethylcellulose (HPMC), hydroxy-propylcellulose (HPC), sucrose, gelatin and acacia. Additionally, lubricating agents such as magnesium stearate, stearic acid, glyceryl behenate and talc may be included.

Solid compositions of a similar type may also be employed as fillers in gelatin capsules.
25 Preferred excipients in this regard include lactose, starch, cellulose, milk sugar or high molecular weight polyethylene glycols. For aqueous suspensions and/or elixirs, the polypeptides may be combined with various sweetening or flavouring agents, colouring matter or dyes, with emulsifying and/or suspending agents and with diluents such as water, ethanol, propylene glycol and glycerin, and combinations thereof.

30 The polypeptides can also be administered parenterally, for example, intravenously, intra-articularly, intra-arterially, intraperitoneally, intra-theccally, intraventricularly, intrasternally, intracranially, intra-muscularly or subcutaneously, or they may be administered by infusion techniques. They are best used in the form of a sterile aqueous
35 solution which may contain other substances, for example, enough salts or glucose to make the solution isotonic with blood. The aqueous solutions should be suitably buffered

(preferably to a pH of from 3 to 9), if necessary. The preparation of suitable parenteral formulations under sterile conditions is readily accomplished by standard pharmaceutical techniques well known to those skilled in the art.

5 Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation isotonic with the blood of the intended recipient; and aqueous and non-aqueous sterile suspensions which may include suspending agents and thickening agents. The formulations may be presented in unit-dose or multi-dose
10 containers, for example sealed ampoules and vials, and may be stored in a freeze-dried (lyophilised) condition requiring only the addition of the sterile liquid carrier, for example water for injections, immediately prior to use. Extemporaneous injection solutions and suspensions may be prepared from sterile powders, granules and tablets of the kind previously described.

15 For oral and parenteral administration to human patients, the daily dosage level of the polypeptides will usually be from 1 to 1000 mg per adult (*i.e.* from about 0.015 to 15 mg/kg), administered in single or divided doses. For example, a dose of 1 to 10 mg/kg may be used, such as 3 mg/kg.

20 In an alternative embodiment of the invention, the pharmaceutical compositions do not comprise the polypeptide itself but instead comprise a nucleic acid molecule capable of expressing said polypeptide. Suitable nucleic acid molecules, expression vectors, and host cells are described in detail above.

25 For example, a recombinant probiotic may be used (LAB strain, *e.g.* *Lactococcus lactis* or a *Lactobacillus sp.*).

In a further embodiment of the invention, the pharmaceutical compositions comprise a
30 bacteriophage capable of expressing a polypeptide according to the first aspect of the invention. For example, the wildtype bacteriophage Φ CD27 may be used to deliver a polypeptide according to the first aspect of the invention. Methods for performing such bacteriophage-based therapies are well known in the art (for example, see Watanabe *et al.*, 2007, *Antimicrobial Agents & Chemotherapy* 51:446-452).

35

Thus, for treatment of bacterial infections described herein, the polypeptide of the invention may be administered as the cognate protein, as a nucleic acid construct, vector or host cell which expresses the cognate protein, as part of a living organism which expresses the cognate protein (including bacteriophages), or by any other convenient method known in the art so as to achieve contact of the lysin with its bacterial target, whether that be a pathogenic bacterium, such as *C. difficile*, or another pathogen or potential pathogen, as further described herein.

Ideally, the protein is delivered to the GI tract in a protected form. This may be achieved by a wide variety of methods known in the art. For example, an appropriate dose of the lysin is microencapsulated in a form that survives the acidic conditions of the stomach, but which releases the protein as it enters the intestine. Delivery by a non-pathogenic microbe which survives GI tract transit, including but not limited to by *Lactococcus lactis*, *Lactobacillus* sp., *Bifidobacterium* sp. or *Bacteroides*. Those skilled in the art are well aware of the options available for use of such means for GI tract delivery of active compounds such as the lysin disclosed herein. These means include intracellular production, secA secretion or secretion by means of another secretion pathway, and delivery by controlled lysis. Preferably the protein is not all released at one time, but is released increasingly as an administered bolus traverses through the GI tract. Alternatively, the lysin is introduced as part of a benign bacterium which expresses the lysin at the appropriate location or upon receipt of an appropriate signal in the GI tract. In a preferred embodiment disclosed herein, a non-pathogenic *Lactococcus* is engineered to express the Φ CD27 lysin upon reaching a particular location in the GI tract. The expression signal may be defined by a pH sensitive promoter, or another means known in the art for this purpose.

Other means of delivery include the following:

- (a) WO 2006/111553 (polyurea and other multilayer encapsulants);
- (b) WO 2006/111570 and EP 1 715 739 (cyclodextrin encapsulation);
- (c) WO 2006/100308 and EP 1 742 728 (for yeast and other microbial cell encapsulation technologies);
- (d) US 5,153,182, EP 1 499 183 and WO 03/092378; US 6,831,070 (therapeutic gene product delivery by intestinal cell expression);
- (e) US 7,202,236 (pharmaceutical formulation for modified release);

- (f) US 5,762,904 (oral delivery of vaccines using polymerized liposomes, which may be modified to deliver the lysin of this invention),
(g) US 7,195,906 (*Bifidobacterium* which may be modified to express the lysin according to this invention); and
5 (h) references cited therein,

all of which are herein incorporated by reference for purposes of enabling those skilled in the art to utilize the present disclosure to achieve the novel methods of delivery and compositions according to the present invention:

10

Thus, in a preferred embodiment of the pharmacological compositions of the invention, the polypeptide, nucleic acid molecule encoding the same, *etc.* is microencapsulated (e.g. within a stable chemical envelope, such as cyclodextrin or a lipid bilayer, or within a living or non-living microbial cell, such as an engineered *Lactococcus* cell). In this way,
15 the polypeptide, nucleic acid molecule, *etc.* may be protected against acidic conditions of stomach en route to its site of action in the GI tract.

20

A seventh aspect of the invention provides polypeptide according to the first aspect of the invention or pharmacological composition according to the sixth aspect of the invention for use in medicine.

25

An eighth aspect of the invention provides the use of a polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile*, or a nucleic acid molecule, vector, host cell or bacteriophage capable of expressing the same, in the preparation of a medicament for killing and/or inhibiting/preventing the growth of microbial cells in a patient, wherein the microbial cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis with said endolysin.

30

It will be appreciated that polypeptides exhibiting cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile* need not necessarily be derived from a bacteriophage of *Clostridium difficile*. For example, the polypeptide may be selected from the following group:

35

- (a) the lysin of bacteriophage Φ CD27;
(b) the lysin of bacteriophage Φ CD119;

- (c) the lysin of bacteriophage Φ C2; and
- (d) the lysin of prophages 1 and 2 of *Clostridium difficile* strain 630 (CD630).

Alternatively, the polypeptide may be derived from (e.g. encoded by) a bacteriophage of
5 a different Clostridial sp..such as *Clostridium bifermentans* or *Clostridium sordelli*.

However, in a preferred embodiment, the polypeptide is derived from a bacteriophage of
Clostridium difficile.

10 Thus, the use of the eighth aspect of the invention is not limited to polypeptides of the first aspect of the invention but encompasses the use of any polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile* (including the lysin of Φ C2, as described in Goh *et al.*, 2007, *Microbiology* **153**:676-685, the disclosures of which are incorporated herein by reference).

15 A related aspect of the invention provides the use of a polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile*, or a nucleic acid molecule, vector, host cell or bacteriophage capable of expressing the same, for killing and/or inhibiting/preventing the growth of microbial cells in a patient, wherein the
20 microbial cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis with said endolysin.

A further aspect of the invention provides the use of a polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile*, or a nucleic acid
25 molecule, vector, host cell or bacteriophage capable of expressing the same, in the preparation of a medicament for the treatment or prevention of a disease or condition associated with microbial cells in a patient, wherein the microbial cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis with said endolysin. A related aspect of the invention provides the use of a
30 polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile* for the treatment or prevention of a disease or condition associated with microbial cells in a patient, wherein the microbial cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis with said endolysin.

By "a disease or condition associated with microbial cells in a patient" we include diseases and conditions arising from or antagonised by infection of a patient with *Clostridium difficile*. Such diseases and conditions include *Clostridium difficile*-associated disease (CDAD).

5

In one embodiment of the above defined uses of the invention, the polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile* is a polypeptide according to the first aspect of the invention, wherein the microbial cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis upon contact with a polypeptide of SEQ ID NO: 1 (see Tables 1 and 2, below).

10

Preferably, the microbial cells comprise or consist of *Clostridium difficile* cells. Thus, the polypeptides having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile* may be used to treat or prevent diseases and conditions associated with infection with *Clostridium difficile* cells (such as *Clostridium difficile*-associated disease, CDAD).

15

Most preferably, the microbial cells comprise or consist of cells are *Clostridium difficile* ribotype 027 cells.

20

Thus, the invention further provides the following:

(a) a method for killing and/or inhibiting/preventing the growth of microbial cells in a patient, the method comprising administering to the patient a polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile*, or a nucleic acid molecule, vector, host cell or bacteriophage capable of expressing the same, wherein the microbial cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis with said endolysin;

25

30

(b) a method for the treatment or prevention a disease or condition associated with microbial cells in a patient, the method comprising administering to the patient a polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile*, or a nucleic acid molecule, vector, host cell or bacteriophage capable of expressing the same, wherein the microbial cells are selected from the

35

group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis with said endolysin.

In one embodiment of the above defined methods of the invention, the polypeptide
5 having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile*
is a polypeptide according to the first aspect of the invention, wherein the microbial cells
are selected from the group consisting of *Clostridium difficile* cells and other bacterial
cells susceptible to lysis upon contact with a polypeptide of SEQ ID NO: 1 (see Tables 1
and 2, below). Preferably, the microbial cells comprise or consist of *Clostridium difficile*
10 cells. Thus, the polypeptides having the cell lysing activity of an endolysin from a
bacteriophage of *Clostridium difficile* may be used to treat or prevent diseases and
conditions associated with infection with *Clostridium difficile* cells (such as *Clostridium*
difficile-associated disease, CDAD). Most preferably, the microbial cells comprise or
consist of cells of *Clostridium difficile* ribotype 027.

15
Persons skilled in the art will further appreciate that the uses and methods of the present
invention have utility in both the medical and veterinary fields. Thus, the medicaments
may be used in the treatment of both human and non-human animals (such as horses,
cows, dogs and cats). Preferably, however, the patient is human.

20
By 'treatment' we include both therapeutic and prophylactic treatment of the patient. The
term 'prophylactic' is used to encompass the use of a polypeptide or formulation
described herein which either prevents or reduces the likelihood of infection with
Clostridium difficile in a patient or subject.

25
As discussed above, the term 'effective amount' is used herein to describe
concentrations or amounts of polypeptides according to the present invention which may
be used to produce a favourable change in a disease or condition treated, whether that
change is a remission, a favourable physiological result, a reversal or attenuation of a
30 disease state or condition treated, the prevention or the reduction in the likelihood of a
condition or disease state occurring, depending upon the disease or condition treated.

It will be appreciated that the medicaments described herein may be administered to
patients in combination with one or more additional therapeutic agents.

35

For example, the medicaments described herein may be administered to patients in combination with:

- 5 (a) one or more conventional antibiotic treatments (such as beta-lactams, aminoglycosides and/or quinolones);
- (b) one or more additional lysins, or nucleic acid molecules, vectors, host cell or bacteriophage capable of expressing the same;
- (c) one or more lantibiotics, or nucleic acid molecules, vectors, host cell or bacteria capable of expressing the same; and/or
- 10 (d) a therapy to neutralise the toxins released upon bacterial lysis of *Clostridium difficile* cells within the gut. Suitable neutralising therapies may include antibodies (see Babcock *et al.*, 2006, *Infect. Immun.* **74**:6339-6347) and toxin-absorbing agents such as tolevamer (see Barker *et al.*, 2006, *Aliment. Pharmacol. Ther.* **24**:1525-1534).

15

A further aspect of the invention provides the use of a polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile*, or a nucleic acid molecule, vector, host cell or bacteriophage capable of expressing the same, for killing and/or inhibiting/preventing the growth of microbial cells *in vitro* and/or *ex vivo*, wherein
20 the microbial cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis with said endolysin. For example, said polypeptides having endolysin activity may be used to clean surfaces, such as those in hospitals, kitchens, etc, which may be susceptible to contamination with such bacterial cells.

25

Preferably, the polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile* is a polypeptide according to the first aspect of the invention, wherein the microbial cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis upon contact with a
30 polypeptide of SEQ ID NO: 1 (see Tables 1 and 2, below). For example, the microbial cells may comprise or consist of *Clostridium difficile* cells. Most preferably, the microbial cells comprise or consist of cells of *Clostridium difficile* ribotype 027.

A further aspect of the present invention provides a kit for detecting the presence of
35 microbial cells in a sample, the kit comprising a polypeptide having the cell lysing activity and/or cell binding specificity of an endolysin from a bacteriophage of *Clostridium*

difficile, or a nucleic acid molecule, vector, host cell or bacteriophage capable of expressing the same, wherein the microbial cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis with said endolysin.

5

In a preferred embodiment, the polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile* is a polypeptide according to the first aspect of the invention, wherein the microbial cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis upon contact with a polypeptide of SEQ ID NO: 1 (see Tables 1 and 2, below). For example, the microbial cells may comprise or consist of *Clostridium difficile* cells. Most preferably, the microbial cells comprise or consist of cells of *Clostridium difficile* ribotype 027..

10

In a further embodiment of the kits of the invention, the polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile* is immobilised on a suitable surface, such as the surface of a multi-well plate.

15

The kits may be used in conjunction with any suitable sample of cells, such as tissue samples, cell culture samples and samples of cells derived from swabs (e.g. taken from a surface to be tested for contamination with microbial cells).

20

Optionally, the kit further comprises a negative control sample (which does not contain cells of the type to be tested for, e.g. *Clostridium difficile* cells) and/or a positive control sample (which contains cells of the type to be tested for).

25

Related aspects of the invention provide:

30

(a) the use of a polypeptide having the cell wall binding activity and/or cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile*, or a nucleic acid molecule, vector, host cell or bacteriophage capable of expressing the same, in the preparation of a diagnostic agent for a disease or condition associated with microbial cells selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis with said endolysin;

35

(b) the use of a polypeptide having the cell wall binding activity and/or cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile*, or a nucleic

acid molecule, vector, host cell or bacteriophage capable of expressing the same, for the diagnosis of a disease or condition associated with microbial cells selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis with said endolysin;

5

(c) the use of a polypeptide having the cell wall binding activity and/or cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile*, or a nucleic acid molecule, vector, host cell or bacteriophage capable of expressing the same, for detecting the presence of microbial cells in a sample *in vitro* and/or *ex vivo*,
10 wherein the microbial cells selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis with said endolysin; and

(d) a method for the diagnosis of a disease or condition associated with microbial cells in a patient, the method comprising contacting a cell sample from a patient
15 to be tested with a polypeptide having the cell wall binding activity and/or cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile*, or a nucleic acid molecule, vector, host cell or bacteriophage capable of expressing the same, and determining whether the cells in the sample have been lysed thereby, wherein the microbial cells are selected from the group consisting of
20 *Clostridium difficile* cells and other bacterial cells susceptible to lysis with said endolysin.

In one embodiment of the above defined uses and methods of the invention, the polypeptide having the cell lysing activity of an endolysin from a bacteriophage of
25 *Clostridium difficile* is a polypeptide according to the first aspect of the invention, wherein the microbial cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis upon contact with a polypeptide of SEQ ID NO: 1 (see Tables 1 and 2, below). Preferably, the microbial cells comprise or consist of *Clostridium difficile* cells. Thus, the polypeptides having the cell lysing activity of an
30 endolysin from a bacteriophage of *Clostridium difficile* may be used to diagnose diseases and conditions associated with infection with *Clostridium difficile* cells (such as *Clostridium difficile*-associated disease, CDAD). Most preferably, the microbial cells comprise or consist of cells of *Clostridium difficile* ribotype 027.

In such diagnostic uses and methods, lysis of cells may be detected using methods well known in the art. For example, levels of ATP may be measured as an indicator of cell lysis.

5 In an alternative embodiment of the above defined uses and methods of the invention, the polypeptide comprises or consists of the cell wall binding domain of an endolysin from a bacteriophage of *Clostridium difficile*. To permit detection, such a polypeptide may be fused to magnetic beads or used as a fusion protein comprising a suitable reporter (for example, green fluorescent protein).

10

Such diagnostic approaches are well established for endolysins from other systems, such as *Listeria* endolysins (for example, see Loessner *et al.*, 2002, *Mol Microbiol* **44**, 335-49; Kretzer *et al.*, 2007, *Applied Environ. Microbiol.* **73**:1992-2000, the disclosures of which are incorporated herein by reference; suitable assays are also available commercially, for example from Profos, Germany [see their website at www.profos.de/content/view/164/69/lang,en/]).

15

Exemplary embodiments of the invention are described in the following non-limiting examples, with reference to the following figures:

20

Figure 1. Electron micrograph of Φ CD27. Samples were negative-stained in saturated uranyl acetate.

25 **Figure 2.** Φ CD27 genome map showing predicted ORFs. Arrows indicate the directions of transcription. Proposed functional modules are marked based on BLAST results and similarity to published sequences of Φ CD119, Φ C2, and *C. difficile* strain 630 prophages.

Figure 3. Nucleotide sequence of Φ CD27 lysin, SEQ. ID. 2.

30

Figure 4. Alignment of Φ CD27 (a) nucleotide and (b) inferred amino acid sequence with published *C. difficile* bacteriophage (Φ C2 (32); Φ CD119 (31)), or prophage (CD630 prophage 1 and 2 from sequenced genome (36)) sequences. Alignment performed with AlignX, Vector NTI Advance 10, Invitrogen. Φ CD27 amino acid sequence is SEQ. ID. 2

35

Figure 5. Cloning site of pET15b vector (Novagen).

Figure 6. (a) Gel analysis of crude protein lysates from *E. coli* expressing Φ CD27 lysin. Lane 1 SeeBlue marker (Invitrogen, sizes 191, 97, 64, 51, 39, 28 and 19kDa), lanes 2-5 BL21(DE3)pET15b Φ CD27L total protein extracts. Lanes 2-4 extracts induced for 3h with IPTG – 2 and 3 extracted with 20 mM Tris-HCl pH 8, 50 mM NaCl, 3 including protease inhibitor (Roche Complete mini EDTA-free) and 4 extracted with denaturing buffer (8M urea, 0.1M NaH₂PO₄, 0.01M Tris-HCl pH 8.0). Lane 5 uninduced control extracted with 20 mM Tris-HCl pH 8, 50 mM NaCl. Lanes 6 and 7 BL21(DE3)pET15bCD630L1 total protein extracts extracted with 20 mM Tris-HCl pH 8, 50 mM NaCl, lane 6 only induced for 3h with IPTG (b) Western analysis of gel (a) with 6xHis antibody.

Figure 7. Gel analysis of NiNTA column-purified His-tagged Φ CD27 lysin. Lane 1 SeeBlue marker (Invitrogen, sizes 191, 97, 64, 51, 39, 28 and 19kDa), lanes 2-5 BL21(DE3)pET15b Φ CD27L total protein extracts after induction with IPTG. Lane 1 crude lysate, lane 2 column flow-through, lane 3 primary wash effluent, lane 4 secondary wash effluent, lane 5 primary eluate (E1, 1 ml), lane 6 secondary eluate (E2).

Figure 8. Bioscreen lysis assay with cells of *C. difficile* 11204 grown to end log, flash frozen in liquid nitrogen then resuspended in PBS. Φ CD27 lysin and CD630 lysin were expressed in *E. coli* and purified using the His tag on a NiNTA column (see Fig. 6). 270 μ l cells were added to 30 μ l of dilutions of E1 extracts. Values are the means of duplicate assays \pm standard deviation. The cell lysis with the CD630L1 extract was equivalent to that seen in the buffer-only control.

Figure 9. Bioscreen lysis assay with cells of *C. difficile* 11204 grown to end log, harvested by centrifugation at 4°C then resuspended in PBS to give an OD of between 1-1.5. Φ CD27 lysin was expressed in *E. coli* and purified using the His tag on a NiNTA column (see Fig. 6). 270 μ l cells were added to 30 μ l samples of eluate 1 (E1) diluted with elution buffer to give a range of concentrations from 10.5 μ g to 0.35 ng per assay. The use of fresh cells gave significantly less lysis in the buffer-only control. No difference to buffer-only control was seen with less than 70 ng NiNTA-purified protein.

Figure 10. Bioscreen lysis assays of Φ CD27 lysin added to *C. difficile* cells to test the spectrum of activity. Cells were incubated with 3.5 μ g NiNTA-purified protein (E1) produced from *E. coli*. Of the 30 strains tested all were sensitive, including the host strain

12727 and bacteriophage Φ CD27-insensitive strains 11208 and hypervirulent ribotype 027 R23 613. Incubations were in duplicate with either buffer (B) or lysin (L).

Figure 11. Activity of Φ CD27 lysin against *Clostridium* species and prevalent gut
 5 bacteria. Cells were harvested at late stationary phase, resuspended in PBS then
 incubated with 7 μ g NiNTA-purified protein (E1) produced from *E. coli*. Results are the
 mean of duplicate assays \pm standard deviation. The Φ CD27 lysin did not produce cell
 lysis in the majority of species (a, and see Table 2). Exceptions (b, and see Table 2)
 included a rapid lysis of *Clostridium bifermentans*, lysis of *Bacillus cereus* and, with a
 10 longer lag phase, *B. subtilis*, and a slight effect on *Listeria ivanovii* (b).

Figure 12. pH profile of Φ CD27 lysin activity. *C. difficile* 11204 cells were resuspended in
 PBS adjusted to a range of pHs and activity of the Ni-NTA-purified lysin E1 produced
 from *E. coli* was measured in the bioscreen as before.

15

Figure 13. (a) Gel analysis of crude protein lysates from *Lactococcus lactis* expressing
 Φ CD27 lysin. Lanes 1 and 10 SeeBlue marker (Invitrogen, sizes 191, 97, 64, 51, 39, 28
 and 19kDa), lanes 2-5 *L. lactis* UKLC10 containing phiCD27LpUK200HIS (2,3) or an
 empty vector pUK200HIS control (4,5), induced for 5 h (2,4) or uninduced (3,5). Lanes
 20 6-9 *E. coli* BL21(DE3) containing phiCD27LpET15b (6, 8, 9) or the empty vector control
 (7) all induced for 4 h (10 μ g per lane). All proteins were extracted in 20 mM Tris-HCl pH
 8, 50 mM NaCl except lanes 8 (20mM sodium phosphate pH 8) and 9 (50mM Tris-HCl
 pH 7.5). (b) Western analysis of gel (a) with 6xHis antibody.

Figure 14. Bioscreen assay of crude protein extracts from phiCD27 lysin-expressing *E.*
coli and *L. lactis* incubated with fresh cells of *C. difficile* strain 11204 compared to
 extracts from empty vector controls. 50 μ g protein was used in each assay, results are
 the mean of duplicate assays \pm standard deviation.

Figure 15. Bioscreen assay of the Ni-NTA-purified lysin E1 produced from *E. coli*
 30 showing the activity of the original extract compared to that of an aliquot which had been
 through a zeba buffer exchange column (Pierce) into 20mM sodium phosphate pH 6.0.
 Lysins and buffer controls were incubated with flash-frozen cells of *C. difficile* strain
 11204 and results are the means of duplicate assays \pm standard deviation.

Figure 16. SDS-PAGE of crude cell extracts of LM4-CD27L (lane 2) and LM4-CD27LE (lane 3) and the corresponding Western blot highlighting the His-tagged proteins. Proteins were extracted in 20mM sodium phosphate pH 6.0 and 10 µg aliquots were electrophoresed on a 10% Bis-Tris NuPage gel in MOPS buffer (Invitrogen). Lane 1, SeeBlue marker.

Figure 17. Bioscreen analysis showing lysis of *C. difficile* strain 11204 cells grown to mid-log then flash frozen in liquid nitrogen. Cells were incubated with 10 µg NiNTA-purified E1 (eluate 1) or elution buffer as a control.

EXAMPLES

Background

5 The exploitation of bacterial viruses as antimicrobial agents has experienced something of a renaissance in recent years. In part, this reflects the need to find alternatives to conventional antibiotics following the continued emergence of drug resistant pathogens. Recent reviews highlight this potential, but also emphasize limitations that are inherent in the use of bacteriophages (7, 8).

10 In general, bacteriophages exhibit significant strain specificity, meaning that they are only active against a restricted range of individual strains. The dynamics of the interaction between a bacteriophage and its bacterial host involve the ready selection of host mutants that are resistant to bacteriophage attack. Other issues of concern include the
15 potential contamination of bacteriophage preparations with viable host bacteria and the potential for bacteriophages to contribute to gene flow and the spread of virulence factors (9). The carriage of toxin genes by bacteriophages is especially well documented, and examples include cholera toxin (10), botulinum toxin (9), shiga toxin (11) and diphtheria toxin (9). Despite these reservations, bacteriophages have been used experimentally to
20 control *E. coli* (12), *Staphylococcus aureus* (13) and vancomycin resistant *Enterococcus faecium* (14) in mouse models. Bacteriophage therapy is being investigated for the control of *Campylobacter* (15) and *E. coli* (16) in chickens. With respect to clostridia, a study that targeted *C. difficile* in the hamster model has been reported (17). Further, the FDA has recently extended GRAS approval to a bacteriophage (LISTEX™, EBI Food
25 Safety) for the control of *Listeria* in all food products (18).

In addition to the use of intact bacteriophages, there is the possibility of using bacteriophage endolysins as antimicrobial agents. The final stage of the bacteriophage
30 life cycle involves the lysis of the bacterial host cell to release the pool of newly replicated intact bacteriophage particles. In general, this is achieved by a two stage process in which the carefully timed production of a membrane disruptive holin allows a cell wall degradative endolysin to access its peptidoglycan target. The endolysin enzyme is not secreted but released from the cell by the action of the holin and by its own capacity to degrade the cell wall. Once released, the endolysin can attack peptidoglycan
35 from outside the cell, a phenomenon that has been observed from the time of early bacteriophage studies: it is referred to as 'lysis from without'. The structure of most

characterised bacteriophage endolysins is modular, with a catalytic domain and a distinct cell wall binding domain (CBD). The catalytic domain can vary and in most cases it is either an amidase or a muramidase. The CBD has a lectin-like ability to recognise sugar motifs on the bacterial cell surface, and the varied specificity involved gives the endolysins their characteristic targeting to a specific taxonomic group (19, 20).

Gasson *et al.* pioneered the exploitation of bacteriophage endolysins both as novel antimicrobial agents and as the basis of a novel detection technology using *Listeria* and *Clostridium* as model systems (21). Subsequently, the potential of endolysins as targeted antimicrobial agents has been widely recognised (22) with published examples that target *Bacillus anthracis* (23), *Streptococcus pneumoniae* (24) and *Enterococcus faecalis* (25). With respect to *Listeria*, significant additional work has been undertaken by Martin Loessner at ETH, Switzerland (19, 20). In addition, an endolysin active against *Clostridium perfringens* has been characterised (26).

Characterization of a novel bacteriophage lysin and methods of use thereof

The temperate bacteriophage Φ CD27 was isolated from *Clostridium difficile* culture collection strain NCTC 12727. Φ CD27 was tested against 25 other *C. difficile* strains and shown to be effective against 4 other strains, including the type strain 11204. The bacteriophage genomic DNA was extracted and sequenced and the endolysin sequence identified by BLAST search. The sequence shows clear amino acid and nucleotide homology to published *C. difficile* bacteriophage endolysins (Φ CD119, Φ C2, prophages 1 and 2 in sequenced *C. difficile* CD630). The lysin was subcloned into pET15b and expressed in *E. coli* with a 6xHis tag. The lysin was partially purified on a nickel column and shown to lyse both phage-sensitive and -insensitive strains, evidenced by a drop in optical density upon incubation at 37°C. Of 30 strains tested all showed lysis, including strains of the virulent ribotype 027. A number of other bacteria from a range of genera showed no susceptibility to the lysin. However some activity was observed against *C. bifermentans*, *C. sordelli*, *Bacillus cereus*, *B. subtilis* and very limited activity against *Listeria ivanovii*. Specific activity of the partially purified lysin varied depending on the *C. difficile* strain. Accordingly, the lysin disclosed herein represents a potent new weapon for the treatment and detection of *C. difficile* pathogenesis.

The lysin identified and characterized herein is a novel composition of matter which may be utilized to treat *C. difficile* infections and other bacterial infections in humans and in

animals. According to this invention, the Φ CD27 lysin may be produced according to methods known in the art. It may be isolated for use from the virus grown for this purpose. Preferably, however, it is produced by recombinant means disclosed herein and by alternate means known to those skilled in the art. Relevant sub-portions of the molecule are characterized for their ability to specifically bind to bacteria and to lyse those bacteria. These molecular sub-portions may be produced and used separately or together as in the native molecule.

Discovery, cloning and activity of Φ CD27 lysin

Lysate production and activity assays were performed as described (27). *C. difficile* strain NCTC 12727 (available from the Health Protection Agency, Colindale, London – deposited by S. Tabaqchali, St. Bart's Hospital, London in 1992 isolated from faeces) was grown for 24 h anaerobically at 37°C in BHI+C (BHI (Oxoid) supplemented with vitamin K (10µl 0.5% v/v /l) hemin (5mg/l), resazurin (1mg/l) and L-cysteine (0.5g/l)). Bacteriophage production was induced for 24 h with mitomycin C (Sigma), at a final concentration of 3 µg/ml. Cultures were centrifuged at 4,000 x g for 20 mins at 4°C and supernatants were filtered through 0.45 µm filter units (Millipore) and stored at 4°C. The supernatant was spotted in 25 µl portions onto BHI plates (1.5% agar) overlaid with BHI soft agar (0.75%) incorporating 150 µl of an overnight *C. difficile* BHI+C culture, and incubated overnight anaerobically at 37°C. Cultures (see Table 1) were tested in duplicate and clear plaque formation from 12727 supernatant was identified on 4 strains – *C. difficile* 11204 (type strain), 11205, 11207 and 11209. Plaques from strain 11204 were picked with a sterile Pasteur pipette into 250 µl BHI+C and incubated overnight at 4°C. The presence of a bacteriophage – Φ CD27 - was confirmed by electron microscopy, which indicated it belonged to the order Caudovirales (28)(Fig. 1). In total 25 strains of *C. difficile* were induced with mitomycin C and the supernatants cross-tested against all 25 strains. Φ CD27 was the only plaque-forming unit discovered by this method. The infrequency of bacteriophage discovery from *C. difficile* has also been noted in previous publications which found 2 bacteriophage producers from 94 isolates (29) or 3 producers from 56 isolates (30).

To increase the titre, 100 µl of the plaque eluate was mixed with 100 µl of a 24h culture of *C. difficile* strain 11204 in 5 ml BHI soft agar and plated onto BHI agar. Overnight anaerobic incubation at 37°C gave near-confluent lysis and elution for 2 h into 5 ml BHI+C gave a titre of 2×10^6 pfu/ml. The titre was increased by consecutive incubations

in 11204 liquid culture, growing the cells in 25ml BHI+C cultures to early to mid-log phase, giving an optical density (OD) to allow a ratio of bacteriophage : cells of at least 4 : 1. This method gave complete clearing of the bacterial suspension and 2 passages gave a titre of 2.5×10^{11} pfu/ml. For DNA extraction, cells at OD 0.3 were inoculated with
 5 filtered lysate to a multiplicity of infection of c. 7. An incubation of 3 h gave complete lysis and the supernatant was harvested and filtered as before and two 50 ml portions were used in a Qiagen λ midikit (Qiagen), giving a yield of c.160 μ g bacteriophage genomic DNA.

10 Sequencing and assembly of the bacteriophage Φ CD27 genome was performed by the Biochemistry DNA Sequencing Facility (University of Cambridge, UK) using the Phred-Phrap program. The circular genome is 50,930 bp and contains 75 proposed open reading frames (orfs) (Fig.2). Many of these show significant homology to identified bacteriophage ORFs, including those from *C. difficile* bacteriophages Φ CD119 (31) and
 15 Φ C2 (32). ORFs were analysed by Artemis (33) with BlastP searches (34, 35) which were run via BITS (Harpندن). The proposed Φ CD27 lysin sequence is 816 bp, coding for a 271 amino acid predicted protein which shows homology to *N*-acetylmuramoyl-L-alanine amidase. Both the nucleotide and amino acid sequences (Fig.3) align to published sequences from *C. difficile* bacteriophages and prophages (Fig. 4), with the
 20 greatest homology (95.9% nucleotide and amino acid identity) being to Φ C2.

The Φ CD27 lysin sequence was amplified from genomic DNA using primers to create an *Nde*I site (CATATG) around the initial Met residue (primer CD27L_NDE, 5'-TTA CAT ATG AAA ATA TGT ATA ACA GTA GG [SEQ ID NO: 3], Sigma Genosys) and a *Xho*I
 25 site (CTCGAG) downstream of the coding sequence (primer CD27L_XHO, 5'-CAA CCA CCT CGA GTT GAT AAC [SEQ ID NO: 4], to facilitate subcloning in the expression vector pET15b (Novagen). Amplification was performed with high fidelity Phusion DNA polymerase (0.02 U/ μ l, Finnzymes) in a 50 μ l reaction containing 1 x Phusion buffer, 200 μ M dNTPs, 0.5 μ M of each primer, 200 ng genomic DNA template. Amplification
 30 conditions were an initial denaturation of 98°C for 30 s followed by 30 cycles of denaturation (98°C 10 s), annealing (58°C 30 s) and extension (72°C 15 s) then a final extension of 72°C for 5 min. Blunt end PCR products were purified using SureClean (Bioline) and given 3' A-overhangs in a 50 μ l reaction containing 1x AmpliTaq buffer, 0.2mM dATP and 1U AmpliTaq DNA polymerase (Applied Biosystems) incubated for 20
 35 min at 72°C. Products were purified with SureClean then ligated into pCR2.1 using the TA cloning kit (Invitrogen). Ligation products were transformed into TOP10 chemically

competent *E. coli* (Invitrogen) and positives were selected on L agar supplemented with 100 µg/ml ampicillin and overlaid with 40 µl of a 40 mg/ml X-gal solution for blue-white selection. Plasmid DNA was extracted using a plasmid mini kit (Qiagen) and inserts were sequenced using vector primers and the BigDye v3.1 sequencing kit (Applied Biosystems). A clone showing 100% sequence homology to the original lysin sequence but with the added *NdeI* and *XhoI* sites was restricted to release the insert. This was gel purified (Qiaex II, Qiagen) and ligated using Fast-Link DNA ligase (Epicentre), into pET15b so that the lysin sequence was expressed downstream of a 6-histidine tag under the control of the high expression T7 promoter with the IPTG-inducible lac operator (Fig.5). Ligation products were transformed first into TOP10 cells for sequence confirmation then into chemically competent BL-21(DE3) cells (Invitrogen) for protein expression. The lysin sequence from prophage 1 of the sequenced *C. difficile* (36) was synthesised by Genscript Corp. (Piscataway, USA) into the vector pUC57 and subcloned for His-tagged expression in the same way using primers CD630L1_NDE (5'-TGC TCA TAT GAA AAT AGG TAT AAA TTG) [SEQ ID NO: 5] and M13 forward (5'-GTA AAA CGA CGG CCA GT) [SEQ ID NO: 6] which amplified the lysin with some vector DNA including a *XhoI* site.

His-tagged lysin was expressed as suggested by the manufacturer in BL-21(DE3) cells grown in 10 ml L broth with 100 µg/ml ampicillin to OD₆₀₀ 0.4 then induced with 0.5 mM IPTG (Melford Biosciences) for 3-4 h. Cells were harvested by centrifugation at 4000 x g and 4°C for 20 min then resuspended in 1 ml buffer (20 mM Tris-HCl pH 8, 50 mM NaCl) and transferred to 2 ml screw cap tubes. Crude protein lysate was obtained by cell disruption with 0.1 mm acid-washed glass beads (Sigma) in a FastPrep FP120 cell disrupter (Savant) with 4 x 30 s bursts (speed 10), incubating on ice for 5-10 min between bursts. Debris was pelleted by centrifugation at 13,000 x g for 20 min at 4°C and the supernatant stored at 4°C. Crude lysates were also produced from cells containing the lysin grown without IPTG induction and cells containing the empty pET15b vector grown with and without induction. Protein content was measured using the Bradford reagent (Bio Rad) and 10 µg portions were electrophoresed on 10% NuPage Novex Bis Tris gels in MOPS buffer (Invitrogen). Presence of the His-tagged lysin was confirmed by Western blotting using an anti His-Tag monoclonal antibody (Novagen). Proteins were transferred to PVDF membrane using NuPage buffer (Invitrogen) and detection was as described by Qiagen (Qiaexpress detection and assay handbook) with anti-mouse IgG as the secondary antibody and colorimetric detection with Sigma Fast BCIP/NBT alkaline phosphatase substrate. This demonstrated high expression of a His

tagged band of c. 33 kDa in IPTG- induced lysates and also lower expression in uninduced lysates (Fig.6).

Lysis of *C. difficile* cells of strains 11204 and 11207 by crude lysates was assessed using the method described by Loessner *et al* (37). Cells of strain 11204 were grown to end-log phase, 1.8 ml aliquots were harvested by centrifugation into screw cap tubes (13,000 x g, 2 min) and pellets were flash-frozen in liquid nitrogen and stored at -20°C. Pellets were resuspended on ice in 900 µl 20 mM Tris-HCl pH 8 and added to a cuvette containing 100 µl crude protein lysate then the drop in OD₆₀₀ was monitored for 1 h with mixing before reading. With this system the *C. difficile* cells showed a certain amount of lysis in the buffer, although lysis with the ΦCD27 lysin crude extract was more rapid and extensive. However, a subsequent test with the induced empty pET15b vector crude lysate demonstrated an equivalent lysis, suggesting the activity of *E. coli* lysozymes. To avoid this problem the ΦCD27 and CD630L1 lysins were affinity-purified using the Qiagen NiNTA kit. BL-21(DE3) cells were grown to OD₆₀₀ 0.6 in 250 ml L broth containing 100 µg/ml ampicillin then induced for 5 h with IPTG at a final concentration of 1 mM. Cells were harvested by centrifugation at 4000 x g and 4°C for 20 min and pellets stored at -20°C. Protein was purified under native conditions and purification was confirmed by NuPage gel analysis (Fig. 7). This method produced partially purified protein of which the majority was lysin, with a yield of 2.3 mg total protein in the first ΦCD27 eluate (E1) and 0.5 mg in the second (E2). Incubation of dilutions of the E1 eluate showed rapid lysis of strain 11204 cells compared to an eluate from cells prepared in the same way but expressing the empty pET15b vector. However, CD630L1 E1 eluate did not lyse strain 11204 and there was no synergistic effect with ΦCD27 lysin.

Lysis assays continued in multiwell plates using the Bioscreen C (LabSystems) and NiNTA-partially purified lysin extract in elution buffer (EB, Qiagen). Initially assays used c.7 µg protein in a total volume of 30 µl EB and 270 µl cells as in the spectrophotometer assays. Assays were set up on ice then transferred to the Bioscreen C pre-heated to 37°C and the program was run as follows – sampling every 2 min with 10 s shake before sampling at an optical density of 600 nm. Each assay was run with two wells of buffer only and 2 wells of lysin, all 4 wells being inoculated from the same bacterial cell suspension. In this system lysis in the lysin wells of sensitive strains was rapid – a difference being notable within 5 min. However, lysis of the cells in buffer-only controls was also obvious, albeit at a much slower rate than the lysin-induced lysis (Fig.8).

When both *C. difficile* and other bacterial cells were grown to end log and harvested onto ice without freezing then assayed as soon as possible the buffer-only lysis was reduced or totally absent (Fig. 9) and lysis of all other species was absent with the notable exceptions of *Clostridium bifermentans*, *Clostridium sordelli*, *Bacillus cereus* and to lesser extents *B. subtilis* and *Listeria ivanovii* (Fig. 11, Table 2). Additional strains representative of the AT rich *Clostridium*-like component of the GI tract microflora were tested for sensitivity to the Φ CD27 lysin. As shown in Table 3, none of those tested were sensitive to the lysin.

Using fresh cells gave a less rapid onset of lysis in *C. difficile* with a notable lag of up to 12 mins (Fig. 9). All *C. difficile* strains were re-tested using 3.5 μ g lysin isolated from a second NINTA column (tested to show equal lysis to the first purification; Fig. 10). In both cases, using fresh or frozen cells, the sensitivity profiles were the same with all 30 strains showing clear sensitivity to the lysin (Table 1).

The pH profile of the Φ CD27 lysin was tested using the sensitive strain 11204 – activity showed very little variation within a fairly large pH range, tested at pH 4.5, 5.8, 6.5, 7.0, 7.3 (usual pH of PBS), 7.6 and 8.3 (Fig. 12). A dilution series showed that although the activity with 10.5 μ g protein in the 300 μ l assay was maximal, good lysis was also seen with 3.5 μ g and 0.7 μ g. However, 0.35 μ g gave a response only slightly below the buffer controls and lower amounts showed no lysis within the 45 min assay.

The delivery of the Φ CD27 lysin to the GI tract could be achieved by the use of physical encapsulation or a recombinant commensal microorganism such as a member of the lactic acid bacteria. *Lactococcus lactis* has established potential in this regard and thus sub-cloning and expression of the Φ CD27 lysin in this species was demonstrated. The Φ CD27 lysin sequence was subcloned into the vector pUK200His. This is a derivative of the *nisA* translational fusion plasmid pUK200 (38) constructed by restriction of pUK200 with *NcoI*, end-filling, then insertion of an oligomer encoding a 6-histidine tag (AGT CAT CAC CAT CAC CAT GC) [SEQ ID NO: 7] downstream from the nisin-inducible promoter. When recircularised, this recreated an *NcoI* site for subcloning (Horn *et al.*, unpublished). Vector pUK200His was restricted with *NcoI* and end-filled with T4 DNA polymerase (Promega) to create the first ATG codon for a translational fusion under control of the *nisA* promoter. The *phicd27I* sequence was amplified from the CD27L-NDE...CD27L-XHO PCR product subcloned in pCR2.1 (see above). Primers CD27LCOD2_F (5'-AAA ATA TGT ATA ACA GTA GGA CAC) [SEQ ID NO: 8] and M13

forward (5'-GTA AAA CGA CGG CCA GT) [SEQ ID NO: 9] amplified the full sequence from the second codon AAA and some of the vector sequence, giving an *EcoRI* site immediately after the lysin coding sequence. Amplification was as described above but with an annealing temperature of 56°C. Both the PCR product and the *NcoI*-cut, end-filled pUK200His vector were restricted with *EcoRI* and ligated together to create the His-tagged translational fusion under control of the *nisA* promoter. Ligation products were transformed into electrocompetent *E.coli* strain MC1022 for sequence verification, with positive transformants being selected on chloramphenicol (15 µg/ml). Purified plasmid preparations were then transformed into electrocompetent *Lactococcus lactis* strain FI10676 and selected on GM17 agar supplemented with 5 µg/ml chloramphenicol.

L. lactis strains expressing pUK200His-phiCD27L or pUK200His empty vector control were grown in 10 ml GM17 broth with 5 µg/ml chloramphenicol at 30°C static. 100 µl of an overnight culture was used to inoculate pre-warmed broth and the culture grown to midlog (OD₆₀₀ 0.5). Expression was induced with 1ng/ml nisin for 5 h at 30°C and crude protein lysates were produced as described for *E. coli* in 20 mM Tris-HCl pH 8.0, 50 mM NaCl. A demonstration of lactococcal expression of ΦCD27 lysin is presented as a protein gel analysis (Fig. 13). The sensitivity of *Clostridium difficile* strain 11204 to the ΦCD27 endolysin expressed in *Lactococcus lactis* was demonstrated using crude protein extracts as is shown in Fig. 14.

Table 1 (overleaf) Strains of *Clostridium difficile* used in bacteriophage and lysin assay tests. Sources a: National Collection of Type Cultures, Central Public Health Laboratory, 61, Colindale Ave, London; b: Dr Jonathan Brazier, Anaerobe Reference Unit, Dept. of Medical Microbiology and PHLS, University Hospital of Wales, Heath Park, Cardiff; c: Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ), GmbH Inhoffenstrasse 7 B, 38124 Braunschweig, Germany. S = shows sensitivity to infection by bacteriophage ΦCD27, R = insensitive to infection by bacteriophage ΦCD27, nt = not tested; L = lysed by ΦCD27 lysin.

Strain	Source	Details	Bacterio-phage	Lysin
NCTC 11204 N1	a	Meconium from neonates, 1970	S	L
NCTC 11205 N2	a	Meconium from neonates, 1970	S	L
NCTC 11206 N3	a	Meconium from neonates, 1970	R	L
NCTC 11207 N4	a		S	L
NCTC 11208 N5	a		R	L
NCTC 11209 N6	a		S	L
NCTC 11223 335722	a	Faeces	R	L
NCTC 12726	a	faeces, 35S methionine protein type A	R	L
NCTC 12727	a	faeces, 35S methionine protein type B	R	L
NCTC 12728	a	faeces, 35S methionine protein type C	R	L
NCTC 13287	a	R7404	nt	L
NCTC 13307	a	Strain 630	nt	L
NCTC 12731	a	faeces, 35S methionine protein type W	R	L
NCTC 13366	a	R2029	nt	L
NCTC 11382 74/1451	a	Blood culture, New Zealand, 1980	R	L
R23 521	b	Ribotype 118	R	L
R23 524	b	Ribotype 001	R	L
R23 613	b	Ribotype 027	R	L
R23 614	b	Ribotype 106	R	L
R23 621	b	Ribotype 179	R	L
R23 635	b	Ribotype 015	R	L
R23 639	b	Ribotype 014	R	L
R23 642	b	Ribotype 012	R	L
R23 720	b	Ribotype 005	R	L
R23 727	b	Ribotype 001	R	L
R23 732	b	Ribotype 027	R	L
R23 737	b	Ribotype 106	R	L
G83/03	b	Ribotype 180	R	L
12056	c	Rumen of new-born lamb	nt	L
12057	c	Rumen of new-born lamb	nt	L

Table 2. Spectrum of activity of Φ CD27 lysin against a range of bacteria. - = no lysis, +++ = clear lysis, + = limited lysis

Bacterium	Strain	Lysin test
<i>Bacillus cereus</i>	ATCC 9139	++
<i>Bacillus subtilis</i>	ATCC 6633	+
<i>Bifidobacterium adolescentis</i>	DSMZ 20083	-
<i>Bifidobacterium angulatum</i>	DSMZ 20098	-
<i>Bifidobacterium bifidum</i>	DSMZ 20082	-
<i>Bifidobacterium longum</i>	DSMZ 20219	-
<i>Bifidobacterium pseudocatenulatum</i>	DSMZ 20438	-
<i>Clostridium coccooides</i>	NCTC 11035	-
<i>Clostridium perfringens</i>	NCTC 3110	-
<i>Clostridium bifermentans</i>	C22/10	+++
<i>Clostridium bifermentans</i>	NCTC 13019	++
<i>Clostridium sordelli</i>	NCTC 13356	++
<i>Clostridium sporogenes</i>	ATCC 17886	-
<i>Enterococcus faecalis</i>	FI10734	-
<i>Enterococcus faecium</i>	FI10735	-
<i>Enterococcus hirae</i>	FI10477	-
<i>Escherichia coli</i> wild type	K12	-
<i>Lactobacillus bulgaricus</i>	FI10643	-
<i>Lactobacillus casei</i>	FI107346	-
<i>Lactobacillus gasseri</i>	NCIMB1171	-
<i>Lactobacillus johnsonii</i>	FI109785	-
<i>Lactobacillus plantarum</i>	FI108595	-
<i>Lactobacillus rhamnosus</i>	FI107347	-
<i>Lactobacillus sakei</i>	FI10645	-
<i>Lactococcus lactis</i>	MG1316	-
<i>Lactococcus garvieae</i>	FI08174	-
<i>Listeria innocua</i>	NCTC 11288	-
<i>Listeria ivanovii</i>	NCTC 11007	+
<i>Listeria monocytogenes</i>	NCTC 5412	-
<i>Micrococcus luteus</i>	FI106340	-
<i>Pediococcus pentosaceus</i>	FI10642	-
<i>Pediococcus acidilactici</i>	FI10738	-
<i>Salmonella enterica</i> serovar Typhimurium	FI10739	-
<i>Salmonella enteritidis</i>	FI10113	-

Bacterium	Strain	Lysin test
<i>Staphylococcus aureus</i>	FI10139	-
<i>Streptococcus anginosus</i>	FI10740	-
<i>Veillonella atypical</i>	FI10741	-

In addition to the above, it is noted that many additional strains representative of commensal strains which are desirably not harmed in order to maintain health gut physiology, are not harmed by contact with the lysin according to this invention. All of the following *Clostridium* species tested against Φ CD27, all from DSMZ, all gave no lysis. These strains were specifically chosen on the basis of being representative of the main *Clostridium* clusters commonly found in the human gut, as references Eckberg *et al.* (2005) *Science* **308** 1635- and supplementary material, and Kikuchi *et al.* (2002) *Microbiol. Immunol.* **46**, 353 and refs therein:

10

Table 3. GI tract *Clostridium* and clostridium-like species not lysed by Φ CD27 lysin.

	Bacterial cells	Deposit	Cluster
15	<i>Clostridium cellobioparum</i>	DSMZ 1351	Cluster III
	<i>Clostridium leptum</i>	DSMZ 753	Cluster IV
	<i>Clostridium nexile</i>	DSMZ 1787	Cluster XIVa
	<i>Clostridium colinum</i>	DSMZ 6011	Cluster XIVb
	<i>Clostridium innocuum</i>	DSMZ 1286	Cluster XIVb
20	<i>Clostridium ramosum</i>	DSMZ 1402	Cluster XVIII
	<i>Eubacterium barkeri</i> (formally <i>C. barkeri</i>)	DSMZ 1223	Cluster XV
	<i>Anaerococcus hydrogenalis</i>	DSMZ 7454	Cluster XIII

All *C. difficile* strains were tested against *C. difficile* strain 630 prophage 1 lysin expressed in *E. coli* by the same method and the CD630L1 lysin gave no lysis.

25

Cell viability

To measure the effect of phiCD27 lysin on cell viability, replicate assays were set up under anaerobic conditions using pre-reduced buffers and media. Cells were grown to end log, harvested by centrifugation in anaerobic conditions then resuspended in PBS buffer at pH 7.3. 10-fold dilutions were made in PBS from c. 1×10^8 cells to c. 1×10^3 cells; 10 μ l aliquots of these were spotted onto BHI agar at time 0 to allow estimation of the number of cells in each assay. Assays were in duplicate and contained either 100 μ g partially-purified endolysin (E1) or an equivalent volume (50 μ l) of buffer (EB) and cells to a final volume of 300 μ l. After 2 h incubation with continuous gentle shaking, 30 μ l samples were taken for 10-fold serial dilutions in PBS; 10 μ l aliquots of these dilutions were spotted onto BHI agar and the remaining 270 μ l assay from one of each duplicate pair was plated to allow cell enumeration.

Assays containing c. 1×10^8 cells at time 0 showed a drop of 1 log after 2 h incubation, while assays to which 1×10^7 cells or 1×10^6 cells had been added showed a drop of 2 log compared to buffer controls. In assays with lower initial cell numbers the lysin was more effective, with only 4 viable colonies being recovered from an assay inoculated with 1×10^5 cells and no live cells remaining in assays of 1×10^4 cells or less.

The above viability assay was then repeated using a 400 μ l aliquot of E1 that had been subjected to a buffer exchange using 2ml Zeba Desalt spin columns (Pierce) to replace the Ni-NTA elution buffer (EB) with 20mM sodium phosphate pH 6 (NP). The lysin in NP buffer showed equivalent activity against frozen cells of *Clostridium difficile* 11204 to the original NiNTA E1 (figure 15). The viability assay was repeated as above using 50 μ g E1-NP or NP buffer control and c. 1×10^6 cells; a 2 h incubation with the lysin produced a drop of 3 log compared to buffer controls.

The above data subsequently formed the basis of a published scientific manuscript (Mayer *et al.*, 2008, *J. Bacteriol.* **190**:6734-6740), the disclosures of which are incorporated herein by reference.

Domain swappingEngineering new enzymatic domains onto the Φ CD27L endolysin by splice overlap PCR

- 5 The endolysin LM-4 from bacteriophage Φ LM4, active against *Listeria monocytogenes*, was demonstrated to cause effective lysis of host cells (see GB 2,255,561 B). The endolysin is 864 bp long, giving a protein of 287 amino acids which shows homology to pfam02557, VanY, D-alanyl-D-alanine carboxypeptidase in the first part of the protein and COG5632, N-acetylmuramoyl-L-alanine amidase over the whole sequence (NCBI
- 10 Blast). The first half of the sequence, encoding the proposed enzyme active domain (EAD), was inserted upstream of either the CD27L cell wall binding domain (CBD, from Asn 180 to the final Arg 270) or the entire 270 amino acid enzyme by splice overlap extension PCR. The LM4 enzymatic domain was amplified by PCR from plasmid pFI567 (Payne *et al.*, 1996 *FEMS Microbiology Letters* **136**: 19-24) using primers LM4Nde 5'-
- 15 GGA TGA TTA CAT ATG GCA TTA ACA G [SEQ ID NO: 10], to create an NdeI site at the ATG of LM4, and one of two splice overlap primers: LM4-splice-CD27LE 5'- TAT ACA TAT TTT CAT GTT TTG TGT CGC AGT [SEQ ID NO: 11], which represents nucleotides 439-453, Thr147 to Asn 151, of the LM4 sequence with a tail that matches the first 15 nucleotides of the CD27L enzyme to give LM4 EAD-CD27L EAD-CBD; or
- 20 LM4-splice-CD27L 5'- TTT AAC TCC CTC ATT GTT TTG TGT CGC AGT [SEQ ID NO: 12], which represents nucleotides 439-453, Thr147 to Asn 151, of the LM4 sequence with a tail that matches the proposed C-terminal binding domain of CD27L from Asn 180 to Arg 270, to give LM4 EAD-CD27L CBD. Similarly, the CD27L entire sequence or CBD were amplified from Φ CD27L-pET15b using a primer from the vector, T7T 5'- GCT AGT
- 25 TAT TGC TCA GCG G [SEQ ID NO: 13] and splicing primers which had tails to match the end of the LM4 EAD sequence - CD27LEspliceLM4 5'-ACT GCG ACA CAA AAC ATG AAA ATA TGT ATA ACA GT [SEQ ID NO: 14] for the entire sequence, where the last 20 nt of the primer encode the beginning of the CD27L sequence from Met 1; and CD27LspliceLM4 5'- CT GCG ACA CAA AAC AAT GAG GGA GTT AAA C [SEQ ID NO:
- 30 15] for the CBD only, where the last 16 nt of the primer encodes the proposed CBD of the CD27L sequence from Asn180. PCR was performed with Phusion (Finnzymes) with the conditions recommended by the manufacturer, using annealing temperatures for 5 cycles to match the portion of the splicing primer which gave 100% match to the original template, then 20 cycles at an annealing temperature to match the entire splicing primer.
- 35 Products were purified using SureClean (Bioline) and resuspended in a volume of 50 μ l. These templates were diluted 100-fold and 1 μ l aliquots used in a PCR reaction using

the original outer primers – LM4Nde and T7T – at an annealing temperature to allow splicing of the two sequences (54°C). The final products were purified with SureClean, restricted with NdeI and XhoI and subcloned into pET15b to produce His-tagged LM4-CD27LE and LM4-CD27L. These plasmids were then transformed into *E. coli* and their sequences confirmed.

Both crude extracts and NiNTA- purified extracts of the composite enzymes were produced, analysed by SDS-PAGE and Western blotting and assayed as described previously (see Figure 16). Both His-tagged LM4-CD27LE and LM4-CD27L were present at high levels in crude extracts. When incubated with frozen cells of *C. difficile* 11204 in PBS buffer pH 5.8, 10 µg NiNTA-purified extracts produced a rapid lysis compared to buffer controls (see Figure 17), with LM4-CD27LE showing a similar speed of lysis to the native CD27L. An equivalent activity was seen using PBS buffer at pH 7.3 as the cell diluent.

In a viability assay, NiNTA-purified eluates of both LM4-CD27LE and LM4-CD27L produced a drop in viable counts (see Figure 17). Using 50 µg NiNTA E1, assays containing c.1x10⁴ cells showed a reduction of at least 1 log after 2 h incubation compared to buffer controls. This drop was not as great as that seen with the native enzyme, but proves the principle that the addition of alternate enzyme domains can produce active novel enzymes which have the capability to kill *C. difficile*.

Nucleotide and amino acid sequences of wildtype LM4 and domain swapped lysins

LM4

ATGGCATTAAACAGAGGCATGGCTAATTGAAAAAGCAAATCGCAAATTGAATACGTCA
GGTATGAATAAAGCTACATCTGATAAGACTCGGAATGTAATTAATAAATGGCAAAA
GAAGGGATTTATCTTTGTGTTGCGCAAGGTTACCGCTCAACAGCGGAACAAAATGC
GCTATATGCACAAGGGAGAACCAAACCTGGAGCGATTGTTACTAATGCTAAAGGTG
GGCAATCTAATCATAATTTCCGGTGTAGCAGTTGATTTGTGCTTGATACGAGCGACG
GAAAAGATGTTATTTGGGAGTCGACAACCTCCCGGTGGAAAAAGGTTGTTGCTGCT
ATGAAAGCGGAAGGATTCGAATGGGGCGGAGATTGGAAAAGTTTTAAAGACTATCC
GCATTTTGAAGTATGTGACGCTGTAAGTGGTGAGAAAATCCCTACTGCGACACAAAA
CACCAATCCAAACAGACATGATGGGAAAATCGTTGACAGCGCGCCACTATTGCCAA
AAATGGACTTTAAATCAAATCCAGCGCGCATGTATAAATCAGGAACTGAGTTCTTAG

TATATGAACATAATCAATATTGGTACAAGACGTACATCAACGACAAATTATACTACAT
GTATAAGAGCTTTTTCGATGTTGTAGCTAAAAAGATGCAAAGGACGCATCAAAGT
TCGAATTAAGCGCGAAAGACTTACGAATTCCAGTTTGAATAACACAAAATTGAA
TTCTGGGAAAATTAAATGGTATGCACCCAATACAAAATTAGCATGGTACAACAACGG
5 AAAAGGATACTTGGAAGCTCTGGTATGAAAAGGATGGCTGGTACTACACAGCGAACT
ACTTCTTAAATAA [SEQ ID NO: 16]

MALTEAWLIEKANRKLNTSGMKNKATSDKTRNVIKKMAKEGIYLCVAQGYRSTAEQNALY
AQGRTPGAIVTNAKGGQSNHNFVAVDLCLYTSKGKDVWESTTSRWKKVVAAMKA
10 EGFEGGWDWKSFKDYPHFELCDAVSKEIPTATQNTNPNRHDGKIVDSAPLLPKMDFK
SNPARMYKSGTEFLVYEHNQYWKTYINDKLYMYKSFCDVVAKKDAKGRIKVRISAK
DLRIPVWNNTKLNKSGIKWYAPNTKLAWYNNKGYLELWYEKDGWYYTANYFLK [SEQ
ID NO: 17]

15 LM4-CD27LE

ATGGCATTAAACAGAGGCATGGCTAATTGAAAAAGCAAATCGCAAATTGAATACGTCA
GGTATGAATAAAGCTACATCTGATAAGACTCGGAATGTAATTAATAAATGGCAAAA
GAAGGGATTTATCTTTGTGTTGCGCAAGGTTACCGCTCAACAGCGGAACAAAATGC
20 GCTATATGCACAAGGGAGAACCAAACCTGGAGCGATTGTTACTAATGCTAAAGGTG
GGCAATCTAATCATAATTTCCGGTGTAGCAGTTGATTTGTGCTTGTATACGAGCGACG
GAAAAGATGTTATTTGGGAGTCGACAACCTCCCGGTGGAAAAAGGTTGTTGCTGCT
ATGAAAGCGGAAGGATTCGAATGGGGCGGAGATTGGAAAAGTTTTAAAGACTATCC
GCATTTTGAAGTATGTGACGCTGTAAGTGGTGAGAAAATCCCTACTGCGACACAAAA
25 CATGAAAATATGTATAACAGTAGGACACAGTATTTTAAAAAGTGGAGCATGTACTTCT
GCTGATGGAGTAGTTAACGAGTATCAATACAACAAATCTCTTGACACAGTATTAGCA
GATACATTTAGAAAAGAAGGGCATAAGGTAGATGTAATAATATGCCAGAAAAGCAG
TTTAAACTAAGAATGAAGAAAAGTCTTATAAAATACCTAGAGTTAATAGTGGAGGAT
ATGATTTACTTATAGAGTTACATTTAAATGCAAGTAACGGTCAAGGTAAAGGTTTCAGA
30 AGTCCTATATTATAGTAATAAAGGCTTAGAGTATGCAACTAGAAATATGTGATAAACTA
GGTACAGTATTTAAAAATAGAGGTGCTAAATTAGATAAAAGATTATATATCTTAAATA
GTTCAAAGCCTACAGCAGTATTAATTGAAAGTTTCTTCTGTGATAATAAGAAGATTA
TGATAAAGCTAAGAACTAGGTCATGAAGGTATTGCTAAGTTAATTGTAGAAGGTGT
ATTAAATAAAATATAAATAATGAGGGAGTTAAACAGATGTACAAACATACAATTGTT
35 TATGATGGAGAAGTTGACAAAATCTCTGCAACTGTAGTTGGTTGGGGTTATAATGAT
GGGAAAATACTGATATGTGATATAAAAGATTACGTGCCAGGTCAGACGCAAAATCTT

TATGTTGTAGGAGGTGGCGCATGTGAAAAGATAAGTTCTATTACTAAAGAAAAATTT
ATTATGATAAAAGGTAATGATAGATTTGATACACTTTATAAAGCATTGGATTTTATTAA
TAGATAG [SEQ ID NO: 18]

5 MALTEAWLIEKANRKLNTSGMNKATSDKTRNVIKKMAKEGIYLCVAQGYRSTAEQNALY
AQGRTPGAIVTNAKGGQSNHNFVAVDLCLYTS DGKDWIWESTTSRWKKVVAAMKA
EGFEWGGDWKSFKDYPHFELCDAVS GEKIPTATQNMKICITVGHSILKSGACTSADGVV
NEYQYNKSLAPVLADTFRKEGHKVDVIICPEKQFKTKNEEKSYKIPRVNSGGYDLLIELH
LNASNGQGKGSEVLYYSNKGLEYATRICDKLGT VFKNRGAKL DKRLYILNSSKPTAVLIE
10 SFFCDNKEDYDKAKKL GHEGIAKLIVEGV LNKNNINNEG VKQMYKHTIVYDGEVDKISATV
VGWGYNDGKILICDIKDYPGQTQNL YVGGGACEKISSITKEKFIMIKGNDRFDTLYKAL
DFINR [SEQ ID NO: 19]

LM4-CD27L

15 ATGGCATTAAACAGAGGCATGGCTAATTGAAAAAGCAAATCGCAAATTGAATACGTCA
GGTATGAATAAAGCTACATCTGATAAGACTCGGAATGTAATTAATAAATGGCAAAA
GAAGGGATTATCTTTGTGTTGCGCAAGGTTACCGCTCAACAGCGGAACAAAATGC
GCTATATGCACAAGGGAGAACCAAACCTGGAGCGATTGTTACTAATGCTAAAGGTG
20 GGCAATCTAATCATAATTTCCGGTGTAGCAGTTGATTTGTGCTTGATACGAGCGACG
GAAAAGATGTTATTTGGGAGTCGACAACCTCCCGGTGGAAAAAGGTTGTTGCTGCT
ATGAAAGCGGAAGGATTGAATGGGGCGGAGATTGGAAAAGTTTTAAAGACTATCC
GCATTTTGAACATATGTGACGCTGTAAGTGGTGAGAAAATCCCTACTGCGACACAAAA
CAATGAGGGAGTTAAACAGATGTACAAACATACAATTGTTTATGATGGAGAAGTTGA
25 CAAAATCTCTGCAACTGTAGTTGGTTGGGGTTATAATGATGGGAAAATACTGATATG
TGATATAAAAGATTACGTGCCAGGTCAGACGCAAAATCTTTATGTTGTAGGAGGTGG
CGCATGTGAAAAGATAAGTTCTATTACTAAAGAAAAATTTATTATGATAAAAGGTAAT
GATAGATTTGATACACTTTATAAAGCATTGGATTTTATTAATAGATAG [SEQ ID NO:
20]

30 MALTEAWLIEKANRKLNTSGMNKATSDKTRNVIKKMAKEGIYLCVAQGYRSTAEQNALY
AQGRTPGAIVTNAKGGQSNHNFVAVDLCLYTS DGKDWIWESTTSRWKKVVAAMKAE
GFEWGGDWKSFKDYPHFELCDAVS GEKIPTATQNNNEG VKQMYKHTIVYDGEVDKISAT
VVGWGYNDGKILICDIKDYPGQTQNL YVGGGACEKISSITKEKFIMIKGNDRFDTLYKA
35 LDFINR [SEQ ID NO: 21]

References

1. Kuijper, E. J., Coignard, B. & Tull, P. (2006) *Clin Microbiol Infect* **12 Suppl 6**, 2-18.
2. Anonymous (2006) *Health Statistics Quarterly* **30**, 56-60.
3. Rupnik, M., Dupuy, B., Fairweather, N. F., Gerding, D. N., Johnson, S., Just, I., Lyster, D. M., Popoff, M. R., Rood, J. I., Sonenshein, A. L., Thelestam, M., Wren, B. W., Wilkins, T. D. & von Eichel-Streiber, C. (2005) *J Med Microbiol* **54**, 113-7.
4. Braun, V., Hundsberger, T., Leukel, P., Sauerborn, M. & von Eichel-Streiber, C. (1996) *Gene* **181**, 29-38.
5. Goncalves, C., Decre, D., Barbut, F., Burghoffer, B. & Petit, J. C. (2004) *J Clin Microbiol* **42**, 1933-9.
6. Popoff, M. R., Rubin, E. J., Gill, D. M. & Boquet, P. (1988) *Infect Immun* **56**, 2299-306.
7. Skurnik, M. & Strauch, E. (2006) *Int J Med Microbiol* **296**, 5-14.
8. Projan, S. (2004) *Nat Biotechnol* **22**, 167-8.
9. Brussow, H., Canchaya, C. & Hardt, W. D. (2004) *Microbiol Mol Biol Rev* **68**, 560-602.
10. Davis, B. M. & Waldor, M. K. (2003) *Curr Opin Microbiol* **6**, 35-42.
11. Strauch, E., Schaudinn, C. & Beutin, L. (2004) *Infect Immun* **72**, 7030-9.
12. Chibani-Chennoufi, S., Sidoti, J., Bruttin, A., Kutter, E., Sarker, S. & Brussow, H. (2004) *Antimicrob Agents Chemother* **48**, 2558-69.
13. Matsuzaki, S., Yasuda, M., Nishikawa, H., Kuroda, M., Ujihara, T., Shuin, T., Shen, Y., Jin, Z., Fujimoto, S., Nasimuzzaman, M. D., Wakiguchi, H., Sugihara, S., Sugiura, T., Koda, S., Muraoka, A. & Imai, S. (2003) *J Infect Dis* **187**, 613-24.
14. Biswas, B., Adhya, S., Washart, P., Paul, B., Trostel, A. N., Powell, B., Carlton, R. & Merrill, C. R. (2002) *Infect Immun* **70**, 204-10.
15. Loc Carrillo, C., Atterbury, R. J., el-Shibiny, A., Connerton, P. L., Dillon, E., Scott, A. & Connerton, I. F. (2005) *Appl Environ Microbiol* **71**, 6554-63.
16. Huff, W. E., Huff, G. R., Rath, N. C., Balog, J. M. & Donoghue, A. M. (2004) *Poult Sci* **83**, 1944-7.
17. Ramesh, V., Fralick, J. A. & Rolfe, R. D. (1999) *Anaerobe* **5**, 69-78.
18. Wray, T. (2007) *National Provisioner* **5th July**
19. Loessner, M. J., Kramer, K., Ebel, F. & Scherer, S. (2002) *Mol Microbiol* **44**, 335-49.

20. Loessner, M. J. (2005) *Curr Opin Microbiol* **8**, 480-7.
21. Gasson (1995-2003) *Patents GB 2255561 B (1995); AU 650737B (1994); US 5763251 (1998); US 6083684 (2000); CA 2066387 (2003); EP 0510907B (2003).*
- 5 22. Fischetti, V. A. (2005) *Trends Microbiol* **13**, 491-6.
23. Schuch, R., Nelson, D. & Fischetti, V. A. (2002) *Nature* **418**, 884-9.
- 10 24. Loeffler, J. M., Djurkovic, S. & Fischetti, V. A. (2003) *Infect Immun* **71**, 6199-204.
25. Yoong, P., Schuch, R., Nelson, D. & Fischetti, V. A. (2004) *J Bacteriol* **186**, 4808-12.
- 15 26. Zimmer, M., Vukov, N., Scherer, S. & Loessner, M. J. (2002) *Appl Environ Microbiol* **68**, 5311-7.
27. Sell, T. L., Schaberg, D. R. & Fekety, F. R. (1983) *J Clin Microbiol* **17**, 1148-52.
28. Nelson, D. (2004) *J Bacteriol* **186**, 7029-31.
- 20 29. Mahony, D. E., Bell, P. D. & Easterbrook, K. B. (1985) *J Clin Microbiol* **21**, 251-4.
30. Goh, S., Riley, T. V. & Chang, B. J. (2005) *Appl Environ Microbiol* **71**, 1079-83.
- 25 31. Govind, R., Fralick, J. A. & Rolfe, R. D. (2006) *J Bacteriol* **188**, 2568-77.
32. Goh, S., Ong, P. F., Song, K. P., Riley, T. V. & Chang, B. J. (2007) *Microbiology* **153**, 676-85.
- 30 33. Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M-A. Rajandream and B. Barrell . (2000) *Bioinformatics* **16**, 944-945.
34. Altschul, S. F., Thomas L. Madden, Alejandro A. Schäffer, & Jinghui Zhang, Z. Z., Webb Miller, and David J. Lipman (1997) *Nucleic Acids Res.* **25**, 3389-3402
- 35 35. Schäffer, A. A., L. Aravind, Thomas L. Madden, Sergei , Shavirin, J. L. S., Yuri I. Wolf, Eugene V. Koonin, and & Altschul, S. F. (2001) *Nucleic Acids Res.* **29**, 2994-3005.
- 40 36. Sebaihia, M., Wren, B. W., Mullany, P., Fairweather, N. F., Minton, N., Stabler, R., Thomson, N. R., Roberts, A. P., Cerdano-Tarraga, A. M., Wang, H., Holden, M. T., Wright, A., Churcher, C., Quail, M. A., Baker, S., Bason, N., Brooks, K., Chillingworth, T., Cronin, A., Davis, P., Dowd, L., Fraser, A., Feltwell, T., Hance, Z., Holroyd, S., Jagels, K., Moule, S., Mungall, K., Price, C., Rabinowitsch, E., Sharp, S., Simmonds, M., Stevens, K., Unwin, L., Whithead, S., Dupuy, B., Dougan, G., Barrell, B. & Parkhill, J. (2006) *Nat Genet* **38**, 779-86.
- 45 37. Loessner, M. J., Wendlinger, G. & Scherer, S. (1995) *Mol Microbiol* **16**, 1231-41.
- 50 38. Wegmann, U., Klein, J. R., Drumm, I., Kuipers, O. P. & Henrich, B. (1999) *Appl Environ Microbiol* **65**, 4729-33.

CLAIMS

1. An isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1, or a fragment, variant, derivative or fusion thereof which is capable of binding specifically to and/or lysing cells of *Clostridium difficile*.
5
2. A polypeptide according to Claim 1 wherein the polypeptide is not the lysin of bacteriophage Φ CD119, bacteriophage Φ C2 or prophages 1 or 2 of *Clostridium difficile* strain 630 (CD630).
10
3. A polypeptide according to Claim 1 or 2 wherein the fragment, variant, derivative or fusion thereof exhibits at least 60% identity to the amino acid sequence of SEQ ID NO: 1.
- 15 4. A polypeptide according to any one of the preceding claims wherein the fragment, variant, derivative or fusion thereof is not a naturally occurring lysin of a bacteriophage of *Clostridium difficile*.
5. An isolated polypeptide according to any one of the preceding claims capable of binding specifically to cells of *Clostridium difficile*.
20
6. An isolated polypeptide according to any one of the preceding claims capable of lysing cells of *Clostridium difficile*.
- 25 7. An isolated polypeptide according to any one of the preceding claims capable of binding specifically to and lysing cells of *Clostridium difficile*.
8. An isolated polypeptide according to any one of the preceding claims comprising the amino acid sequence of SEQ ID NO:1.
30
9. An isolated polypeptide according to Claim 8 consisting of the amino acid sequence of SEQ ID NO:1
10. An isolated polypeptide according to any one of the preceding claims comprising or consisting of a fragment of the amino acid sequence of SEQ ID NO:1.
35

11. An isolated polypeptide according to Claim 10 wherein the fragment comprises at least 50 contiguous amino acids of SEQ ID NO: 1, for example at least 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260 or 265 contiguous amino acids of SEQ ID NO: 1.
- 5 12. An isolated polypeptide according to Claim 10 or 11 wherein the fragment comprises or consists of the enzymatic (lytic) domain of SEQ ID NO: 1.
- 10 13. An isolated polypeptide according to any one of Claims 10 to 12 wherein the fragment comprises or consists of the cell wall binding domain of SEQ ID NO: 1.
14. An isolated polypeptide according to any one of the preceding claims comprising or consisting of a variant of the amino acid sequence of SEQ ID NO:1, or of a fragment thereof.
- 15 15. An isolated polypeptide according to Claim 14 wherein the variant comprises or consists of an amino acid sequence with at least 60% identity to the amino acid sequence of SEQ ID NO: 1, or to a fragment thereof, more preferably at least 70% or 80% or 85% or 90% identity to said sequence, and most preferably at least 95%, 96%, 97%, 98% or 99% identity to said amino acid sequence.
- 20 16. An isolated polypeptide according to any one of the preceding claims comprising or consisting of a derivative of the amino acid sequence of SEQ ID NO:1, or of a fragment or variant thereof.
- 25 17. An isolated polypeptide according to any one of the preceding claims comprising or consisting of a fusion of the amino acid sequence of SEQ ID NO:1, or of a fragment, variant or derivative thereof.
- 30 18. An isolated polypeptide according to Claim 17 comprising or consisting of one or more additional amino acids inserted at the N- and/or C-termini of the amino acid sequence of SEQ ID NO:1, or of a fragment, variant or derivative thereof.
- 35 19. An isolated polypeptide according to Claim 17 or 18 comprising or consisting of the cell wall binding domain of SEQ ID NO:1 and an enzymatic (lytic) domain different to that in SEQ ID NO:1.

20. An isolated polypeptide according to any one of the preceding claims wherein the polypeptide is capable of lysing cells of multiple strains of *Clostridium difficile*.
- 5 21. An isolated polypeptide according to any one of the preceding claims wherein the polypeptide is capable of lysing one or more cell types selected from the group consisting of cells of *Bacillus* sp. (e.g. *Bacillus cereus*, *Bacillus subtilis* and *Bacillus anthracis*), other *Clostridium* sp. (e.g. *Clostridium sordelli* and *Clostridium bifermentans*) and *Listeria* sp. (e.g. *Listeria ivanovii*).
- 10 22. An isolated polypeptide according to any one of the preceding claims wherein the polypeptide does not lyse cells of *Clostridium leptum*, *Clostridium nexile*, *Clostridium coccoides*, *Clostridium innocuum*, *Clostridium ramosum* and/or *Anaerococcus hydrogenalis*.
- 15 23. An isolated polypeptide according to any one of the preceding claims wherein the polypeptide is capable of lysing cells of *Clostridium difficile* ribotype 027.
- 20 24. An isolated polypeptide according to Claim 23 wherein the polypeptide exhibits at least 10% of the lysis activity of the polypeptide of SEQ ID NO: 1 on cells of *Clostridium difficile* ribotype 027, for example at least 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100% or more.
- 25 25. An isolated polypeptide according to Claim 24 wherein the polypeptide exhibits at least 100% of the lysis activity of the polypeptide of SEQ ID NO: 1 on cells of *Clostridium difficile* ribotype 027, for example at least 120%, 130%, 140%, 150%, 160%, 170%, 180%, 190%, 200%, 250%, 300%, 500% or more.
- 30 26. An isolated polypeptide according to any one of the preceding claims wherein the polypeptide is capable of lysing cells of pathogenic bacteria selectively.
27. An isolated polypeptide according to any one of the preceding claims wherein the polypeptide is a recombinant polypeptide.
- 35 28. An isolated nucleic acid molecule encoding a polypeptide according to any one of Claims 1 to 27.

29. A nucleic acid molecule according to Claim 28 wherein the nucleic acid molecule comprises or consists of the nucleotide sequence of SEQ ID NO:2.
- 5 30. A vector comprising a nucleic acid molecule according to Claim 28 or 29.
31. A vector according to Claim 230 wherein the vector is an expression vector.
- 10 32. A vector according to Claim 30 or 31 wherein the vector is selected from the group consisting of pET15b and pACYC184.
33. A host cell comprising a nucleic acid molecule according to Claim 28 or 29 or a vector according to any one of Claims 30 to 31.
- 15 34. A host cell according to Claim 33 wherein the host cell is capable of expressing a polypeptide according to any one of Claims 1 to 27.
35. A host cell according to Claim 33 or 34 wherein the host cell is a microbial cell.
- 20 36. A host cell according to any one of Claims 33 to 35 wherein the host cell is a bacterial cell.
37. A host cell according to Claim 35 or 36 wherein the host cell is non-pathogenic.
- 25 38. A host cell according to any one of Claim 33 to 37 wherein the host cell is selected from the group consisting of cells of *Escherichia coli*, *Lactococcus sp.*, *Bacteroides sp*, *Lactobacillus sp.*, *Enterococcus sp.* and *Bacillus sp.*
- 30 39. A host cell according to Claim 38 wherein the host cell is a *Lactococcus lactis* cell.
40. A method for producing a polypeptide according to any one of Claims 1 to 27 comprising culturing a population of host cells comprising a nucleic acid molecule according to Claim 28 or 29 or a vector according to any one of Claims 30 to 32 under conditions in which the polypeptide is expressed, and isolating the polypeptide therefrom.
- 35

41. A pharmacological composition comprising:
- (a) a polypeptide according to any one of Claims 1 to 27;
 - (b) a nucleic acid molecule according to Claim 28 or 29;
 - (c) a vector according to any one of Claims 30 to 32;
 - (d) a host according to any one of Claims 33 to 39; and/or
 - (e) a bacteriophage capable of expressing a polypeptide according to the first aspect of the invention
- and a pharmaceutically acceptable carrier, diluent or excipient.
42. A pharmacological composition according to Claim 41 comprising a polypeptide according to any one of Claims 1 to 27.
43. A pharmacological composition according to Claim 41 or 42 for oral administration.
44. A pharmacological composition according to Claim 41 to 43 wherein the polypeptide is microencapsulated.
45. A pharmacological composition according to any one of Claims 41 to 44 capable of delivering the polypeptide to the GI tract.
46. A pharmacological composition according to any one of Claims 41 to 45 comprising a nucleic acid molecule according to Claim 28 or 29 and/or a vector according any one of Claims 30 to 32.
47. A pharmacological composition according to any one of Claims 41 to 46 comprising a host cell according to any one of Claims 33 to 39.
48. A pharmacological composition according to Claim 47 comprising a non-pathogenic bacterial host cell which is genetically engineered to express a polypeptide according to any one of Claims 1 to 27 and to release said polypeptide upon reaching a predetermined location within the GI tract.

49. A pharmacological composition according to any one of Claims 41 to 48 comprising a bacteriophage capable of expressing a polypeptide according to any one of Claims 1 to 27.
- 5 50. A pharmacological composition according to any one of Claims 41 to 49 wherein the composition allows sustained or slow-release of the polypeptide within the GI tract.
- 10 51. A polypeptide according to any one of Claims 1 to 27 or pharmacological composition according to any one of Claims 41 to 50 for use in medicine.
- 15 52. Use of a polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile*, or a nucleic acid molecule, vector, host cell or bacteriophage capable of expressing the same, in the preparation of a medicament for killing and/or inhibiting/preventing the growth of microbial cells in a patient, wherein the microbial cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis with said endolysin.
- 20 53. A polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile*, or a nucleic acid molecule, vector, host cell or bacteriophage capable of expressing the same, for use in killing and/or inhibiting/preventing the growth of microbial cells in a patient, wherein the microbial cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis with said endolysin.
- 25 54. Use of a polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile*, or a nucleic acid molecule, vector, host cell or bacteriophage capable of expressing the same, in the preparation of a medicament for the treatment or prevention of a disease or condition associated with microbial cells in a patient, wherein the microbial cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis with said endolysin.
- 30 55. A polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile*, or a nucleic acid molecule, vector, host cell or,
- 35

bacteriophage capable of expressing the same, for use in the treatment or prevention of a disease or condition associated with microbial cells in a patient, wherein the microbial cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis with said endolysin.

5

56. The use according to any one of Claims 52 to 55 wherein the polypeptide having the cell lysing activity is selected from the group consisting of

- (a) the lysin of bacteriophage Φ CD27;
- 10 (b) the lysin of bacteriophage Φ CD119;
- (c) the lysin of bacteriophage Φ C2; and
- (d) the lysin of prophages 1 and 2 of *Clostridium difficile* strain 630 (CD630).

15

57. The use according to any one of Claims 52 to 56 wherein the polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile* is a polypeptide according to any one of Claims 1 to 27, and wherein the microbial cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis upon contact with a polypeptide of SEQ ID NO: 1.

20

58. The use according to any one of Claims 52 or 57 wherein the microbial cells are *Clostridium difficile* cells.

25

59. The use according to Claim 58 wherein the microbial cells are *Clostridium difficile* ribotype 027 cells.

30

60. A method for killing and/or inhibiting/preventing the growth of microbial cells in a patient, the method comprising administering to the patient a polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile*, or a nucleic acid molecule, vector, host cell or bacteriophage capable of expressing the same, wherein the microbial cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis with said endolysin.

35

61. A method for the treatment or prevention a disease or condition associated with microbial cells in a patient, the method comprising administering to the patient a

polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile*, or a nucleic acid molecule, vector, host cell or bacteriophage capable of expressing the same, wherein the microbial cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis with said endolysin.

62. A method according to Claim 60 or 61 wherein the polypeptide having the cell lysing activity is selected from the group consisting of
- (a) the lysin of bacteriophage Φ CD27;
 - (b) the lysin of bacteriophage Φ CD119;
 - (c) the lysin of bacteriophage Φ C2; and
 - (d) the lysin of prophages 1 and 2 of *Clostridium difficile* strain 630 (CD630).
63. A method according to any one of Claims 60 to 62 wherein the polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile* is a polypeptide according to any one of Claims 1 to 27, and wherein the microbial cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis upon contact with a polypeptide of SEQ ID NO: 1.
64. A method according to any one of Claims 60 to 63 wherein the microbial cells are *Clostridium difficile* cells.
65. A method according to Claim 64 wherein the microbial cells are *Clostridium difficile* ribotype 027 cells.
66. Use of a polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile*, or a nucleic acid molecule, vector, host cell or bacteriophage capable of expressing the same, for killing and/or inhibiting/preventing the growth of microbial cells *in vitro* and/or *ex vivo*, wherein the microbial cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis with said endolysin.
67. The use according to Claim 66 wherein the polypeptide having the cell lysing activity is selected from the group consisting of

- (a) the lysin of bacteriophage Φ CD27;
(b) the lysin of bacteriophage Φ CD119;
(c) the lysin of bacteriophage Φ C2; and
5 (d) the lysin of prophages 1 and 2 of *Clostridium difficile* strain 630.
68. The use according to Claim 66 or 67 wherein the polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile* is a polypeptide according to any one of Claims 1 to 27, and wherein the microbial
10 cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis upon contact with a polypeptide of SEQ ID NO: 1.
69. The use according to Claim 66 to 68 wherein the microbial cells are *Clostridium*
15 *difficile* cells.
70. The use according to Claim 69 wherein the microbial cells are *Clostridium difficile* ribotype 027 cells.
- 20 71. A kit for detecting the presence of microbial cells in a sample, the kit comprising a polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile*, or a nucleic acid molecule, vector, host cell or bacteriophage capable of expressing the same, wherein the microbial cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible
25 to lysis with said endolysin.
72. A kit according to Claim 71 wherein the polypeptide having the cell lysing activity is selected from the group consisting of
30 (a) the lysin of bacteriophage Φ CD27;
(b) the lysin of bacteriophage Φ CD119;
(c) the lysin of bacteriophage Φ C2; and
(d) the lysin of prophages 1 and 2 of *Clostridium difficile* strain 630.
- 35 73. A kit according to Claim 71 or 72 wherein the polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile* is a

polypeptide according to any one of Claims 1 to 27, and wherein the microbial cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis upon contact with a polypeptide of SEQ ID NO: 1.

5

74. A kit according to Claim 71 to 73 wherein the microbial cells are *Clostridium difficile* cells.

10

75. A kit according to Claim 74 wherein the microbial cells are *Clostridium difficile* ribotype 027 cells.

76. A kit according to any one of Claims 71 to 75 wherein the polypeptide is immobilised on a surface.

15

77. A kit according to any one of Claims 71 to 76 wherein the sample is a cell sample.

78. A kit according to any one of Claims 71 to 77 wherein the sample is derived from a swab taken from a surface to be tested for contamination with microbial cells.

20

79. A kit according to any one of Claims 71 to 78 further comprising a negative control sample.

80. A kit according to any one of Claims 71 to 79 further comprising a positive control sample.

25

81. Use of a polypeptide having the cell wall binding activity and/or cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile*, or a nucleic acid molecule, vector, host cell or bacteriophage capable of expressing the same, in the preparation of a diagnostic agent for a disease or condition associated with microbial cells selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis with said endolysin.

30

82. A polypeptide having the cell wall binding activity and/or cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile*, or a nucleic acid molecule, vector, host cell or bacteriophage capable of expressing the same, for use in the

35

diagnosis of a disease or condition associated with microbial cells selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis with said endolysin.

- 5 83. Use of a polypeptide having the cell wall binding activity and/or cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile*, or a nucleic acid molecule, vector, host cell or bacteriophage capable of expressing the same, for detecting the presence of microbial cells in a sample *in vitro* and/or *ex vivo*, wherein the microbial cells selected from the group consisting of *Clostridium*
10 *difficile* cells and other bacterial cells susceptible to lysis with said endolysin.
84. The use according to any one of Claims 81 to 83 wherein the polypeptide having the cell wall binding activity and/or cell lysing activity is selected from the group consisting of
15 (a) the lysin of bacteriophage Φ CD27;
(b) the lysin of bacteriophage Φ CD119;
(c) the lysin of bacteriophage Φ C2; and
(d) the lysin of prophages 1 and 2 of *Clostridium difficile* strain 630 (CD630).
20
85. The use according to any one of Claims 81 to 84 wherein the polypeptide having the cell wall binding activity and/or cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile* is a polypeptide according to any one of
25 Claims 1 to 24, and wherein the microbial cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis upon contact with a polypeptide of SEQ ID NO: 1.
86. The use according to any one of Claims 81 to 85 wherein the microbial cells are
30 *Clostridium difficile* cells.
87. The use according to Claim 86 wherein the microbial cells are *Clostridium difficile* ribotype 027 cells.
- 35 88. A method for the diagnosis of a disease or condition associated with microbial cells in a patient, the method comprising contacting a cell sample from a patient

to be tested with a polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile*, or a nucleic acid molecule, vector, host cell or bacteriophage capable of expressing the same, and determining whether the cells in the sample have been lysed thereby, wherein the microbial cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis with said endolysin.

89. A method according to Claim 88 wherein the polypeptide having the cell lysing activity is selected from the group consisting of

- (a) the lysin of bacteriophage Φ CD27;
- (b) the lysin of bacteriophage Φ CD119;
- (c) the lysin of bacteriophage Φ C2; and
- (d) the lysin of prophages 1 and 2 of *Clostridium difficile* strain 630.

90. A method according to Claim 88 or 89 wherein the polypeptide having the cell lysing activity of an endolysin from a bacteriophage of *Clostridium difficile* is a polypeptide according to any one of Claims 1 to 27, and wherein the microbial cells are selected from the group consisting of *Clostridium difficile* cells and other bacterial cells susceptible to lysis upon contact with a polypeptide of SEQ ID NO: 1.

91. A method according to any one of Claims 88 to 90 wherein the microbial cells are *Clostridium difficile* cells.

92. A method according to Claim 91 wherein the microbial cells are *Clostridium difficile* ribotype 027 cells.

93. An isolated polypeptide substantially as herein described with reference to the examples.

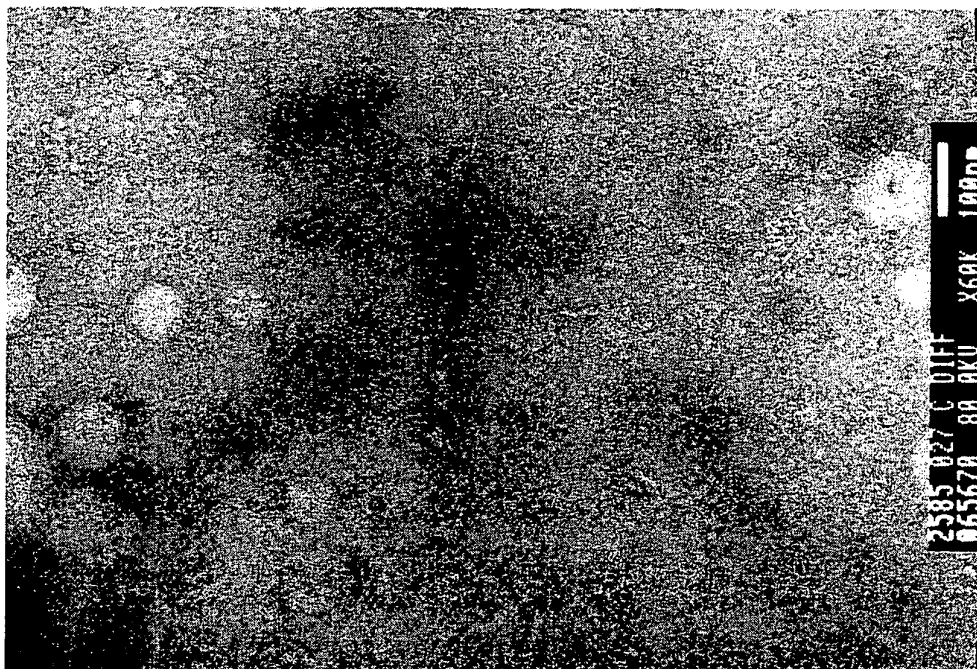
94. An isolated nucleic acid molecule substantially as herein described with reference to the examples.

95. A vector substantially as herein described with reference to the examples.

96. A host cell substantially as herein described with reference to the examples.
97. A method for producing a polypeptide according to any one of Claims 1 to 27 substantially as herein described with reference to the examples.
- 5 98. A pharmacological composition substantially as herein described with reference to the examples.
- 10 99. Use of a polypeptide for killing and/or inhibiting/preventing the growth of microbial cells substantially as herein described with reference to the examples.
100. Use of a polypeptide in the preparation of a medicament for the treatment or prevention of a disease or condition associated with microbial cells substantially as herein described with reference to the examples.
- 15 101. A kit for detecting the presence of microbial cells in a sample substantially as herein described with reference to the examples.
- 20 102. Use of a polypeptide in the preparation of a diagnostic agent for a disease or condition associated with microbial cells substantially as herein described with reference to the examples.
103. Use of a polypeptide for detecting the presence of microbial cells in a sample substantially as herein described with reference to the examples.

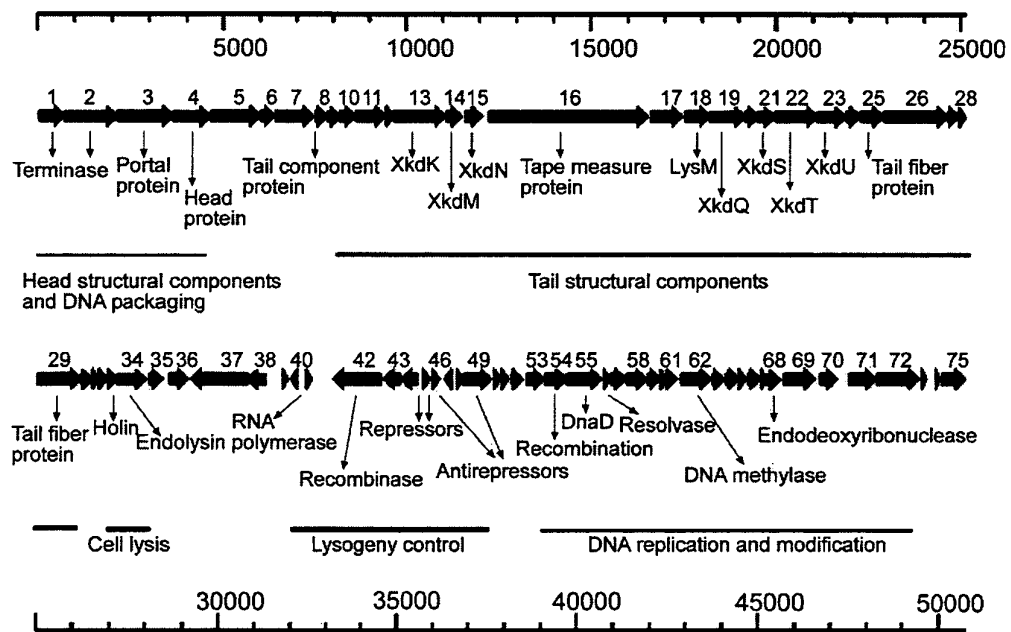
1/31

Figure 1



2/31

Figure 2



3/31

Figure 3

ATGAAAATATGTATAACAGTAGGACACAGTATTTTAAAAAGTGGAGCATGTA
CTTCTGCTGATGGAGTAGTTAACGAGTATCAATACAACAAATCTCTTGCACC
AGTATTAGCAGATACATTTAGAAAAGAAGGGCATAAGGTAGATGTAATAATA
TGCCCAGAAAAGCAGTTTAAACTAAGAATGAAGAAAAGTCTTATAAAATAC
CTAGAGTTAATAGTGGAGGATATGATTTACTTATAGAGTTACATTTAAATGCA
AGTAACGGTCAAGGTAAAGGTTCAGAAGTCCTATATTATAGTAATAAAGGCT
TAGAGTATGCAACTAGAATATGTGATAAACTAGGTACAGTATTTAAAAATAGA
GGTGCTAAATTAGATAAAAGATTATATATCTTAAATAGTTCAAAGCCTACAGC
AGTATTAATTGAAAGTTTCTTCTGTGATAATAAAGAAGATTATGATAAAGCTA
AGAACTAGGTCATGAAGGTATTGCTAAGTTAATTGTAGAAGGTGTATTAAA
TAAAAATATAAATAATGAGGGAGTTAAACAGATGTACAAACATACAATTGTTT
ATGATGGAGAAGTTGACAAAATCTCTGCAACTGTAGTTGGTTGGGGTTATAA
TGATGGGAAAATACTGATATGTGATATAAAAGATTACGTGCCAGGTCAGACG
CAAAATCTTTATGTTGTAGGAGGTGGCGCATGTGAAAAGATAAGTTCTATTA
CTAAAGAAAAATTTATTATGATAAAAGGTAATGATAGATTTGATACACTTTAT
AAAGCATTGGATTTTATTAATAGATAG

[SEQ ID NO: 2]

4/31

Figure 4 (a)

```

      *      20      *      40      *      60      *
phiCD27_L : ATGAAAATATCTATTACAGTAGGACACCGTATTTTAAAAAGTGGACGCTCTTCTCTGCTGATGGAGTGGTTTAC : 75
phiC2_L : ATGAAAATATCTATTACAGTAGGACACCGTATTTTAAAAAGTGGTCCATGCGCTTCTCTGCTGATGGAGTGGTTTAC : 75
QCD-32g58_L : ATGAAAATATCTATTACAGTAGGACACCGTATTTTAAAAAGTGGTCCATGCGCTTCTCTGCTGATGGAGTGGTTTAC : 75
phiCD119_L : ATGAAAATAGCTGTAAATTGTGGACCTTC-----AAAAACAGCGCCAGCAAGCTGGAGCTATAGGTAAATATAT : 69
Cd630P1_L : ATGAAAATAGCTGTAAATTGTGGACCTTC-----AAAAACAGCGCCAGCAAGCTGGAGCTATAGGTAAATATAT : 69
Cd630P2_L : ATGAAAATAGCTGTAAATTGTGGACCTTC-----AAAAACAGCGCCAGCAAGCTGGAGCTATAGGTAAATATAT : 69
      ATGAAAATA GTat A      GGACA A      AAAAA GG GCA G A T      GCT GG      A T AA

      80      *      100      *      120      *      140      *
phiCD27_L : GAGTATCAATTCGAACCAATCTCTTGCACCAGTATTAGCAATATCCTTTGGAAGGAGAGGGCGTTAAAGCTAGATTC : 149
phiC2_L : GAAATACCAATTCGAACCAATCTCTTGCACCAGTATTAGCAATATCCTTTGGAAGGAGAGGGCGTTAAAGCTAGATTC : 149
QCD-32g58_L : GAAATACCAATTCGAACCAATCTCTTGCACCAGTATTAGCAATATCCTTTGGAAGGAGAGGGCGTTAAAGCTAGATTC : 149
phiCD119_L : GAAATCAATTCGAACCTAGGAATCTAGGCTATATAAGTAATAGATAATTAATAAACTCTAGGTATCAATCTAGTTTCAT : 144
Cd630P1_L : GAGTCAATTCGAACCTAGGAATCTAGGCTATATAAGTAATTCATTAATTAATAAAATTAGGCAATAATCTAGTTCAT : 144
Cd630P2_L : GAGTCAATTCGAACCTAGGAATCTAGGCTATATAAGTAATTCATTAATTAATAAAATTAGGCAATAATCTAGTTCAT : 144
      GA T      A A A tA      T T G A      A TA      GATA ATT A Aaaa aGG AtAA GtAG G T

      160      *      180      *      200      *      220
phiCD27_L : AATTAATATGCCCCAGAAAAGCAGTTTAAAACTAAAGAAATGAAAGAAAAGTCTTTATTAATATCCCTAGAGTTAATAGTGG : 224
phiC2_L : AATTAATATGCCCCAGAAAAGCAGTTTAAAACTAAAGAAATGAAAGAAAAGTCTTTATTAATATCCCTAGAGTTAATAGTGG : 224
QCD-32g58_L : AATTAATATGTCCTTGAAGAAGCAGTTTAAAACTAAAGCAGAAAGAAAAGTCTTTATTAATATCCCTAGAGTTAATAGTGG : 224
phiCD119_L : TGTACTAT-----AGTAAAGCATCT---ACACAAAGTCAGTGTTTCTCTAATATGCAACACAGCAATAGACA : 212
Cd630P1_L : TGTACTAT-----AGTAAAGCATCT---ACACAAATCAGATGTTTATCTAATATATCAGACACAGCTAATAGACA : 212
Cd630P2_L : TGTACTAT-----AGTAAAGCATCT---ACACAAATCAGATGTTTATCTAATATATCAGACACAGCTAATAGACA : 212
      TA TAT      aGA AA      gT T      Act A      GAa      tCT A A Aa A C      AG tAATAG

      *      240      *      260      *      280      *      300
phiCD27_L : AGGATATGATTTACCTATAGAGTTACATTTAAATGCAAGTAACGGTCAAGGTAAAGGTTCAGAAGTCCCTATATTA : 299
phiC2_L : AGGATATGATTTACCTATAGAGTTACATTTAAATGCAAGTAACGGTCAAGGTAAAGGTTCAGAAGTCCCTATATTA : 299
QCD-32g58_L : AGGATATGATTTACCTATAGAACTACATCTAAATGCAAGTGATGGTCAAGGAAAGGTTCAGAAGTCCCTATATTA : 299
phiCD119_L : AGATTAGATTGGTTTATAGCATACATTTTAATGCA-----GGAAAAGSACGAGGATGCGAAGTTTACACATA : 281
Cd630P1_L : AGATTAGATTGGTTTATATCAATACATTTCAATGCA-----GGTGGTGGCAAGGATGCGAAGTTTACACATA : 281
Cd630P2_L : AGATTAGATTGGTTTATATCAATACATTTCAATGCA-----GGTGGTGGCAAGGATGCGAAGTTTACACATA : 281
      AG T GATT      T ATA      TACATTT AATGCA      Ggt GG aaAGG T GAAGT      TA

      *      320      *      340      *      360      *
phiCD27_L : TAGTAATAAAGGCTAGAGTATGCACCTAGAA-TATGCTATAAATCTA---GCTACAGTATTTAATAATAGAGGTG : 370
phiC2_L : TAGTAATAAAGGCTAGAGTATGCACCTAGAA-TATGCTATAAATCTA---GCTACAGTATTTAATAATAGAGGTG : 370
QCD-32g58_L : TAGTAATAAAGGCTAGAGTATGCACCTAGAA-TATGCTATAAATCTA---GCTACAGTATTTAATAATAGAGGAG : 370
phiCD119_L : CAAGGCACACAGTATCAGATGCTT-TAGATGTTTGTAAATAATCTCTCATTTAGGATTTAATAATCAGGGGG : 355
Cd630P1_L : CAAGGCACACAGTATCAGATGCTT-TAGATGTTTGTAAATAATCTCTCATTTAGGATTTAATAATCAGGGGG : 355
Cd630P2_L : CAAGGCACACAGTATCAGATGCTT-TAGATGTTTGTAAATAATCTCTCATTTAGGATTTAATAATCAGGGGG : 355
      tA tAA      T A      ATGC A TAGA      T TGT A Aa T      G T      AG ATTTA AAAT GAgG G

```

5/31

Figure 4 (a) - continued

```

      380          *          400          *          420          *          440          *
phiCD27_L : CTAAATTAGATGAAA GATTATATATCTTAATTGTTCAAGGCTACG CAGTATTAATTGAAAGTTTCTTCTGTG : 445
phiC2_L : CTAAATTAGATGAAA GATTATATATCTTAATTGTTCAAGGCTACG CAGTATTAATTGAAAGTTTCTTCTGTG : 445
QCD-32g58_ : CTAAATTAGATGAAAGATTATATATCTTAATTGTTCAAACTCTACG CAGTATTAATTGAAAGTTTCTTCTGTG : 445
phiCD119_L : TAAAAAGATGGAAGTGGATTATATCTAGTAAAGGAAA CAAAAGCTAAAGTAAGCTAAAGAA-CTAGCTTTTGTG : 429
Cd630P1_L : TAAAAAGATGGAAGTGGATTATATCTAGTAAAGGAAA CAAAAGCTAAAGTAAGCTAAAGAA-CTAGCTTTTGTG : 429
CD630P2_L : TAAAAAGATGGAAGTGGATTATATCTAGTAAAGGAAA CAAAAGCTAAAGTAAGCTAAAGAA-CTAGCTTTTGTG : 429
      AAA G A GATTATAT T TAAA A CAAA CTA A T TAAT GAA GT T tT TGT

      460          *          480          *          500          *          520
phiCD27_L : ATAAATAAGAGATATCTTAAGGCTAAGGAACTAGGTCATGAAGGTATTCCTAAGTTAATTGTAGAAAGGCTCTAT : 520
phiC2_L : ATAAATAAGAGATATCTTAAGGCTAAGGAACTAGGTCATGAAGGTATTCCTAAGTTAATTGTAGAAAGGCTCTAT : 520
QCD-32g58_ : ATAAATAAGAGATATCTTAAGGCTAAGGAACTAGGTCATGAAGGTATTCCTAAGTTAATTGTAGAAAGGCTCTAT : 520
phiCD119_L : G--ACCGTGAAGATGCAAAATAAGTATTGAGTTTAGGAGCTGATAAATTAGCTACTGCAATASTAGAAAGCTATA- : 501
Cd630P1_L : G--ACCGTGAAGATGCAAAATAAGTATTGAGTTTAGGAGCTGATAAATTAGCTACTGCAATASTAGAAAGCTATA- : 501
CD630P2_L : G--ACCGTGAAGATGCAAAATAAGTATTGAGTTTAGGAGCTGATAAATTAGCTACTGCAATASTAGAAAGCTATA- : 501
      A A GAAGAT AtAA T GA TAGG TGA T GCTA AAT GTAGAAG T TA

      *          540          *          560          *          580          *          600
phiCD27_L : TAAATAAAAATATATAAATAATGAGGAGTTTAAACG-----TTGTACAAACATACAAATCTTTATGATGGAGAAG : 589
phiC2_L : TAAATAAAAATATATAAATAATGAGGAGTTTAAACG-----TTGTACAAACATACAAATCTTTATGATGGAGAAG : 589
QCD-32g58_ : TAAATAAAAATATATAAATAATGAGGAGTTTAAACG-----TTGTACAAACATACAAATCTTTATGATGGAGAAG : 589
phiCD119_L : --ACTAAACATATAAGTTCACACAGAGAAACCAATTATAATGATATAAACATACATACTCTCTATTCGGAGA-- : 572
Cd630P1_L : --ACTAAACATATAAGTTCACACAGAGAAACCAATTATAATGATATAAACATACATACTCTCTATTCGGAGA-- : 572
CD630P2_L : --ACTAAACATATAAGTTCACACAGAGAAACCAATTATAATGATATAAACATACATACTCTCTATTCGGAGA-- : 572
      A TAAA ATATAA T G G AG AA A A TA AAACATAC AT GT TAT TGGAGA

      *          620          *          640          *          660          *
phiCD27_L : TTGACAAATCTCTGCGACTGTAAGTTGGTTGGGGCTATATG---TTGGGAAAATA--CTGATGTGTCATATAA : 658
phiC2_L : TTGATAAGATCTCTGCGAATGTGCTAGTTGGGGCTATATGTC---CAAGCAAACTT--TTAGCTTTGTGATATAA : 658
QCD-32g58_ : TTGACAAATCTCTGCGACTGTAAGTTGGTTGGGGCTATATG---TTGGGAAAATA--CTGATGTGTCATATAA : 658
phiCD119_L : -TGATAAAGTATCAGCGACCACTTTAGGACTATATTATAGAGAGAAAGAAAGAACTTACTTTAGTACACAGATATAA : 646
Cd630P1_L : -TGATAAAGTATCAGCGACCACTTTAGGACTATATTATAGAGAGAAAGAAAGAACTTACTTTAGTACACAGATATAA : 646
CD630P2_L : -TGATAAAGTATCAGCGACCACTTTAGGACTATATTATAGAGAGAAAGAAAGAACTTACTTTAGTACACAGATATAA : 646
      TGA AAa T tc GCa T T gG tTATAa a gAAA T T Ta GATATAA

      680          *          700          *          720          *          740          *
phiCD27_L : AAGAATTACGTGCCCGGTCAGAGCGCAAAATCTTTATGTTGTAAGGAGTGGCCGATCTGCAAAAGTAAAGTTCTCTT : 733
phiC2_L : AAGAATTACGTGCCCGGTCAGAGCGCAAAATCTTTATGTTGTAAGGAGTGGCCGATCTGCAAAAGTAAAGTTCTCTT : 733
QCD-32g58_ : AAGAATTACGTGCCCGGTCAGAGCGCAAAATCTTTATGTTGTAAGGAGTGGCCGATCTGCAAAAGTAAAGTTCTCTT : 733
phiCD119_L : AAGACTATAAACCACTAGAGACAGAAATCTTATGTAATCGGTGCACTAACTTCTAATTAAGTGAAGGAAATCTA : 721
Cd630P1_L : AAGACTATAAACCACTAGAGACAGAAATCTTATGTAATCGGTGCACTAACTTCTAATTAAGTGAAGGAAATCTA : 721
CD630P2_L : AAGACTATAAACCACTAGAGACAGAAATCTTATGTAATCGGTGCACTAACTTCTAATTAAGTGAAGGAAATCTA : 721
      AAGA TA CCA T AC CAAAATcT TATGT T GG GG G C TGT A AA AT a AT A

```

6/31

Figure 4 (a) – continued

```

760          *          780          *          800          *          820
phiCD27_L : CTAA-----AGAAAAATTTATATGATAAAGGTAATGATAGATTGATACACTTTATAAAGCATGGAT : 799
phiC2_L   : CTAA-----AGAAAAATTTATATGATAAAGGTAATGATAGATTGATACACTTTATAAAGCATGGAT : 799
QCD-32g58_ : CTAA-----AGAAAAATTTATATGATAAAGGTAATGATAGATTGATACACTTTATAAAGCATGGAT : 799
phiCD119_L : GTAA GACTACAGGAGAAAAATTTAGCTCAGCTATATGCTAATGATGTATGGTCAACCAATGGATAAAGCTATAGCAAT : 796
Cd630P1_L : GTAA GACTACAGGAGAAAAATTTAGCTCAGCTATATGCTAATGATGTATGGTCAACCAATGGATAAAGCTATAGCAAT : 796
Cd630P2_L : GTAA GACTACAGGAGAAAAATTTAGCTCAGCTATATGCTAATGATGTATGGTCAACCAATGGATAAAGCTATAGCAAT : 796
          TAA          AGAAAAATTTA T G TA A GGTAATGAT AT          ACA T ATAAaGC T GA T

          *          840
phiCD27_L : TTAATTAAATGATG----- : 813          [SEQ ID NO: 22]
phiC2_L   : TTAATTAAATGATG----- : 813          [SEQ ID NO: 23]
QCD-32g58_ : TTAATTAAATGATG----- : 813          [SEQ ID NO: 24]
phiCD119_L : TTCTAAAT-AGCAATAATTGTAG : 816          [SEQ ID NO: 25]
Cd630P1_L : TTCTAAAT-AGCAATAATTGTAG : 816          [SEQ ID NO: 26]
Cd630P2_L : TTCTAAAT-AGCAATAATTGTAG : 816          [SEQ ID NO: 27]
          TT T AA AGA A          [SEQ ID NO: 28]

```

7/31

Figure 4 (b)

```

      *           20           *           40           *           60           *           80
phiCD27  MKICITVGHSLKSGACTSADGVVNEYQYNKSLAPVLEADTFRKEGHKVDVLIICPEKQFKTKNEEKSYPKIPRVNSGGYDLL
phiC2    MKICITVGHSLKSGACTSADGVVNEYQYNKSLAPVLEADTFRKEGHKVDVLIICPEKQFKTKNEEKSYPKIPRVNSGGYDLL
QCD-32g58 MKICITVGHSLKSGACTSADGVVNEYQYNKSLAPVLEADTFRKEGHKADVLIICPEKQFKTKAEKTYKIPRVNSGGYDLL
phiCD119 MKIGVNCGH--TKTGAGSGAIGKINESIETRNVGYKVIDKLKILGN--NVVDCTIDKASTQSECLSKIATQANRQDLDFW
CD630P1  MKIGVNCGH--TKTGAGSGAIGKINESIETRNVGYKVIDKLKILGN--NVVDCTIDKASTQSECLSKIATQANRQDLDFW
CD630P2  MKIGVNCGH--TKTGAGSGAIGKINESIETRNVGYKVIDKLKILGN--NVVDCTIDKASTQSECLSKIATQANRQDLDFW

      *           100          *           120          *           140          *           160
phiCD27  IEEHLNASNGQGKGSEVLYSNKGLVYATRIGDNLGTV-FKNRGAKLDKRLYITNSSKPTAVLIESFFCDNKEDYDKAKK
phiC2    IEEHLNASNGQGKGSEVLYSNKGLVYATRIGDNLGTV-FKNRGAKLDKRLYITNSSKPTAVLIESFFCDNKEDYDKAKK
QCD-32g58 IEEHLNASDGQKGSEVLYSNKGLVYATRIGDNLGTV-FRNRRAKLDKGLYITNSSNPTAVLIESFFCDNKEDYDKAKK
phiCD119 ISIHFNAGKG--KGCEVYTYKGRQYDAIDVCKKISDLGFTNRGVNDGSGLYVYKKTAKSMLIEVCFVD-SEDANKYLS
CD630P1  ISIHFNAGGG--KGCEVYTYKGRQYDAIDVCKKISDLGFTNRGVNDGSGLYVYKKTAKSMLIEVCFVD-TEDANKYLS
CD630P2  ISIHFNAGGG--KGCEVYTYKGRQYDAIDVCKKISDLGFTNRGVNDGSGLYVYKKTAKSMLIEVCFVD-TEDANKYLS

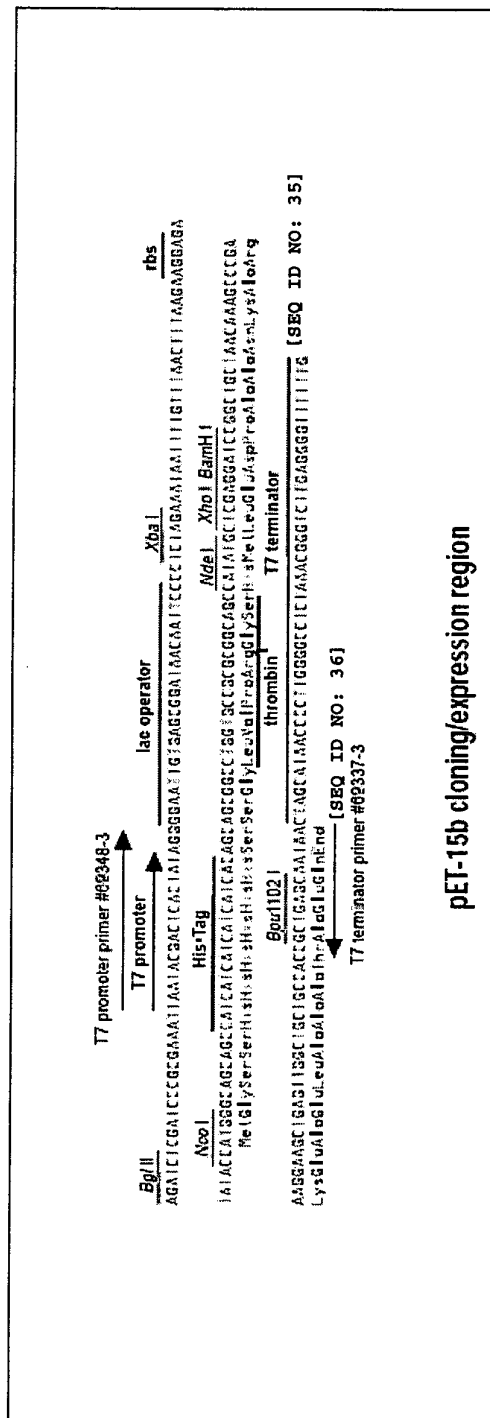
      *           180          *           200          *           220          *           240
phiCD27  LGHECHAKLIVEGHLNKNINNE-GVKQMYKHTIVYDGEVDKLSATVVGWGYNDGN--ILHCDIKDYVEGQTQNLVVIWGGG
phiC2    LGHECHAKLIVEGHLNKNINNE-GVKQMYKHTIVYDGEVDKILANVLSWGYSPSK--VLWCDIKDYVEGQTENLYVVIWGG
QCD-32g58 LGHECHAKLIVEGHLNKNINNE-GVKQMYKHTIVYDGEVDKLSATVVGWGYNDGN--ILHCDIKDYVEGQTQNLVVIWGA
phiCD119 LGADKLATAIVEATTKHISSAEENNYNRYKHTIVYSGD-DKVSADILGLYYKRKNESYLVTDIKDYKEHRTQNLVVIWGGV
CD630P1  LGADKLATAIVEATTKHISSAEENNYNRYKHTIVYSGD-DKVSADILGLYYKRKNESYLVTDIKDYKEHRTQNLVVIWGGV
CD630P2  LGADKLATAIVEATTKHISSAEENNYNRYKHTIVYSGD-DKVSADILGLYYKRKNESYLVTDIKDYKEHRTQNLVVIWGGV

      *           260          *           280
phiCD27  ACEKHSSTTK---EKFIMIKCNDREDTLYKALDFINR--- [SEQ ID NO: 29]
phiC2    ACEKHSSTTK---EKFIMIKCNDREDTLYKALDFINR--- [SEQ ID NO: 30]
QCD-32g58 ACEKHGSMTK---EKFTIMIKCNDREDTLYKALDFINR--- [SEQ ID NO: 31]
phiCD119 TCNKMKEMSKTTGEKFTQLYGNDVSTMDKALBFVKEKL- [SEQ ID NO: 32]
CD630P1  TCNKMKEMSKTTGEKFTQLYGNDVSTMDKALBFVKEKL- [SEQ ID NO: 33]
CD630P2  TCNKMKEMSKTTGEKFTQLYGNDVSTMDKALBFVKEKL- [SEQ ID NO: 34]

```

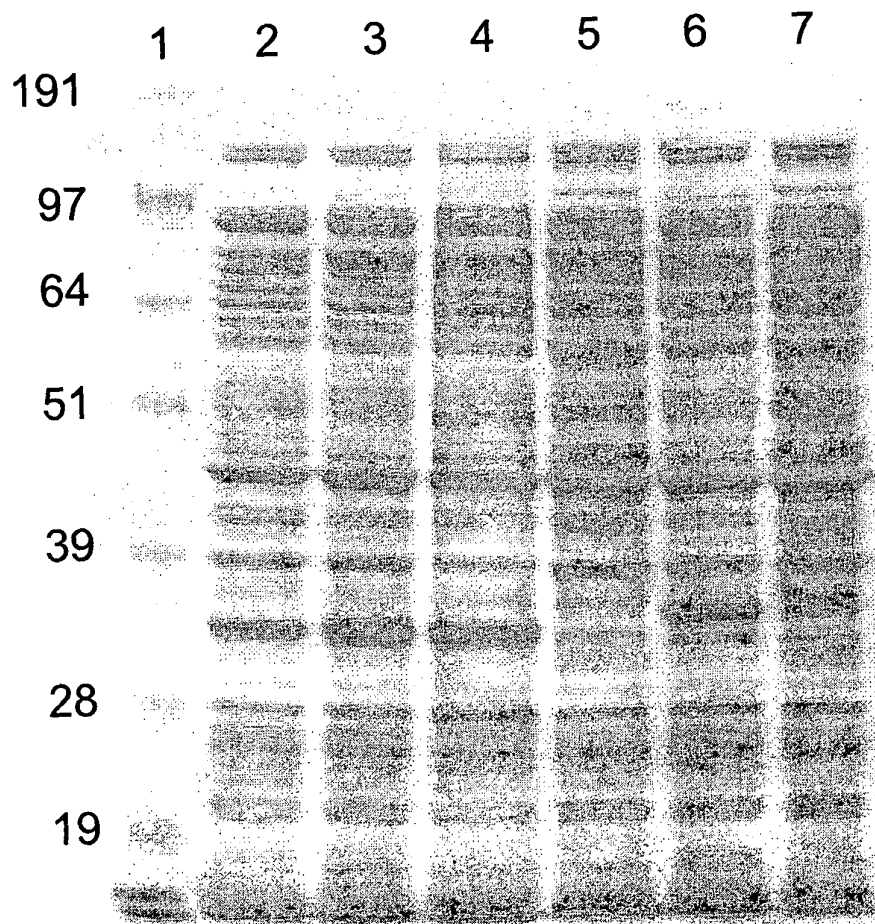
8/31

Figure 5



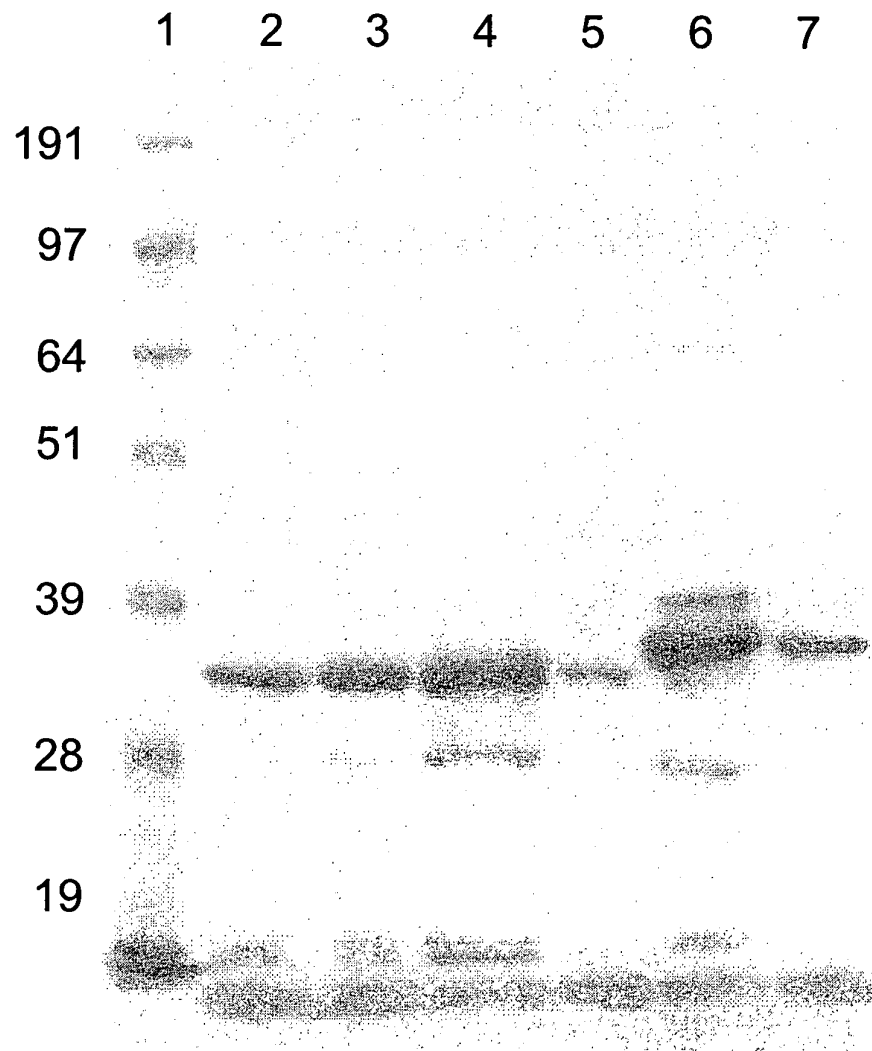
9/31

Figure 6 (a)



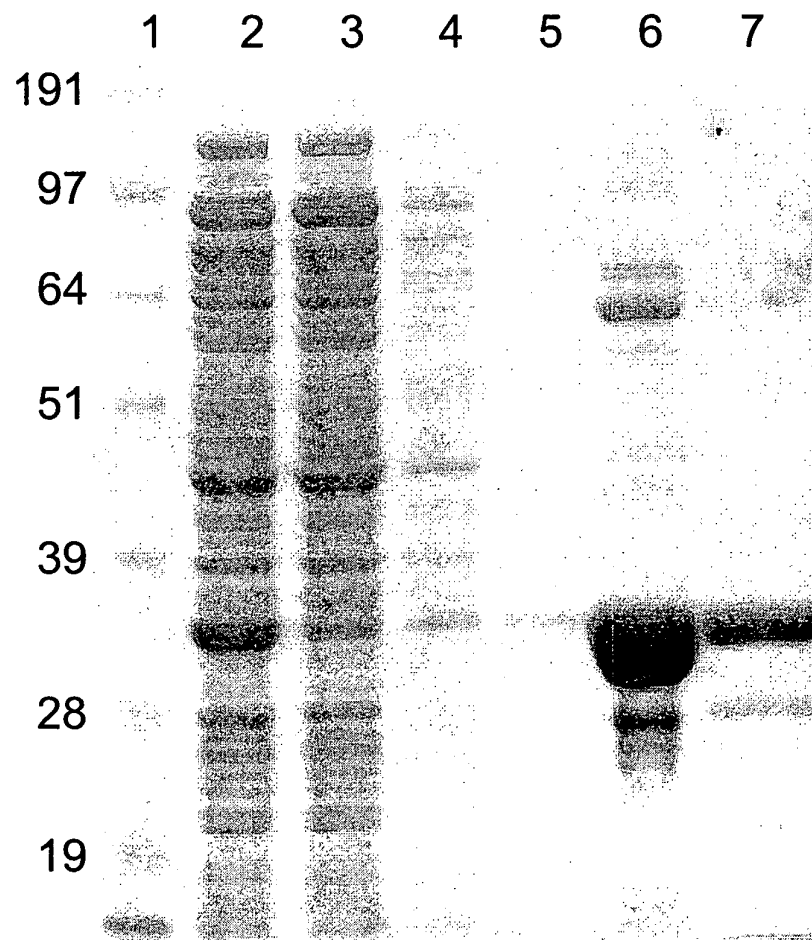
10/31

Figure 6(b)



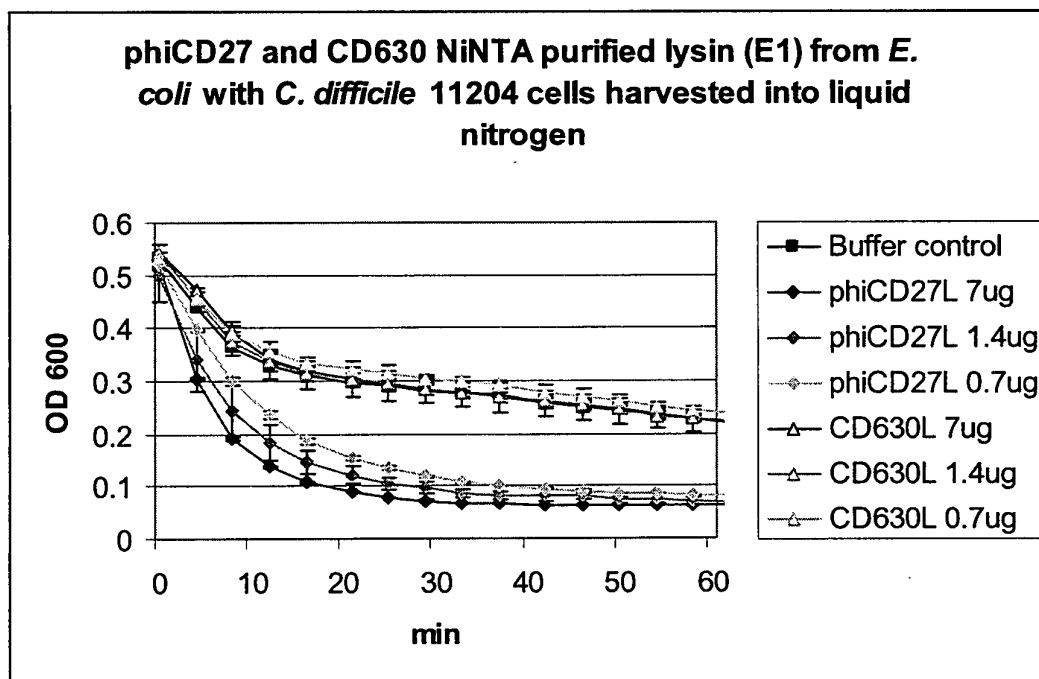
11/31

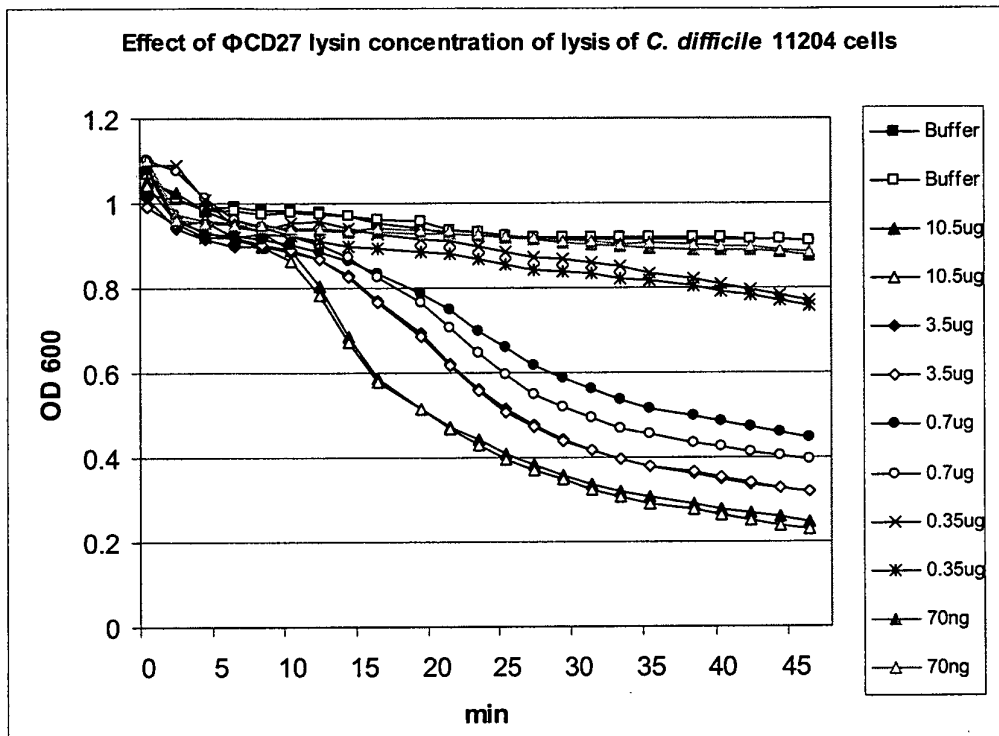
Figure 7



12/31

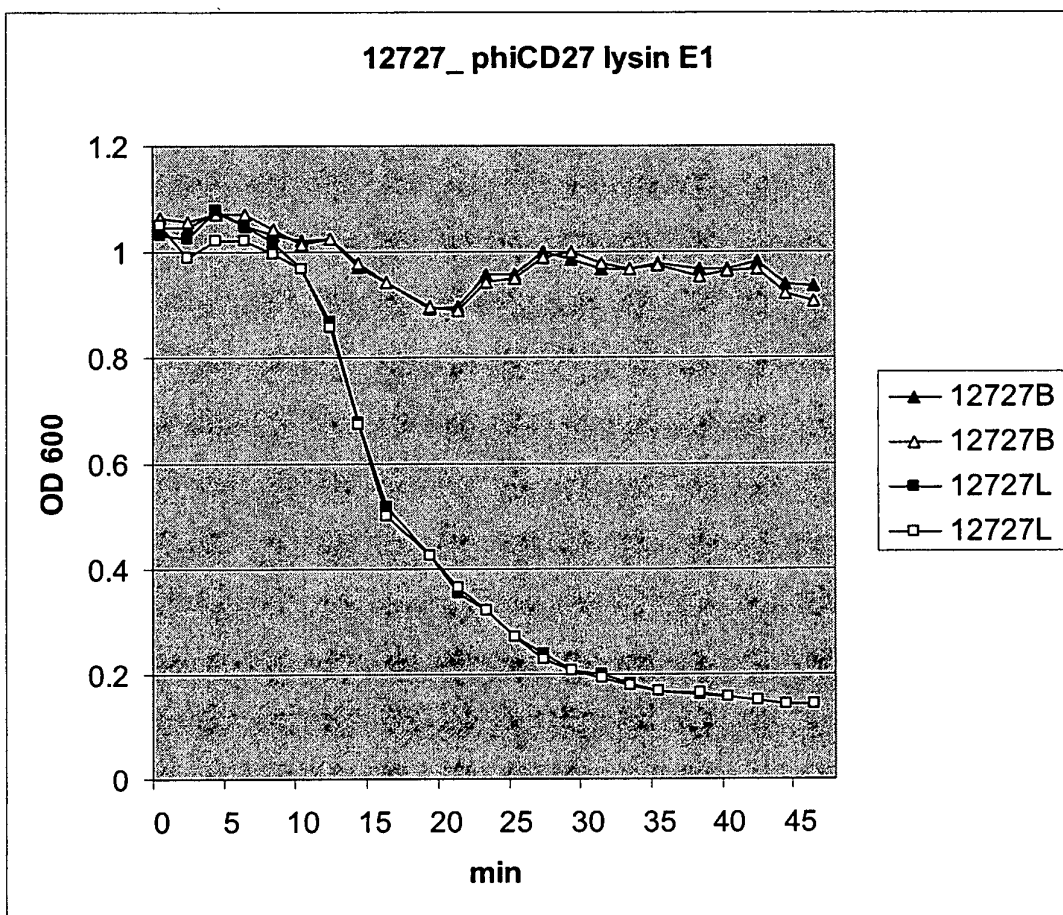
Figure 8





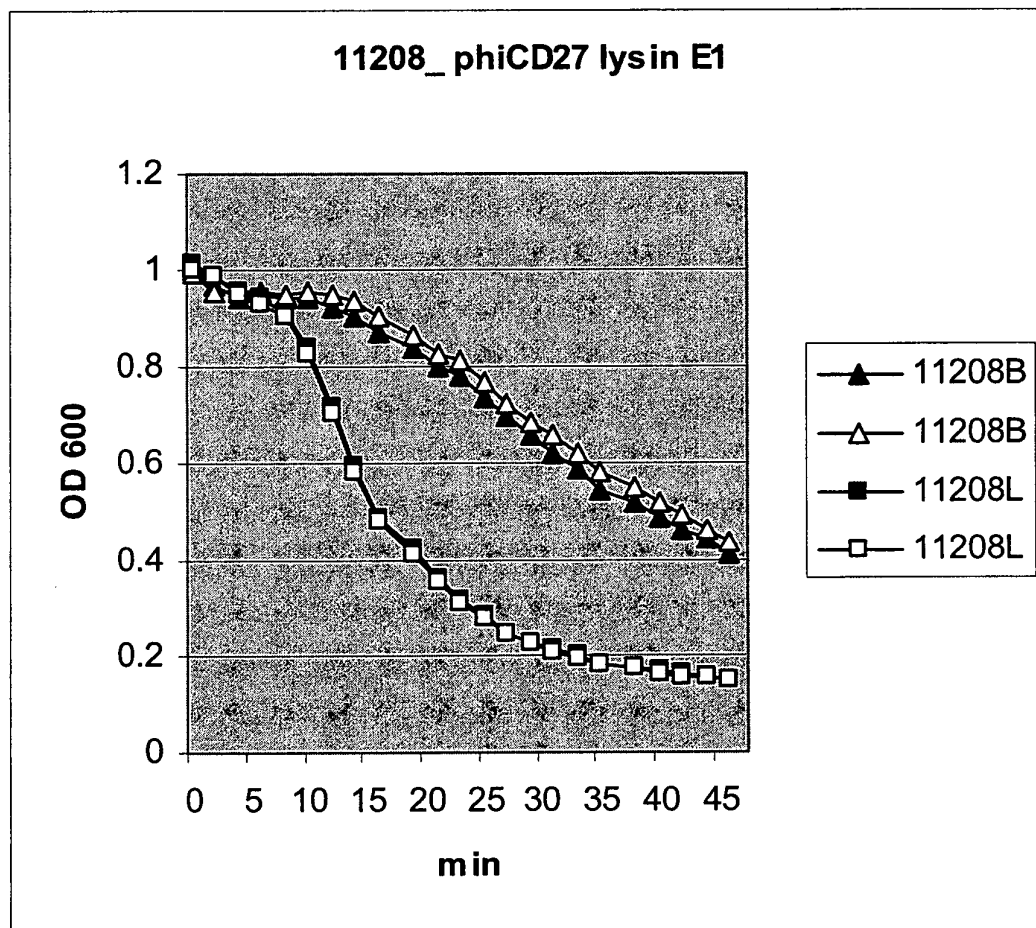
14/31

Figure 10 (a)



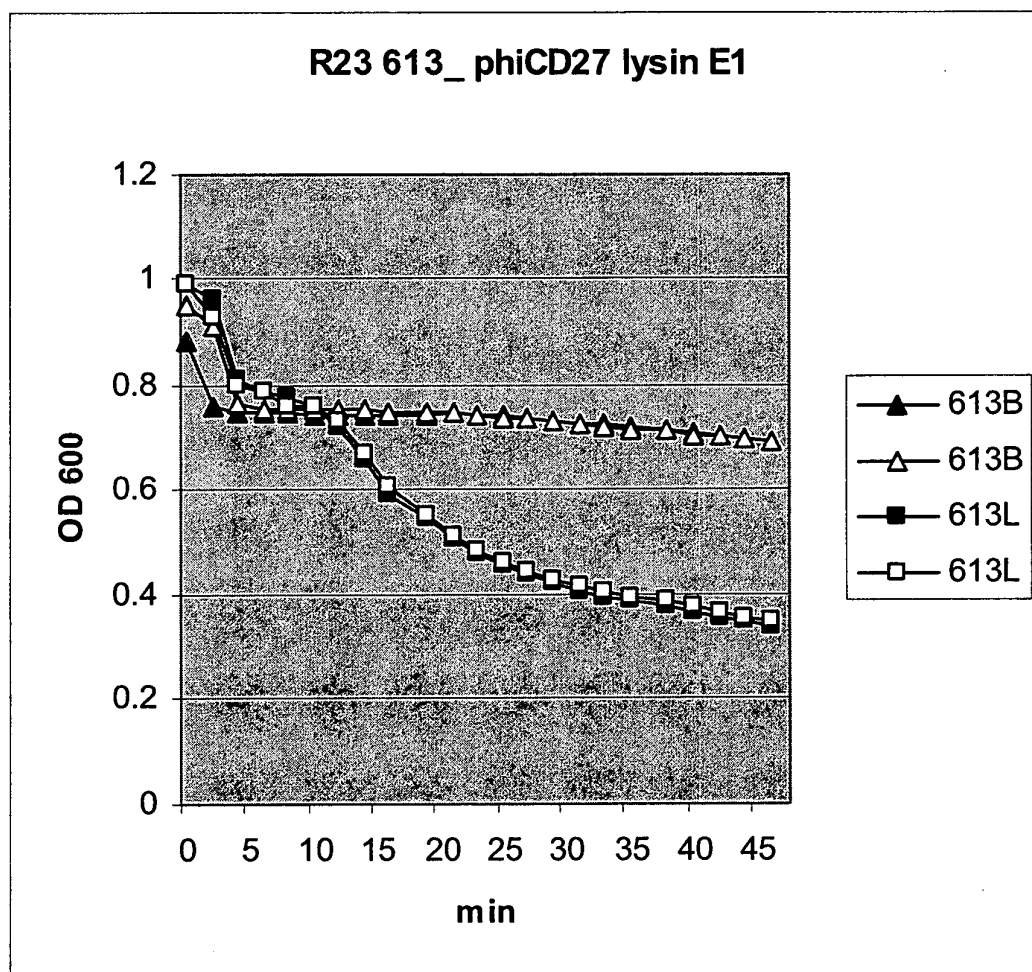
15/31

Figure 10 (b)



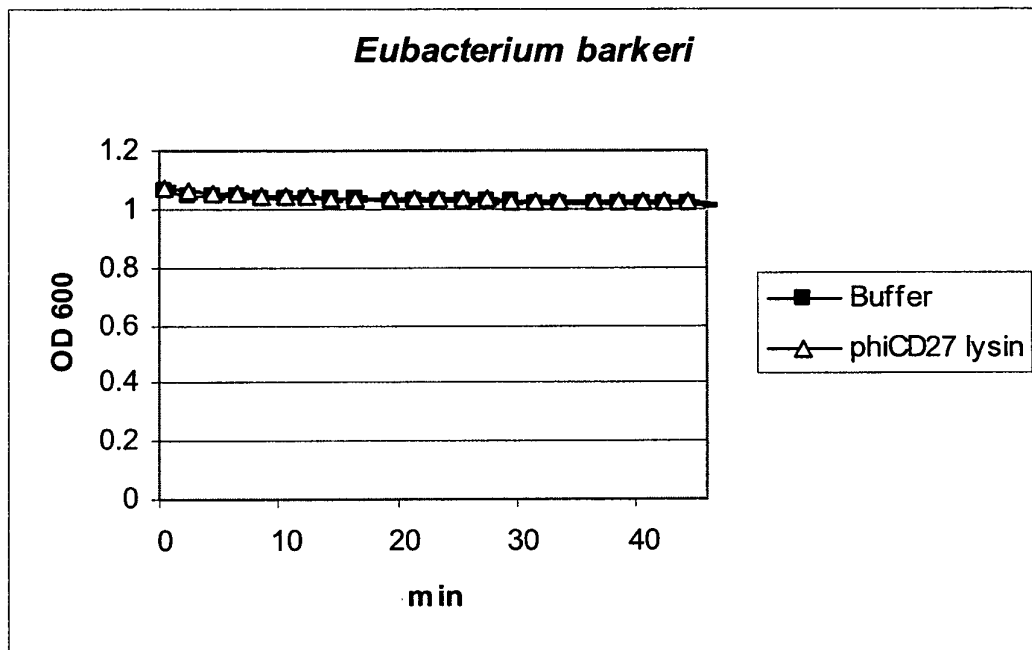
16/31

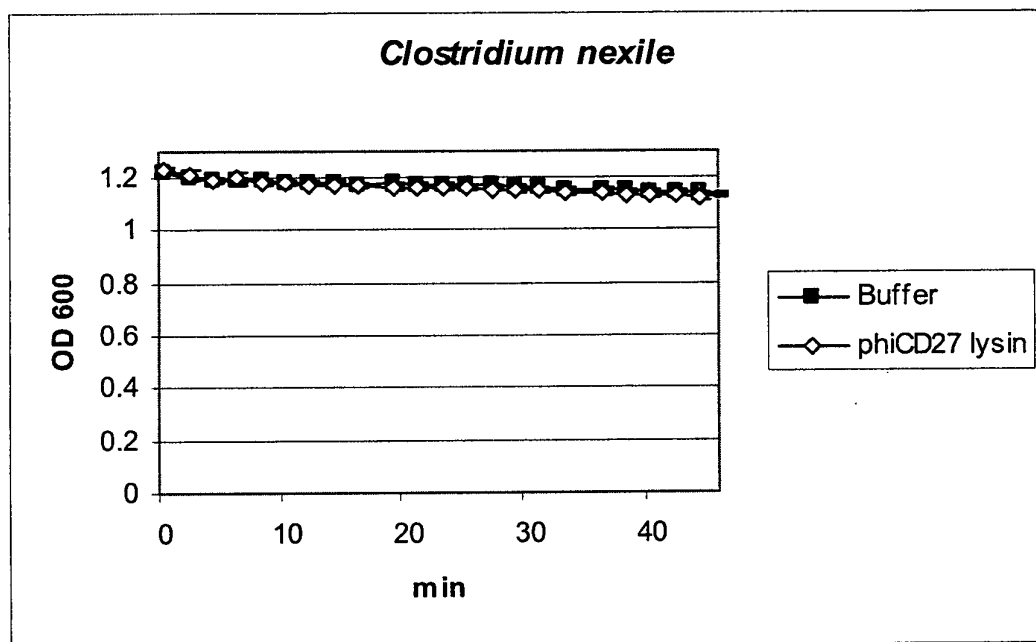
Figure 10 (c)

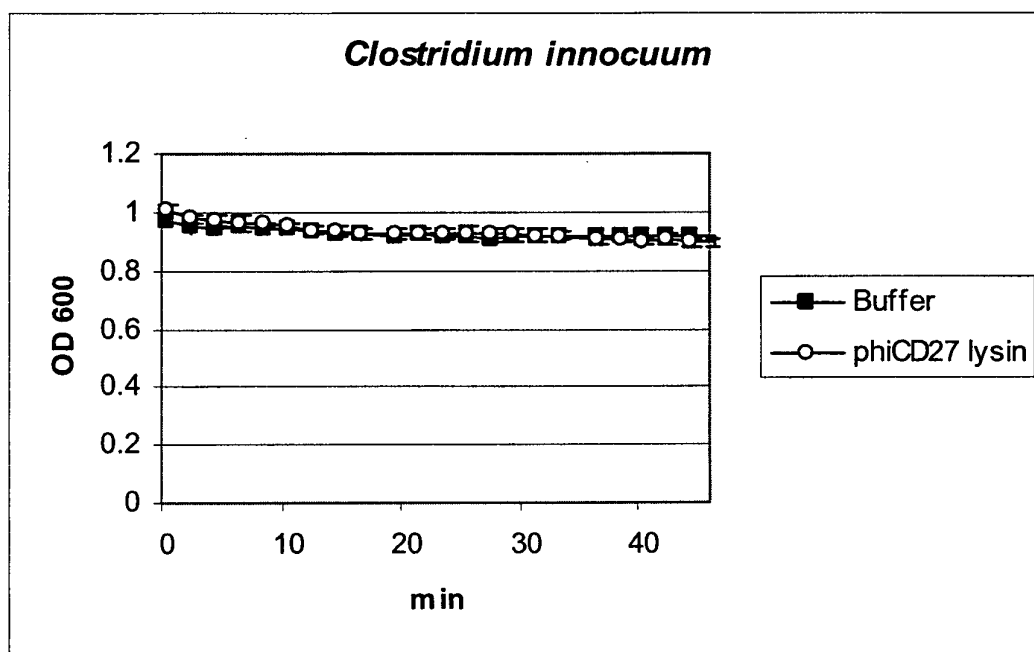


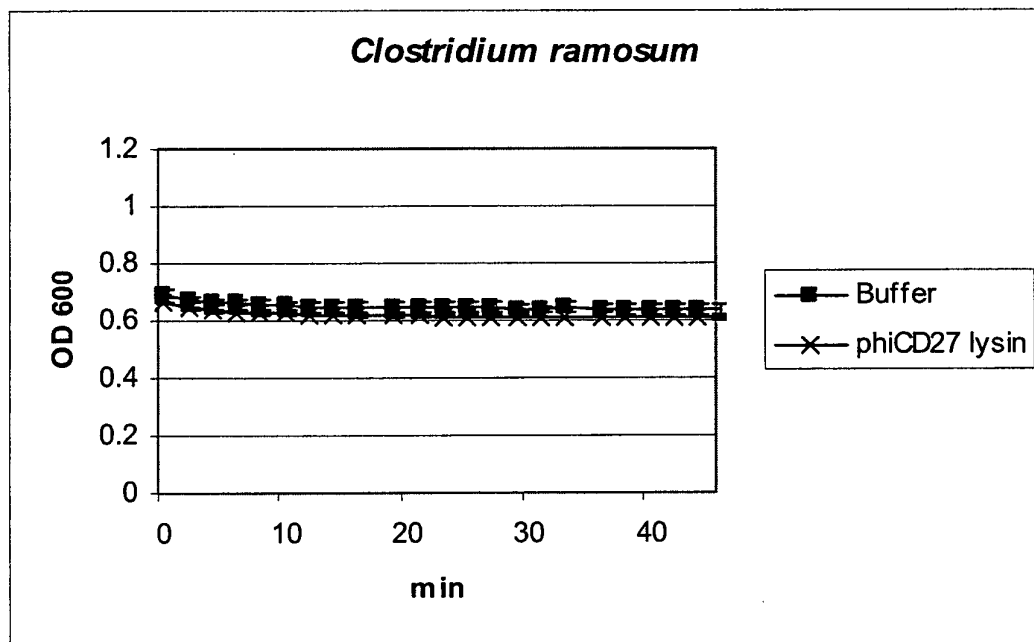
17/31

Figure 11 (a)



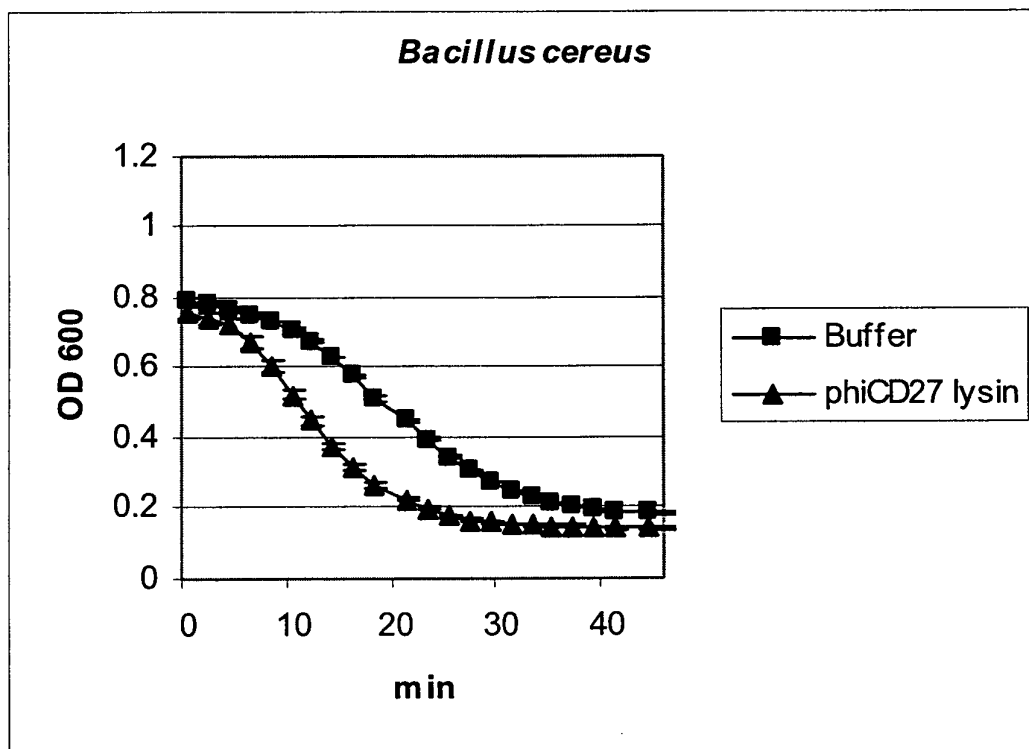
18/31Figure 11 (a) - *continued*

19/31Figure 11 (a) - *continued*

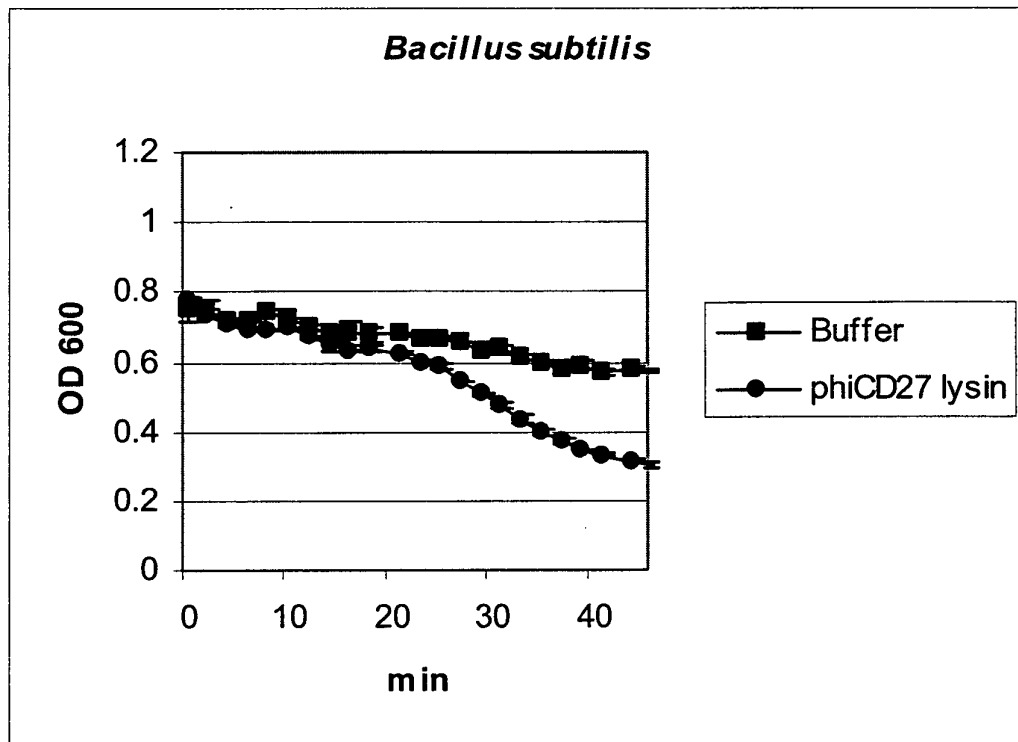
20/31Figure 11 (a) - *continued*

21/31

Figure 11 (b)

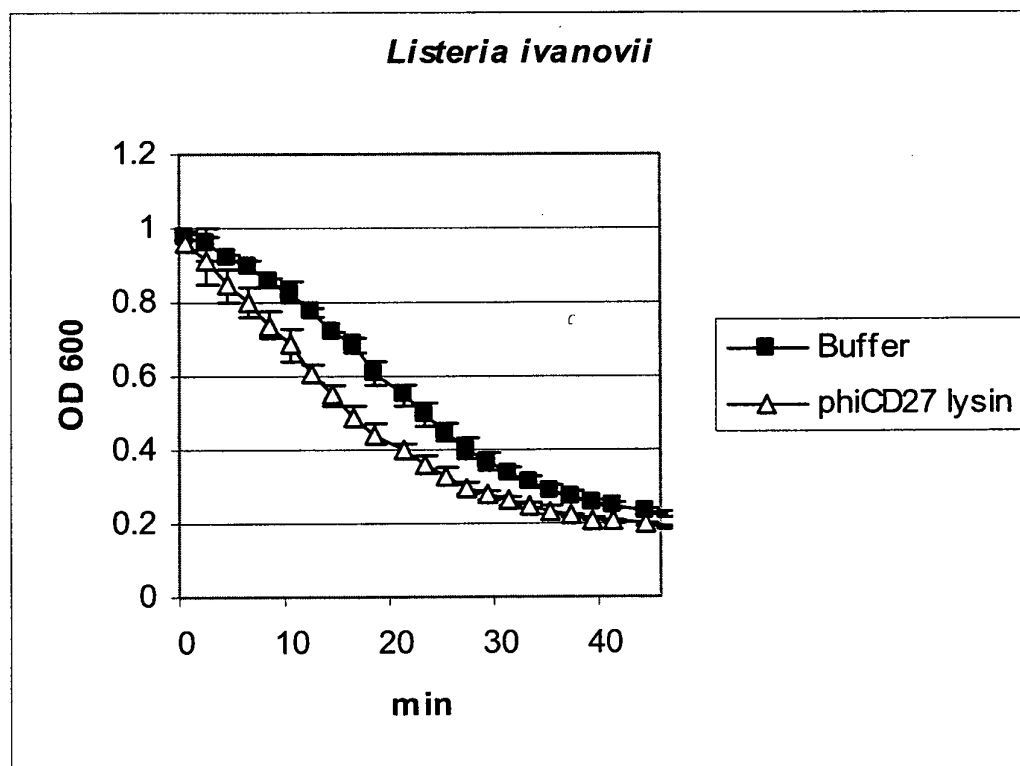


22/31

Figure 11 (b) - *continued*

23/31

Figure 11 (b) - continued



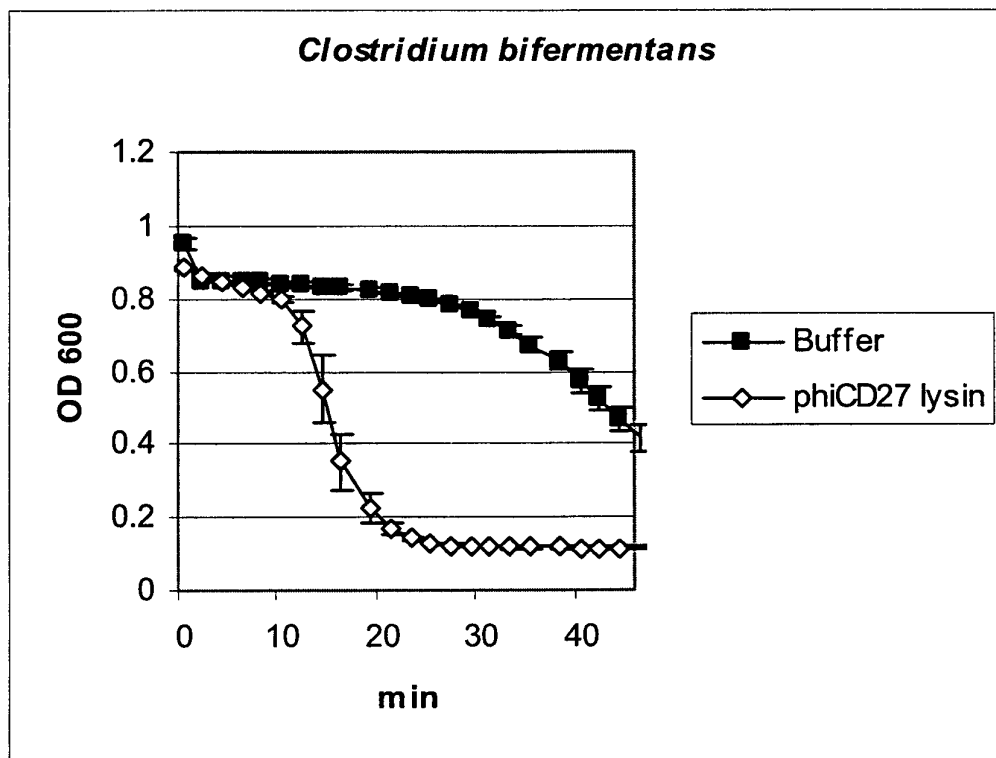
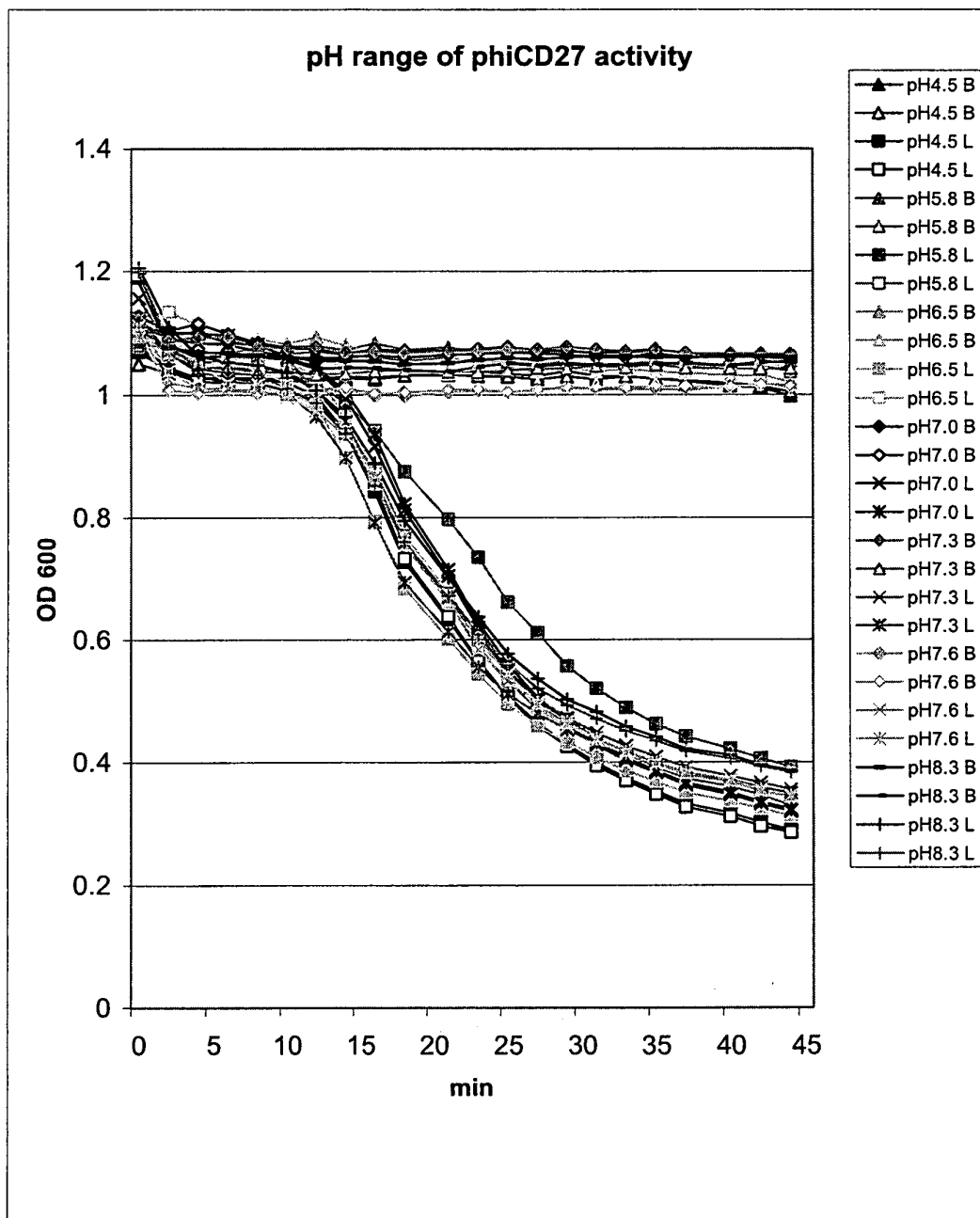
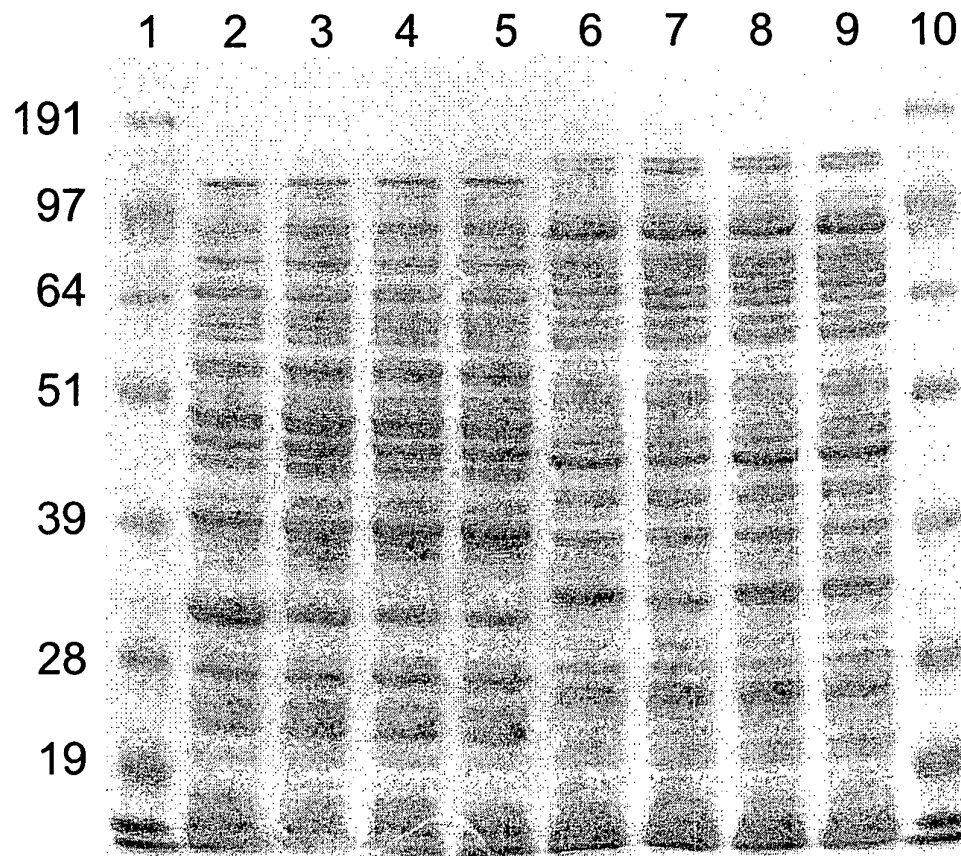
24/31Figure 11 (b) - *continued*

Figure 12



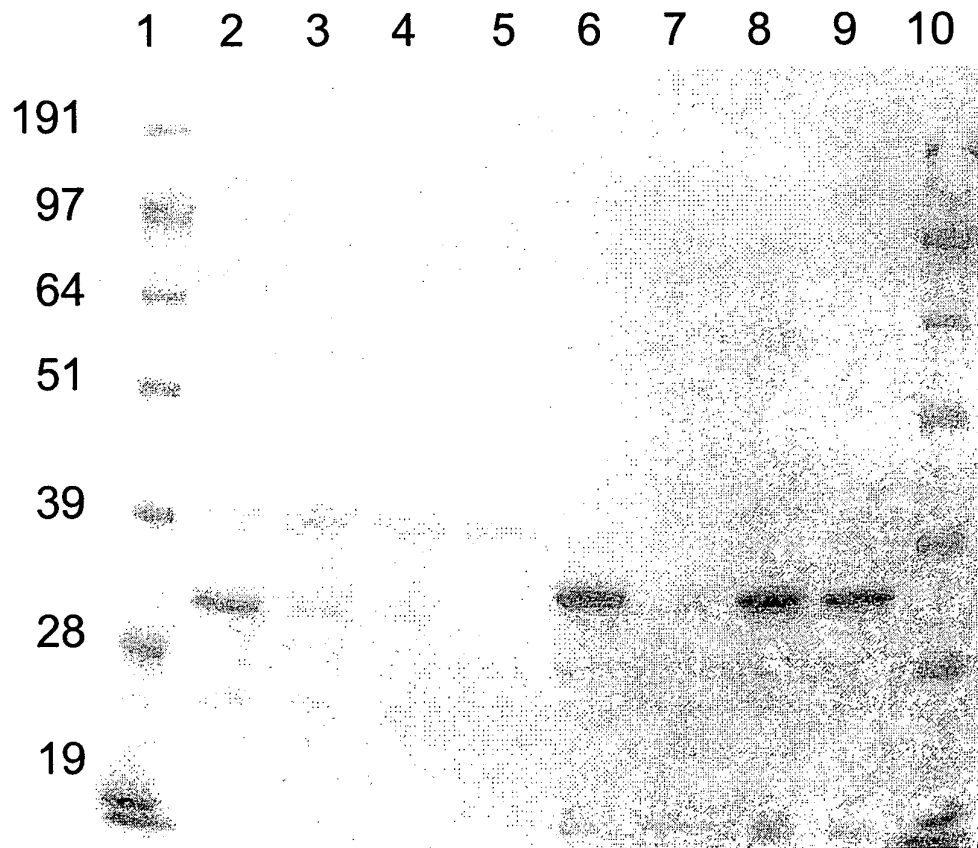
26/31

Figure 13(A)



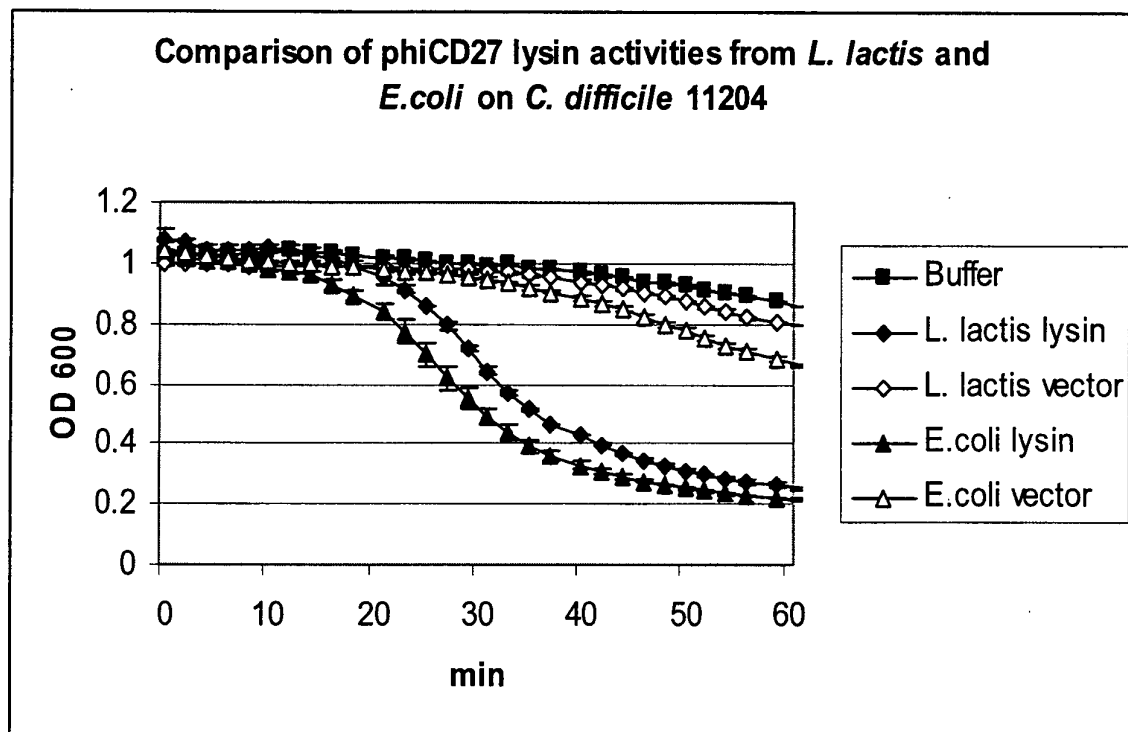
27/31

Figure 13(B)



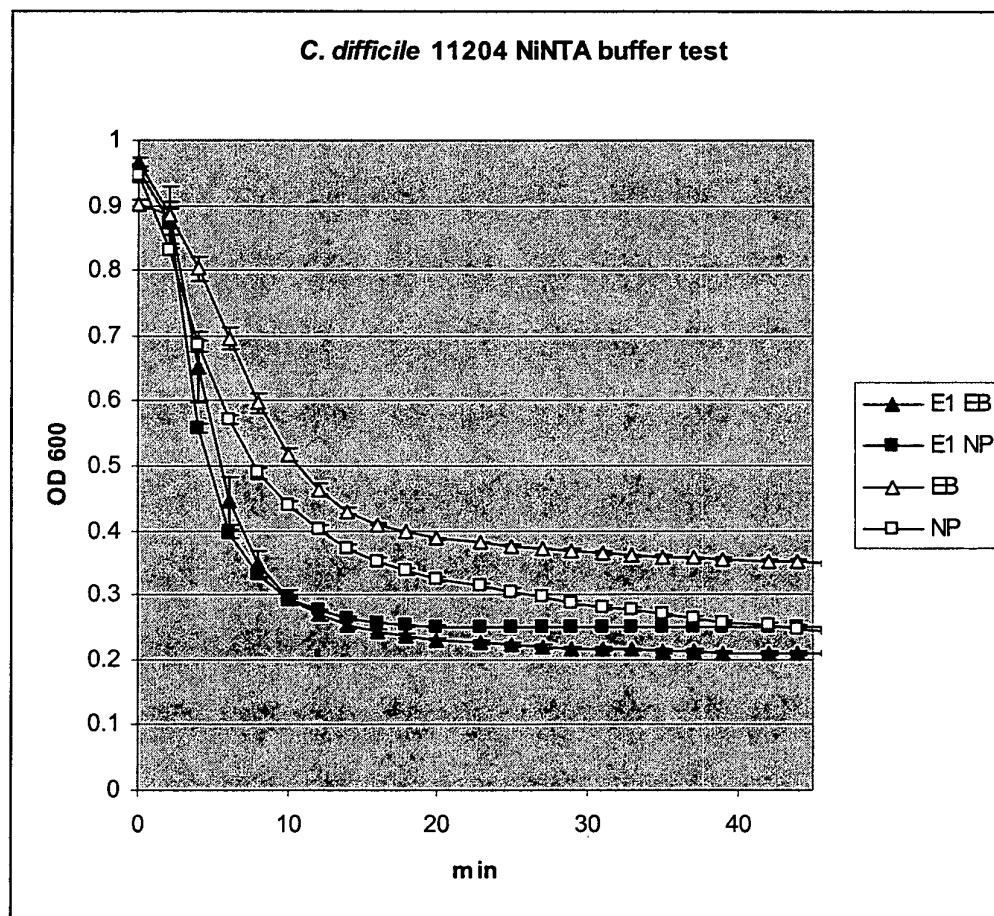
28/31

Figure 14



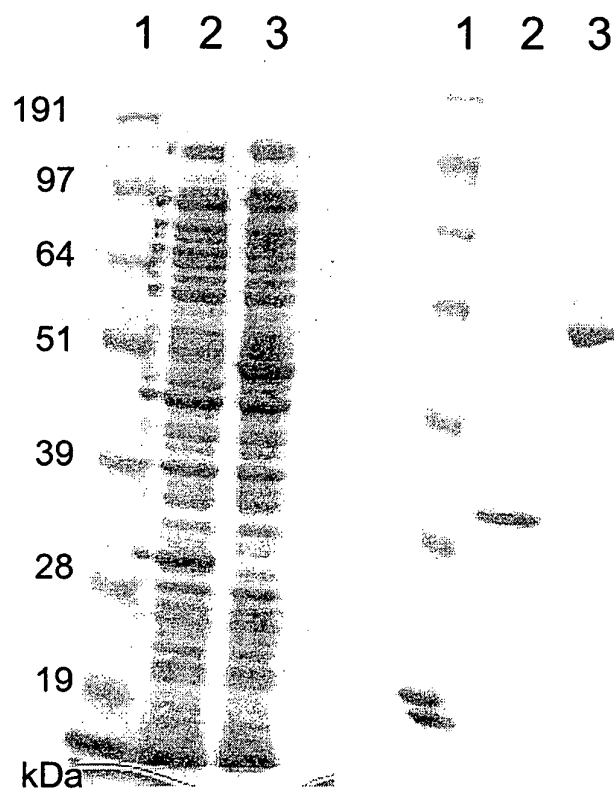
29/31

Figure 15



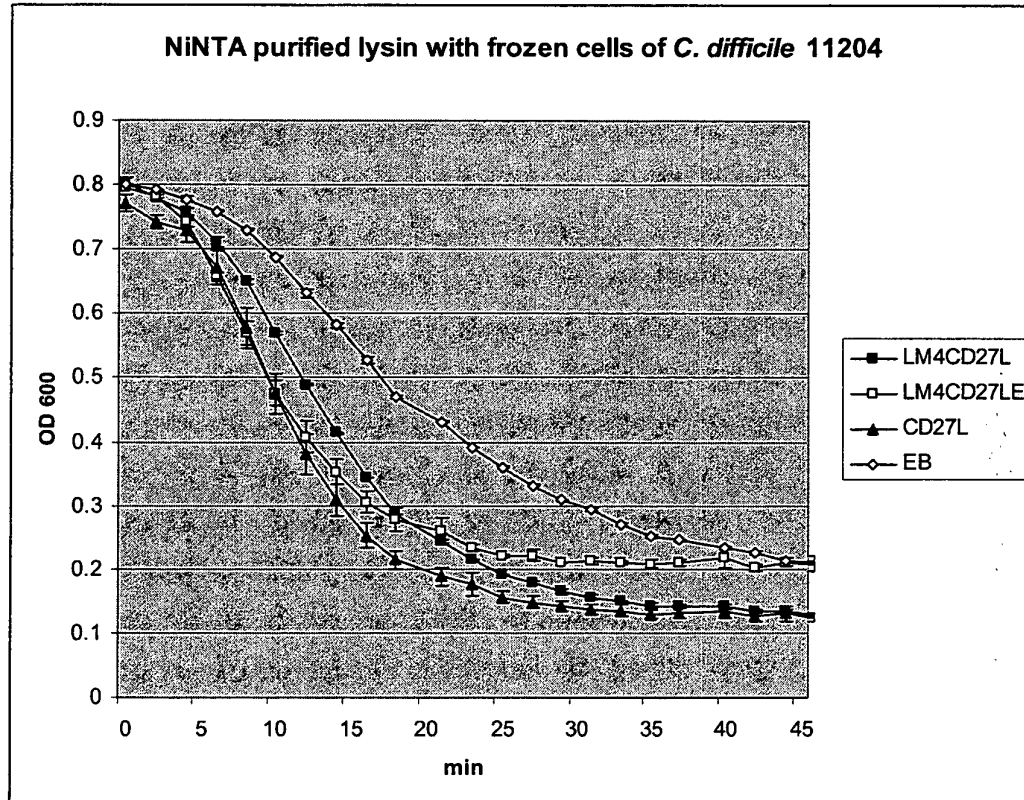
30/31

Figure 16



31/31

Figure 17



SEQUENCE LISTING

<120> Treatment of Interstitial Cystitis

<130> PC33649

<160> 77

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 120

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 1

```

Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Lys Pro Ser Glu
 1         5         10        15
Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Phe Ser Leu Ile Gly Tyr
 20        25        30
Asp Leu Asn Trp Ile Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp Ile
 35        40        45
Gly Ile Ile Trp Gly Asp Gly Thr Thr Asp Tyr Asn Ser Ala Val Lys
 50        55        60
Ser Arg Val Thr Ile Ser Lys Asp Thr Ser Lys Asn Gln Phe Ser Leu
 65        70        75        80
Lys Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala
 85        90        95
Arg Gly Gly Tyr Trp Tyr Ala Thr Ser Tyr Tyr Phe Asp Tyr Trp Gly
100       105       110
Gln Gly Thr Leu Val Thr Val Ser
115       120

```

<210> 2

<211> 109

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 2

```

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly

```

1 5 10 15
 Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Ser Ile Ser Asn Asn
 20 25 30
 Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
 35 40 45
 Tyr Tyr Thr Ser Arg Phe His Ser Gly Val Pro Ser Arg Phe Ser Gly
 50 55 60
 Ser Gly Ser Gly Thr Asp Phe Thr Phe Thr Ile Ser Ser Leu Gln Pro
 65 70 75 80
 Glu Asp Ile Ala Thr Tyr Tyr Cys Gln Gln Glu His Thr Leu Pro Tyr
 85 90 95
 Thr Phe Gly Gln Gly Thr Lys Leu Glu Ile Lys Arg Thr
 100 105

<210> 3
 <211> 10
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 3
 Gly Phe Ser Leu Ile Gly Tyr Asp Leu Asn
 1 5 10

<210> 4
 <211> 16
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 4
 Ile Ile Trp Gly Asp Gly Thr Thr Asp Tyr Asn Ser Ala Val Lys Ser
 1 5 10 15

<210> 5
 <211> 13
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 5

Gly Gly Tyr Trp Tyr Ala Thr Ser Tyr Tyr Phe Asp Tyr
1 5 10

<210> 6

<211> 11

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 6

Arg Ala Ser Gln Ser Ile Ser Asn Asn Leu Asn
1 5 10

<210> 7

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 7

Tyr Thr Ser Arg Phe His Ser
1 5

<210> 8

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 8

Gln Gln Glu His Thr Leu Pro Tyr Thr
1 5

<210> 9

<211> 10

<212> PRT

<213> Mus musculus

<400> 9

Gly Phe Ser Leu Ile Gly Tyr Asp Ile Asn
1 5 10

<210> 10

<211> 16

<212> PRT

<213> Mus musculus

<400> 10

Met Ile Trp Gly Asp Gly Thr Thr Asp Tyr Asn Ser Ala Leu Lys Ser
1 5 10 15

<210> 11

<211> 13

<212> PRT

<213> Mus musculus

<400> 11

Gly Gly Tyr Tyr Tyr Gly Thr Ser Tyr Tyr Phe Asp Tyr
1 5 10

<210> 12

<211> 11

<212> PRT

<213> Mus musculus

<400> 12

Arg Ala Ser Gln Asp Ile Ser Asn His Leu Asn
1 5 10

<210> 13

<211> 7

<212> PRT

<213> Mus musculus

<400> 13

Tyr Ile Ser Arg Phe His Ser
1 5

<210> 14
 <211> 9
 <212> PRT
 <213> Mus musculus

<400> 14
 Gln Gln Ser Lys Thr Leu Pro Tyr Thr
 1 5

<210> 15
 <211> 15
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 15
 Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
 1 5 10 15

<210> 16
 <211> 447
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 16
 Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Lys Pro Ser Glu
 1 5 10 15
 Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Phe Ser Leu Ile Gly Tyr
 20 25 30
 Asp Leu Asn Trp Ile Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp Ile
 35 40 45
 Gly Ile Ile Trp Gly Asp Gly Thr Thr Asp Tyr Asn Ser Ala Val Lys
 50 55 60
 Ser Arg Val Thr Ile Ser Lys Asp Thr Ser Lys Asn Gln Phe Ser Leu
 65 70 75 80
 Lys Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala
 85 90 95
 Arg Gly Gly Tyr Trp Tyr Ala Thr Ser Tyr Tyr Phe Asp Tyr Trp Gly
 100 105 110
 Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser
 115 120 125
 Val Phe Pro Leu Ala Pro Cys Ser Arg Ser Thr Ser Glu Ser Thr Ala

130 135 140
 Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val
 145 150 155 160
 Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala
 165 170 175
 Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val
 180 185 190
 Pro Ser Ser Asn Phe Gly Thr Gln Thr Tyr Thr Cys Asn Val Asp His
 195 200 205
 Lys Pro Ser Asn Thr Lys Val Asp Lys Thr Val Glu Arg Lys Cys Cys
 210 215 220
 Val Glu Cys Pro Pro Cys Pro Ala Pro Pro Val Ala Gly Pro Ser Val
 225 230 235 240
 Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr
 245 250 255
 Pro Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu
 260 265 270
 Val Gln Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys
 275 280 285
 Thr Lys Pro Arg Glu Glu Gln Phe Asn Ser Thr Phe Arg Val Val Ser
 290 295 300
 Val Leu Thr Val Val His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys
 305 310 315 320
 Cys Lys Val Ser Asn Lys Gly Leu Pro Ser Ser Ile Glu Lys Thr Ile
 325 330 335
 Ser Lys Thr Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro
 340 345 350
 Pro Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu
 355 360 365
 Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn
 370 375 380
 Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Met Leu Asp Ser
 385 390 395 400
 Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg
 405 410 415
 Trp Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu
 420 425 430
 His Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys
 435 440 445

<210> 17

<211> 214

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 17

```

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
1         5         10        15
Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Ser Ile Ser Asn Asn
20        25        30
Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
35        40        45
Tyr Tyr Thr Ser Arg Phe His Ser Gly Val Pro Ser Arg Phe Ser Gly
50        55        60
Ser Gly Ser Gly Thr Asp Phe Thr Phe Thr Ile Ser Ser Leu Gln Pro
65        70        75        80
Glu Asp Ile Ala Thr Tyr Tyr Cys Gln Gln Glu His Thr Leu Pro Tyr
85        90        95
Thr Phe Gly Gln Gly Thr Lys Leu Glu Ile Lys Arg Thr Val Ala Ala
100       105       110
Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser Gly
115       120       125
Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg Glu Ala
130       135       140
Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln
145       150       155       160
Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser
165       170       175
Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr
180       185       190
Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser
195       200       205
Phe Asn Arg Gly Glu Cys
210

```

<210> 18

<211> 11

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 18

```

Arg Ala Ser Gln Ser Ile Ser Asn Asn Leu Asn
1         5         10

```

<210> 19

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 19

Tyr Thr Ser Arg Phe His Ser
1 5

<210> 20

<211> 11

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 20

Arg Ala Ser Gln Tyr Ile Ser Asn His Leu Asn
1 5 10

<210> 21

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 21

Tyr Thr Ser Arg Phe His Ser
1 5

<210> 22

<211> 11

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 22

Arg Ala Ser Gln Ser Ile Ser Asn Gln Leu Asn
1 5 10

<210> 23

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 23

Tyr Val Ser Arg Phe His Ser

1 5

<210> 24

<211> 11

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 24

Arg Ala Phe Gln Ala Ile Ser Asn Gln Leu Asn

1 5 10

<210> 25

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 25

Tyr Ile Ser Arg Phe His Thr

1 5

<210> 26

<211> 11

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 26

Arg Ala Phe Gln Ser Ile Ser Asn Gln Leu Asn

1 5 10

<210> 27
 <211> 7
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 27
 Tyr Ala Ser Arg Phe His Ser
 1 5

<210> 28
 <211> 10
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 28
 Gly Phe Ser Leu Ile Gly Tyr Asp Ser Asn
 1 5 10

<210> 29
 <211> 14
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 29
 Ile Ile Trp Gly Asp Gly Thr Thr Asp Tyr Asn Ser Ala Leu
 1 5 10

<210> 30
 <211> 10
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 30
 Gly Phe Ser Leu Ile Gly Tyr Asp Leu Asn
 1 5 10

<210> 31
 <211> 14
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 31
 Ile Ile Trp Gly Asp Gly Thr Thr Asp Tyr Asn Ser Ala Val
 1 5 10

<210> 32
 <211> 10
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 32
 Gly Phe Ser Leu Ile Gly Tyr Asp Val Thr
 1 5 10

<210> 33
 <211> 14
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 33
 Gly Ile Trp Gly Asp Gly Thr Thr Asp Tyr Asn Ser Ala Val
 1 5 10

<210> 34
 <211> 10
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 34

Gly Phe Ser Leu Ile Gly Tyr Asp Val Thr
 1 5 10

<210> 35
 <211> 14
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 35
 Gly Ile Trp Gly Asp Gly Thr Thr Asp Tyr Asn Ser Ser Val
 1 5 10

<210> 36
 <211> 10
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 36
 Gly Phe Ser Leu Ile Gly Tyr Asp Ala Thr
 1 5 10

<210> 37
 <211> 14
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 37
 Gly Ile Trp Gly Asp Gly Thr Thr Asp Tyr Asn Ser Ala Val
 1 5 10

<210> 38
 <211> 10
 <212> PRT
 <213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 38

Gly Phe Ser Leu Ile Gly Tyr Asp Val Ser
1 5 10

<210> 39

<211> 14

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 39

Ile Ile Trp Gly Asp Gly Thr Thr Asp Tyr Asn Ser Ser Val
1 5 10

<210> 40

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 40

Gly Phe Ser Leu Ile Gly Tyr Asp Ile Ser
1 5 10

<210> 41

<211> 14

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 41

Gln Ile Trp Gly Asp Gly Thr Thr Asp Tyr Asn Ser Ser Val
1 5 10

<210> 42

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 42

Gly Phe Ser Leu Ile Gly Tyr Asp Ala Ser
1 5 10

<210> 43

<211> 14

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 43

Gly Ile Trp Gly Asp Gly Thr Thr Asp Tyr Asn Ser Ser Val
1 5 10

<210> 44

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 44

Gly Phe Ser Leu Ile Gly Tyr Asp Ser Thr
1 5 10

<210> 45

<211> 14

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 45

Ser Ile Trp Gly Asp Gly Thr Thr Asp Tyr Asn Ser Ala Leu
1 5 10

<210> 46

<211> 13
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 46
 Gly Gly Tyr Trp Tyr Gly Thr Ser Tyr Tyr Phe Asp Tyr
 1 5 10

<210> 47
 <211> 13
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 47
 Gly Gly Tyr Tyr Tyr Gly Thr Ala Tyr Tyr Phe Asp Tyr
 1 5 10

<210> 48
 <211> 13
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 48
 Gly Gly Tyr Tyr Tyr Gly Thr Thr Tyr Tyr Phe Asp Tyr
 1 5 10

<210> 49
 <211> 13
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 49
 Gly Gly Tyr Tyr Tyr Ala Thr Ser Tyr Tyr Phe Asp Tyr
 1 5 10

<210> 50
 <211> 9
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 50
 Gln Gln Glu Lys Thr Leu Pro Tyr Thr
 1 5

<210> 51
 <211> 9
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 51
 Gln Gln Glu Ala Thr Leu Pro Tyr Thr
 1 5

<210> 52
 <211> 13
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 52
 Gly Gly Tyr Trp Tyr Ala Thr Ser Tyr Tyr Phe Asp Tyr
 1 5 10

<210> 53
 <211> 9
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 53

Gln Gln Glu Arg Thr Leu Pro Tyr Thr
 1 5

<210> 54
 <211> 13
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 54
 Gly Gly Tyr Trp Tyr Ala Thr Ser Tyr Tyr Phe Asp Tyr
 1 5 10

<210> 55
 <211> 9
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 55
 Gln Gln Glu His Thr Leu Pro Tyr Thr
 1 5

<210> 56
 <211> 13
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 56
 Gly Gly Tyr Trp Tyr Ala Thr Ser Tyr Tyr Phe Asp Tyr
 1 5 10

<210> 57
 <211> 9
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 57

Gln Gln Glu Ser Thr Leu Pro Tyr Thr

1 5

<210> 58

<211> 13

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 58

Gly Gly Tyr Trp Tyr Ser Thr Ser Tyr Tyr Phe Asp Tyr

1 5 10

<210> 59

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 59

Gln Gln Glu Lys Thr Leu Pro Tyr Thr

1 5

<210> 60

<211> 13

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 60

Gly Gly Tyr Tyr Tyr Ala Thr Ser Tyr Tyr Phe Asp Tyr

1 5 10

<210> 61

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 61

Gln Gln Glu Arg Thr Leu Pro Tyr Thr
1 5

<210> 62

<211> 13

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 62

Gly Gly Tyr Trp Tyr Ala Thr Ser Tyr Tyr Phe Asp Tyr
1 5 10

<210> 63

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 63

Gln Gln Glu Arg Thr Leu Pro Tyr Thr
1 5

<210> 64

<211> 13

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 64

Gly Gly Tyr Tyr Tyr Ala Thr Ser Tyr Tyr Phe Asp Tyr
1 5 10

<210> 65

<211> 1344

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 65

```
caggtgcagc tgcaggagtc tggcccagga ctggtgaagc cttccgagac cctgtccctc 60
acctgcaactg tctctgggtt ctaacttate ggctatgac ttaactggat cgcacagcct 120
ccagggaagg gactggagtg gattgggatt atctgggggtg atggaaccac agactataat 180
tcagctgtca aatcccgcgt caccatctca aaagacacct ccaagaacca gttctccctg 240
aagctgagct ctgtgaccgc cgcggacacg gccgtgtatt actgtgagag aggaggttat 300
tggtacgcca ctagtacta ctttgactac tggggccagg gcacctggt caccgtctcc 360
tcagcctcca ccaagggccc atctgtcttc cactggccc catgtcccg cagcacctcc 420
gagagcacag ccgccttggg ctgctgtgac aaggactact tcccagaacc tgtgaccgtg 480
tcttgaact ctggcgtct gaccagcggc gtgcacacct tccagctgt cctgcagtcc 540
tcaggtctct actcctcag cagcgtggtg accgtgcat ccagcaactt cggcaccag 600
acctacacct gcaacgtaga tcacaagcca agcaacacca aggtcgacaa gaccgtggag 660
agaaagtgtt gtgtggagtg tccacctgt ccagccctc cagtggccgg accatccgtg 720
ttctgttcc ctcaaagcc aaaggacacc ctgatgatct ccagaacccc agaggtgacc 780
tgtgtggtgg tggacgtgc ccacgaggac ccagaggtgc agttcaactg gtatgtggac 840
ggagtggagg tgcacaacgc caagaccaag ccaagagagg agcagttcaa ctccacctc 900
agagtgggtg gctgtctgac cgtggtgcac caggactggc tgaacggaaa ggagtataag 960
tgaaggtgt ccaacaaggg actgccatcc agcatcgaga agaccatctc caagaccaag 1020
ggacagccaa gagagccaca ggtgtatacc ctgccacat ccagagagga gatgaccaag 1080
aaccaggtgt cctgacctg tctggtgaag ggattctatc catccgacat cgccgtggag 1140
tggtgagcca acggacagcc agagaacaac tataagacca cccctccaat gctggactcc 1200
gacggatcct tcttctgta ttccaagctg accgtggaca agtccagatg gcagcaggga 1260
aacgtgttct cttgtccgt gatgcacgag gcctgcaca accactatac ccagaagagc 1320
ctgtccctgt ctccaggaaa gtaa                                     1344
```

<210> 66

<211> 363

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 66

```
caggtgcagc tgcaggagtc tggcccagga ctggtgaagc cttccgagac cctgtccctc 60
acctgcaactg tctctgggtt ctaacttate ggctatgac ttaactggat cgcacagcct 120
ccagggaagg gactggagtg gattgggatt atctgggggtg atggaaccac agactataat 180
tcagctgtca aatcccgcgt caccatctca aaagacacct ccaagaacca gttctccctg 240
aagctgagct ctgtgaccgc cgcggacacg gccgtgtatt actgtgagag aggaggttat 300
tggtacgcca ctagtacta ctttgactac tggggccagg gcacctggt caccgtctcc 360
tca                                     363
```

<210> 67

<211> 645

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 67

```

gatatccaga tgacacagtc cccatcctcc ctgtctgcct ctgtgggtga ccgcgtcacc 60
atcacctgcc ggcacatca gtccattagc aataatctga actggtatca gcagaagcca 120
ggcaaagccc caaaactcct gatctactac acctcacgct tccactcagg tgccecatca 180
cgcttcagtg gcagtggctc tggtagatg ttcaccttca ccattagcag cctgcaacca 240
gaagatattg ccacttatta ctgccaacag gagcataccc ttccatatac cttecggtcaa 300
ggcaccaagc tggagatcaa acgcactgtg gctgcacat ctgtcttcat ctttctcca 360
tctgatgagc agttgaaatc cggaactgcc tctgtgtgtg gctgctgaa taacttctat 420
ccacgcgagg ccaaagtaca gtggaagggtg gataacgccc tccaatccgg taactcccag 480
gagagtgtca cagagcagga cagcaaggac agcacctaca gcctcagcag caccctgacc 540
ctgagcaaag cagactacga gaaacacmaa gtctacgcct gcgaagtcac ccacagggc 600
ctgagttctc cagtcacaaa gagcttcaac cgcggtgagt gctaa          645

```

<210> 68

<211> 324

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 68

```

gatatccaga tgacacagtc cccatcctcc ctgtctgcct ctgtgggtga ccgcgtcacc 60
atcacctgcc ggcacatca gtccattagc aataatctga actggtatca gcagaagcca 120
ggcaaagccc caaaactcct gatctactac acctcacgct tccactcagg tgccecatca 180
cgcttcagtg gcagtggctc tggtagatg ttcaccttca ccattagcag cctgcaacca 240
gaagatattg ccacttatta ctgccaacag gagcataccc ttccatatac cttecggtcaa 300
ggcaccaagc tggagatcaa acgc          324

```

<210> 69

<211> 98

<212> PRT

<213> Homo sapiens

<400> 69

```

Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Lys Pro Ser Glu
 1         5         10        15
Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Gly Ser Ile Ser Ser Tyr
 20        25        30
Tyr Trp Ser Trp Ile Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp Ile
 35        40        45
Gly Tyr Ile Tyr Tyr Ser Gly Ser Thr Asn Tyr Asn Pro Ser Leu Lys
 50        55        60

```

Ser Arg Val Thr Ile Ser Val Asp Thr Ser Lys Asn Gln Phe Ser Leu
65 70 75 80
Lys Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala
 85 90 95
Arg Gly

<210> 70

<211> 120

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 70

Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Lys Pro Ser Glu
1 5 10 15
Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Phe Ser Leu Ile Gly Tyr
 20 25 30
Asp Ile Asn Trp Ile Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp Ile
 35 40 45
Gly Met Ile Trp Gly Asp Gly Thr Thr Asp Tyr Asn Ser Ala Leu Lys
 50 55 60
Ser Arg Val Thr Ile Ser Val Asp Thr Ser Lys Asn Gln Phe Ser Leu
65 70 75 80
Lys Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala
 85 90 95
Arg Gly Gly Tyr Tyr Tyr Gly Thr Ser Tyr Tyr Phe Asp Tyr Trp Gly
 100 105 110
Gln Gly Thr Leu Val Thr Val Ser
 115 120

<210> 71

<211> 120

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Construct

<400> 71

Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Lys Pro Ser Glu
1 5 10 15
Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Phe Ser Leu Ile Gly Tyr
 20 25 30
Asp Ile Asn Trp Ile Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp Ile
 35 40 45
Gly Met Ile Trp Gly Asp Gly Thr Thr Asp Tyr Asn Ser Ala Leu Lys

50 55 60
 Ser Arg Val Thr Ile Ser Lys Asp Thr Ser Lys Asn Gln Phe Ser Leu
 65 70 75 80
 Lys Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala
 85 90 95
 Arg Gly Gly Tyr Tyr Tyr Gly Thr Ser Tyr Tyr Phe Asp Tyr Trp Gly
 100 105 110
 Gln Gly Thr Leu Val Thr Val Ser
 115 120

<210> 72
 <211> 120
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 72
 Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Lys Pro Ser Glu
 1 5 10 15
 Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Phe Ser Leu Ile Gly Tyr
 20 25 30
 Asp Leu Asn Trp Ile Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp Ile
 35 40 45
 Gly Ile Ile Trp Gly Asp Gly Thr Thr Asp Tyr Asn Ser Ala Val Lys
 50 55 60
 Ser Arg Val Thr Ile Ser Lys Asp Thr Ser Lys Asn Gln Phe Ser Leu
 65 70 75 80
 Lys Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala
 85 90 95
 Arg Gly Gly Tyr Tyr Tyr Gly Thr Ser Tyr Tyr Phe Asp Tyr Trp Gly
 100 105 110
 Gln Gly Thr Leu Val Thr Val Ser
 115 120

<210> 73
 <211> 95
 <212> PRT
 <213> Homo sapiens

<400> 73
 Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
 1 5 10 15

Asp Arg Val Thr Ile Thr Cys Gln Ala Ser Gln Asp Ile Ser Asn Tyr
 20 25 30
 Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
 35 40 45
 Tyr Asp Ala Ser Asn Leu Glu Thr Gly Val Pro Ser Arg Phe Ser Gly
 50 55 60
 Ser Gly Ser Gly Thr Asp Phe Thr Phe Thr Ile Ser Ser Leu Gln Pro
 65 70 75 80
 Glu Asp Ile Ala Thr Tyr Tyr Cys Gln Gln Tyr Asp Asn Leu Pro
 85 90 95

<210> 74
 <211> 109
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 74
 Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
 1 5 10 15
 Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Asp Ile Ser Asn His
 20 25 30
 Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
 35 40 45
 Tyr Tyr Ile Ser Arg Phe His Ser Gly Val Pro Ser Arg Phe Ser Gly
 50 55 60
 Ser Gly Ser Gly Thr Asp Phe Thr Phe Thr Ile Ser Ser Leu Gln Pro
 65 70 75 80
 Glu Asp Ile Ala Thr Tyr Tyr Cys Gln Gln Ser Lys Thr Leu Pro Tyr
 85 90 95
 Thr Phe Gly Gln Gly Thr Lys Leu Glu Ile Lys Arg Thr
 100 105

<210> 75
 <211> 109
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Synthetic Construct

<400> 75
 Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
 1 5 10 15
 Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Ser Ile Ser Asn Asn
 20 25 30

Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
 35 40 45
 Tyr Tyr Thr Ser Arg Phe His Ser Gly Val Pro Ser Arg Phe Ser Gly
 50 55 60
 Ser Gly Ser Gly Thr Asp Phe Thr Phe Thr Ile Ser Ser Leu Gln Pro
 65 70 75 80
 Glu Asp Ile Ala Thr Tyr Tyr Cys Gln Gln Ser Lys Thr Leu Pro Tyr
 85 90 95
 Thr Phe Gly Gln Gly Thr Lys Leu Glu Ile Lys Arg Thr
 100 105

<210> 76

<211> 1429

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct

<400> 76

atggccaccg actccagaac ctectggctg ctgacagtgt cctgctgtg tctgctgtgg 60
 ccacaggagg ccagcgtca ggtgcagctg caggagtctg gccaggact ggtgaagcct 120
 tccgagacc tgctcctcac ctgactgtc tctgggtct cacttatcgg ctatgatctt 180
 aactggatcc gacagcctcc aggggaagga ctggagtga ttgggattat ctgggggtgat 240
 ggaaccacag actataattc agctgtcaaa tcccgctca ccatctcaaa agacacctcc 300
 aagaaccagt tctcctgaa gctgagctct gtgaccgcc cggacacggc cgtgtattac 360
 tgtcgagag gaggttattg gtacccact agctactact ttgactactg gggccagggc 420
 acctgggtca cgtctcctc agcctccacc aagggcccat ctgtctccc actggcccca 480
 tgctcccgca gcacctccga gagcacagcc gccctgggct gcctggtaa ggactactc 540
 ccagaacctg tgacctgtc ctggaactct ggcgtctga ccagcggcgt gcacacctc 600
 ccagctgtcc tgcagctc aggtctctac tccctcagca gcgtgggtgac cgtgccatcc 660
 agcaactcg gcaccagac ctacacctgc aacgtagatc acaagccaag caacaccaag 720
 gtgcacaaga cgtggagag aaagtgtgt gtggagtgtc cacctgtcc agcccctcca 780
 gtggccggac catcctgtt cctgttccct ccaaagccaa aggacacct gatgatctcc 840
 agaaccccag aggtgacctg tgtggtggtg gacgtgtccc acgaggacct agaggtgcag 900
 ttcaactggt atgtggacgg agtggaggtg cacaacgcca agaccaagcc aagagaggag 960
 cagttcaact ccacctcag agtggtgagc gtgctgaccg tggtgacca ggactggctg 1020
 aacggaaagg agtataagt taaggtgtcc aacaaggac tgccatccag catcgagaag 1080
 accatctcca agaccaaggg acagccaaga gagccacagg tgtataacct gccaccatcc 1140
 agagaggaga tgaccaagaa ccaggtgtcc ctgacctgtc tggtaaggg attctatcca 1200
 tccgacatcg ccgtggagt ggagtccaac ggacagccag agaacaacta taagaccacc 1260
 cctccaatgc tggactccga cggatcctc tctctgatt ccaagctgac cgtggacaag 1320
 tccagatggc agcagggaaa cgtgttctc ttgtccgtga tgcacaggc cctgcacaac 1380
 cactataccc agaagagcct gtccctgtct ccaggaaagt aattctaga 1429

<210> 77

<211> 729

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct

<220>

<221> misc_feature

<222> 646

<223> m = A or C

<400> 77

atggccaccg actccagaac ctctggctg ctgacagtgt ccctgctgtg tctgctgtgg 60
 ccacaggagg ccagcgctga tatccagatg acacagtcct cctctccct gtctgcctct 120
 gtgggtgacc gcgtcccat cacctgccgc gcatctcagt ccattagcaa taatctgaac 180
 tggtatcagc agaagccagg caaagcccca aaactcctga tctactacac ctcacgttc 240
 cactcaggtg tcccatcacg ctccagtggc agtggctctg gtacagattt cacctcacc 300
 attagcagcc tgcaaccaga agatattgcc acttattact gccaacagga gcataccctt 360
 ccatatacct tcggtcaagg caccaagctg gagatcaaac gcactgtggc tgcaccatct 420
 gtcttcattt ttctccatc tgatgagcag ttgaaatccg gaactgcctc tgttgtgtgc 480
 ctgtgaata acttctatcc acgcgaggcc aaagtacagt ggaaggtgga taacgcctc 540
 caatccggtg actcccagga gagtgtcaca gagcaggaca gcaaggacag cacctacagc 600
 ctcagcagca ccctgaccct gagcaaagca gactacgaga aacacmaagt ctacgcctgc 660
 gaagtcaccc atcagggcct gaggttctca gtcacaaaga gcttcaaccg cggtgagtgc 720
 taattctag

729