(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 March 2013 (21.03.2013)

WIPO | PCT

(10) International Publication Number
WO 2013/038385 A2

(51) International Patent Classification:
C07K 14/245 (2006.01) *A61P 31/04* (2006.01)
A61K 39/108 (2006.01)

(21) International Application Number:
PCT/IB2012/054825

(22) International Filing Date:
14 September 2012 (14.09.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
1115906.8 14 September 2011 (14.09.2011) GB
1213251.0 25 July 2012 (25.07.2012) GB

(71) Applicant *(for all designated States except US)*: **NOVARTIS AG** [CH/CH]; Lichtstrasse 35, CH-4056 Basel (CH).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **SERINO, Laura** [IT/IT]; Novartis Vaccines and Diagnostics, Inc., Via Fiorentina 1, I-53100 Siena (IT). **PIZZA, Mariagrazia** [IT/IT]; Novartis Vaccines and Diagnostics, Inc., Via Fiorentina 1, I-53100 Siena (IT).

(74) Agents: **MARSHALL, Cameron John** et al.; Carpmaels & Ransford, One Southampton Row, London WC1B 5HA (GB).

(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, BN, BR, BW, BY, BZ, CA, CH, CL, 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, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

Published:
— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*
— *with sequence listing part of description (Rule 5.2(a))*

(54) **Title**: ESCHERICHIA COLI VACCINE COMBINATION

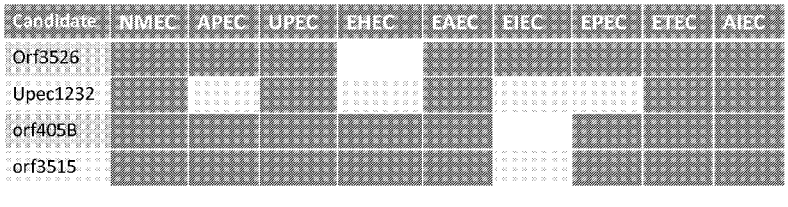


Fig. 1A

Candidate	ExPEC	EHEC	EPEC	ETEC	AIEC
Orf3526	51/59	6/25	5/6	9/9	4/4
orf405B	54/59	24/25	6/6	8/9	3/4
Upec1232	46/59	4/25	2/4	8/9	1/4

Fig. 1B

(57) **Abstract**: The invention provides an immunogenic composition comprising a combination of (i) bacterial Ig-like domain protein fragment (orf405B) having the amino acid sequence set forth in SEQ ID NO:2 or a protein having at least 80% similarity thereto, and (ii) putative Lipoprotein (orf3526) having the amino acid sequence set forth in SEQ ID NO:8 or a protein having at least 80% similarity thereto.

WO 2013/038385 A2

ESCHERICHIA COLI VACCINE COMBINATION

TECHNICAL FIELD

The present invention relates generally to the field of immunisation against *E. coli* and *E. coli* vaccines. More specifically, the invention relates to combinations of polypeptides useful in the preparation of prophylactic and therapeutic vaccine combinations for use in immunisation against pathogenic *E. coli* pathotypes. In particular, it relates to a vaccine useful in protecting humans against a broad spectrum of *E. coli* strains.

BACKGROUND TO THE INVENTION

Several publications and patent documents are referenced in this application in order to describe more fully the state of the art to which this invention pertains. The disclosure of each of these publications is incorporated herein by reference.

Escherichia coli is a common colonizer of the human gastrointestinal tract and although *E. coli* strains are largely regarded as commensal, some isolates have the potential to cause diseases. Two distinct pathogenic categories of *E. coli* are recognized depending on whether they cause intestinal or extraintestinal infections. The extraintestinal pathogenic *E. coli* (ExPEC) group includes human pathogenic strains causing urinary tract infections (UPEC), neonatal meningitis (NMEC) and septicemia. Other ExPEC strains are instead pathogenic for avian species (APEC). The intestinal pathogenic *E. coli* (InPEC) group includes many pathotypes such as enterotoxigenic (ETEC), enteropathogenic (EPEC), enterohemorrhagic (EHEC), enteroinvasive, adherent invasive, and diffusely adherent *E. coli*, all causing infections to the human intestinal tract. These pathogenic strains of *Escherichia coli* are the most common cause of bacterial infections presenting a recurrent global threat that kills over two million of people in the world every year.

Vaccination would undoubtedly be the most cost-effective preventive measure against morbidity and death from pathogenic *E. coli* strains. However, primary studies have focused attention on the identification of candidates for use in vaccines protective against individual pathogenic categories, for example against UPEC alone or NMEC alone. In an earlier publication based on a comparative genome analysis and using the complete genome sequences of three ExPEC strains (IHE3034, CFT073 and 536), the Inventors identified nine potential vaccine candidates able to confer protection from sepsis (PNAS 2010; 107:20, p9072-9077).

Whilst vaccines for use against individual diseases or illnesses are useful, it would be desirable to provide broad spectrum vaccines that provide protective immunity in animals, particularly humans, against all, or a large number, of infections caused by *E. coli*. For example, once vaccinated an individual could be covered or protected against all, or a high percentage, of the different diseases that *E. coli* can cause. A

broadly protective vaccine would be of further benefit due to the spread of antibiotic resistant bacteria (CTX-M β-lactamase and carbepenemases) in hospitals and communities as a whole. However, the development of such a 'universal' or 'pan *E. coli*' vaccine is challenging because of the need to selectively prevent against subtypes of *E. coli* strains that are not normally part of the commensal flora. There is thus a need for improved *E. coli* vaccines, including a need to move away from crude cell lysates and towards better-defined molecules, and a need to identify antigens that are suitable for inclusion in a 'universal' vaccine, particularly antigens that are prevalent among clinically relevant strains without also being found in commensal strains.

In addition, a needle-free or mucosally administered vaccine would be preferable for reasons of improved patient comfort and ease of administration, as well as reducing the risk of contamination and other adverse effects while promoting patient compliance and increasing the safety of vaccination.

The present Inventors have discovered a combination of antigens suitable for use in the preparation of a broad spectrum vaccine against pathogenic *E. coli*. Surprisingly, the vaccine combination provides protection against disease/illness caused by pathogenic strains of *E. coli* from both ExPEC and InPEC groups.

BRIEF DESCRIPTION OF THE INVENTION

In a first aspect, the invention provides an immunogenic composition comprising a combination of (i) bacterial Ig-like domain protein fragment (orf405B) having the amino acid sequence set forth in SEQ ID NO:2 or a protein having at least 80% similarity thereto, and (ii) putative Lipoprotein (orf3526) having the amino acid sequence set forth in SEQ ID NO:8 or a protein having at least 80% similarity thereto.

In certain embodiments the immunogenic compositions further comprise (iii) upec1232 having the amino acid sequence set forth in SEQ ID NO:4 or a protein having at least 80% similarity thereto.

In further embodiments the immunogenic compositions further comprise (iv) gspK (orf3515) having the amino acid sequence set forth in SEQ ID NO:30 or a protein having at least 80% similarity thereto.

In other embodiments the immunogenic compositions further comprise at least one bacterial toxin. Particularly, the bacterial toxin is an Escherichia coli toxin. More particularly, the bacterial toxin is modified heat-labile toxin of Escherichia coli (LTK63), yet more particularly detoxified heat-labile toxin of Escherichia coli (LTK63).

Protein components of the compositions of the invention may be fragments of the proteins or amino acid sequences mentioned herein.

In certain embodiments, the putative Lipoprotein (orf3526) utilised in the immunogenic compositions is a mutant protein wherein at least one amino acid (*e.g.*, 1, 2, 3, 4 or 5 amino acids) at positions 1304, 1305, 1306, 1307 and/or 1308 with reference to SEQ ID 8 is/are substituted by another amino acid. In certain embodiments, the putative Lipoprotein (orf3526) is a mutant orf3526 protein wherein the zinc binding activity is reduced by at least 50% relative to wild-type orf3526. In certain embodiments, the mutant has a zinc content which is at least 50% lower than the content of an equivalent amount of wild-type orf3526. The mutant polypeptides may be lipidated *e.g.* at an N-terminal cysteine. The mutant polypeptides may be prepared having a reduced zinc ion content or substantially free from zinc ions, relative to other variants of orf3526 polypeptide or fragments thereof, for example relative to wild-type orf3526 polypeptide. A mutant orf3526 protein having such reduced the zinc binding activity may have one or more (*e.g.*, 2, 3, 4 or 5) of the aforementioned amino acid substitutions at positions 1304, 1305, 1306, 1307 and/or 1308. Particular mutant orf3526 proteins comprise the amino acid sequence of SEQ ID: 31, or immunogenic fragments thereof, which include mutations at positions 1304, 1305 and 1308. One such fragment includes mutations at positions 1304, 1305 and 1308 and comprises amino acid residues 24-1520, or residues 34-1520, of SEQ ID 31. The mutant polypeptides, or immunogenic fragments thereof, may be prepared having a reduced zinc ion content or substantially free from zinc ions, relative to other variants of orf3526 polypeptide or fragments thereof, for example relative to wild-type orf3526 polypeptide.

In certain embodiments, isoform B (corresponding to peak B in Figure 9a) of orf3526, or an immunogenic fragment thereof, is preferred. One such exemplary immunogenic fragments is isoform B of a polypeptide that includes mutations at positions 1304, 1305 and 1308 and comprises amino acid residues 24-1520, or residues 34-1520, of SEQ ID 31.

In other embodiments, isoform A (corresponding to peak A in Figure 9a) of orf3526, or an immunogenic fragment thereof, is preferred. One such exemplary immunogenic fragments is isoform A of a polypeptide that includes mutations at positions 1304, 1305 and 1308 and comprises amino acid residues 24-1520, or residues 34-1520, of SEQ ID 31.

In other embodiments, isoform C (corresponding to peak C in Figure 9a) of orf3526, or an immunogenic fragment thereof, is preferred.

In further embodiments, a combination of at least two of isoforms A, B and C of orf3526, or immunogenic fragments thereof, is preferred. In particular, a combination of isoform A and B of orf3526, or immunogenic fragments thereof, is preferred. For example, a combination of isoforms A and B of a polypeptide that includes mutations at positions 1304, 1305 and 1308 and comprises amino acid residues 24-1520, or residues 34-1520, of SEQ ID 31, is preferred.

The immunogenic compositions of the invention may comprise one or more pharmaceutically acceptable carriers, diluents and/or adjuvants. The immunogenic compositions of the invention may comprise propane-1,2,3-triol (glycerol). The immunogenic compositions of the invention may be vaccines, or vaccine compositions.

In other aspects there is provided a method for treating or preventing E. coli infection in a mammal, which comprises administering to said mammal an effective amount of an immunogenic composition according to the invention. In certain embodiments the immunogenic composition will be administered to a mucosal surface such as nasal epithelium, oral mucosa or luminal surface of a gastrointestinal organ selected from the group consisting of: stomach, small intestine, large intestine, and rectum. Preferably, immunogenic compositions of the present invention are administered by parenteral administration.

In other aspects there is provided the use of immunogenic compositions of the invention in medicine, *e.g.* for treating or preventing *E. coli* infections in a mammal, in particular for providing broad protection against pathogenic *E.coli*, e.g. extraintestinal or intraintestinal pathogenic *E. coli*, in particular for treating or preventing infections by more than one *E. coli* pathotype, e.g. infections by both extraintestinal and intraintestinal pathogenic *E. coli, i.e.* both ExPEC and InPEC pathotypes, such as NMEC, APEC, UPEC, EHEC, AIEC, EPEC, EAEC, EIEC, ETEC and DAEC pathotypes. Thus the subject may be protected against diseases including, but not limited to peritonitis, pyelonephritis, cystitis, endocarditis, prostatitis, urinary tract infections (UTIs), meningitis (particularly neonatal meningitis), sepsis (or SIRS), dehydration, pneumonia, diarrhea (infantile, travellers', acute, persistent, *etc.*), bacillary dysentery, hemolytic uremic syndrome (HUS), pericarditis, bacteriuria, *etc.* Thus, the invention provides the use of immunogenic compositions of the invention for the manufacture of a medicament for treating or preventing *E. coli* infections, e.g. extraintestinal or intraintestinal pathogenic *E. coli*, in particular infections by more than one *E.coli* pathotype, e.g. infections by both extraintestinal and intraintestinal pathogenic *E. coli, i.e.* both ExPEC and InPEC pathotypes, such as NMEC, APEC, UPEC, EHEC, AIEC, EPEC, EAEC, EIEC, ETEC and DAEC pathotypes, or any of the aforementioned diseases.

The invention also provides orf3526 mutant polypeptides wherein at least one amino acid (*e.g.*, 1, 2, 3, 4 or 5 amino acids) at positions 1304, 1305, 1306, 1307 and/or 1308 (numbered with reference to SEQ ID: 8) is/are substituted by another amino acid. In certain embodiments, the mutant has a zinc content which is at least 50% lower than the content of an equivalent amount of wild-type orf3526. Particular mutant orf3526 polypeptides comprise the amino acid sequence of SEQ ID: 31, or immunogenic fragments thereof which include said amino acid positions 1304, 1305 and 1308. One such fragment comprises amino acid residues 24-1520, or residues 34-1520, of SEQ ID 31. The mutant polypeptides may be lipidated *e.g.* at an N-terminal cysteine. The mutant polypeptides may be prepared having a reduced zinc ion content or substantially free from zinc ions, relative to other variants of orf3526 polypeptide or fragments thereof, for example relative to wild-type orf3526 polypeptide.

The invention also provides isoform B of the 3526 polypeptide (corresponding to peak B in Figure 9a), or an immunogenic fragment thereof, obtainable by purification using size exclusion chromatography (e.g. CaptoQ and/or butyl sepharose chromatography) from a composition comprising recombinant 3526 polypeptide(s) of a fraction that elutes at around 17 mins, or that elutes after isoform A and/or before isoform C. In further embodiments, isoform B is of a 3526 polypeptide that includes mutations at positions 1304, 1305 and 1308 and, optionally, lacks amino acid residues 1-23, or residues 1-33, of SEQ ID 31.

The invention also provides isoform A of the 3526 polypeptide (corresponding to peak A in Figure 9a), or an immunogenic fragment thereof, obtainable by purification using size exclusion chromatography (e.g. CaptoQ and/or butyl sepharose chromatography) from a composition comprising recombinant 3526 polypeptide(s) of a fraction that elutes at around 16 mins, or of a fraction that elutes before isoform B and/or before isoform C. In further embodiments, isoform A is of a 3526 polypeptide that includes mutations at positions 1304, 1305 and 1308 and, optionally, lacks amino acid residues 1-23, or residues 1-33, of SEQ ID 31.

The invention also provides isoform C of the 3526 polypeptide (corresponding to peak C in Figure 9a), or an immunogenic fragment thereof, obtainable by purification using size exclusion chromatography (e.g. CaptoQ and/or butyl sepharose chromatography) from a composition comprising recombinant 3526 polypeptide(s) of a fraction that elutes at around 19 mins, or that elutes after isoform A and/or after isoform B. In further embodiments, isoform C is of a 3526 polypeptide that includes mutations at positions 1304, 1305 and 1308 and, optionally, lacks amino acid residues 1-23, or residues 1-33, of SEQ ID 31.

In further embodiments, the invention provides a combination of at least two of isoforms A, B and C of orf3526, or immunogenic fragments thereof, obtainable by purification using size exclusion chromatography (e.g. CaptoQ and/or butyl sepharose chromatography) from a composition comprising recombinant 3526 polypeptide(s) of a fraction that elutes at around 16 mins for isoform A, at around 17 mins for isoform B, and around 19 mins for isoform C; or of a fraction that elutes before isoform B and/or before isoform C for isoform A, or that elutes after isoform A and/or before isoform C for isoform B, or that elutes after isoform A and/or after isoform B for isoform C. For example, a combination of isoform A and B of orf3526, or immunogenic fragments thereof, such as a combination of isoforms A and B of a polypeptide that includes mutations at positions 1304, 1305 and 1308 and comprises amino acid residues 24-1520, or residues 34-1520, of SEQ ID 31.

"Isoform C" is used as a synonym for "fragment C".

The invention also provides a protein which binds to an antibody which antibody does bind to isoform B of the 3526 polypeptide obtainable by purification using size exclusion chromatography (e.g. CaptoQ and/or butyl sepharose chromatography) from a composition comprising recombinant 3526 polypeptide(s) of a fraction that elutes at around 17 mins, or that elutes after isoform A and/or before isoform C, but which antibody does not bind to isoform A of the 3526 polypeptide obtainable by purification using size exclusion chromatography (e.g. CaptoQ and/or butyl sepharose chromatography) from a composition comprising recombinant 3526 polypeptide(s) of a fraction that elutes at around 16 mins, or of a fraction that elutes before isoform B and/or before isoform C, or to isoform C of the 3526 polypeptide, , obtainable by purification using size exclusion chromatography (e.g. CaptoQ and/or butyl sepharose chromatography) from a composition comprising recombinant 3526 polypeptide(s) of a fraction that elutes at around 19 mins, or that elutes after isoform A and/or after isoform B. In further embodiments, isoform B is of a 3526 polypeptide that includes mutations at positions 1304, 1305 and 1308 and, optionally, lacks amino acid residues 1-23, or residues 1-33, of SEQ ID 31.

The invention also provides a protein which binds to an antibody which antibody does bind to isoform A of the 3526 polypeptide obtainable by purification using size exclusion chromatography (e.g. CaptoQ and/or butyl sepharose chromatography) from a composition comprising recombinant 3526 polypeptide(s) of a fraction that elutes at around 16 mins, or of a fraction that elutes before isoform B and/or before isoform C, but which antibody does not bind to isoform B of the 3526 polypeptide obtainable by purification using size exclusion chromatography (e.g. CaptoQ and/or butyl sepharose chromatography) from a composition comprising recombinant 3526 polypeptide(s) of a fraction that elutes at around 17 mins, or that elutes after isoform A and/or before isoform C, or to isoform C of the 3526 polypeptide obtainable by purification using size exclusion chromatography (e.g. CaptoQ and/or butyl sepharose chromatography) from a composition comprising recombinant 3526 polypeptide(s) of a fraction that elutes at around 19 mins, or that elutes after isoform A and/or after isoform B. In further embodiments, isoform B is of a 3526 polypeptide that includes mutations at positions 1304, 1305 and 1308 and, optionally, lacks amino acid residues 1-23, or residues 1-33, of SEQ ID 31.

The invention also provides a protein which binds to an antibody which antibody does bind to isoform C of the 3526 polypeptide obtainable by purification using size exclusion chromatography (e.g. CaptoQ and/or butyl sepharose chromatography) from a composition comprising recombinant 3526 polypeptide(s) of a fraction that elutes at around 19 mins, or that elutes after isoform A and/or after isoform B, but which antibody does not bind to isoform A of the 3526 polypeptide obtainable by purification using size exclusion chromatography (e.g. CaptoQ and/or butyl sepharose chromatography) from a composition comprising recombinant 3526 polypeptide(s) of a fraction that elutes at around 16 mins, or of a fraction that elutes before isoform B and/or before isoform C, or to obtainable by purification using size exclusion chromatography (e.g. CaptoQ and/or butyl sepharose chromatography) from a composition comprising recombinant 3526 polypeptide(s) of a fraction that elutes at around 17

mins, or that elutes after isoform A and/or before isoform C. In further embodiments, isoform B is of a 3526 polypeptide that includes mutations at positions 1304, 1305 and 1308 and, optionally, lacks amino acid residues 1-23, or residues 1-33, of SEQ ID 31.

Such antibodies can be prepared by screening methods known in the art (e.g. chromatography using isoforms for positive and negative selection; phage display).

The invention also provides immunogenic compositions comprising one or more of these isoforms, and their use in a method for treating or preventing E. coli infection, for example in a mammal.

BRIEF DESCRIPTION OF FIGURES

Figure 1: (A) Presence of each of the antigens utilised in compositions of the present invention were determined in most pathogenic strains, specifically NMEC, APEC, UPEC, EHEC, EAEC, EIEC, EPEC, ETEC and AIEC; (B) Gene presence of protective candidates was evaluated in sequenced genomes (based on >85% sequence homology) and in clinical isolates (by PCR amplification).

Figure 2: Gene distribution analysis - Gene presence of best protective candidates was evaluated in 603 E.coli strains by PCR amplification and blast searches on sequenced genomes. ExPEC: APEC=Avian, UPEC=Uropathogenic, NMEC=Newborn meningitic, SEPEC=Septicaemia, pathogenic E.coli. EPEC = Enteropathogenic E.coli. ETEC = Enterotoxigenic E.coli. EHEC = Enterohaemorrhagic E.coli.. Faecal E.coli and Other Pathovars = Lab.Strains, STEC = Toxinproducing E.coli, EIEC = enteroinvasive e.coli, AIEC = adhesive invasive E.coli, EAEC = enteroagregative E.coli. AREC = ampicillin resistent E.coli . orf3526 (ECOK1_3385) expression was assessed by immunoblotting analysis on supernatant fractions using polyclonal rabbit serum.

Figure 3: Study of molecular diversity of orf3526: sequence alignment analysis of 217 sequences of orf3526 protein revealed a sequence identity ranging from 86% to 100%; the derived phylogenetic tree revealed the presence of 6 major variants. The phylogenetic tree was inferred using the Maximum likelihood method implemented in MEGA5 package.

Figure 4: Antibodies to antigen orf3526 were found to protect in passive immunization in an avian (chicken) model of sepsis.

Figure 5: Intranasal immunization with antigen orf3526 (ECOK1_3385) reduced colonization in ileum tract following challenge with ETEC strain GL53.

Figure 6: Antigen 405B (ECOK1_0290) prevents kidney colonisation in a UTI model of infection.

Figure 7: Antigen upec1232 (c1275) prevents kidney colonisation in a UTI model of infection.

Figure 8: Analysis of the orf3526 sequence revealed several conserved motifs.

Figure 9(a): Isoforms of DG3526TL were separated by SE-HPLC on a Tosoh G3000SWxl column. The fraction that binds to butyl Sepharose appears less compact in SEC and forms peak A, while the form that binds to butyl Sepharose forms peak B. Peak C corresponds to an N-terminally truncated form.

Figure 9(b): To evaluate possible functional differences between DG3526TL and native 3526, the latter was purified from culture supernatant of ExPEC IHE3034. Running 3526 along with the isoforms of DG3526TL reveals that the native form coelutes with peak B, suggesting that native 3526 is i) compact and ii) a monomer.

Figure 10: Cartoon illustrating the seven mutants/variants of orf3526 utilised.

Figure 11: Zinc content of various orf3526 derivatives was determined by atomic absorption spectroscopy. Results suggest the presence of a single zinc ion per protein molecule. The unexpected low zinc content of 3526 B his, actually containing the zinc binding motif, could be explained by misfolding of this truncated derivative, while the single amino acid exchange in the E1305 mutant apparently is not sufficient to completely abolish zinc binding (red boxes). In contrast, zinc affinity is completely lost in the TL3M triple mutant.

Figure 12: orf3526 75% consensus sequence. Each specified residue is found at that position in at least 75% of the orf3526 sequences used to generate the consensus sequence. X represents any amino acid. The MO60-Like domain is highlighted.

Figure 13: orf3526 100% consensus sequence. Each specified residue is found at that position in all of the orf3526 sequences used to generate the consensus sequence. X represents any amino acid. The MO60-Like domain is highlighted.

Figure 14: Bacterial titres were significantly reduced after immunisation with 3526 + alum, or after immunisation with 3526 + MF59, compared to adjuvant alone. The results are confirmed in an experiment were mice were immunised on days 0 and 21 only (Figure 14(b)).

Figure 15: 405B+3526+LTK63 and 3526+1232+LTK63 significantly reduces intestinal colonization by GL53 in the caecum, compared to LTK63 alone.

DETAILED DESCRIPTION OF THE INVENTION

The invention provides immunogenic compositions comprising an immunogenic component of *Escherichia coli* wherein the immunogenic component is selected from the group consisting of bacterial Ig-like domain protein fragment (orf405B) having the amino acid sequence set forth in SEQ ID NO:2 or a protein having at least 80% similarity thereto, upec1232 having the amino acid sequence set forth in SEQ ID NO:4 or a protein having at least 80% similarity thereto, putative Lipoprotein (orf3526) having the amino acid sequence set forth in SEQ ID NO:8 or a protein having at least 80% similarity thereto and gspK (orf3515) having the amino acid sequence set forth in SEQ ID NO:30 or a protein having at least 80% similarity thereto. A composition of the invention may comprise e.g. one, two, three or four of the aforementioned components, e.g.

orf405B and upec1232; orf405B and orf3526; orf405B and orf3515;

upec1232 and orf3526; upec1232 and orf3515;

orf3526 and orf3515;

orf405B, upec1232 and orf3526; orf405B, orf3526 and orf3515; or upec1232, orf3526s and orf3515;

or proteins having at least 80% similarity to any thereof.

Preferably, an immunogenic composition of the invention comprises one, two or three components selected from orf405B, upec1232, orf3526, or proteins having at least 80% similarity to any thereof. For example, an immunogenic composition of the invention may comprise the three components orf405B, upec1232 and orf3526, or proteins having at least 80% similarity to any thereof. More preferably, an immunogenic composition of the invention comprises one or two components selected from orf405B and orf3526, or proteins having at least 80% similarity to any thereof. For example, an immunogenic composition of the invention may comprise the two components orf405B and orf3526, or proteins having at least 80% similarity to any thereof. Alternatively, an immunogenic composition of the invention may comprise orf405B, upec1232, orf3526, and orf3515, or proteins having at least 80% similarity to any thereof.

Components of compositions of the invention may be isolated or purified.

As used herein, the term "immunogenic" means that, for example the polypeptide(s), composition and the like, is/are capable of eliciting a humoral or cellular immune response, and preferably both. For example, the term "immunogenic composition" refers to any composition able, once it has been administered to a subject, such as an animal for example a human, to induce or stimulate an immune response against *E. coli*.

An immunogenic polypeptide is also antigenic. A molecule is "antigenic" when it is capable of specifically interacting with an antigen recognition molecule of the immune system, such as an immunoglobulin (antibody) or T cell antigen receptor. An antigenic polypeptide contains an epitope of at least about five, and particularly at least about 10, at least 15, at least 20 or at least 50 amino acids. An antigenic portion of a polypeptide, also referred to as an epitope, can be that portion that is immunodominant for antibody or T cell receptor recognition, or it can be a portion used to generate an antibody to the molecule by conjugating the antigenic portion to a carrier polypeptide for immunization. The skilled person will recognise that a molecule that is antigenic need not be itself immunogenic, for example, some antigens require the presence of an adjuvant or carrier to render them capable of eliciting an immune response.

The term "antigen" refers to a molecule against which a subject can initiate a humoral and/or cellular immune response. An "immunological response" to a composition or vaccine is the development in the host of a cellular and/or antibody-mediated immune response to a composition or vaccine of interest. Usually, an "immunological response" includes but is not limited to one or more of the following effects: the production of antibodies, B cells, helper T cells, and/or cytotoxic T cells, directed specifically to an antigen or antigens included in the composition or vaccine of interest. Preferably, the subject will display

either a therapeutic or protective immunological response such that resistance to new infection will be enhanced and/or the clinical severity of the disease reduced. Such protection will be demonstrated by either a reduction or lack of symptoms normally displayed by an infected subject, a quicker recovery time and/or a lowered viral titre in the infected host. The term "immunogenic" protein or polypeptide as used herein also refers to an amino acid sequence which elicits an immunological response as described above.

When said immunogenic compositions prevent, ameliorate, palliate or eliminate disease from an animal then the immunogenic composition may optionally be referred to as a vaccine.

The term "vaccine" as used herein refers to a vaccine composition that comprises either purified antigenic determinants, nucleic acids encoding the purified antigenic determinants or fragments thereof, in the absence of the disease-causing organism. Such vaccines may also be referred to as a "sub-unit vaccine". The terms are not intended to encompass "whole-cell vaccines", for example those derived from whole bacterial cells that have been killed and which may contain the antigenic determinants in un-purified form as part of a complex and uncharacterised composition.

As used herein, the term "multivalent", means that the vaccine contains structurally similar or 'related' antigenic determinants from at least two strains or isolates, the antigenic determinants being homologues having minor differences between their amino acid sequences.

The terms "variant", "homologue", "derivative" or "fragment" in relation to polypeptides or antigens include any substitution of, variation of, modification of, replacement of, deletion of or addition of one (or more) amino acid or nucleotide from or to a sequence. Unless the context admits otherwise, references to particular antigens includes references to such variants, homologues, derivatives and fragments.

Preferred variants of an antigen can elicit antibodies which bind to that antigen. In particular, the antibodies bind to wild-type antigens as present in an *E. coli* cell.

In particular, the term "homologue" covers identity with respect to structure and/or function providing the resultant amino acid sequence has antigenic or immunogenic activity. With respect to sequence identity (i.e. similarity), there may be at least 70%, at least 75%, at least 80%, at least 85%, at least 90% sequence identity. There may be, e.g., at least 91%, 92%, 93%, or 94%, sequence identity. There may also be at least 95%, 96%, 97%, 98%, 99%, or 99.5% sequence identity. These terms also encompass polypeptides derived from amino acids which are allelic variations of nucleic acid or amino acid sequence(s). Percentage sequence identity and similarity between a sequence A and a sequence B is calculated as (x/y)*100, wherein x is the number of amino acids that are identical between A and B and y is the number of amino acids of the longest sequence selected from A and B. For example, in the case of 10 identical

residues between a first sequence A consisting of 50 amino acids and a second sequence B consisting of 200 amino acids, the sequence identity between the two sequences is 5%.

Where reference is made to the "activity" or "biological activity" of a polypeptide, these terms are intended to refer to the antigenic and immunogenic activities of the polypeptide. Examples of such activities, and methods of assaying and quantifying these activities, are known in the art, and are described in detail elsewhere in this document.

As used herein a "deletion" is defined as a change in either nucleotide or amino acid sequence in which one or more nucleotides or amino acid residues, respectively, are absent. As used herein an "insertion" or "addition" is that change in a nucleotide or amino acid sequence which has resulted in the addition of one or more nucleotides or amino acid residues, respectively, as compared to the naturally occurring substance. As used herein "substitution" results from the replacement of one or more nucleotides or amino acids by different nucleotides or amino acids, respectively.

The terms "antigen" and "amino acid sequence", as they are used in this document, should be taken to include reference to each of the above sequences, as well as to their fragments, homologues, derivatives and variants.

The term "fragment" as used herein refers to partial nucleotide or amino acid sequences according to the present invention. In certain embodiments amino acid sequence or polypeptide fragments may include polypeptides comprising an amino acid sequence of at least 'n' consecutive amino acids derived from the listed sequence identifiers, for example at least 5 amino acid residues, at least 10 amino acid residues, at least 15 amino acid residues, at least 20 amino acid residues, at least 25 amino acid residues, at least 30 amino acid residues, at least 35 amino acid residues, at least 40 amino acid residues, at least 45 amino acid residues, at least 50 amino acid residues, at least 60 amino residues, at least 70 amino acid residues, at least 80 amino acid residues, at least 90 amino acid residues, at least 100, at least 150, at least 200 or at least 250 amino acid residues of the amino acid sequence. In certain embodiments amino acid fragments may include polypeptides comprising an amino acid sequence of no more than 50, no more than 60, no more than 75, no more than 100, no more than 150, no more than 200, no more than 250, no more than 300, no more than 350, no more than 400 amino acid residues. Preferred fragments comprise an epitope or are immunogenic fragments. Preferred fragments lack an amino-terminal portion of the polypeptides of the invention, such as residues 1-23 or residues 1-33 of SEQ ID NO: 8 or SEQ ID NO:31, or corresponding residues in other orf3526 polypeptides of the invention. Sequence identity and similarity between a fragment and a longer sequence is calculated according to the same method as described above, i.e. based on identical residues relative to the longest sequence.

As used herein, the term "purified" or "to purify" refers to the removal of contaminants from a sample. For example, antigens are purified by removal of contaminating proteins. The removal of contaminants results in an increase in the percent of antigen (e.g., antigen of the present invention) in the sample.

"Isolated" and "purified" as used herein describe certain molecules, proteins, polysaccharides, lipids, antigens, and the like, and refers to a state beyond that in which the molecules, proteins, polysaccharides, lipids, or antigens exist naturally in cells. Particularly the term as used herein means removed from its naturally occurring environment such as a cell, for example. In preferred embodiments, the isolated molecules, proteins, polysaccharides, lipids, antigens, and the like, are separated from greater than 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% or 99% of the proteins and/or the lipids with which the molecules, proteins, polysaccharides, lipids, antigens, and the like are normally associated naturally in cells. If the isolated molecules, proteins, polysaccharides, lipids, antigens, and the like are synthesized, they are contaminated with less than 50%, 40%, 30%, 20%, 10%, 5%, 1% or 0.1% of the chemical precursors or synthesis reagents used to synthesize the lipid antigen. In preferred embodiments, the molecules, proteins, polysaccharides, lipids or antigens are at least 1% pure, 5% pure, 10% pure, 20% pure, 30% pure, 40% pure, 50% pure, 60% pure, 70% pure, 80% pure, 90% pure, 95% pure, 99% pure, or 100% pure. As used herein, the term "% pure" indicates the percentage of a composition that is made up of the molecule of interest, by weight. Thus, a composition of 100 grams containing 50 grams of a molecule of interest is 50% pure with respect to the molecule of interest.

The term "treating" includes both therapeutic treatment and prophylactic or preventative treatment, wherein the object is to prevent or lessen infection. For example, treating may include directly affecting or curing, suppressing, inhibiting, preventing, reducing the severity of, delaying the onset of, reducing symptoms associated with, for example, infection, or a combination thereof. "Preventing" may refer, *inter alia*, to delaying the onset of symptoms, preventing relapse to a disease, and the like. Treating may also include "suppressing" or "inhibiting" an infection or illness, for example reducing severity, number, incidence or latency of symptoms, ameliorating symptoms, reducing secondary symptoms, reducing secondary infections, prolonging patient survival, or combinations thereof.

Polypeptides used in the invention can be prepared in many ways *e.g.* by chemical synthesis (in whole or in part), by digesting longer polypeptides using proteases, by translation from RNA, by purification from cell culture (*e.g.* from recombinant expression), from the organism itself (*e.g.* after bacterial culture, or direct from patients), *etc.* A preferred method for production of peptides <40 amino acids long involves *in vitro* chemical synthesis [1,2]. Solid-phase peptide synthesis is particularly preferred, such as methods based on tBoc or Fmoc [3] chemistry. Enzymatic synthesis [4] may also be used in part or in full. As an alternative to chemical synthesis, biological synthesis may be used *e.g.* the polypeptides may be produced by translation. This may be carried out *in vitro* or *in vivo*. Biological methods are in general restricted to the production of polypeptides based on L-amino acids, but manipulation of translation machinery (*e.g.* of

aminoacyl tRNA molecules) can be used to allow the introduction of D-amino acids (or of other non natural amino acids, such as iodotyrosine or methylphenylalanine, azidohomoalanine, etc.) [5]. Where D-amino acids are included, however, it is preferred to use chemical synthesis. Polypeptides may have covalent modifications at the C-terminus and/or N-terminus.

Polypeptides can take various forms (*e.g.* native, fusions, glycosylated, non-glycosylated, lipidated, non-lipidated, phosphorylated, non-phosphorylated, myristoylated, non-myristoylated, monomeric, multimeric, particulate, denatured, *etc.*).

Polypeptides are preferably provided in purified or substantially purified form. Polypeptides may be attached to a solid support. Polypeptides may comprise a detectable label (*e.g.* a radioactive or fluorescent label, or a biotin label).

The term "polypeptide" refers to amino acid polymers of any length. The polymer may be linear or branched, it may comprise modified amino acids, and it may be interrupted by non-amino acids. The terms also encompass an amino acid polymer that has been modified naturally or by intervention; for example, disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, or any other manipulation or modification, such as conjugation with a labeling component. Also included within the definition are, for example, polypeptides containing one or more analogs of an amino acid (including, for example, unnatural amino acids, *etc.*), as well as other modifications known in the art. Polypeptides can occur as single chains or associated chains. Polypeptides can be naturally or non-naturally glycosylated (*i.e.* the polypeptide has a glycosylation pattern that differs from the glycosylation pattern found in the corresponding naturally occurring polypeptide).

Polypeptides used in the present invention may be produced by culturing a host cell under conditions which induce polypeptide expression. Expression of the polypeptide may take place in a heterologous host for expression. The heterologous host may be prokaryotic (*e.g.* a bacterium) or eukaryotic. Suitable hosts include *Bacillus subtilis, Vibrio cholerae, Salmonella typhi, Salmonella typhimurium, Neisseria lactamica, Neisseria cinerea, Mycobacteria* (*e.g. M.tuberculosis*), yeasts, *etc.*

Particular polypeptides used in combinations of the invention may comprise an amino acid sequence that is derived from bacterial Ig-like domain (group 1) protein fragment (orf405B), gspK (orf3515), upec-1232 and putative Lipoprotein (orf3526) each as more fully described herein.

<u>Bacterial Ig-like domain protein fragment 'orf405B'</u>

Bacterial Ig-like domain (group 1) protein, from E. coli NMEC, is disclosed in WO2006/089264 (SEQ IDs 809 and 810) and is referred to therein as 'orf405', the protein is also referred to as 'orf284' from E. coli NMEC strain IHE3034, 'c0415' from CFT073 and 'ecp_0367' from 536. Fragments of this orf405

protein were first disclosed in WO2011/004263 (for example in SEQ IDs 641 and 642). Compositions according to the present invention preferably comprise bacterial Ig-like domain protein fragment 'orf405B'. The nucleotide and amino acid sequences of this protein fragment, referred to herein, as SEQ IDs 1 and 2 are:

>orf405B (SEQ ID 1)

GTTGCTGATGGTCAGCAAGCCTACACGCTGACACTGACAGCGGTGGACTCCGAGGGTAAT
CCGGTGACGGGAGAAGCCAGCCGCCTGCGACTTGTTCCGCAAGACACTAATGGTGTAACC
GTTGGTGCCATTTCGGAAATAAAACCAGGGGTTTACAGCGCCACGGTTTCTTCGACCCGT
GCCGGAAACGTTGTTGTGCGTGCCTTCAGCGAGCAGTATCAGCTGGGCACATTACAACAA
ACGCTGAAGTTTGTTGCCGGGCCGCTTGATGCAGCACATTCGTCCATCACACTGAATCCT
GATAAACCGGTGGTTGGCGGTACAGTTACGGCAATCTGGACGGCAAAAGATGCTAATGAC
AACCCTGTAACTGGCCTCAATCCGGATGCACCGTCATTATCGGGCGCAGCTGCTGCTGGT
TCTACGGCATCAGGCTGGACGGATAATGGCGACGGGACCTGGACTGCGCAGATTTCTCTC
GGCACTACGGCGGGTGAATTAGACGTTATGCCGAAGCTCAATGGGCAGGACGCGGCAGCA
AATGCGGCAAAAGTAACCGTGGTGGCTGATGCATTATCTTCAAACCAGTCGAAAGTCTCT
GTCGCAGAAGATCACGTAAAAGCCGGTGAAAGCACAACCGTAACGCTGGTGGCGAAAGAT
GCGCATGGCAACGCTATCAGTGGTCTTTCGTTGTCGGCAAGTTTGACGGGGACCGCCTCT
GAAGGGGCGACCGTTTCCAGTTGGACCGAAAAAGGTGACGGTTCCTATGTTGCTACGTTA
ACTACAGGCGGAAAGACGGGCGAGCTTCGTGTCATGCCGCTCTTCAACGGCCAGCCTGCA
GCCACCGAAGCCGCGCAGCTGACTGTTATTGCCGGAGAGATGTCATCAGCGAACTCTACG
CTTGTTGCGGACAATAAAACTCCAACGGTTAAAACGACGACGGAACTCACCTTCACCATG
AAGGATGCGTACGGGAATCCGGTCACCGGGCTGAAGCCAGATGCACCAGTGTTTAGTGGT
GCCGCCAGCACGGGGAGTGAGCGTCCTTCAGCAGGAAACTGGACAGAGAAAGGTAATGGG
GTCTACGTGTCGACCTTAACGCTGGGATCTGCCGCGGGTCAGTTGTCTGTGATGCCGCGA
GTGAACGGCCAAAATGCCGTTGCTCAGCCACTGGTGCTGAATGTTGCAGGTGACGCATCT
AAGGCTGAGATTCGTGATATGACAGTGAAGGTTAATAACCAA

>orf405B (SEQ ID 2)

VADGQQAYTLTLTAVDSEGNPVTGEASRLRLVPQDTNGVTVGAISEIKPGVYSATVSSTR
AGNVVVRAFSEQYQLGTLQQTLKFVAGPLDAAHSSITLNPDKPVVGGTVTAIWTAKDAND
NPVTGLNPDAPSLSGAAAAGSTASGWTDNGDGTWTAQISLGTTAGELDVMPKLNGQDAAA
NAAKVTVVADALSSNQSKVSVAEDHVKAGESTTVTLVAKDAHGNAISGLSLSASLTGTAS
EGATVSSWTEKGDGSYVATLTTGGKTGELRVMPLFNGQPAATEAAQLTVIAGEMSSANST
LVADNKTPTVKTTTELTFTMKDAYGNPVTGLKPDAPVFSGAASTGSERPSAGNWTEKGNG
VYVSTLTLGSAAGQLSVMPRVNGQNAVAQPLVLNVAGDASKAEIRDMTVKVNNQ

When used according to the present invention, orf405B protein may take various forms. Particular orf405B sequences have 80% or more identity (e.g. 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more) to SEQ ID NOs 1 and/or 2. This includes variants (e.g. allelic variants, homologs, orthologs, paralogs, mutants etc).

Upec1232 protein

'Upec1232' protein from *E. coli* UPEC is disclosed in WO2006/091517 (SEQ ID 138) and is also known as: 'c1275' from CFT073. When used according to the present invention, upec1232 protein may take various forms. Preferred upec1232 sequences have 80% or more identity (e.g. 80%, 85%, 90%, 95%,

96%, 97%, 98%, 99% or more) to SEQ ID NOs 3, 4, 5 or 6. This includes variants (e.g. allelic variants, homologs, orthologs, paralogs, mutants etc).

>upec-1232 (SEQ ID 3)

ATGATTCACCTGTTCAAAACCTGCATGATTACCGCCTTCATTCTGGGGTTAACGTGGTCT
GCCCCACTCCGGGCACAGGATCAACGTTACATCAGTATACGCAATACAGATACGATATGG
CTCCCGGGAAATATTTGTGCTTACCAGTTCCGGCTGGATAATGGCGGAAACGATGAAGGA
TTTGGCCCCCTCACCATCACTCTGCAACTCAAAGACAAATATGGTCAGACGCTGGTGACC
AGAAAAATGGAAACGGAAGCCTTTGGTGACAGTAATGCCACGCGAACCACAGACGCATTT
CTGGAAACGGAGTGCGTGGAAAATGTCGCCACAACCGAAATCATTAAAGCAACTGAAGAA
AGTAACGGCCATCGTGTCAGTCTGCCGTTATCGGTTTTCGATCCCCAGGACTACCATCCA
CTGCTGATTACCGTTTCCGGAAAAAACGTTAAC

>upec-1232 (SEQ ID 4)

MIHLFKTCMITAFILGLTWSAPLRAQDQRYISIRNTDTIWLPGNICAYQFRLDNGGNDEG
FGPLTITLQLKDKYGQTLVTRKMETEAFGDSNATRTTDAFLETECVENVATTEIIKATEE
SNGHRVSLPLSVFDPQDYHPLLITVSGKNVN

>pCFT-1232 (SEQ ID 5)

CAGGATCAACGTTACATCAGTATACGCAATACAGATACGATATGGCTCCCGGGAAATATT
TGTGCTTACCAGTTCCGGCTGGATAATGGCGGAAACGATGAAGGATTTGGCCCCCTCACC
ATCACTCTGCAACTCAAAGACAAATATGGTCAGACGCTGGTGACCAGAAAAATGGAAACG
GAAGCCTTTGGTGACAGTAATGCCACGCGAACCACAGACGCATTTCTGGAAACGGAGTGC
GTGGAAAATGTCGCCACAACCGAAATCATTAAAGCAACTGAAGAAAGTAACGGCCATCGT
GTCAGTCTGCCGTTATCGGTTTTCGATCCCCAGGACTACCATCCACTGCTGATTACCGTT
TCCGGAAAAAACGTTAAC

>pCFT-1232 (SEQ ID 6)

QDQRYISIRNTDTIWLPGNICAYQFRLDNGGNDEGFGPLTITLQLKDKYGQTLVTRKMET
EAFGDSNATRTTDAFLETECVENVATTEIIKATEESNGHRVSLPLSVFDPQDYHPLLITV
SGKNVN

Putative Lipoprotein orf3526

Accessory colonization factor D (AcfD) precursor, also known as 'ECOK1_3385', also known as 'putative lipoprotein orf3526', also referred to as 'orf3526' protein from *E. coli* NMEC strain IHE3034 is disclosed in WO2006/089264. When used according to the present invention, orf3526 protein may take various forms. Preferred orf3526 sequences have 80% or more identity (e.g. 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more) to SEQ ID NOs 7 - 28. This includes variants (e.g. allelic variants, homologs, orthologs, paralogs, mutants etc).

>orf03526 (SEQ ID 7)

ATGAATAAGAAATTTAAATATAAGAAATCGCTTTTAGCGGCTATTTTAAGCGCAACCCTG
TTAGCCGGTTGTGATGGTGGTGGTTCAGGATCGTCCTCCGATACGCCGTCTGTAGATTCT
GGATCAGGGACTTTGCCGGAAGTGAAACCCGATCCAACACCAACCCCGGAGCCGACACCT
GAGCCGACGCCGGACCCAGAACCTACGCCGGATCCAACACCTGATCCTGAGCCGACACCA
GAACCGGAGCCAGAACCTGTTCCTACGAAAACGGGTTATCTGACCCTGGGCGGAAGCCAG

CGGGTAACTGGTGCTACCTGTAATGGTGAATCCAGCGATGGCTTTACCTTTACGCCAGGC
AATACCGTGAGTTGTGTGGTGGGCAGTACGACCATTGCAACATTCAACACCCAGTCAGAA
GCTGCGCGTAGCCTGCGTGCGGTTGACAAAGTGTCGTTTAGCCTGGAGGACGCGCAGGAG
CTGGCGAATTCTGAAAATAAGAAAACCAACGCCATCTCTCTGGTGACGTCCAGCGACAGT
TGCCCCGCAGATGCAGAACAGCTTTGTCTTACTTTCTCGTCAGTGGTTGATCGCGCGCGA
TTTGAAAAACTGTATAAGCAAATTGATCTGGCAACAGACAATTTCAGCAAGCTGGTCAAT
GAAGAGGTGGAAAACAATGCTGCGACTGATAAAGCGCCGTCCACCCATACCTCAACGGTA
GTGCCAGTCACGACAGAGGGAACAAAACCGGATCTGAACGCGTCCTTCGTGTCGGCTAAC
GCGGAACAGTTTTATCAGTATCAACCCACTGAAATCATTCTTTCCGAAGGCCAACTGGTG
GATAGCCTGGGGAACGGTGTTGCTGGCGTTGACTACTACACCAATTCAGGCCGTGGCGTA
ACTGACGAAAACGGTAAATTTTCCTTTAGCTGGGGCGAAACCATCTCCTTTGGTATCGAT
ACCTTTGAACTGGGCTCAGTACGTGGCAATAAGTCGACCATTGCGCTGACTGAATTGGGT
GATGAAGTTCGCGGGGCAAATATCGATCAGCTCATTCATCGTTATTCGACGACTGGTCAA
AATAATACTCGTGTTGTTCCGGACGATGTACGCAAGGTCTTTGCCGAATATCCCAACGTG
ATCAACGAGATAATCAATCTTTCGTTATCCAACGGTGCGACGCTGGATGAAGGCGATCAA
AACGTTGTGCTGCCTAACGAATTTATCGAGCAGTTTAAGACGGGTCAGGCCAAAGAGATC
GATACCGCGATTTGTGCGAAAACCGACGGTTGTAACGAGGCTCGCTGGTTCTCGCTGACA
ACGCGCAATGTTAATGACGGCCAGATTCAGGGCGTTATTAACAAGCTGTGGGGCGTGGAT
ACGAACTATCAGTCTGTCAGCAAGTTCCACGTCTTCCATGACTCTACCAACTTCTATGGC
AGCACCGGTAACGCGCGCGGTCAGGCGGTGGTAAATATCTCCAACTCGGCATTCCCGATT
CTGATGGCGCGTAATGATAAAAACTACTGGCTGGCGTTTGGCGAAAAACGCGCCTGGGAT
AAAAATGAGCTGGCGTACATTACGGAAGCGCCTTCCATTGTGCAGCCAGAGAACGTTACG
CGCGATACTGCGACTTTCAACCTGCCGTTTATTTCGCTGGGGCAAGTCGGTGAAGGCAAA
CTGATGGTTATCGGTAACCCGCACTACAACAGCATCCTGCGTTGCCCGAACGGTTACAGT
TGGGGCGGTGGTGTTAATAGTAAAGGTGAGTGTACGCTCAGCGGTGATTCTGATGACATG
AAGCACTTTATGCAGAACGTACTGCGCTACTTGTCAAATGACATCTGGCAGCCAAATACC
AAGAGCATCATGACTGTCGGCACCAACCTGGAGAACGTTTATTTCAAAAAAGCGGGCCAG
GTATTGGGAAATAGTGCACCATTTGCTTTCCATGAGGATTTCACTGGTATCACGGTTAAA
CAGTTGACCAGCTATGGCGATCTGAATCCGGAAGAGATTCCGTTGCTGATCCTCAACGGC
TTTGAATATGTGACTCAGTGGTCTGGCGATCCCTATGCTGTGCCTCTGCGTGCAGATACC
AGCAAACCGAAGCTGACTCAGCAGGATGTGACCGATCTGATCGCTTATCTGAACAAAGGT
GGCTCGGTGCTGATCATGGAAAACGTGATGAGCAATCTTAAGGAAGAGAGCGCGTCCAGT
TTTGTGCGTCTGCTGGATGCCGCGGGTCTGTCAATGGCTCTGAACAAATCGGTGGTGAAC
AACGATCCGCAAGGGTATCCGGATCGCGTTCGTCAGCGTCGCGCGACTGGCATTTGGGTT
TATGAACGTTATCCTGCTGCAGACGGCGCGCAACCGCCGTACACCATCGACCCAAATACA
GGGGAAGTGACCTGGAAATACCAGCAAGACAACAAGCCTGATGACAAGCCGAAACTGGAA
GTTGCGAGCTGGCAGGAGGAAGTTGAGGGCAAACAGGTAACGCGTTATGCCTTTATTGAT
GAAGCGGAATACACAACAGAAGAATCTCTGGAAGCGGCAAAGGCAAAAATCTTTGAGAAG
TTTCCTGGGTTACAGGAGTGTAAGGACTCGACTTACCATTACGAGATTAACTGTTTGGAG
CGCCGCCCAGGCACGGATGTTCCGGTAACAGGTGGCATGTATGTTCCGCGCTATACGCAA
CTGAATCTTGACGCCGACACCGCGAAAGCGATGGTGCAGGCGGCGGATTTAGGCACCAAC
ATTCAGCGCCTGTATCAGCATGAGCTTTATTTCCGTACCAAAGGCAGTAAAGGTGAGCGT
CTGAACAGTGTTGATCTGGAACGTCTGTACCAGAACATGTCGGTCTGGCTGTGGAACGAT
ACGAAATATCGTTACGAAGAGGGCAAGGAAGATGAGCTGGGCTTTAAAACGTTCACCGAG
TTCCTGAACTGCTACGCCAATGATGCCTATGCAGGCGGCACCAAGTGCTCCGCAGATCTG
AAAAAATCGCTGGTCGATAACAACATGATCTACGGTGACGGTAGCAGCAAAGCGGGCATG
ATGAACCCAAGCTATCCGCTCAACTATATGGAAAAACCGCTGACGCGTCTGATGCTGGGC
CGTTCCTGGTGGGATCTGAACATTAAGGTTGATGTGGAGAAGTACCCAGGATCCGTATCG
GCAAAGGGTGAGAGCGTTACGGAAAACATCAGCCTGTACTCGAATCCGACCAAATGGTTT
GCGGGTAACATGCAGTCAACCGGCCTGTGGGCACCGGCCCAGCAGGACGTCACCATTAAG
TCTTCGGCGTCAGTCCCAGTGACTGTTACCGTGGCGCTGGCTGACGACCTGACTGGACGT
GAGAAGCATGAAGTTGCGCTGAACCGTCCGCCAAGAGTGACTAAAACGTATACTCTGGAG
GCTAACGGTGAAGTGACCTTCAAGGTGCCTTATGGTGGTCTGATTTATATCAAGGGCGAC
AGTAAGGATGATGTTTCTGCTAACTTCACCTTTACCGGTGTAGTAAAAGCGCCGTTCTAT
AAAGACGGCGAATGGAAAAACGATCTGGACTCACCGGCGCCGCTGGGCGAGCTGGAGTCT
GCGTCGTTCGTCTATACCACGCCGAAGAAGAACCTTGAGGCCAGCAATTTCACTGGTGGT

GTAGCAGAATTCGCTAAAGATCTGGATACCTTTGCCAGCTCGATGAATGACTTCTACGGT
CGTAATGATGAAGACGGTAAGCACCGGATGTTTACCTATAAAAACTTGACGGGGCACAAG
CATCGTTTCACCAACGATGTGCAGATCTCCATCGGTGATGCGCACTCGGGTTATCCGGTA
ATGAACAGCAGCTTCTCGACGAACAGCACCACGCTGCCGACGACGCCGCTGAACGACTGG
CTGATTTGGCACGAAGTCGGTCATAACGCTGCAGAAACACCGCTGAACGTACCGGGTGCA
ACTGAAGTGGCGAACAACGTGCTGGCGCTGTACATGCAGGATCGCTATCTCGGTAAGATG
AACCGTGTCGCTGACGACATTACCGTCGCGCCGGAATATCTGGACGAGAGCAACGGTCAG
GCCTGGGCGCGCGGCGGTGCGGGTGACCGTCTGCTGATGTACGCACAGTTGAAGGAGTGG
GCAGAGGAAAACTTTGATATCAAACAGTGGTATCCAGATGGTGAGCTGCCTAAGTTCTAC
AGCGATCGTAAAGGGATGAAGGGCTGGAACCTGTTCCAGTTGATGCACCGTAAAGCGCGC
GGCGATGATGTTGGTAACAGCACCTTTGGTGGCAAGAATTACTGTGCTGAATCCAATGGT
AACGCTGCCGACACGCTGATGCTGTGTGCATCCTGGGTCGCTCAGGCGGATCTTTCGGAA
TTCTTTAAGAAATGGAATCCGGGTGCAAGTGCTTACCAGTTGCCGGGAGCAACGGAGATG
AGTTTCCAGGGCGGTGTGAGCTCTTCGGCTTACAGCACGCTGGCGTCACTCAAGCTGCCG
AAACCGGAAAAAGGGCCGGAAACCATTAACAAGGTTACCGAGCATAAGATGTCTGCCGAG

>orf03526 (SEQ ID 8)

MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPSVDSGSGTLPEVKPDPTPTPEPTP
EPTPDPEPTPDPTPDPEPTPEPEPEPVPTKTGYLTLGGSQRVTGATCNGESSDGFTFTPG
NTVSCVVGSTTIATFNTQSEAARSLRAVDKVSFSLEDAQELANSENKKTNAISLVTSSDS
CPADAEQLCLTFSSVVDRARFEKLYKQIDLATDNFSKLVNEEVENNAATDKAPSTHTSTV
VPVTTEGTKPDLNASFVSANAEQFYQYQPTEIILSEGQLVDSLGNGVAGVDYYTNSGRGV
TDENGKFSFSWGETISFGIDTFELGSVRGNKSTIALTELGDEVRGANIDQLIHRYSTTGQ
NNTRVVPDDVRKVFAEYPNVINEIINLSLSNGATLDEGDQNVVLPNEFIEQFKTGQAKEI
DTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYQSVSKFHVFHDSTNFYG
STGNARGQAVVNISNSAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSIVQPENVT
RDTATFNLPFISLGQVGEGKLMVIGNPHYNSILRCPNGYSWGGGVNSKGECTLSGDSDDM
KHFMQNVLRYLSNDIWQPNTKSIMTVGTNLENVYFKKAGQVLGNSAPFAFHEDFTGITVK
QLTSYGDLNPEEIPLLILNGFEYVTQWSGDPYAVPLRADTSKPKLTQQDVTDLIAYLNKG
GSVLIMENVMSNLKEESASSFVRLLDAAGLSMALNKSVVNNDPQGYPDRVRQRRATGIWV
YERYPAADGAQPPYTIDPNTGEVTWKYQQDNKPDDKPKLEVASWQEEVEGKQVTRYAFID
EAEYTTEESLEAAKAKIFEKFPGLQECKDSTYHYEINCLERRPGTDVPVTGGMYVPRYTQ
LNLDADTAKAMVQAADLGTNIQRLYQHELYFRTKGSKGERLNSVDLERLYQNMSVWLWND
TKYRYEEGKEDELGFKTFTEFLNCYANDAYAGGTKCSADLKKSLVDNNMIYGDGSSKAGM
MNPSYPLNYMEKPLTRLMLGRSWWDLNIKVDVEKYPGSVSAKGESVTENISLYSNPTKWF
AGNMQSTGLWAPAQQDVTIKSSASVPVTVTVALADDLTGREKHEVALNRPPRVTKTYTLE
ANGEVTFKVPYGGLIYIKGDSKDDVSANFTFTGVVKAPFYKDGEWKNDLDSPAPLGELES
ASFVYTTPKKNLEASNFTGGVAEFAKDLDTFASSMNDFYGRNDEDGKHRMFTYKNLTGHK
HRFTNDVQISIGDAHSGYPVMNSSFSTNSTTLPTTPLNDWLIWHEVGHNAAETPLNVPGA
TEVANNVLALYMQDRYLGKMNRVADDITVAPEYLDESNGQAWARGGAGDRLLMYAQLKEW
AEENFDIKQWYPDGELPKFYSDRKGMKGWNLFQLMHRKARGDDVGNSTFGGKNYCAESNG
NAADTLMLCASWVAQADLSEFFKKWNPGASAYQLPGATEMSFQGGVSSSAYSTLASLKLP
KPEKGPETINKVTEHKMSAE

>pK1-3526A (SEQ ID 9)

TGTGATGGTGGTGGTTCAGGATCGTCCTCCGATACGCCGTCTGTAGATTCTGGATCAGGG
ACTTTGCCGGAAGTGAAACCCGATCCAACACCAACCCCGGAGCCGACACCTGAGCCGACG
CCGGACCCAGAACCTACGCCGGATCCAACACCTGATCCTGAGCCGACACCAGAACCGGAG
CCAGAACCTGTTCCTACGAAAACGGGTTATCTGACCCTGGGCGGAAGCCAGCGGGTAACT
GGTGCTACCTGTAATGGTGAATCCAGCGATGGCTTTACCTTTACGCCAGGCAATACCGTG
AGTTGTGTGGTGGGCAGTACGACCATTGCAACATTCAACACCCAGTCAGAAGCTGCGCGT
AGCCTGCGTGCGGTTGACAAAGTGTCGTTTAGCCTGGAGGACGCGCAGGAGCTGGCGAAT
TCTGAAAATAAGAAAACCAACGCCATCTCTCTGGTGACGTCCAGCGACAGTTGCCCCGCA
GATGCAGAACAGCTTTGTCTTACTTTCTCGTCAGTGGTTGATCGCGCGCGATTTGAAAAA

CTGTATAAGCAAATTGATCTGGCAACAGACAATTTCAGCAAGCTGGTCAATGAAGAGGTG
GAAAACAATGCTGCGACTGATAAAGCGCCGTCCACCCATACCTCAACGGTAGTGCCAGTC
ACGACAGAGGGAACAAAACCGGATCTGAACGCGTCCTTCGTGTCGGCTAACGCGGAACAG
TTTTATCAGTATCAACCCACTGAAATCATTCTTTCCGAAGGCCAACTGGTGGATAGCCTG
GGGAACGGTGTTGCTGGCGTTGACTACTACACCAATTCAGGCCGTGGCGTAACTGACGAA
AACGGTAAATTTTCCTTTAGCTGGGGCGAAACCATCTCCTTTGGTATCGATACCTTTGAA
CTGGGCTCAGTACGTGGCAATAAGTCGACCATTGCGCTGACTGAATTGGGTGATGAAGTT
CGCGGGGCAAATATCGATCAGCTCATTCATCGTTATTCGACGACTGGTCAAAATAATACT
CGTGTTGTTCCGGACGATGTACGCAAGGTCTTTGCCGAATATCCCAACGTGATCAACGAG
ATAATCAATCTTTCGTTATCCAACGGTGCGACGCTGGATGAAGGCGATCAAAACGTTGTG
CTGCCTAACGAATTTATCGAGCAGTTTAAGACGGGTCAGGCCAAAGAGATCGATACCGCG
ATTTGTGCGAAAACCGACGGTTGTAACGAGGCTCGCTGGTTCTCGCTGACAACGCGCAAT
GTTAATGACGGCCAGATTCAGGGCGTTATTAACAAGCTGTGGGGCGTGGATACGAACTAT
CAGTCTGTCAGCAAGTTCCACGTCTTCCATGACTCTACCAACTTCTATGGCAGCACCGGT
AACGCGCGCGGTCAGGCGGTGGTAAATATCTCCAACTCGGCATTCCCGATTCTGATGGCG
CGTAATGATAAAAACTACTGGCTGGCGTTTGGCGAAAAACGCGCCTGGGATAAAAATGAG
CTGGCGTACATTACGGAAGCGCCTTCCATTGTGCAGCCAGAGAACGTTACGCGCGATACT
GCGACTTTCAACCTGCCGTTTATTTCGCTGGGGCAAGTCGGTGAAGGCAAACTGATGGTT
ATCGGTAACCCGCACTACAACAGCATCCTGCGTTGCCCGAACGGTTACAGTTGGGGCGGT
GGTGTTAATAGTAAAGGTGAGTGTACGCTCAGCGGTGATTCTGATGACATGAAGCACTTT
ATGCAGAACGTACTGCGCTACTTGTCAAATGACATCTGGCAGCCAAATACCAAGAGCATC
ATGACTGTCGGCACCAACCTGGAGAACGTTTATTTCAAAAAAGCGGGCCAGGTATTGGGA
AATAGTGCACCATTTGCTTTCCATGAGGATTTCACTGGTATCACGGTTAAACAGTTGACC
AGCTATGGCGATCTGAATCCGGAAGAGATTCCGTTGCTGATCCTCAACGGCTTTGAATAT
GTGACTCAGTGGTCTGGCGATCCCTATGCTGTGCCTCTGCGTGCAGATACCAGCAAACCG
AAGCTGACTCAGCAGGATGTGACCGATCTGATCGCTTATCTGAACAAAGGTGGCTCGGTG
CTGATCATGGAAAACGTGATGAGCAATCTTAAGGAAGAGAGCGCGTCCAGTTTTGTGCGT
CTGCTGGATGCCGCGGGTCTGTCAATGGCTCTGAACAAATCGGTGGTGAACAAC

>pK1-3526A (SEQ ID 10)

CDGGGSGSSSDTPSVDSGSGTLPEVKPDPTPTPEPTPEPTPDPEPTPDPTPDPEPTPEPE
PEPVPTKTGYLTLGGSQRVTGATCNGESSDGFTFTPGNTVSCVVGSTTIATFNTQSEAAR
SLRAVDKVSFSLEDAQELANSENKKTNAISLVTSSDSCPADAEQLCLTFSSVVDRARFEK
LYKQIDLATDNFSKLVNEEVENNAATDKAPSTHTSTVVPVTTEGTKPDLNASFVSANAEQ
FYQYQPTEIILSEGQLVDSLGNGVAGVDYYTNSGRGVTDENGKFSFSWGETISFGIDTFE
LGSVRGNKSTIALTELGDEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINE
IINLSLSNGATLDEGDQNVVLPNEFIEQFKTGQAKEIDTAICAKTDGCNEARWFSLTTRN
VNDGQIQGVINKLWGVDTNYQSVSKFHVFHDSTNFYGSTGNARGQAVVNISNSAFPILMA
RNDKNYWLAFGEKRAWDKNELAYITEAPSIVQPENVTRDTATFNLPFISLGQVGEGKLMV
IGNPHYNSILRCPNGYSWGGGVNSKGECTLSGDSDDMKHFMQNVLRYLSNDIWQPNTKSI
MTVGTNLENVYFKKAGQVLGNSAPFAFHEDFTGITVKQLTSYGDLNPEEIPLLILNGFEY
VTQWSGDPYAVPLRADTSKPKLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASSFVR
LLDAAGLSMALNKSVVNN

>pK1-3526B (SEQ ID 11)

GATCCGCAAGGGTATCCGGATCGCGTTCGTCAGCGTCGCGCGACTGGCATTTGGGTTTAT
GAACGTTATCCTGCTGCAGACGGCGCGCAACCGCCGTACACCATCGACCCAAATACAGGG
GAAGTGACCTGGAAATACCAGCAAGACAACAAGCCTGATGACAAGCCGAAACTGGAAGTT
GCGAGCTGGCAGGAGGAAGTTGAGGGCAAACAGGTAACGCGTTATGCCTTTATTGATGAA
GCGGAATACACAACAGAAGAATCTCTGGAAGCGGCAAAGGCAAAAATCTTTGAGAAGTTT
CCTGGGTTACAGGAGTGTAAGGACTCGACTTACCATTACGAGATTAACTGTTTGGAGCGC
CGCCCAGGCACGGATGTTCCGGTAACAGGTGGCATGTATGTTCCGCGCTATACGCAACTG
AATCTTGACGCCGACACCGCGAAAGCGATGGTGCAGGCGGCGGATTTAGGCACCAACATT
CAGCGCCTGTATCAGCATGAGCTTTATTTCCGTACCAAAGGCAGTAAAGGTGAGCGTCTG

AACAGTGTTGATCTGGAACGTCTGTACCAGAACATGTCGGTCTGGCTGTGGAACGATACG
AAATATCGTTACGAAGAGGGCAAGGAAGATGAGCTGGGCTTTAAAACGTTCACCGAGTTC
CTGAACTGCTACGCCAATGATGCCTATGCAGGCGGCACCAAGTGCTCCGCAGATCTGAAA
AAATCGCTGGTCGATAACAACATGATCTACGGTGACGGTAGCAGCAAAGCGGGCATGATG
AACCCAAGCTATCCGCTCAACTATATGGAAAAACCGCTGACGCGTCTGATGCTGGGCCGT
TCCTGGTGGGATCTGAACATTAAGGTTGATGTGGAGAAGTACCCAGGATCCGTATCGGCA
AAGGGTGAGAGCGTTACGGAAAACATCAGCCTGTACTCGAATCCGACCAAATGGTTTGCG
GGTAACATGCAGTCAACCGGCCTGTGGGCACCGGCCCAGCAGGACGTCACCATTAAGTCT
TCGGCGTCAGTCCCAGTGACTGTTACCGTGGCGCTGGCTGACGACCTGACTGGACGTGAG
AAGCATGAAGTTGCGCTGAACCGTCCGCCAAGAGTGACTAAAACGTATACTCTGGAGGCT
AACGGTGAAGTGACCTTCAAGGTGCCTTATGGTGGTCTGATTTATATCAAGGGCGACAGT
AAGGATGATGTTTCTGCTAACTTCACCTTTACCGGTGTAGTAAAAGCGCCGTTCTATAAA
GACGGCGAATGGAAAAACGATCTGGACTCACCGGCGCCGCTGGGCGAGCTGGAGTCTGCG
TCGTTCGTCTATACCACGCCGAAGAAGAACCTTGAGGCCAGCAATTTCACTGGTGGTGTA
GCAGAATTCGCTAAAGATCTGGATACCTTTGCCAGCTCGATGAATGACTTCTACGGTCGT
AATGATGAAGACGGTAAGCACCGGATGTTTACCTATAAAAACTTGACGGGGCACAAGCAT
CGTTTCACCAACGATGTGCAGATCTCCATCGGTGATGCGCACTCGGGTTATCCGGTAATG
AACAGCAGCTTCTCGACGAACAGCACCACGCTGCCGACGACGCCGCTGAACGACTGGCTG
ATTTGGCACGAAGTCGGTCATAACGCTGCAGAAACACCGCTGAACGTACCGGGTGCAACT
GAAGTGGCGAACAACGTGCTGGCGCTGTACATGCAGGATCGCTATCTCGGTAAGATGAAC
CGTGTCGCTGACGACATTACCGTCGCGCCGGAATATCTGGACGAGAGCAACGGTCAGGCC
TGGGCGCGCGGCGGTGCGGGTGACCGTCTGCTGATGTACGCACAGTTGAAGGAGTGGGCA
GAGGAAAACTTTGATATCAAACAGTGGTATCCAGATGGTGAGCTGCCTAAGTTCTACAGC
GATCGTAAAGGGATGAAGGGCTGGAACCTGTTCCAGTTGATGCACCGTAAAGCGCGCGGC
GATGATGTTGGTAACAGCACCTTTGGTGGCAAGAATTACTGTGCTGAATCCAATGGTAAC
GCTGCCGACACGCTGATGCTGTGTGCATCCTGGGTCGCTCAGGCGGATCTTTCGGAATTC
TTTAAGAAATGGAATCCGGGTGCAAGTGCTTACCAGTTGCCGGGAGCAACGGAGATGAGT
TTCCAGGGCGGTGTGAGCTCTTCGGCTTACAGCACGCTGGCGTCACTCAAGCTGCCGAAA
CCGGAAAAAGGGCCGGAAACCATTAACAAGGTTACCGAGCATAAGATGTCTGCCGAG

>pK1-3526B (SEQ ID 12)

DPQGYPDRVRQRRATGIWVYERYPAADGAQPPYTIDPNTGEVTWKYQQDNKPDDKPKLEV
ASWQEEVEGKQVTRYAFIDEAEYTTEESLEAAKAKIFEKFPGLQECKDSTYHYEINCLER
RPGTDVPVTGGMYVPRYTQLNLDADTAKAMVQAADLGTNIQRLYQHELYFRTKGSKGERL
NSVDLERLYQNMSVWLWNDTKYRYEEGKEDELGFKTFTEFLNCYANDAYAGGTKCSADLK
KSLVDNNMIYGDGSSKAGMMNPSYPLNYMEKPLTRLMLGRSWWDLNIKVDVEKYPGSVSA
KGESVTENISLYSNPTKWFAGNMQSTGLWAPAQQDVTIKSSASVPVTVTVALADDLTGRE
KHEVALNRPPRVTKTYTLEANGEVTFKVPYGGLIYIKGDSKDDVSANFTFTGVVKAPFYK
DGEWKNDLDSPAPLGELESASFVYTTPKKNLEASNFTGGVAEFAKDLDTFASSMNDFYGR
NDEDGKHRMFTYKNLTGHKHRFTNDVQISIGDAHSGYPVMNSSFSTNSTTLPTTPLNDWL
IWHEVGHNAAETPLNVPGATEVANNVLALYMQDRYLGKMNRVADDITVAPEYLDESNGQA
WARGGAGDRLLMYAQLKEWAEENFDIKQWYPDGELPKFYSDRKGMKGWNLFQLMHRKARG
DDVGNSTFGGKNYCAESNGNAADTLMLCASWVAQADLSEFFKKWNPGASAYQLPGATEMS
FQGGVSSSAYSTLASLKLPKPEKGPETINKVTEHKMSAE

>pK1-3526dG (SEQ ID 13)

GATACGCCGTCTGTAGATTCTGGATCAGGGACTTTGCCGGAAGTGAAACCCGATCCAACA
CCAACCCCGGAGCCGACACCTGAGCCGACGCCGGACCCAGAACCTACGCCGGATCCAACA
CCTGATCCTGAGCCGACACCAGAACCGGAGCCAGAACCTGTTCCTACGAAAACGGGTTAT
CTGACCCTGGGCGGAAGCCAGCGGGTAACTGGTGCTACCTGTAATGGTGAATCCAGCGAT
GGCTTTACCTTTACGCCAGGCAATACCGTGAGTTGTGTGGTGGGCAGTACGACCATTGCA
ACATTCAACACCCAGTCAGAAGCTGCGCGTAGCCTGCGTGCGGTTGACAAAGTGTCGTTT
AGCCTGGAGGACGCGCAGGAGCTGGCGAATTCTGAAAATAAGAAAACCAACGCCATCTCT
CTGGTGACGTCCAGCGACAGTTGCCCCGCAGATGCAGAACAGCTTTGTCTTACTTTCTCG

TCAGTGGTTGATCGCGCGCGATTTGAAAAACTGTATAAGCAAATTGATCTGGCAACAGAC
AATTTCAGCAAGCTGGTCAATGAAGAGGTGGAAAACAATGCTGCGACTGATAAAGCGCCG
TCCACCCATACCTCAACGGTAGTGCCAGTCACGACAGAGGGAACAAAACCGGATCTGAAC
GCGTCCTTCGTGTCGGCTAACGCGGAACAGTTTTATCAGTATCAACCCACTGAAATCATT
CTTTCCGAAGGCCAACTGGTGGATAGCCTGGGGAACGGTGTTGCTGGCGTTGACTACTAC
ACCAATTCAGGCCGTGGCGTAACTGACGAAAACGGTAAATTTTCCTTTAGCTGGGGCGAA
ACCATCTCCTTTGGTATCGATACCTTTGAACTGGGCTCAGTACGTGGCAATAAGTCGACC
ATTGCGCTGACTGAATTGGGTGATGAAGTTCGCGGGGCAAATATCGATCAGCTCATTCAT
CGTTATTCGACGACTGGTCAAAATAATACTCGTGTTGTTCCGGACGATGTACGCAAGGTC
TTTGCCGAATATCCCAACGTGATCAACGAGATAATCAATCTTTCGTTATCCAACGGTGCG
ACGCTGGATGAAGGCGATCAAAACGTTGTGCTGCCTAACGAATTTATCGAGCAGTTTAAG
ACGGGTCAGGCCAAAGAGATCGATACCGCGATTTGTGCGAAAACCGACGGTTGTAACGAG
GCTCGCTGGTTCTCGCTGACAACGCGCAATGTTAATGACGGCCAGATTCAGGGCGTTATT
AACAAGCTGTGGGGCGTGGATACGAACTATCAGTCTGTCAGCAAGTTCCACGTCTTCCAT
GACTCTACCAACTTCTATGGCAGCACCGGTAACGCGCGCGGTCAGGCGGTGGTAAATATC
TCCAACTCGGCATTCCCGATTCTGATGGCGCGTAATGATAAAAACTACTGGCTGGCGTTT
GGCGAAAAACGCGCCTGGGATAAAAATGAGCTGGCGTACATTACGGAAGCGCCTTCCATT
GTGCAGCCAGAGAACGTTACGCGCGATACTGCGACTTTCAACCTGCCGTTTATTTCGCTG
GGGCAAGTCGGTGAAGGCAAACTGATGGTTATCGGTAACCCGCACTACAACAGCATCCTG
CGTTGCCCGAACGGTTACAGTTGGGGCGGTGGTGTTAATAGTAAAGGTGAGTGTACGCTC
AGCGGTGATTCTGATGACATGAAGCACTTTATGCAGAACGTACTGCGCTACTTGTCAAAT
GACATCTGGCAGCCAAATACCAAGAGCATCATGACTGTCGGCACCAACCTGGAGAACGTT
TATTTCAAAAAAGCGGGCCAGGTATTGGGAAATAGTGCACCATTTGCTTTCCATGAGGAT
TTCACTGGTATCACGGTTAAACAGTTGACCAGCTATGGCGATCTGAATCCGGAAGAGATT
CCGTTGCTGATCCTCAACGGCTTTGAATATGTGACTCAGTGGTCTGGCGATCCCTATGCT
GTGCCTCTGCGTGCAGATACCAGCAAACCGAAGCTGACTCAGCAGGATGTGACCGATCTG
ATCGCTTATCTGAACAAAGGTGGCTCGGTGCTGATCATGGAAAACGTGATGAGCAATCTT
AAGGAAGAGAGCGCGTCCAGTTTTGTGCGTCTGCTGGATGCCGCGGGTCTGTCAATGGCT
CTGAACAAATCGGTGGTGAACAACGATCCGCAAGGGTATCCGGATCGCGTTCGTCAGCGT
CGCGCGACTGGCATTTGGGTTTATGAACGTTATCCTGCTGCAGACGGCGCGCAACCGCCG
TACACCATCGACCCAAATACAGGGGAAGTGACCTGGAAATACCAGCAAGACAACAAGCCT
GATGACAAGCCGAAACTGGAAGTTGCGAGCTGGCAGGAGGAAGTTGAGGGCAAACAGGTA
ACGCGTTATGCCTTTATTGATGAAGCGGAATACACAACAGAAGAATCTCTGGAAGCGGCA
AAGGCAAAAATCTTTGAGAAGTTTCCTGGGTTACAGGAGTGTAAGGACTCGACTTACCAT
TACGAGATTAACTGTTTGGAGCGCCGCCCAGGCACGGATGTTCCGGTAACAGGTGGCATG
TATGTTCCGCGCTATACGCAACTGAATCTTGACGCCGACACCGCGAAAGCGATGGTGCAG
GCGGCGGATTTAGGCACCAACATTCAGCGCCTGTATCAGCATGAGCTTTATTTCCGTACC
AAAGGCAGTAAAGGTGAGCGTCTGAACAGTGTTGATCTGGAACGTCTGTACCAGAACATG
TCGGTCTGGCTGTGGAACGATACGAAATATCGTTACGAAGAGGGCAAGGAAGATGAGCTG
GGCTTTAAAACGTTCACCGAGTTCCTGAACTGCTACGCCAATGATGCCTATGCAGGCGGC
ACCAAGTGCTCCGCAGATCTGAAAAAATCGCTGGTCGATAACAACATGATCTACGGTGAC
GGTAGCAGCAAAGCGGGCATGATGAACCCAAGCTATCCGCTCAACTATATGGAAAAACCG
CTGACGCGTCTGATGCTGGGCCGTTCCTGGTGGGATCTGAACATTAAGGTTGATGTGGAG
AAGTACCCAGGATCCGTATCGGCAAAGGGTGAGAGCGTTACGGAAAACATCAGCCTGTAC
TCGAATCCGACCAAATGGTTTGCGGGTAACATGCAGTCAACCGGCCTGTGGGCACCGGCC
CAGCAGGACGTCACCATTAAGTCTTCGGCGTCAGTCCCAGTGACTGTTACCGTGGCGCTG
GCTGACGACCTGACTGGACGTGAGAAGCATGAAGTTGCGCTGAACCGTCCGCCAAGAGTG
ACTAAAACGTATACTCTGGAGGCTAACGGTGAAGTGACCTTCAAGGTGCCTTATGGTGGT
CTGATTTATATCAAGGGCGACAGTAAGGATGATGTTTCTGCTAACTTCACCTTTACCGGT
GTAGTAAAAGCGCCGTTCTATAAAGACGGCGAATGGAAAAACGATCTGGACTCACCGGCG
CCGCTGGGCGAGCTGGAGTCTGCGTCGTTCGTCTATACCACGCCGAAGAAGAACCTTGAG
GCCAGCAATTTCACTGGTGGTGTAGCAGAATTCGCTAAAGATCTGGATACCTTTGCCAGC
TCGATGAATGACTTCTACGGTCGTAATGATGAAGACGGTAAGCACCGGATGTTTACCTAT
AAAAACTTGACGGGGCACAAGCATCGTTTCACCAACGATGTGCAGATCTCCATCGGTGAT
GCGCACTCGGGTTATCCGGTAATGAACAGCAGCTTCTCGACGAACAGCACCACGCTGCCG
ACGACGCCGCTGAACGACTGGCTGATTTGGCACGAAGTCGGTCATAACGCTGCAGAAACA

CCGCTGAACGTACCGGGTGCAACTGAAGTGGCGAACAACGTGCTGGCGCTGTACATGCAG
GATCGCTATCTCGGTAAGATGAACCGTGTCGCTGACGACATTACCGTCGCGCCGGAATAT
CTGGACGAGAGCAACGGTCAGGCCTGGGCGCGCGGCGGTGCGGGTGACCGTCTGCTGATG
TACGCACAGTTGAAGGAGTGGGCAGAGGAAAACTTTGATATCAAACAGTGGTATCCAGAT
GGTGAGCTGCCTAAGTTCTACAGCGATCGTAAAGGGATGAAGGGCTGGAACCTGTTCCAG
TTGATGCACCGTAAAGCGCGCGGCGATGATGTTGGTAACAGCACCTTTGGTGGCAAGAAT
TACTGTGCTGAATCCAATGGTAACGCTGCCGACACGCTGATGCTGTGTGCATCCTGGGTC
GCTCAGGCGGATCTTTCGGAATTCTTTAAGAAATGGAATCCGGGTGCAAGTGCTTACCAG
TTGCCGGGAGCAACGGAGATGAGTTTCCAGGGCGGTGTGAGCTCTTCGGCTTACAGCACG
CTGGCGTCACTCAAGCTGCCGAAACCGGAAAAAGGGCCGGAAACCATTAACAAGGTTACC
GAGCATAAGATGTCTGCCGAG

>pK1-3526dG (SEQ ID 14)

DTPSVDSGSGTLPEVKPDPTPTPEPTPEPTPDPEPTPDPTPDPEPTPEPEPEPVPTKTGY
LTLGGSQRVTGATCNGESSDGFTFTPGNTVSCVVGSTTIATFNTQSEAARSLRAVDKVSF
SLEDAQELANSENKKTNAISLVTSSDSCPADAEQLCLTFSSVVDRARFEKLYKQIDLATD
NFSKLVNEEVENNAATDKAPSTHTSTVVPVTTEGTKPDLNASFVSANAEQFYQYQPTEII
LSEGQLVDSLGNGVAGVDYYTNSGRGVTDENGKFSFSWGETISFGIDTFELGSVRGNKST
IALTELGDEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINEIINLSLSNGA
TLDEGDQNVVLPNEFIEQFKTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVI
NKLWGVDTNYQSVSKFHVFHDSTNFYGSTGNARGQAVVNISNSAFPILMARNDKNYWLAF
GEKRAWDKNELAYITEAPSIVQPENVTRDTATFNLPFISLGQVGEGKLMVIGNPHYNSIL
RCPNGYSWGGGVNSKGECTLSGDSDDMKHFMQNVLRYLSNDIWQPNTKSIMTVGTNLENV
YFKKAGQVLGNSAPFAFHEDFTGITVKQLTSYGDLNPEEIPLLILNGFEYVTQWSGDPYA
VPLRADTSKPKLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASSFVRLLDAAGLSMA
LNKSVVNNDPQGYPDRVRQRRATGIWVYERYPAADGAQPPYTIDPNTGEVTWKYQQDNKP
DDKPKLEVASWQEEVEGKQVTRYAFIDEAEYTTEESLEAAKAKIFEKFPGLQECKDSTYH
YEINCLERRPGTDVPVTGGMYVPRYTQLNLDADTAKAMVQAADLGTNIQRLYQHELYFRT
KGSKGERLNSVDLERLYQNMSVWLWNDTKYRYEEGKEDELGFKTFTEFLNCYANDAYAGG
TKCSADLKKSLVDNNMIYGDGSSKAGMMNPSYPLNYMEKPLTRLMLGRSWWDLNIKVDVE
KYPGSVSAKGESVTENISLYSNPTKWFAGNMQSTGLWAPAQQDVTIKSSASVPVTVTVAL
ADDLTGREKHEVALNRPPRVTKTYTLEANGEVTFKVPYGGLIYIKGDSKDDVSANFTFTG
VVKAPFYKDGEWKNDLDSPAPLGELESASFVYTTPKKNLEASNFTGGVAEFAKDLDTFAS
SMNDFYGRNDEDGKHRMFTYKNLTGHKHRFTNDVQISIGDAHSGYPVMNSSFSTNSTTLP
TTPLNDWLIWHEVGHNAAETPLNVPGATEVANNVLALYMQDRYLGKMNRVADDITVAPEY
LDESNGQAWARGGAGDRLLMYAQLKEWAEENFDIKQWYPDGELPKFYSDRKGMKGWNLFQ
LMHRKARGDDVGNSTFGGKNYCAESNGNAADTLMLCASWVAQADLSEFFKKWNPGASAYQ
LPGATEMSFQGGVSSSAYSTLASLKLPKPEKGPETINKVTEHKMSAE

>pK1-3526dP (SEQ ID 15)

AAAACGGGTTATCTGACCCTGGGCGGAAGCCAGCGGGTAACTGGTGCTACCTGTAATGGT
GAATCCAGCGATGGCTTTACCTTTACGCCAGGCAATACCGTGAGTTGTGTGGTGGGCAGT
ACGACCATTGCAACATTCAACACCCAGTCAGAAGCTGCGCGTAGCCTGCGTGCGGTTGAC
AAAGTGTCGTTTAGCCTGGAGGACGCGCAGGAGCTGGCGAATTCTGAAAATAAGAAAACC
AACGCCATCTCTCTGGTGACGTCCAGCGACAGTTGCCCCGCAGATGCAGAACAGCTTTGT
CTTACTTTCTCGTCAGTGGTTGATCGCGCGCGATTTGAAAAACTGTATAAGCAAATTGAT
CTGGCAACAGACAATTTCAGCAAGCTGGTCAATGAAGAGGTGGAAAACAATGCTGCGACT
GATAAAGCGCCGTCCACCCATACCTCAACGGTAGTGCCAGTCACGACAGAGGGAACAAAA
CCGGATCTGAACGCGTCCTTCGTGTCGGCTAACGCGGAACAGTTTTATCAGTATCAACCC
ACTGAAATCATTCTTTCCGAAGGCCAACTGGTGGATAGCCTGGGGAACGGTGTTGCTGGC
GTTGACTACTACACCAATTCAGGCCGTGGCGTAACTGACGAAAACGGTAAATTTTCCTTT
AGCTGGGGCGAAACCATCTCCTTTGGTATCGATACCTTTGAACTGGGCTCAGTACGTGGC
AATAAGTCGACCATTGCGCTGACTGAATTGGGTGATGAAGTTCGCGGGGCAAATATCGAT
CAGCTCATTCATCGTTATTCGACGACTGGTCAAAATAATACTCGTGTTGTTCCGGACGAT

GTACGCAAGGTCTTTGCCGAATATCCCAACGTGATCAACGAGATAATCAATCTTTCGTTA
TCCAACGGTGCGACGCTGGATGAAGGCGATCAAAACGTTGTGCTGCCTAACGAATTTATC
GAGCAGTTTAAGACGGGTCAGGCCAAAGAGATCGATACCGCGATTTGTGCGAAAACCGAC
GGTTGTAACGAGGCTCGCTGGTTCTCGCTGACAACGCGCAATGTTAATGACGGCCAGATT
CAGGGCGTTATTAACAAGCTGTGGGGCGTGGATACGAACTATCAGTCTGTCAGCAAGTTC
CACGTCTTCCATGACTCTACCAACTTCTATGGCAGCACCGGTAACGCGCGCGGTCAGGCG
GTGGTAAATATCTCCAACTCGGCATTCCCGATTCTGATGGCGCGTAATGATAAAAACTAC
TGGCTGGCGTTTGGCGAAAAACGCGCCTGGGATAAAAATGAGCTGGCGTACATTACGGAA
GCGCCTTCCATTGTGCAGCCAGAGAACGTTACGCGCGATACTGCGACTTTCAACCTGCCG
TTTATTTCGCTGGGGCAAGTCGGTGAAGGCAAACTGATGGTTATCGGTAACCCGCACTAC
AACAGCATCCTGCGTTGCCCGAACGGTTACAGTTGGGGCGGTGGTGTTAATAGTAAAGGT
GAGTGTACGCTCAGCGGTGATTCTGATGACATGAAGCACTTTATGCAGAACGTACTGCGC
TACTTGTCAAATGACATCTGGCAGCCAAATACCAAGAGCATCATGACTGTCGGCACCAAC
CTGGAGAACGTTTATTTCAAAAAAGCGGGCCAGGTATTGGGAAATAGTGCACCATTTGCT
TTCCATGAGGATTTCACTGGTATCACGGTTAAACAGTTGACCAGCTATGGCGATCTGAAT
CCGGAAGAGATTCCGTTGCTGATCCTCAACGGCTTTGAATATGTGACTCAGTGGTCTGGC
GATCCCTATGCTGTGCCTCTGCGTGCAGATACCAGCAAACCGAAGCTGACTCAGCAGGAT
GTGACCGATCTGATCGCTTATCTGAACAAAGGTGGCTCGGTGCTGATCATGGAAAACGTG
ATGAGCAATCTTAAGGAAGAGAGCGCGTCCAGTTTTGTGCGTCTGCTGGATGCCGCGGGT
CTGTCAATGGCTCTGAACAAATCGGTGGTGAACAACGATCCGCAAGGGTATCCGGATCGC
GTTCGTCAGCGTCGCGCGACTGGCATTTGGGTTTATGAACGTTATCCTGCTGCAGACGGC
GCGCAACCGCCGTACACCATCGACCCAAATACAGGGGAAGTGACCTGGAAATACCAGCAA
GACAACAAGCCTGATGACAAGCCGAAACTGGAAGTTGCGAGCTGGCAGGAGGAAGTTGAG
GGCAAACAGGTAACGCGTTATGCCTTTATTGATGAAGCGGAATACACAACAGAAGAATCT
CTGGAAGCGGCAAAGGCAAAAATCTTTGAGAAGTTTCCTGGGTTACAGGAGTGTAAGGAC
TCGACTTACCATTACGAGATTAACTGTTTGGAGCGCCGCCCAGGCACGGATGTTCCGGTA
ACAGGTGGCATGTATGTTCCGCGCTATACGCAACTGAATCTTGACGCCGACACCGCGAAA
GCGATGGTGCAGGCGGCGGATTTAGGCACCAACATTCAGCGCCTGTATCAGCATGAGCTT
TATTTCCGTACCAAAGGCAGTAAAGGTGAGCGTCTGAACAGTGTTGATCTGGAACGTCTG
TACCAGAACATGTCGGTCTGGCTGTGGAACGATACGAAATATCGTTACGAAGAGGGCAAG
GAAGATGAGCTGGGCTTTAAAACGTTCACCGAGTTCCTGAACTGCTACGCCAATGATGCC
TATGCAGGCGGCACCAAGTGCTCCGCAGATCTGAAAAAATCGCTGGTCGATAACAACATG
ATCTACGGTGACGGTAGCAGCAAAGCGGGCATGATGAACCCAAGCTATCCGCTCAACTAT
ATGGAAAAACCGCTGACGCGTCTGATGCTGGGCCGTTCCTGGTGGGATCTGAACATTAAG
GTTGATGTGGAGAAGTACCCAGGATCCGTATCGGCAAAGGGTGAGAGCGTTACGGAAAAC
ATCAGCCTGTACTCGAATCCGACCAAATGGTTTGCGGGTAACATGCAGTCAACCGGCCTG
TGGGCACCGGCCCAGCAGGACGTCACCATTAAGTCTTCGGCGTCAGTCCCAGTGACTGTT
ACCGTGGCGCTGGCTGACGACCTGACTGGACGTGAGAAGCATGAAGTTGCGCTGAACCGT
CCGCCAAGAGTGACTAAAACGTATACTCTGGAGGCTAACGGTGAAGTGACCTTCAAGGTG
CCTTATGGTGGTCTGATTTATATCAAGGGCGACAGTAAGGATGATGTTTCTGCTAACTTC
ACCTTTACCGGTGTAGTAAAAGCGCCGTTCTATAAAGACGGCGAATGGAAAAACGATCTG
GACTCACCGGCGCCGCTGGGCGAGCTGGAGTCTGCGTCGTTCGTCTATACCACGCCGAAG
AAGAACCTTGAGGCCAGCAATTTCACTGGTGGTGTAGCAGAATTCGCTAAAGATCTGGAT
ACCTTTGCCAGCTCGATGAATGACTTCTACGGTCGTAATGATGAAGACGGTAAGCACCGG
ATGTTTACCTATAAAAACTTGACGGGGCACAAGCATCGTTTCACCAACGATGTGCAGATC
TCCATCGGTGATGCGCACTCGGGTTATCCGGTAATGAACAGCAGCTTCTCGACGAACAGC
ACCACGCTGCCGACGACGCCGCTGAACGACTGGCTGATTTGGCACGAAGTCGGTCATAAC
GCTGCAGAAACACCGCTGAACGTACCGGGTGCAACTGAAGTGGCGAACAACGTGCTGGCG
CTGTACATGCAGGATCGCTATCTCGGTAAGATGAACCGTGTCGCTGACGACATTACCGTC
GCGCCGGAATATCTGGACGAGAGCAACGGTCAGGCCTGGGCGCGCGGCGGTGCGGGTGAC
CGTCTGCTGATGTACGCACAGTTGAAGGAGTGGGCAGAGGAAAACTTTGATATCAAACAG
TGGTATCCAGATGGTGAGCTGCCTAAGTTCTACAGCGATCGTAAAGGGATGAAGGGCTGG
AACCTGTTCCAGTTGATGCACCGTAAAGCGCGCGGCGATGATGTTGGTAACAGCACCTTT
GGTGGCAAGAATTACTGTGCTGAATCCAATGGTAACGCTGCCGACACGCTGATGCTGTGT
GCATCCTGGGTCGCTCAGGCGGATCTTTCGGAATTCTTTAAGAAATGGAATCCGGGTGCA
AGTGCTTACCAGTTGCCGGGAGCAACGGAGATGAGTTTCCAGGGCGGTGTGAGCTCTTCG

GCTTACAGCACGCTGGCGTCACTCAAGCTGCCGAAACCGGAAAAAGGGCCGGAAACCATT
AACAAGGTTACCGAGCATAAGATGTCTGCCGAG

>pK1-3526dP (SEQ ID 16)

KTGYLTLGGSQRVTGATCNGESSDGFTFTPGNTVSCVVGSTTIATFNTQSEAARSLRAVD
KVSFSLEDAQELANSENKKTNAISLVTSSDSCPADAEQLCLTFSSVVDRARFEKLYKQID
LATDNFSKLVNEEVENNAATDKAPSTHTSTVVPVTTEGTKPDLNASFVSANAEQFYQYQP
TEIILSEGQLVDSLGNGVAGVDYYTNSGRGVTDENGKFSFSWGETISFGIDTFELGSVRG
NKSTIALTELGDEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINEIINLSL
SNGATLDEGDQNVVLPNEFIEQFKTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQI
QGVINKLWGVDTNYQSVSKFHVFHDSTNFYGSTGNARGQAVVNISNSAFPILMARNDKNY
WLAFGEKRAWDKNELAYITEAPSIVQPENVTRDTATFNLPFISLGQVGEGKLMVIGNPHY
NSILRCPNGYSWGGGVNSKGECTLSGDSDDMKHFMQNVLRYLSNDIWQPNTKSIMTVGTN
LENVYFKKAGQVLGNSAPFAFHEDFTGITVKQLTSYGDLNPEEIPLLILNGFEYVTQWSG
DPYAVPLRADTSKPKLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASSFVRLLDAAG
LSMALNKSVVNNDPQGYPDRVRQRRATGIWVYERYPAADGAQPPYTIDPNTGEVTWKYQQ
DNKPDDKPKLEVASWQEEVEGKQVTRYAFIDEAEYTTEESLEAAKAKIFEKFPGLQECKD
STYHYEINCLERRPGTDVPVTGGMYVPRYTQLNLDADTAKAMVQAADLGTNIQRLYQHEL
YFRTKGSKGERLNSVDLERLYQNMSVWLWNDTKYRYEEGKEDELGFKTFTEFLNCYANDA
YAGGTKCSADLKKSLVDNNMIYGDGSSKAGMMNPSYPLNYMEKPLTRLMLGRSWWDLNIK
VDVEKYPGSVSAKGESVTENISLYSNPTKWFAGNMQSTGLWAPAQQDVTIKSSASVPVTV
TVALADDLTGREKHEVALNRPPRVTKTYTLEANGEVTFKVPYGGLIYIKGDSKDDVSANF
TFTGVVKAPFYKDGEWKNDLDSPAPLGELESASFVYTTPKKNLEASNFTGGVAEFAKDLD
TFASSMNDFYGRNDEDGKHRMFTYKNLTGHKHRFTNDVQISIGDAHSGYPVMNSSFSTNS
TTLPTTPLNDWLIWHEVGHNAAETPLNVPGATEVANNVLALYMQDRYLGKMNRVADDITV
APEYLDESNGQAWARGGAGDRLLMYAQLKEWAEENFDIKQWYPDGELPKFYSDRKGMKGW
NLFQLMHRKARGDDVGNSTFGGKNYCAESNGNAADTLMLCASWVAQADLSEFFKKWNPGA
SAYQLPGATEMSFQGGVSSSAYSTLASLKLPKPEKGPETINKVTEHKMSAE

>pK1-3526 (SEQ ID 17)

TGTGATGGTGGTGGTTCAGGATCGTCCTCCGATACGCCGTCTGTAGATTCTGGATCAGGG
ACTTTGCCGGAAGTGAAACCCGATCCAACACCAACCCCGGAGCCGACACCTGAGCCGACG
CCGGACCCAGAACCTACGCCGGATCCAACACCTGATCCTGAGCCGACACCAGAACCGGAG
CCAGAACCTGTTCCTACGAAAACGGGTTATCTGACCCTGGGCGGAAGCCAGCGGGTAACT
GGTGCTACCTGTAATGGTGAATCCAGCGATGGCTTTACCTTTACGCCAGGCAATACCGTG
AGTTGTGTGGTGGGCAGTACGACCATTGCAACATTCAACACCCAGTCAGAAGCTGCGCGT
AGCCTGCGTGCGGTTGACAAAGTGTCGTTTAGCCTGGAGGACGCGCAGGAGCTGGCGAAT
TCTGAAAATAAGAAAACCAACGCCATCTCTCTGGTGACGTCCAGCGACAGTTGCCCCGCA
GATGCAGAACAGCTTTGTCTTACTTTCTCGTCAGTGGTTGATCGCGCGCGATTTGAAAAA
CTGTATAAGCAAATTGATCTGGCAACAGACAATTTCAGCAAGCTGGTCAATGAAGAGGTG
GAAAACAATGCTGCGACTGATAAAGCGCCGTCCACCCATACCTCAACGGTAGTGCCAGTC
ACGACAGAGGGAACAAAACCGGATCTGAACGCGTCCTTCGTGTCGGCTAACGCGGAACAG
TTTTATCAGTATCAACCCACTGAAATCATTCTTTCCGAAGGCCAACTGGTGGATAGCCTG
GGGAACGGTGTTGCTGGCGTTGACTACTACACCAATTCAGGCCGTGGCGTAACTGACGAA
AACGGTAAATTTTCCTTTAGCTGGGGCGAAACCATCTCCTTTGGTATCGATACCTTTGAA
CTGGGCTCAGTACGTGGCAATAAGTCGACCATTGCGCTGACTGAATTGGGTGATGAAGTT
CGCGGGGCAAATATCGATCAGCTCATTCATCGTTATTCGACGACTGGTCAAAATAATACT
CGTGTTGTTCCGGACGATGTACGCAAGGTCTTTGCCGAATATCCCAACGTGATCAACGAG
ATAATCAATCTTTCGTTATCCAACGGTGCGACGCTGGATGAAGGCGATCAAAACGTTGTG
CTGCCTAACGAATTTATCGAGCAGTTTAAGACGGGTCAGGCCAAAGAGATCGATACCGCG
ATTTGTGCGAAAACCGACGGTTGTAACGAGGCTCGCTGGTTCTCGCTGACAACGCGCAAT
GTTAATGACGGCCAGATTCAGGGCGTTATTAACAAGCTGTGGGGCGTGGATACGAACTAT
CAGTCTGTCAGCAAGTTCCACGTCTTCCATGACTCTACCAACTTCTATGGCAGCACCGGT
AACGCGCGCGGTCAGGCGGTGGTAAATATCTCCAACTCGGCATTCCCGATTCTGATGGCG

CGTAATGATAAAAACTACTGGCTGGCGTTTGGCGAAAAACGCGCCTGGGATAAAAATGAG
CTGGCGTACATTACGGAAGCGCCTTCCATTGTGCAGCCAGAGAACGTTACGCGCGATACT
GCGACTTTCAACCTGCCGTTTATTTCGCTGGGGCAAGTCGGTGAAGGCAAACTGATGGTT
ATCGGTAACCCGCACTACAACAGCATCCTGCGTTGCCCGAACGGTTACAGTTGGGGCGGT
GGTGTTAATAGTAAAGGTGAGTGTACGCTCAGCGGTGATTCTGATGACATGAAGCACTTT
ATGCAGAACGTACTGCGCTACTTGTCAAATGACATCTGGCAGCCAAATACCAAGAGCATC
ATGACTGTCGGCACCAACCTGGAGAACGTTTATTTCAAAAAAGCGGGCCAGGTATTGGGA
AATAGTGCACCATTTGCTTTCCATGAGGATTTCACTGGTATCACGGTTAAACAGTTGACC
AGCTATGGCGATCTGAATCCGGAAGAGATTCCGTTGCTGATCCTCAACGGCTTTGAATAT
GTGACTCAGTGGTCTGGCGATCCCTATGCTGTGCCTCTGCGTGCAGATACCAGCAAACCG
AAGCTGACTCAGCAGGATGTGACCGATCTGATCGCTTATCTGAACAAAGGTGGCTCGGTG
CTGATCATGGAAAACGTGATGAGCAATCTTAAGGAAGAGAGCGCGTCCAGTTTTGTGCGT
CTGCTGGATGCCGCGGGTCTGTCAATGGCTCTGAACAAATCGGTGGTGAACAACGATCCG
CAAGGGTATCCGGATCGCGTTCGTCAGCGTCGCGCGACTGGCATTTGGGTTTATGAACGT
TATCCTGCTGCAGACGGCGCGCAACCGCCGTACACCATCGACCCAAATACAGGGGAAGTG
ACCTGGAAATACCAGCAAGACAACAAGCCTGATGACAAGCCGAAACTGGAAGTTGCGAGC
TGGCAGGAGGAAGTTGAGGGCAAACAGGTAACGCGTTATGCCTTTATTGATGAAGCGGAA
TACACAACAGAAGAATCTCTGGAAGCGGCAAAGGCAAAAATCTTTGAGAAGTTTCCTGGG
TTACAGGAGTGTAAGGACTCGACTTACCATTACGAGATTAACTGTTTGGAGCGCCGCCCA
GGCACGGATGTTCCGGTAACAGGTGGCATGTATGTTCCGCGCTATACGCAACTGAATCTT
GACGCCGACACCGCGAAAGCGATGGTGCAGGCGGCGGATTTAGGCACCAACATTCAGCGC
CTGTATCAGCATGAGCTTTATTTCCGTACCAAAGGCAGTAAAGGTGAGCGTCTGAACAGT
GTTGATCTGGAACGTCTGTACCAGAACATGTCGGTCTGGCTGTGGAACGATACGAAATAT
CGTTACGAAGAGGGCAAGGAAGATGAGCTGGGCTTTAAAACGTTCACCGAGTTCCTGAAC
TGCTACGCCAATGATGCCTATGCAGGCGGCACCAAGTGCTCCGCAGATCTGAAAAAATCG
CTGGTCGATAACAACATGATCTACGGTGACGGTAGCAGCAAAGCGGGCATGATGAACCCA
AGCTATCCGCTCAACTATATGGAAAAACCGCTGACGCGTCTGATGCTGGGCCGTTCCTGG
TGGGATCTGAACATTAAGGTTGATGTGGAGAAGTACCCAGGATCCGTATCGGCAAAGGGT
GAGAGCGTTACGGAAAACATCAGCCTGTACTCGAATCCGACCAAATGGTTTGCGGGTAAC
ATGCAGTCAACCGGCCTGTGGGCACCGGCCCAGCAGGACGTCACCATTAAGTCTTCGGCG
TCAGTCCCAGTGACTGTTACCGTGGCGCTGGCTGACGACCTGACTGGACGTGAGAAGCAT
GAAGTTGCGCTGAACCGTCCGCCAAGAGTGACTAAAACGTATACTCTGGAGGCTAACGGT
GAAGTGACCTTCAAGGTGCCTTATGGTGGTCTGATTTATATCAAGGGCGACAGTAAGGAT
GATGTTTCTGCTAACTTCACCTTTACCGGTGTAGTAAAAGCGCCGTTCTATAAAGACGGC
GAATGGAAAAACGATCTGGACTCACCGGCGCCGCTGGGCGAGCTGGAGTCTGCGTCGTTC
GTCTATACCACGCCGAAGAAGAACCTTGAGGCCAGCAATTTCACTGGTGGTGTAGCAGAA
TTCGCTAAAGATCTGGATACCTTTGCCAGCTCGATGAATGACTTCTACGGTCGTAATGAT
GAAGACGGTAAGCACCGGATGTTTACCTATAAAAACTTGACGGGGCACAAGCATCGTTTC
ACCAACGATGTGCAGATCTCCATCGGTGATGCGCACTCGGGTTATCCGGTAATGAACAGC
AGCTTCTCGACGAACAGCACCACGCTGCCGACGACGCCGCTGAACGACTGGCTGATTTGG
CACGAAGTCGGTCATAACGCTGCAGAAACACCGCTGAACGTACCGGGTGCAACTGAAGTG
GCGAACAACGTGCTGGCGCTGTACATGCAGGATCGCTATCTCGGTAAGATGAACCGTGTC
GCTGACGACATTACCGTCGCGCCGGAATATCTGGACGAGAGCAACGGTCAGGCCTGGGCG
CGCGGCGGTGCGGGTGACCGTCTGCTGATGTACGCACAGTTGAAGGAGTGGGCAGAGGAA
AACTTTGATATCAAACAGTGGTATCCAGATGGTGAGCTGCCTAAGTTCTACAGCGATCGT
AAAGGGATGAAGGGCTGGAACCTGTTCCAGTTGATGCACCGTAAAGCGCGCGGCGATGAT
GTTGGTAACAGCACCTTTGGTGGCAAGAATTACTGTGCTGAATCCAATGGTAACGCTGCC
GACACGCTGATGCTGTGTGCATCCTGGGTCGCTCAGGCGGATCTTTCGGAATTCTTTAAG
AAATGGAATCCGGGTGCAAGTGCTTACCAGTTGCCGGGAGCAACGGAGATGAGTTTCCAG
GGCGGTGTGAGCTCTTCGGCTTACAGCACGCTGGCGTCACTCAAGCTGCCGAAACCGGAA
AAAGGGCCGGAAACCATTAACAAGGTTACCGAGCATAAGATGTCTGCCGAG

>pK1-3526 (SEQ ID 18)

CDGGGSGSSSDTPSVDSGSGTLPEVKPDPTPTPEPTPEPTPDPEPTPDPTPDPEPTPEPE
PEPVPTKTGYLTLGGSQRVTGATCNGESSDGFTFTPGNTVSCVVGSTTIATFNTQSEAAR

SLRAVDKVSFSLEDAQELANSENKKTNAISLVTSSDSCPADAEQLCLTFSSVVDRARFEK
LYKQIDLATDNFSKLVNEEVENNAATDKAPSTHTSTVVPVTTEGTKPDLNASFVSANAEQ
FYQYQPTEIILSEGQLVDSLGNGVAGVDYYTNSGRGVTDENGKFSFSWGETISFGIDTFE
LGSVRGNKSTIALTELGDEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINE
IINLSLSNGATLDEGDQNVVLPNEFIEQFKTGQAKEIDTAICAKTDGCNEARWFSLTTRN
VNDGQIQGVINKLWGVDTNYQSVSKFHVFHDSTNFYGSTGNARGQAVVNISNSAFPILMA
RNDKNYWLAFGEKRAWDKNELAYITEAPSIVQPENVTRDTATFNLPFISLGQVGEGKLMV
IGNPHYNSILRCPNGYSWGGGVNSKGECTLSGDSDDMKHFMQNVLRYLSNDIWQPNTKSI
MTVGTNLENVYFKKAGQVLGNSAPFAFHEDFTGITVKQLTSYGDLNPEEIPLLILNGFEY
VTQWSGDPYAVPLRADTSKPKLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASSFVR
LLDAAGLSMALNKSVVNNDPQGYPDRVRQRRATGIWVYERYPAADGAQPPYTIDPNTGEV
TWKYQQDNKPDDKPKLEVASWQEEVEGKQVTRYAFIDEAEYTTEESLEAAKAKIFEKFPG
LQECKDSTYHYEINCLERRPGTDVPVTGGMYVPRYTQLNLDADTAKAMVQAADLGTNIQR
LYQHELYFRTKGSKGERLNSVDLERLYQNMSVWLWNDTKYRYEEGKEDELGFKTFTEFLN
CYANDAYAGGTKCSADLKKSLVDNNMIYGDGSSKAGMMNPSYPLNYMEKPLTRLMLGRSW
WDLNIKVDVEKYPGSVSAKGESVTENISLYSNPTKWFAGNMQSTGLWAPAQQDVTIKSSA
SVPVTVTVALADDLTGREKHEVALNRPPRVTKTYTLEANGEVTFKVPYGGLIYIKGDSKD
DVSANFTFTGVVKAPFYKDGEWKNDLDSPAPLGELESASFVYTTPKKNLEASNFTGGVAE
FAKDLDTFASSMNDFYGRNDEDGKHRMFTYKNLTGHKHRFTNDVQISIGDAHSGYPVMNS
SFSTNSTTLPTTPLNDWLIWHEVGHNAAETPLNVPGATEVANNVLALYMQDRYLGKMNRV
ADDITVAPEYLDESNGQAWARGGAGDRLLMYAQLKEWAEENFDIKQWYPDGELPKFYSDR
KGMKGWNLFQLMHRKARGDDVGNSTFGGKNYCAESNGNAADTLMLCASWVAQADLSEFFK
KWNPGASAYQLPGATEMSFQGGVSSSAYSTLASLKLPKPEKGPETINKVTEHKMSAE

>pK1-3526E1305A (SEQ ID 19)

TGTGATGGTGGTGGTTCAGGATCGTCCTCCGATACGCCGTCTGTAGATTCTGGATCAGGG
ACTTTGCCGGAAGTGAAACCCGATCCAACACCAACCCCGGAGCCGACACCTGAGCCGACG
CCGGACCCAGAACCTACGCCGGATCCAACACCTGATCCTGAGCCGACACCAGAACCGGAG
CCAGAACCTGTTCCTACGAAAACGGGTTATCTGACCCTGGGCGGAAGCCAGCGGGTAACT
GGTGCTACCTGTAATGGTGAATCCAGCGATGGCTTTACCTTTACGCCAGGCAATACCGTG
AGTTGTGTGGTGGGCAGTACGACCATTGCAACATTCAACACCCAGTCAGAAGCTGCGCGT
AGCCTGCGTGCGGTTGACAAAGTGTCGTTTAGCCTGGAGGACGCGCAGGAGCTGGCGAAT
TCTGAAAATAAGAAAACCAACGCCATCTCTCTGGTGACGTCCAGCGACAGTTGCCCCGCA
GATGCAGAACAGCTTTGTCTTACTTTCTCGTCAGTGGTTGATCGCGCGCGATTTGAAAAA
CTGTATAAGCAAATTGATCTGGCAACAGACAATTTCAGCAAGCTGGTCAATGAAGAGGTG
GAAAACAATGCTGCGACTGATAAAGCGCCGTCCACCCATACCTCAACGGTAGTGCCAGTC
ACGACAGAGGGAACAAAACCGGATCTGAACGCGTCCTTCGTGTCGGCTAACGCGGAACAG
TTTTATCAGTATCAACCCACTGAAATCATTCTTTCCGAAGGCCAACTGGTGGATAGCCTG
GGGAACGGTGTTGCTGGCGTTGACTACTACACCAATTCAGGCCGTGGCGTAACTGACGAA
AACGGTAAATTTTCCTTTAGCTGGGGCGAAACCATCTCCTTTGGTATCGATACCTTTGAA
CTGGGCTCAGTACGTGGCAATAAGTCGACCATTGCGCTGACTGAATTGGGTGATGAAGTT
CGCGGGGCAAATATCGATCAGCTCATTCATCGTTATTCGACGACTGGTCAAAATAATACT
CGTGTTGTTCCGGACGATGTACGCAAGGTCTTTGCCGAATATCCCAACGTGATCAACGAG
ATAATCAATCTTTCGTTATCCAACGGTGCGACGCTGGATGAAGGCGATCAAAACGTTGTG
CTGCCTAACGAATTTATCGAGCAGTTTAAGACGGGTCAGGCCAAAGAGATCGATACCGCG
ATTTGTGCGAAAACCGACGGTTGTAACGAGGCTCGCTGGTTCTCGCTGACAACGCGCAAT
GTTAATGACGGCCAGATTCAGGGCGTTATTAACAAGCTGTGGGGCGTGGATACGAACTAT
CAGTCTGTCAGCAAGTTCCACGTCTTCCATGACTCTACCAACTTCTATGGCAGCACCGGT
AACGCGCGCGGTCAGGCGGTGGTAAATATCTCCAACTCGGCATTCCCGATTCTGATGGCG
CGTAATGATAAAAACTACTGGCTGGCGTTTGGCGAAAAACGCGCCTGGGATAAAAATGAG
CTGGCGTACATTACGGAAGCGCCTTCCATTGTGCAGCCAGAGAACGTTACGCGCGATACT
GCGACTTTCAACCTGCCGTTTATTTCGCTGGGGCAAGTCGGTGAAGGCAAACTGATGGTT
ATCGGTAACCCGCACTACAACAGCATCCTGCGTTGCCCGAACGGTTACAGTTGGGGCGGT
GGTGTTAATAGTAAAGGTGAGTGTACGCTCAGCGGTGATTCTGATGACATGAAGCACTTT
ATGCAGAACGTACTGCGCTACTTGTCAAATGACATCTGGCAGCCAAATACCAAGAGCATC

ATGACTGTCGGCACCAACCTGGAGAACGTTTATTTCAAAAAAGCGGGCCAGGTATTGGGA
AATAGTGCACCATTTGCTTTCCATGAGGATTTCACTGGTATCACGGTTAAACAGTTGACC
AGCTATGGCGATCTGAATCCGGAAGAGATTCCGTTGCTGATCCTCAACGGCTTTGAATAT
GTGACTCAGTGGTCTGGCGATCCCTATGCTGTGCCTCTGCGTGCAGATACCAGCAAACCG
AAGCTGACTCAGCAGGATGTGACCGATCTGATCGCTTATCTGAACAAAGGTGGCTCGGTG
CTGATCATGGAAAACGTGATGAGCAATCTTAAGGAAGAGAGCGCGTCCAGTTTTGTGCGT
CTGCTGGATGCCGCGGGTCTGTCAATGGCTCTGAACAAATCGGTGGTGAACAACGATCCG
CAAGGGTATCCGGATCGCGTTCGTCAGCGTCGCGCGACTGGCATTTGGGTTTATGAACGT
TATCCTGCTGCAGACGGCGCGCAACCGCCGTACACCATCGACCCAAATACAGGGGAAGTG
ACCTGGAAATACCAGCAAGACAACAAGCCTGATGACAAGCCGAAACTGGAAGTTGCGAGC
TGGCAGGAGGAAGTTGAGGGCAAACAGGTAACGCGTTATGCCTTTATTGATGAAGCGGAA
TACACAACAGAAGAATCTCTGGAAGCGGCAAAGGCAAAAATCTTTGAGAAGTTTCCTGGG
TTACAGGAGTGTAAGGACTCGACTTACCATTACGAGATTAACTGTTTGGAGCGCCGCCCA
GGCACGGATGTTCCGGTAACAGGTGGCATGTATGTTCCGCGCTATACGCAACTGAATCTT
GACGCCGACACCGCGAAAGCGATGGTGCAGGCGGCGGATTTAGGCACCAACATTCAGCGC
CTGTATCAGCATGAGCTTTATTTCCGTACCAAAGGCAGTAAAGGTGAGCGTCTGAACAGT
GTTGATCTGGAACGTCTGTACCAGAACATGTCGGTCTGGCTGTGGAACGATACGAAATAT
CGTTACGAAGAGGGCAAGGAAGATGAGCTGGGCTTTAAAACGTTCACCGAGTTCCTGAAC
TGCTACGCCAATGATGCCTATGCAGGCGGCACCAAGTGCTCCGCAGATCTGAAAAAATCG
CTGGTCGATAACAACATGATCTACGGTGACGGTAGCAGCAAAGCGGGCATGATGAACCCA
AGCTATCCGCTCAACTATATGGAAAAACCGCTGACGCGTCTGATGCTGGGCCGTTCCTGG
TGGGATCTGAACATTAAGGTTGATGTGGAGAAGTACCCAGGATCCGTATCGGCAAAGGGT
GAGAGCGTTACGGAAAACATCAGCCTGTACTCGAATCCGACCAAATGGTTTGCGGGTAAC
ATGCAGTCAACCGGCCTGTGGGCACCGGCCCAGCAGGACGTCACCATTAAGTCTTCGGCG
TCAGTCCCAGTGACTGTTACCGTGGCGCTGGCTGACGACCTGACTGGACGTGAGAAGCAT
GAAGTTGCGCTGAACCGTCCGCCAAGAGTGACTAAAACGTATACTCTGGAGGCTAACGGT
GAAGTGACCTTCAAGGTGCCTTATGGTGGTCTGATTTATATCAAGGGCGACAGTAAGGAT
GATGTTTCTGCTAACTTCACCTTTACCGGTGTAGTAAAAGCGCCGTTCTATAAAGACGGC
GAATGGAAAAACGATCTGGACTCACCGGCGCCGCTGGGCGAGCTGGAGTCTGCGTCGTTC
GTCTATACCACGCCGAAGAAGAACCTTGAGGCCAGCAATTTCACTGGTGGTGTAGCAGAA
TTCGCTAAAGATCTGGATACCTTTGCCAGCTCGATGAATGACTTCTACGGTCGTAATGAT
GAAGACGGTAAGCACCGGATGTTTACCTATAAAAACTTGACGGGGCACAAGCATCGTTTC
ACCAACGATGTGCAGATCTCCATCGGTGATGCGCACTCGGGTTATCCGGTAATGAACAGC
AGCTTCTCGACGAACAGCACCACGCTGCCGACGACGCCGCTGAACGACTGGCTGATTTGG
CACGCAGTCGGTCATAACGCTGCAGAAACACCGCTGAACGTACCGGGTGCAACTGAAGTG
GCGAACAACGTGCTGGCGCTGTACATGCAGGATCGCTATCTCGGTAAGATGAACCGTGTC
GCTGACGACATTACCGTCGCGCCGGAATATCTGGACGAGAGCAACGGTCAGGCCTGGGCG
CGCGGCGGTGCGGGTGACCGTCTGCTGATGTACGCACAGTTGAAGGAGTGGGCAGAGGAA
AACTTTGATATCAAACAGTGGTATCCAGATGGTGAGCTGCCTAAGTTCTACAGCGATCGT
AAAGGGATGAAGGGCTGGAACCTGTTCCAGTTGATGCACCGTAAAGCGCGCGGCGATGAT
GTTGGTAACAGCACCTTTGGTGGCAAGAATTACTGTGCTGAATCCAATGGTAACGCTGCC
GACACGCTGATGCTGTGTGCATCCTGGGTCGCTCAGGCGGATCTTTCGGAATTCTTTAAG
AAATGGAATCCGGGTGCAAGTGCTTACCAGTTGCCGGGAGCAACGGAGATGAGTTTCCAG
GGCGGTGTGAGCTCTTCGGCTTACAGCACGCTGGCGTCACTCAAGCTGCCGAAACCGGAA
AAAGGGCCGGAAACCATTAACAAGGTTACCGAGCATAAGATGTCTGCCGAG

>pK1-3526E1305A (SEQ ID 20)

CDGGGSGSSSDTPSVDSGSGTLPEVKPDPTPTPEPTPEPTPDPEPTPDPTPDPEPTPEPE
PEPVPTKTGYLTLGGSQRVTGATCNGESSDGFTFTPGNTVSCVVGSTTIATFNTQSEAAR
SLRAVDKVSFSLEDAQELANSENKKTNAISLVTSSDSCPADAEQLCLTFSSVVDRARFEK
LYKQIDLATDNFSKLVNEEVENNAATDKAPSTHTSTVVPVTTEGTKPDLNASFVSANAEQ
FYQYQPTEIILSEGQLVDSLGNGVAGVDYYTNSGRGVTDENGKFSFSWGETISFGIDTFE
LGSVRGNKSTIALTELGDEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINE
IINLSLSNGATLDEGDQNVVLPNEFIEQFKTGQAKEIDTAICAKTDGCNEARWFSLTTRN
VNDGQIQGVINKLWGVDTNYQSVSKFHVFHDSTNFYGSTGNARGQAVVNISNSAFPILMA

RNDKNYWLAFGEKRAWDKNELAYITEAPSIVQPENVTRDTATFNLPFISLGQVGEGKLMV
IGNPHYNSILRCPNGYSWGGGVNSKGECTLSGDSDDMKHFMQNVLRYLSNDIWQPNTKSI
MTVGTNLENVYFKKAGQVLGNSAPFAFHEDFTGITVKQLTSYGDLNPEEIPLLILNGFEY
VTQWSGDPYAVPLRADTSKPKLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASSFVR
LLDAAGLSMALNKSVVNNDPQGYPDRVRQRRATGIWVYERYPAADGAQPPYTIDPNTGEV
TWKYQQDNKPDDKPKLEVASWQEEVEGKQVTRYAFIDEAEYTTEESLEAAKAKIFEKFPG
LQECKDSTYHYEINCLERRPGTDVPVTGGMYVPRYTQLNLDADTAKAMVQAADLGTNIQR
LYQHELYFRTKGSKGERLNSVDLERLYQNMSVWLWNDTKYRYEEGKEDELGFKTFTEFLN
CYANDAYAGGTKCSADLKKSLVDNNMIYGDGSSKAGMMNPSYPLNYMEKPLTRLMLGRSW
WDLNIKVDVEKYPGSVSAKGESVTENISLYSNPTKWFAGNMQSTGLWAPAQQDVTIKSSA
SVPVTVTVALADDLTGREKHEVALNRPPRVTKTYTLEANGEVTFKVPYGGLIYIKGDSKD
DVSANFTFTGVVKAPFYKDGEWKNDLDSPAPLGELESASFVYTTPKKNLEASNFTGGVAE
FAKDLDTFASSMNDFYGRNDEDGKHRMFTYKNLTGHKHRFTNDVQISIGDAHSGYPVMNS
SFSTNSTTLPTTPLNDWLIWHAVGHNAAETPLNVPGATEVANNVLALYMQDRYLGKMNRV
ADDITVAPEYLDESNGQAWARGGAGDRLLMYAQLKEWAEENFDIKQWYPDGELPKFYSDR
KGMKGWNLFQLMHRKARGDDVGNSTFGGKNYCAESNGNAADTLMLCASWVAQADLSEFFK
KWNPGASAYQLPGATEMSFQGGVSSSAYSTLASLKLPKPEKGPETINKVTEHKMSAE

>pK1-3526D1422A (SEQ ID 21)

TGTGATGGTGGTGGTTCAGGATCGTCCTCCGATACGCCGTCTGTAGATTCTGGATCAGGG
ACTTTGCCGGAAGTGAAACCCGATCCAACACCAACCCCGGAGCCGACACCTGAGCCGACG
CCGGACCCAGAACCTACGCCGGATCCAACACCTGATCCTGAGCCGACACCAGAACCGGAG
CCAGAACCTGTTCCTACGAAAACGGGTTATCTGACCCTGGGCGGAAGCCAGCGGGTAACT
GGTGCTACCTGTAATGGTGAATCCAGCGATGGCTTTACCTTTACGCCAGGCAATACCGTG
AGTTGTGTGGTGGGCAGTACGACCATTGCAACATTCAACACCCAGTCAGAAGCTGCGCGT
AGCCTGCGTGCGGTTGACAAAGTGTCGTTTAGCCTGGAGGACGCGCAGGAGCTGGCGAAT
TCTGAAAATAAGAAAACCAACGCCATCTCTCTGGTGACGTCCAGCGACAGTTGCCCCGCA
GATGCAGAACAGCTTTGTCTTACTTTCTCGTCAGTGGTTGATCGCGCGCGATTTGAAAAA
CTGTATAAGCAAATTGATCTGGCAACAGACAATTTCAGCAAGCTGGTCAATGAAGAGGTG
GAAAACAATGCTGCGACTGATAAAGCGCCGTCCACCCATACCTCAACGGTAGTGCCAGTC
ACGACAGAGGGAACAAAACCGGATCTGAACGCGTCCTTCGTGTCGGCTAACGCGGAACAG
TTTTATCAGTATCAACCCACTGAAATCATTCTTTCCGAAGGCCAACTGGTGGATAGCCTG
GGGAACGGTGTTGCTGGCGTTGACTACTACACCAATTCAGGCCGTGGCGTAACTGACGAA
AACGGTAAATTTTCCTTTAGCTGGGGCGAAACCATCTCCTTTGGTATCGATACCTTTGAA
CTGGGCTCAGTACGTGGCAATAAGTCGACCATTGCGCTGACTGAATTGGGTGATGAAGTT
CGCGGGGCAAATATCGATCAGCTCATTCATCGTTATTCGACGACTGGTCAAAATAATACT
CGTGTTGTTCCGGACGATGTACGCAAGGTCTTTGCCGAATATCCCAACGTGATCAACGAG
ATAATCAATCTTTCGTTATCCAACGGTGCGACGCTGGATGAAGGCGATCAAAACGTTGTG
CTGCCTAACGAATTTATCGAGCAGTTTAAGACGGGTCAGGCCAAAGAGATCGATACCGCG
ATTTGTGCGAAAACCGACGGTTGTAACGAGGCTCGCTGGTTCTCGCTGACAACGCGCAAT
GTTAATGACGGCCAGATTCAGGGCGTTATTAACAAGCTGTGGGGCGTGGATACGAACTAT
CAGTCTGTCAGCAAGTTCCACGTCTTCCATGACTCTACCAACTTCTATGGCAGCACCGGT
AACGCGCGCGGTCAGGCGGTGGTAAATATCTCCAACTCGGCATTCCCGATTCTGATGGCG
CGTAATGATAAAAACTACTGGCTGGCGTTTGGCGAAAAACGCGCCTGGGATAAAAATGAG
CTGGCGTACATTACGGAAGCGCCTTCCATTGTGCAGCCAGAGAACGTTACGCGCGATACT
GCGACTTTCAACCTGCCGTTTATTTCGCTGGGGCAAGTCGGTGAAGGCAAACTGATGGTT
ATCGGTAACCCGCACTACAACAGCATCCTGCGTTGCCCGAACGGTTACAGTTGGGGCGGT
GGTGTTAATAGTAAAGGTGAGTGTACGCTCAGCGGTGATTCTGATGACATGAAGCACTTT
ATGCAGAACGTACTGCGCTACTTGTCAAATGACATCTGGCAGCCAAATACCAAGAGCATC
ATGACTGTCGGCACCAACCTGGAGAACGTTTATTTCAAAAAAGCGGGCCAGGTATTGGGA
AATAGTGCACCATTTGCTTTCCATGAGGATTTCACTGGTATCACGGTTAAACAGTTGACC
AGCTATGGCGATCTGAATCCGGAAGAGATTCCGTTGCTGATCCTCAACGGCTTTGAATAT
GTGACTCAGTGGTCTGGCGATCCCTATGCTGTGCCTCTGCGTGCAGATACCAGCAAACCG
AAGCTGACTCAGCAGGATGTGACCGATCTGATCGCTTATCTGAACAAAGGTGGCTCGGTG
CTGATCATGGAAAACGTGATGAGCAATCTTAAGGAAGAGAGCGCGTCCAGTTTTGTGCGT

CTGCTGGATGCCGCGGGTCTGTCAATGGCTCTGAACAAATCGGTGGTGAACAACGATCCG
CAAGGGTATCCGGATCGCGTTCGTCAGCGTCGCGCGACTGGCATTTGGGTTTATGAACGT
TATCCTGCTGCAGACGGCGCGCAACCGCCGTACACCATCGACCCAAATACAGGGGAAGTG
ACCTGGAAATACCAGCAAGACAACAAGCCTGATGACAAGCCGAAACTGGAAGTTGCGAGC
TGGCAGGAGGAAGTTGAGGGCAAACAGGTAACGCGTTATGCCTTTATTGATGAAGCGGAA
TACACAACAGAAGAATCTCTGGAAGCGGCAAAGGCAAAAATCTTTGAGAAGTTTCCTGGG
TTACAGGAGTGTAAGGACTCGACTTACCATTACGAGATTAACTGTTTGGAGCGCCGCCCA
GGCACGGATGTTCCGGTAACAGGTGGCATGTATGTTCCGCGCTATACGCAACTGAATCTT
GACGCCGACACCGCGAAAGCGATGGTGCAGGCGGCGGATTTAGGCACCAACATTCAGCGC
CTGTATCAGCATGAGCTTTATTTCCGTACCAAAGGCAGTAAAGGTGAGCGTCTGAACAGT
GTTGATCTGGAACGTCTGTACCAGAACATGTCGGTCTGGCTGTGGAACGATACGAAATAT
CGTTACGAAGAGGGCAAGGAAGATGAGCTGGGCTTTAAAACGTTCACCGAGTTCCTGAAC
TGCTACGCCAATGATGCCTATGCAGGCGGCACCAAGTGCTCCGCAGATCTGAAAAAATCG
CTGGTCGATAACAACATGATCTACGGTGACGGTAGCAGCAAAGCGGGCATGATGAACCCA
AGCTATCCGCTCAACTATATGGAAAAACCGCTGACGCGTCTGATGCTGGGCCGTTCCTGG
TGGGATCTGAACATTAAGGTTGATGTGGAGAAGTACCCAGGATCCGTATCGGCAAAGGGT
GAGAGCGTTACGGAAAACATCAGCCTGTACTCGAATCCGACCAAATGGTTTGCGGGTAAC
ATGCAGTCAACCGGCCTGTGGGCACCGGCCCAGCAGGACGTCACCATTAAGTCTTCGGCG
TCAGTCCCAGTGACTGTTACCGTGGCGCTGGCTGACGACCTGACTGGACGTGAGAAGCAT
GAAGTTGCGCTGAACCGTCCGCCAAGAGTGACTAAAACGTATACTCTGGAGGCTAACGGT
GAAGTGACCTTCAAGGTGCCTTATGGTGGTCTGATTTATATCAAGGGCGACAGTAAGGAT
GATGTTTCTGCTAACTTCACCTTTACCGGTGTAGTAAAAGCGCCGTTCTATAAAGACGGC
GAATGGAAAAACGATCTGGACTCACCGGCGCCGCTGGGCGAGCTGGAGTCTGCGTCGTTC
GTCTATACCACGCCGAAGAAGAACCTTGAGGCCAGCAATTTCACTGGTGGTGTAGCAGAA
TTCGCTAAAGATCTGGATACCTTTGCCAGCTCGATGAATGACTTCTACGGTCGTAATGAT
GAAGACGGTAAGCACCGGATGTTTACCTATAAAAACTTGACGGGGCACAAGCATCGTTTC
ACCAACGATGTGCAGATCTCCATCGGTGATGCGCACTCGGGTTATCCGGTAATGAACAGC
AGCTTCTCGACGAACAGCACCACGCTGCCGACGACGCCGCTGAACGACTGGCTGATTTGG
CACGAAGTCGGTCATAACGCTGCAGAAACACCGCTGAACGTACCGGGTGCAACTGAAGTG
GCGAACAACGTGCTGGCGCTGTACATGCAGGATCGCTATCTCGGTAAGATGAACCGTGTC
GCTGACGACATTACCGTCGCGCCGGAATATCTGGACGAGAGCAACGGTCAGGCCTGGGCG
CGCGGCGGTGCGGGTGACCGTCTGCTGATGTACGCACAGTTGAAGGAGTGGGCAGAGGAA
AACTTTGATATCAAACAGTGGTATCCAGATGGTGAGCTGCCTAAGTTCTACAGCGATCGT
AAAGGGATGAAGGGCTGGAACCTGTTCCAGTTGATGCACCGTAAAGCGCGCGGCGCTGAT
GTTGGTAACAGCACCTTTGGTGGCAAGAATTACTGTGCTGAATCCAATGGTAACGCTGCC
GACACGCTGATGCTGTGTGCATCCTGGGTCGCTCAGGCGGATCTTTCGGAATTCTTTAAG
AAATGGAATCCGGGTGCAAGTGCTTACCAGTTGCCGGGAGCAACGGAGATGAGTTTCCAG
GGCGGTGTGAGCTCTTCGGCTTACAGCACGCTGGCGTCACTCAAGCTGCCGAAACCGGAA
AAAGGGCCGGAAACCATTAACAAGGTTACCGAGCATAAGATGTCTGCCGAG

>pK1-3526D1422A (SEQ ID 22)

CDGGGSGSSSDTPSVDSGSGTLPEVKPDPTPTPEPTPEPTPDPEPTPDPTPDPEPTPEPE
PEPVPTKTGYLTLGGSQRVTGATCNGESSDGFTFTPGNTVSCVVGSTTIATFNTQSEAAR
SLRAVDKVSFSLEDAQELANSENKKTNAISLVTSSDSCPADAEQLCLTFSSVVDRARFEK
LYKQIDLATDNFSKLVNEEVENNAATDKAPSTHTSTVVPVTTEGTKPDLNASFVSANAEQ
FYQYQPTEIILSEGQLVDSLGNGVAGVDYYTNSGRGVTDENGKFSFSWGETISFGIDTFE
LGSVRGNKSTIALTELGDEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINE
IINLSLSNGATLDEGDQNVVLPNEFIEQFKTGQAKEIDTAICAKTDGCNEARWFSLTTRN
VNDGQIQGVINKLWGVDTNYQSVSKFHVFHDSTNFYGSTGNARGQAVVNISNSAFPILMA
RNDKNYWLAFGEKRAWDKNELAYITEAPSIVQPENVTRDTATFNLPFISLGQVGEGKLMV
IGNPHYNSILRCPNGYSWGGGVNSKGECTLSGDSDDMKHFMQNVLRYLSNDIWQPNTKSI
MTVGTNLENVYFKKAGQVLGNSAPFAFHEDFTGITVKQLTSYGDLNPEEIPLLILNGFEY
VTQWSGDPYAVPLRADTSKPKLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASSFVR
LLDAAGLSMALNKSVVNNDPQGYPDRVRQRRATGIWVYERYPAADGAQPPYTIDPNTGEV
TWKYQQDNKPDDKPKLEVASWQEEVEGKQVTRYAFIDEAEYTTEESLEAAKAKIFEKFPG

LQECKDSTYHYEINCLERRPGTDVPVTGGMYVPRYTQLNLDADTAKAMVQAADLGTNIQR
LYQHELYFRTKGSKGERLNSVDLERLYQNMSVWLWNDTKYRYEEGKEDELGFKTFTEFLN
CYANDAYAGGTKCSADLKKSLVDNNMIYGDGSSKAGMMNPSYPLNYMEKPLTRLMLGRSW
WDLNIKVDVEKYPGSVSAKGESVTENISLYSNPTKWFAGNMQSTGLWAPAQQDVTIKSSA
SVPVTVTVALADDLTGREKHEVALNRPPRVTKTYTLEANGEVTFKVPYGGLIYIKGDSKD
DVSANFTFTGVVKAPFYKDGEWKNDLDSPAPLGELESASFVYTTPKKNLEASNFTGGVAE
FAKDLDTFASSMNDFYGRNDEDGKHRMFTYKNLTGHKHRFTNDVQISIGDAHSGYPVMNS
SFSTNSTTLPTTPLNDWLIWHEVGHNAAETPLNVPGATEVANNVLALYMQDRYLGKMNRV
ADDITVAPEYLDESNGQAWARGGAGDRLLMYAQLKEWAEENFDIKQWYPDGELPKFYSDR
KGMKGWNLFQLMHRKARGADVGNSTFGGKNYCAESNGNAADTLMLCASWVAQADLSEFFK
KWNPGASAYQLPGATEMSFQGGVSSSAYSTLASLKLPKPEKGPETINKVTEHKMSAE

>pK1-3526AdG (SEQ ID 23)

GATACGCCGTCTGTAGATTCTGGATCAGGGACTTTGCCGGAAGTGAAACCCGATCCAACA
CCAACCCCGGAGCCGACACCTGAGCCGACGCCGGACCCAGAACCTACGCCGGATCCAACA
CCTGATCCTGAGCCGACACCAGAACCGGAGCCAGAACCTGTTCCTACGAAAACGGGTTAT
CTGACCCTGGGCGGAAGCCAGCGGGTAACTGGTGCTACCTGTAATGGTGAATCCAGCGAT
GGCTTTACCTTTACGCCAGGCAATACCGTGAGTTGTGTGGTGGGCAGTACGACCATTGCA
ACATTCAACACCCAGTCAGAAGCTGCGCGTAGCCTGCGTGCGGTTGACAAAGTGTCGTTT
AGCCTGGAGGACGCGCAGGAGCTGGCGAATTCTGAAAATAAGAAAACCAACGCCATCTCT
CTGGTGACGTCCAGCGACAGTTGCCCCGCAGATGCAGAACAGCTTTGTCTTACTTTCTCG
TCAGTGGTTGATCGCGCGCGATTTGAAAAACTGTATAAGCAAATTGATCTGGCAACAGAC
AATTTCAGCAAGCTGGTCAATGAAGAGGTGGAAAACAATGCTGCGACTGATAAAGCGCCG
TCCACCCATACCTCAACGGTAGTGCCAGTCACGACAGAGGGAACAAAACCGGATCTGAAC
GCGTCCTTCGTGTCGGCTAACGCGGAACAGTTTTATCAGTATCAACCCACTGAAATCATT
CTTTCCGAAGGCCAACTGGTGGATAGCCTGGGGAACGGTGTTGCTGGCGTTGACTACTAC
ACCAATTCAGGCCGTGGCGTAACTGACGAAAACGGTAAATTTTCCTTTAGCTGGGGCGAA
ACCATCTCCTTTGGTATCGATACCTTTGAACTGGGCTCAGTACGTGGCAATAAGTCGACC
ATTGCGCTGACTGAATTGGGTGATGAAGTTCGCGGGGCAAATATCGATCAGCTCATTCAT
CGTTATTCGACGACTGGTCAAAATAATACTCGTGTTGTTCCGGACGATGTACGCAAGGTC
TTTGCCGAATATCCCAACGTGATCAACGAGATAATCAATCTTTCGTTATCCAACGGTGCG
ACGCTGGATGAAGGCGATCAAAACGTTGTGCTGCCTAACGAATTTATCGAGCAGTTTAAG
ACGGGTCAGGCCAAAGAGATCGATACCGCGATTTGTGCGAAAACCGACGGTTGTAACGAG
GCTCGCTGGTTCTCGCTGACAACGCGCAATGTTAATGACGGCCAGATTCAGGGCGTTATT
AACAAGCTGTGGGGCGTGGATACGAACTATCAGTCTGTCAGCAAGTTCCACGTCTTCCAT
GACTCTACCAACTTCTATGGCAGCACCGGTAACGCGCGCGGTCAGGCGGTGGTAAATATC
TCCAACTCGGCATTCCCGATTCTGATGGCGCGTAATGATAAAAACTACTGGCTGGCGTTT
GGCGAAAAACGCGCCTGGGATAAAAATGAGCTGGCGTACATTACGGAAGCGCCTTCCATT
GTGCAGCCAGAGAACGTTACGCGCGATACTGCGACTTTCAACCTGCCGTTTATTTCGCTG
GGGCAAGTCGGTGAAGGCAAACTGATGGTTATCGGTAACCCGCACTACAACAGCATCCTG
CGTTGCCCGAACGGTTACAGTTGGGGCGGTGGTGTTAATAGTAAAGGTGAGTGTACGCTC
AGCGGTGATTCTGATGACATGAAGCACTTTATGCAGAACGTACTGCGCTACTTGTCAAAT
GACATCTGGCAGCCAAATACCAAGAGCATCATGACTGTCGGCACCAACCTGGAGAACGTT
TATTTCAAAAAAGCGGGCCAGGTATTGGGAAATAGTGCACCATTTGCTTTCCATGAGGAT
TTCACTGGTATCACGGTTAAACAGTTGACCAGCTATGGCGATCTGAATCCGGAAGAGATT
CCGTTGCTGATCCTCAACGGCTTTGAATATGTGACTCAGTGGTCTGGCGATCCCTATGCT
GTGCCTCTGCGTGCAGATACCAGCAAACCGAAGCTGACTCAGCAGGATGTGACCGATCTG
ATCGCTTATCTGAACAAAGGTGGCTCGGTGCTGATCATGGAAAACGTGATGAGCAATCTT
AAGGAAGAGAGCGCGTCCAGTTTTGTGCGTCTGCTGGATGCCGCGGGTCTGTCAATGGCT
CTGAACAAATCGGTGGTGAACAAC

>pK1-3526AdG (SEQ ID 24)

DTPSVDSGSGTLPEVKPDPTPTPEPTPEPTPDPEPTPDPTPDPEPTPEPEPEPVPTKTGY
LTLGGSQRVTGATCNGESSDGFTFTPGNTVSCVVGSTTIATFNTQSEAARSLRAVDKVSF

SLEDAQELANSENKKTNAISLVTSSDSCPADAEQLCLTFSSVVDRARFEKLYKQIDLATD
NFSKLVNEEVENNAATDKAPSTHTSTVVPVTTEGTKPDLNASFVSANAEQFYQYQPTEII
LSEGQLVDSLGNGVAGVDYYTNSGRGVTDENGKFSFSWGETISFGIDTFELGSVRGNKST
IALTELGDEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINEIINLSLSNGA
TLDEGDQNVVLPNEFIEQFKTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVI
NKLWGVDTNYQSVSKFHVFHDSTNFYGSTGNARGQAVVNISNSAFPILMARNDKNYWLAF
GEKRAWDKNELAYITEAPSIVQPENVTRDTATFNLPFISLGQVGEGKLMVIGNPHYNSIL
RCPNGYSWGGGVNSKGECTLSGDSDDMKHFMQNVLRYLSNDIWQPNTKSIMTVGTNLENV
YFKKAGQVLGNSAPFAFHEDFTGITVKQLTSYGDLNPEEIPLLILNGFEYVTQWSGDPYA
VPLRADTSKPKLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASSFVRLLDAAGLSMA
LNKSVVNN

>pK1-3526CdG (SEQ ID 25)

GATACGCCGTCTGTAGATTCTGGATCAGGGACTTTGCCGGAAGTGAAACCCGATCCAACA
CCAACCCCGGAGCCGACACCTGAGCCGACGCCGGACCCAGAACCTACGCCGGATCCAACA
CCTGATCCTGAGCCGACACCAGAACCGGAGCCAGAACCTGTTCCTACGAAAACGGGTTAT
CTGACCCTGGGCGGAAGCCAGCGGGTAACTGGTGCTACCTGTAATGGTGAATCCAGCGAT
GGCTTTACCTTTACGCCAGGCAATACCGTGAGTTGTGTGGTGGGCAGTACGACCATTGCA
ACATTCAACACCCAGTCAGAAGCTGCGCGTAGCCTGCGTGCGGTTGACAAAGTGTCGTTT
AGCCTGGAGGACGCGCAGGAGCTGGCGAATTCTGAAAATAAGAAAACCAACGCCATCTCT
CTGGTGACGTCCAGCGACAGTTGCCCCGCAGATGCAGAACAGCTTTGTCTTACTTTCTCG
TCAGTGGTTGATCGCGCGCGATTTGAAAAACTGTATAAGCAAATTGATCTGGCAACAGAC
AATTTCAGCAAGCTGGTCAATGAAGAGGTGGAAAACAATGCTGCGACTGATAAAGCGCCG
TCCACCCATACCTCAACGGTAGTGCCAGTCACGACAGAGGGAACAAAACCGGATCTGAAC
GCGTCCTTCGTGTCGGCTAACGCGGAACAGTTTTATCAGTATCAACCCACTGAAATCATT
CTTTCCGAAGGCCAACTGGTGGATAGCCTGGGGAACGGTGTTGCTGGCGTTGACTACTAC
ACCAATTCAGGCCGTGGCGTAACTGACGAAAACGGTAAATTTTCCTTTAGCTGGGGCGAA
ACCATCTCCTTTGGTATCGATACCTTTGAACTGGGCTCAGTACGTGGCAATAAGTCGACC
ATTGCGCTGACTGAATTGGGTGATGAAGTTCGCGGGGCAAATATCGATCAGCTCATTCAT
CGTTATTCGACGACTGGTCAAAATAATACTCGTGTTGTTCCGGACGATGTACGCAAGGTC
TTTGCCGAATATCCCAACGTGATCAACGAGATAATCAATCTTTCGTTATCCAACGGTGCG
ACGCTGGATGAAGGCGATCAAAACGTTGTGCTGCCTAACGAATTTATCGAGCAGTTTAAG
ACGGGTCAGGCCAAAGAGATCGATACCGCGATTTGTGCGAAAACCGACGGTTGTAACGAG
GCTCGCTGGTTCTCGCTGACAACGCGCAATGTTAATGACGGCCAGATTCAGGGCGTTATT
AACAAGCTGTGGGGCGTGGATACGAACTATCAGTCTGTCAGCAAGTTCCACGTCTTCCAT
GACTCTACCAACTTCTATGGCAGCACCGGTAACGCGCGCGGTCAGGCGGTGGTAAATATC
TCCAACTCGGCATTCCCGATTCTGATGGCGCGTAATGATAAAAACTACTGGCTGGCGTTT
GGCGAAAAACGCGCCTGGGATAAAAATGAGCTGGCGTACATTACGGAAGCGCCTTCCATT
GTGCAGCCAGAGAACGTTACGCGCGATACTGCGACTTTCAACCTGCCGTTTATTTCGCTG
GGGCAAGTCGGTGAAGGCAAACTGATGGTTATCGGTAACCCGCACTACAACAGCATCCTG
CGTTGCCCGAACGGTTACAGTTGGGGCGGTGGTGTTAATAGTAAAGGTGAGTGTACGCTC
AGCGGTGATTCTGATGACATGAAGCACTTTATGCAGAACGTACTGCGCTACTTGTCAAAT
GACATCTGGCAGCCAAATACCAAGAGCATCATGACTGTCGGCACCAACCTGGAGAACGTT
TATTTCAAAAAAGCGGGCCAGGTATTGGGAAATAGTGCACCATTTGCTTTCCATGAGGAT
TTCACTGGTATCACGGTTAAACAGTTGACCAGCTATGGCGATCTGAATCCGGAAGAGATT
CCGTTGCTGATCCTCAACGGCTTTGAATATGTGACTCAGTGGTCTGGCGATCCCTATGCT
GTGCCTCTGCGTGCAGATACCAGCAAACCGAAGCTGACTCAGCAGGATGTGACCGATCTG
ATCGCTTATCTGAACAAAGGTGGCTCGGTGCTGATCATGGAAAACGTGATGAGCAATCTT
AAGGAAGAGAGCGCGTCCAGTTTTGTGCGTCTGCTGGATGCCGCGGGTCTGTCAATGGCT
CTGAACAAATCGGTGGTGAACAACGATCCGCAAGGGTATCCGGATCGCGTTCGTCAGCGT
CGCGCGACTGGCATTTGGGTTTATGAACGTTATCCTGCTGCAGACGGCGCGCAACCGCCG
TACACCATCGACCCAAATACAGGGGAAGTGACCTGGAAATACCAGCAAGACAACAAGCCT
GATGACAAGCCGAAACTGGAAGTTGCGAGCTGGCAGGAGGAAGTTGAGGGCAAACAGGTA
ACGCGTTATGCCTTTATTGATGAAGCGGAATACACAACAGAAGAATCTCTGGAAGCGGCA
AAGGCAAAAATCTTTGAGAAGTTTCCTGGGTTACAGGAGTGTAAGGACTCGACTTACCAT

TACGAGATTAACTGTTTGGAGCGCCGCCCAGGCACGGATGTTCCGGTAACAGGTGGCATG
TATGTTCCGCGCTATACGCAACTGAATCTTGACGCCGACACCGCGAAAGCGATGGTGCAG
GCGGCGGATTTAGGCACCAACATTCAGCGCCTGTATCAGCATGAGCTTTATTTCCGTACC
AAAGGCAGTAAAGGTGAGCGTCTGAACAGTGTTGATCTGGAACGTCTGTACCAGAACATG
TCGGTCTGGCTGTGGAACGATACGAAATATCGTTACGAAGAGGGCAAGGAAGATGAGCTG
GGCTTTAAAACGTTCACCGAGTTCCTGAACTGCTACGCCAATGATGCCTATGCAGGCGGC
ACCAAGTGCTCCGCAGATCTGAAAAAATCGCTGGTCGATAACAACATGATCTACGGTGAC
GGTAGCAGCAAAGCGGGCATGATGAACCCAAGCTATCCGCTCAACTATATGGAAAAACCG
CTGACGCGTCTGATGCTGGGCCGTTCCTGGTGGGATCTGAACATTAAGGTTGATGTGGAG
AAGTACCCAGGATCCGTATCGGCAAAGGGTGAGAGCGTTACGGAAAACATCAGCCTGTAC
TCGAATCCGACCAAATGGTTTGCGGGTAACATGCAGTCAACCGGCCTGTGGGCACCGGCC
CAGCAGGACGTCACCATTAAGTCTTCGGCGTCAGTCCCAGTGACTGTTACCGTGGCGCTG
GCTGACGACCTGACTGGACGTGAGAAGCATGAAGTTGCGCTGAACCGTCCGCCAAGAGTG
ACTAAAACGTATACTCTGGAGGCTAACGGTGAAGTGACCTTCAAGGTGCCTTATGGTGGT
CTGATTTATATCAAGGGCGACAGTAAGGATGATGTTTCTGCTAACTTCACCTTTACCGGT
GTAGTAAAAGCGCCGTTCTATAAAGACGGCGAATGGAAAAACGATCTGGACTCACCGGCG
CCGCTGGGCGAGCTGGAGTCTGCGTCGTTCGTCTATACCACGCCGAAGAAGAACCTTGAG
GCCAGCAATTTCACTGGTGGTGTAGCAGAATTCGCTAAAGATCTGGATACCTTTGCCAGC
TCGATGAATGACTTCTACGGTCGTAATGATGAAGACGGTAAGCACCGGATGTTTACCTAT
AAAAACTTGACGGGGCACAAGCATCGTTTCACCAACGATGTGCAGATCTCCATCGGTGAT
GCGCACTCGGGTTATCCGGTAATGAACAGCAGCTTCTCGACGAACAGCACCACGCTGCCG
ACGACGCCGCTGAACGACTGGCTGATTTGG

pK1-3526CdG (SEQ ID 26)

DTPSVDSGSGTLPEVKPDPTPTPEPTPEPTPDPEPTPDPTPDPEPTPEPEPEPVPTKTGY
LTLGGSQRVTGATCNGESSDGFTFTPGNTVSCVVGSTTIATFNTQSEAARSLRAVDKVSF
SLEDAQELANSENKKTNAISLVTSSDSCPADAEQLCLTFSSVVDRARFEKLYKQIDLATD
NFSKLVNEEVENNAATDKAPSTHTSTVVPVTTEGTKPDLNASFVSANAEQFYQYQPTEII
LSEGQLVDSLGNGVAGVDYYTNSGRGVTDENGKFSFSWGETISFGIDTFELGSVRGNKST
IALTELGDEVRGANIDQLIHRYSTTGQNNTRVVPDDVRKVFAEYPNVINEIINLSLSNGA
TLDEGDQNVVLPNEFIEQFKTGQAKEIDTAICAKTDGCNEARWFSLTTRNVNDGQIQGVI
NKLWGVDTNYQSVSKFHVFHDSTNFYGSTGNARGQAVVNISNSAFPILMARNDKNYWLAF
GEKRAWDKNELAYITEAPSIVQPENVTRDTATFNLPFISLGQVGEGKLMVIGNPHYNSIL
RCPNGYSWGGGVNSKGECTLSGDSDDMKHFMQNVLRYLSNDIWQPNTKSIMTVGTNLENV
YFKKAGQVLGNSAPFAFHEDFTGITVKQLTSYGDLNPEEIPLLILNGFEYVTQWSGDPYA
VPLRADTSKPKLTQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASSFVRLLDAAGLSMA
LNKSVVNNDPQGYPDRVRQRRATGIWVYERYPAADGAQPPYTIDPNTGEVTWKYQQDNKP
DDKPKLEVASWQEEVEGKQVTRYAFIDEAEYTTEESLEAAKAKIFEKFPGLQECKDSTYH
YEINCLERRPGTDVPVTGGMYVPRYTQLNLDADTAKAMVQAADLGTNIQRLYQHELYFRT
KGSKGERLNSVDLERLYQNMSVWLWNDTKYRYEEGKEDELGFKTFTEFLNCYANDAYAGG
TKCSADLKKSLVDNNMIYGDGSSKAGMMNPSYPLNYMEKPLTRLMLGRSWWDLNIKVDVE
KYPGSVSAKGESVTENISLYSNPTKWFAGNMQSTGLWAPAQQDVTIKSSASVPVTVTVAL
ADDLTGREKHEVALNRPPRVTKTYTLEANGEVTFKVPYGGLIYIKGDSKDDVSANFTFTG
VVKAPFYKDGEWKNDLDSPAPLGELESASFVYTTPKKNLEASNFTGGVAEFAKDLDTFAS
SMNDFYGRNDEDGKHRMFTYKNLTGHKHRFTNDVQISIGDAHSGYPVMNSSFSTNSTTLP
TTPLNDWLIW

>pK1-L3526-2stop (SEQ ID 27)

ATGAATAAGAAATTTAAATATAAGAAATCGCTTTTAGCGGCTATTTTAAGCGCAACCCTG
TTAGCCGGTTGTGATGGTGGTGGTTCAGGATCGTCCTCCGATACGCCGTCTGTAGATTCT
GGATCAGGGACTTTGCCGGAAGTGAAACCCGATCCAACACCAACCCCGGAGCCGACACCT
GAGCCGACGCCGGACCCAGAACCTACGCCGGATCCAACACCTGATCCTGAGCCGACACCA
GAACCGGAGCCAGAACCTGTTCCTACGAAAACGGGTTATCTGACCCTGGGCGGAAGCCAG
CGGGTAACTGGTGCTACCTGTAATGGTGAATCCAGCGATGGCTTTACCTTTACGCCAGGC

AATACCGTGAGTTGTGTGGTGGGCAGTACGACCATTGCAACATTCAACACCCAGTCAGAA
GCTGCGCGTAGCCTGCGTGCGGTTGACAAAGTGTCGTTTAGCCTGGAGGACGCGCAGGAG
CTGGCGAATTCTGAAAATAAGAAAACCAACGCCATCTCTCTGGTGACGTCCAGCGACAGT
TGCCCCGCAGATGCAGAACAGCTTTGTCTTACTTTCTCGTCAGTGGTTGATCGCGCGCGA
TTTGAAAAACTGTATAAGCAAATTGATCTGGCAACAGACAATTTCAGCAAGCTGGTCAAT
GAAGAGGTGGAAAACAATGCTGCGACTGATAAAGCGCCGTCCACCCATACCTCAACGGTA
GTGCCAGTCACGACAGAGGGAACAAAACCGGATCTGAACGCGTCCTTCGTGTCGGCTAAC
GCGGAACAGTTTTATCAGTATCAACCCACTGAAATCATTCTTTCCGAAGGCCAACTGGTG
GATAGCCTGGGGAACGGTGTTGCTGGCGTTGACTACTACACCAATTCAGGCCGTGGCGTA
ACTGACGAAAACGGTAAATTTTCCTTTAGCTGGGGCGAAACCATCTCCTTTGGTATCGAT
ACCTTTGAACTGGGCTCAGTACGTGGCAATAAGTCGACCATTGCGCTGACTGAATTGGGT
GATGAAGTTCGCGGGGCAAATATCGATCAGCTCATTCATCGTTATTCGACGACTGGTCAA
AATAATACTCGTGTTGTTCCGGACGATGTACGCAAGGTCTTTGCCGAATATCCCAACGTG
ATCAACGAGATAATCAATCTTTCGTTATCCAACGGTGCGACGCTGGATGAAGGCGATCAA
AACGTTGTGCTGCCTAACGAATTTATCGAGCAGTTTAAGACGGGTCAGGCCAAAGAGATC
GATACCGCGATTTGTGCGAAAACCGACGGTTGTAACGAGGCTCGCTGGTTCTCGCTGACA
ACGCGCAATGTTAATGACGGCCAGATTCAGGGCGTTATTAACAAGCTGTGGGGCGTGGAT
ACGAACTATCAGTCTGTCAGCAAGTTCCACGTCTTCCATGACTCTACCAACTTCTATGGC
AGCACCGGTAACGCGCGCGGTCAGGCGGTGGTAAATATCTCCAACTCGGCATTCCCGATT
CTGATGGCGCGTAATGATAAAAACTACTGGCTGGCGTTTGGCGAAAAACGCGCCTGGGAT
AAAAATGAGCTGGCGTACATTACGGAAGCGCCTTCCATTGTGCAGCCAGAGAACGTTACG
CGCGATACTGCGACTTTCAACCTGCCGTTTATTTCGCTGGGGCAAGTCGGTGAAGGCAAA
CTGATGGTTATCGGTAACCCGCACTACAACAGCATCCTGCGTTGCCCGAACGGTTACAGT
TGGGGCGGTGGTGTTAATAGTAAAGGTGAGTGTACGCTCAGCGGTGATTCTGATGACATG
AAGCACTTTATGCAGAACGTACTGCGCTACTTGTCAAATGACATCTGGCAGCCAAATACC
AAGAGCATCATGACTGTCGGCACCAACCTGGAGAACGTTTATTTCAAAAAAGCGGGCCAG
GTATTGGGAAATAGTGCACCATTTGCTTTCCATGAGGATTTCACTGGTATCACGGTTAAA
CAGTTGACCAGCTATGGCGATCTGAATCCGGAAGAGATTCCGTTGCTGATCCTCAACGGC
TTTGAATATGTGACTCAGTGGTCTGGCGATCCCTATGCTGTGCCTCTGCGTGCAGATACC
AGCAAACCGAAGCTGACTCAGCAGGATGTGACCGATCTGATCGCTTATCTGAACAAAGGT
GGCTCGGTGCTGATCATGGAAAACGTGATGAGCAATCTTAAGGAAGAGAGCGCGTCCAGT
TTTGTGCGTCTGCTGGATGCCGCGGGTCTGTCAATGGCTCTGAACAAATCGGTGGTGAAC
AACGATCCGCAAGGGTATCCGGATCGCGTTCGTCAGCGTCGCGCGACTGGCATTTGGGTT
TATGAACGTTATCCTGCTGCAGACGGCGCGCAACCGCCGTACACCATCGACCCAAATACA
GGGGAAGTGACCTGGAAATACCAGCAAGACAACAAGCCTGATGACAAGCCGAAACTGGAA
GTTGCGAGCTGGCAGGAGGAAGTTGAGGGCAAACAGGTAACGCGTTATGCCTTTATTGAT
GAAGCGGAATACACAACAGAAGAATCTCTGGAAGCGGCAAAGGCAAAAATCTTTGAGAAG
TTTCCTGGGTTACAGGAGTGTAAGGACTCGACTTACCATTACGAGATTAACTGTTTGGAG
CGCCGCCCAGGCACGGATGTTCCGGTAACAGGTGGCATGTATGTTCCGCGCTATACGCAA
CTGAATCTTGACGCCGACACCGCGAAAGCGATGGTGCAGGCGGCGGATTTAGGCACCAAC
ATTCAGCGCCTGTATCAGCATGAGCTTTATTTCCGTACCAAAGGCAGTAAAGGTGAGCGT
CTGAACAGTGTTGATCTGGAACGTCTGTACCAGAACATGTCGGTCTGGCTGTGGAACGAT
ACGAAATATCGTTACGAAGAGGGCAAGGAAGATGAGCTGGGCTTTAAAACGTTCACCGAG
TTCCTGAACTGCTACGCCAATGATGCCTATGCAGGCGGCACCAAGTGCTCCGCAGATCTG
AAAAAATCGCTGGTCGATAACAACATGATCTACGGTGACGGTAGCAGCAAAGCGGGCATG
ATGAACCCAAGCTATCCGCTCAACTATATGGAAAAACCGCTGACGCGTCTGATGCTGGGC
CGTTCCTGGTGGGATCTGAACATTAAGGTTGATGTGGAGAAGTACCCAGGATCCGTATCG
GCAAAGGGTGAGAGCGTTACGGAAAACATCAGCCTGTACTCGAATCCGACCAAATGGTTT
GCGGGTAACATGCAGTCAACCGGCCTGTGGGCACCGGCCCAGCAGGACGTCACCATTAAG
TCTTCGGCGTCAGTCCCAGTGACTGTTACCGTGGCGCTGGCTGACGACCTGACTGGACGT
GAGAAGCATGAAGTTGCGCTGAACCGTCCGCCAAGAGTGACTAAAACGTATACTCTGGAG
GCTAACGGTGAAGTGACCTTCAAGGTGCCTTATGGTGGTCTGATTTATATCAAGGGCGAC
AGTAAGGATGATGTTTCTGCTAACTTCACCTTTACCGGTGTAGTAAAAGCGCCGTTCTAT
AAAGACGGCGAATGGAAAAACGATCTGGACTCACCGGCGCCGCTGGGCGAGCTGGAGTCT
GCGTCGTTCGTCTATACCACGCCGAAGAAGAACCTTGAGGCCAGCAATTTCACTGGTGGT
GTAGCAGAATTCGCTAAAGATCTGGATACCTTTGCCAGCTCGATGAATGACTTCTACGGT

CGTAATGATGAAGACGGTAAGCACCGGATGTTTACCTATAAAAACTTGACGGGGCACAAG
CATCGTTTCACCAACGATGTGCAGATCTCCATCGGTGATGCGCACTCGGGTTATCCGGTA
ATGAACAGCAGCTTCTCGACGAACAGCACCACGCTGCCGACGACGCCGCTGAACGACTGG
CTGATTTGGCACGAAGTCGGTCATAACGCTGCAGAAACACCGCTGAACGTACCGGGTGCA
ACTGAAGTGGCGAACAACGTGCTGGCGCTGTACATGCAGGATCGCTATCTCGGTAAGATG
AACCGTGTCGCTGACGACATTACCGTCGCGCCGGAATATCTGGACGAGAGCAACGGTCAG
GCCTGGGCGCGCGGCGGTGCGGGTGACCGTCTGCTGATGTACGCACAGTTGAAGGAGTGG
GCAGAGGAAAACTTTGATATCAAACAGTGGTATCCAGATGGTGAGCTGCCTAAGTTCTAC
AGCGATCGTAAAGGGATGAAGGGCTGGAACCTGTTCCAGTTGATGCACCGTAAAGCGCGC
GGCGATGATGTTGGTAACAGCACCTTTGGTGGCAAGAATTACTGTGCTGAATCCAATGGT
AACGCTGCCGACACGCTGATGCTGTGTGCATCCTGGGTCGCTCAGGCGGATCTTTCGGAA
TTCTTTAAGAAATGGAATCCGGGTGCAAGTGCTTACCAGTTGCCGGGAGCAACGGAGATG
AGTTTCCAGGGCGGTGTGAGCTCTTCGGCTTACAGCACGCTGGCGTCACTCAAGCTGCCG
AAACCGGAAAAAGGGCCGGAAACCATTAACAAGGTTACCGAGCATAAGATGTCTGCCGAG
TAA

>pK1-L3526-2stop (SEQ ID 28)

MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPSVDSGSGTLPEVKPDPTPTPEPTP
EPTPDPEPTPDPTPDPEPTPEPEPEPVPTKTGYLTLGGSQRVTGATCNGESSDGFTFTPG
NTVSCVVGSTTIATFNTQSEAARSLRAVDKVSFSLEDAQELANSENKKTNAISLVTSSDS
CPADAEQLCLTFSSVVDRARFEKLYKQIDLATDNFSKLVNEEVENNAATDKAPSTHTSTV
VPVTTEGTKPDLNASFVSANAEQFYQYQPTEIILSEGQLVDSLGNGVAGVDYYTNSGRGV
TDENGKFSFSWGETISFGIDTFELGSVRGNKSTIALTELGDEVRGANIDQLIHRYSTTGQ
NNTRVVPDDVRKVFAEYPNVINEIINLSLSNGATLDEGDQNVVLPNEFIEQFKTGQAKEI
DTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYQSVSKFHVFHDSTNFYG
STGNARGQAVVNISNSAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSIVQPENVT
RDTATFNLPFISLGQVGEGKLMVIGNPHYNSILRCPNGYSWGGGVNSKGECTLSGDSDDM
KHFMQNVLRYLSNDIWQPNTKSIMTVGTNLENVYFKKAGQVLGNSAPFAFHEDFTGITVK
QLTSYGDLNPEEIPLLILNGFEYVTQWSGDPYAVPLRADTSKPKLTQQDVTDLIAYLNKG
GSVLIMENVMSNLKEESASSFVRLLDAAGLSMALNKSVVNNDPQGYPDRVRQRRATGIWV
YERYPAADGAQPPYTIDPNTGEVTWKYQQDNKPDDKPKLEVASWQEEVEGKQVTRYAFID
EAEYTTEESLEAAKAKIFEKFPGLQECKDSTYHYEINCLERRPGTDVPVTGGMYVPRYTQ
LNLDADTAKAMVQAADLGTNIQRLYQHELYFRTKGSKGERLNSVDLERLYQNMSVWLWND
TKYRYEEGKEDELGFKTFTEFLNCYANDAYAGGTKCSADLKKSLVDNNMIYGDGSSKAGM
MNPSYPLNYMEKPLTRLMLGRSWWDLNIKVDVEKYPGSVSAKGESVTENISLYSNPTKWF
AGNMQSTGLWAPAQQDVTIKSSASVPVTVTVALADDLTGREKHEVALNRPPRVTKTYTLE
ANGEVTFKVPYGGLIYIKGDSKDDVSANFTFTGVVKAPFYKDGEWKNDLDSPAPLGELES
ASFVYTTPKKNLEASNFTGGVAEFAKDLDTFASSMNDFYGRNDEDGKHRMFTYKNLTGHK
HRFTNDVQISIGDAHSGYPVMNSSFSTNSTTLPTTPLNDWLIWHEVGHNAAETPLNVPGA
TEVANNVLALYMQDRYLGKMNRVADDITVAPEYLDESNGQAWARGGAGDRLLMYAQLKEW
AEENFDIKQWYPDGELPKFYSDRKGMKGWNLFQLMHRKARGDDVGNSTFGGKNYCAESNG
NAADTLMLCASWVAQADLSEFFKKWNPGASAYQLPGATEMSFQGGVSSSAYSTLASLKLP
KPEKGPETINKVTEHKMSAE*

Preferred orf3526 sequences of the invention match the consensus sequence as recited in SEQ ID NO:54, or are immunogenic fragments thereof. Other preferred orf3526 sequences of the invention match the consensus sequence as recited in SEQ ID NO:55, or are immunogenic fragments thereof. X represents any amino acid. Other preferred orf3526 sequences further contain a sequence motif at those positions that correspond to positions 1304-1308 of SEQ ID NO:8, selected from: XEVGH, XXVGH, XEVGX, HXVGX and XXVGX, wherein in any such sequence motif X is not H or E; or X is not H, E or D; or X

is not H, E, D, N, Q or C; or X is a non-polar amino acid, or X is selected from A or G, or X is preferably A. In further embodiments, residues 1-23 or 1-33 of SEQ ID NO:54 or SEQ ID NO:55 are lacking.

>consensus_sequence_75% (SEQ ID NO:54)

MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPXVDSGXGXLPEVKPDPTPXPEPTPEPTPD
PEPTPXPTPDPEPTPEPEPEPVPTKTGYLTLGGSQRVTGATCNGESSDGFTFTPGXXVXCVXGXXT
TIATFBTQSEAARSLRAVXKVSFSLEDAQELAXSXBKKXNAXSLVTSXBSCPABXEQXCLXFSSV
XXXXRFXXLYKQIDLAXXXFXKLVNEEVENNAATDKAPSTHTSXVVPVTTXGTKPDLNASFVS
ANAEQFYQYQPTEIILSEGXLVDSXGXGVXGVBYYTXSGRGVTXENGXFSFSWGETISFGIDTFE
LGSVRGNKSTIALTELGDEVRGANIDQLIHRYSXXGXNBTRVVPDXVRKVFAEYPNVINEIINLSL
SNGATLXEGXQXVXLPNEFIEQFXTGQAKEIDTAICAKTXGCNEARWFSLTTRNVNDGQIQGVIN
KLWGVDXBYXSVXKFHVFHDSTNFYGSTGNARGQAVVNISNAAFPILMARNDKNYWLAFGEK
RAWDKNELAYITEAPSJVZPENVTRDTATFNLPFISLGQVGXGKLMVIGNPHYNSILRCPNGYSW
XGGVNXXGZCTLXXDXDDMKXFMZNVLRYLSBDXWXPBXKXXMTVGTNLXXVYFKXHGQV
XGNSAXFXFHXDFXGIXVXXLXSYGDLBPZEXPLLILNGFEYVTQXGXDPYAXPLRADTSKPKL
TQQDVTDLIAYLNKGGSVLIMENVMSNLKEESASXFVRLLDAAGLSMALNKSVVNNDPQGYPB
RVRQXRATGIWVYERYPAXDGXXXPYTIDXXTGEVXWKYQXXNKPDDKPKLEVASWXEXVX
GKQXTRYAFIDEAXXXTXXSLXAAKXKIXXXFPGLXECKDXXYHYEXNCLEXRPGTXVPVTGG
MYVPXYTQLXLXADTAKAMVQAADLGTNIQRLYQHELYFRTNGXKGERLXSVDLERLYQNMS
VWLWNXXXYRYEXXKXDELGFKTFTEFLNCYANDAYXXGTXCSAXLKKSLVDNXMIYGXXS
XKAGMMNPSYPLNYMEKPLTRLMLGRSWWDLNIKVDVEKYPGXVSXXGZXVTEXISLYSNPT
KWFAGNMQSTGLWAPAQXEVTIXSXAXVPVTVTVALADDLTGREKHEVALNRPPXVTKTYXL
XAXGXVXFKVPYGGLIYIKGBSXXBXSAXFTFTGVVKAPFYKDGXWKNXLBSPAPLGELESXXF
VYTTPKKNLXASNXTGGXXZFAXDLDTFASSMNDFYGRBXEXGKHRMFTYKXLTGHKHRFTN
DVQISIGDAHSGYPVMNSSFSTNSTTLPTTPLNDWLIW**HEVGH**NAAETPLXVPGATEVANNVLA
LYMQDRYLGKMNRVADDITVAPEYLXESNGQAWARGGAGDRLLMYAQLKEWAEKNFDIKXW
YPXGXLPXFYSXREGMKGWNLFQLMHRKARGDXVGXXXFGXXNYCAESNGNAADTLMLCAS
WVAQTDLSEFFKKWNPGANAYQLPGAXEMSFEGGVSQSAYXTLAXLXLPKPZXGPETINXVTE
HKMSAE

>consensus_sequence_100% (SEQ ID NO:55)

MNKKFKYKKSLLAAILSATLLAGCDGGGSGXSSDTPXXDSGXGXLPXVKPDPTPXXXXXXXXXX
XXXXXXXXXXXXXXXXXXPXXXPEXXXXPVXTKTGYLXLGGSXRXTGXXXCNXEXSDGFTF
XXGXXVXCVXGXXTTIATFBTQSEAARSLRAVXKVSFSLEDAXELAXSXBKKXNAXSLVTXXB
SCPABXEQXCLXFSSVXXXXRFXXLYKQIXLXXXXFXKLVNEEVENNAATDKAPSTHTSXXVP
XTTXGTXPDLNASFVSANAEQXYQYQPXEIIXSEGXLVXSXGXGVXGVBYYTXXGRGVTXENG
XFXFSWGEXJSFGIDTFELGSVRGNKSTIALTELGDEVRGANIDQLIHRXSXXXXNXXRXVPXXV
RXVFAXYPNVINEIINLSLSNGXXLXEGXQXXXXXNXFJEQXXXGQXXEIDTAICXXXXGCNXX
RWFSLTXRNVNXGXIQXVINKLWGVDXXYXSVXKFHVFHDSTNFYGSTGNARGQAVXNISNX
AFPILMARNDKNYWLAFGEKRAWDKNXLAYITEAPSJVXXENVTRXTAXFNLPFISLGQVGXGK
LMVIGNPHYNSILRCPNGYSWXGXVBXXGZCTXXXDXBDMKXFMZNVLRYLSBXXWXPBXK
XXMTVGTNLXXVYFKXXGQVXGXXAXFXFHXDFXGIXXXXXXSYGBLBPXXXPLLILNGFEY
VTQXXXDPYXXPLRADTSKPKLXQQDVTDLIAYXNKGGXVLIMENVMSNLKEESASXFVRLLD
AAXLSMALNKSVVNXDPQGYPBRXRQXRXXXIWVYERYPXXXXXXXPYTIBXXTXXVXWKY
QXXXKXDXKPKLEVASWXEXVXGXQXXXXAFIDXAXXXTXXXLXAAKXXIXXXFXGLXXCX
BXXYHYEXNCLEXRXGXXVPXTXXXXXGMXVPXYTQLXLXADTAKAMXQAADXGTNIQRLY
QHELYFRTXGXXGERLXXVDLERLYQNXSVWLWNXXXYXYXXXKXDELGFKTFTEFLNCYXN
BAYXXGTXCSXXLKXSLXDNXMIYGXXXXXKAGMMNPXYPLNYMEKPLTRLMLGRSWWDL
NIKVDVEXYPGXVXXXGZXVTZXIXLYSXPTKWFAGNMQSTGLXAPAXXXVXIXSXXXVXV
TVTVAXADDLTGREKHEVXLNRPPXVTKTYXLXAXXXVXFXVPYGGLIYIKXBSXXXXXSAXF
TFXGVVKAPFYKBGXWXXXXXSPAPLGELESXXFVYTXPKXNLXAXXXSNXXXGXXZFAXXL
DTFAXSMNDFXGRBXXXGXHXMFTXXXLXGHKHRFXNDVQISIGDAHSGYPVMNSSFSXBSXT
LPTXPLNDWLIW**HEXGH**NAAETPLXVPGATEVANXVLALYMQDRYLGKMNRVADDITVAPEY

LXESNXQAWARGGAGDRLLMYAQLKEWAEXNFDIXXWYPXGXXLPXFXSXRXGMKGWNLF
QLMHRKAXGDXVXXXXFGXXNYCAESNGNXADXLMLCASWVAQXDLSXFFKKWNPGAXAY
QLPGXXEMXFXXGVXXSAYXTLAXXXLXKPXXGPEXXNXVTEXXMXXE

gspK (orf3515)

gspK general secretion pathway protein is referred to herein as 'orf3515.' 'orf3515' protein from *E. coli* NMEC is disclosed in WO2006/089264 (SEQ IDs 7029 & 7030) is also known as: 'orf3332' from *E. coli* NMEC strain IHE3034, 'c3702' from CFT073 and ecp_3039 from 536.

When used according to the present invention, orf3515 protein may take various forms. Preferred orf3515 sequences have 80% or more identity (*e.g.* 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more) to SEQ ID NOs 29 and 30. This includes variants (*e.g.* allelic variants, homologs, orthologs, paralogs, mutants *etc*).

>orf03515 (SEQ ID 29)

ATGATCACCTCACCACCAAAACGCGGAATGGCACTGGTCGTGGTGCTGGTATTGCTGGCG
GTTATGATGCTGGTGACCATCACGCTTTCCGGGCGGATGCAGCAACAACTTGGGCGAACG
CGCAGCCAGCAGGAGTACCAGCAGGCGCTGTGGTACAGCGCCAGTGCAGAAAGCCTGGCG
CTGAGCGCGCTCAGTCTGAGCCTGAAAAATGAAAAGCGTGTGCATCTGGCACAACCGTGG
GCTTCTGGCCCGCGTTTTTTCCCACTGCCGCAGGGGCAAATTGCCGTCACTCTGCGTGAC
GCACAGGCCTGCTTTAACCTGAATGCCCTCGCTCAGCCGACGACGGCGTCGCGTCCGCTC
GCGGTACAACAACTGATTGCCCTGATCTCGCGCCTCGATGTGCCTGCTTATCGGGCCGAA
CTGATAGCCGAAAGCCTGTGGGAGTTTATTGACGAAGACCGCAGCGTGCAGACGCGTCTG
GGTCGTGAAGACAGCGAGTATCTCGCCCGCTCGGTGCCGTTCTACGCCGCTAATCAACCG
CTGGCTGATATCAGCGAGATGCGCGTGGTGCAGGGAATGGACGCCGGGCTTTATCAAAAA
CTGAAACCGTTGGTCTGTGCGCTGCCGATGGCCCGCCAGCAAATCAACATCAATACATTA
GATGTCACGCAAAGTGTGATTCTTGAGGCGCTGTTTGACCCGTGGTTAAGCCCTGTTCAG
GCGCGGGCATTATTACAACAACGTCCGGCGAAGGGCTGGGAAGATGTCGATCAGTTTCTT
GCTCAGCCGCTACTTGCAGACGTCGATGAGCGTACTAAAAAACAGCTAAAAACCATCCTG
AGCGTGGACAGCAATTACTTCTGGCTGCGTTCAGATATCACCGTGAATGAGATTGAACTG
ACGATGAATTCGTTAATTGTCCGCATGGGCCCACAACACTTTTCTGTTCTCTGGCATCAG
ACAGGAGAAAGTGAG

>orf03515 (SEQ ID 30)

MITSPPKRGMALVVVLVLLAVMMLVTITLSGRMQQQLGRTRSQQEYQQALWYSASAESLA
LSALSLSLKNEKRVHLAQPWASGPRFFPLPQGQIAVTLRDAQACFNLNALAQPTTASRPL
AVQQLIALISRLDVPAYRAELIAESLWEFIDEDRSVQTRLGREDSEYLARSVPFYAANQP
LADISEMRVVQGMDAGLYQKLKPLVCALPMARQQININTLDVTQSVILEALFDPWLSPVQ
ARALLQQRPAKGWEDVDQFLAQPLLADVDERTKKQLKTILSVDSNYFWLRSDITVNEIEL
TMNSLIVRMGPQHFSVLWHQTGESE

Particular compositions of the invention will comprise a combination of (i) bacterial Ig-like domain protein (orf405B) having the amino acid sequence set forth in SEQ ID NO:2 or a protein having at least 80% similarity thereto, and (ii) putative Lipoprotein (orf3526) having the amino acid sequence set forth in SEQ ID NO:8 or a protein having at least 80% similarity thereto.

Other particular compositions of the invention will further comprise (iii) upec1232 having the amino acid sequence set forth in SEQ ID NO:4 or a protein having at least 80% similarity thereto.

More particularly the immunogenic components of a composition of the invention will consist essentially of (i) bacterial Ig-like domain protein (orf405B) having the amino acid sequence set forth in SEQ ID NO:2 or a protein having at least 80% similarity thereto, and (ii) putative Lipoprotein (orf3526) having the amino acid sequence set forth in SEQ ID NO:8 or a protein having at least 80% similarity thereto. The composition can additionally include non-immunogenic components.

Other particular immunogenic components of a composition of the invention will consist essentially of (i) bacterial Ig-like domain protein (orf405B) having the amino acid sequence set forth in SEQ ID NO:2 or a protein having at least 80% similarity thereto, and (ii) putative Lipoprotein (orf3526) having the amino acid sequence set forth in SEQ ID NO:8 or a protein having at least 80% similarity thereto, and (iii) upec1232 having the amino acid sequence set forth in SEQ ID NO:4 or a protein having at least 80% similarity thereto. The composition can additionally include non-immunogenic components.

Particularly, the compositions of the invention may further comprise at least one bacterial toxin. Particularly, the toxin is derived from *E. coli* (i.e., E. coli heat labile enterotoxin "LT"), cholera ("CT"), or pertussis ("PT").

The use of detoxified ADP-ribosylating toxins as mucosal adjuvants is described in WO95/17211 and as parenteral adjuvants in WO98/42375.

Particular detoxified LT mutants include LT-K63, LT-R72, and LTR192G. Preferably, the bacterial toxin will be a mutant or modified bacterial toxin. In a preferred embodiment the bacterial toxin is the modified heat-labile toxin of *Escherichia coli* (LTK63).

The use of ADP-ribosylating toxins and detoxified derivaties thereof, particularly LT-K63 and LT-R72, as adjuvants can be found in the following references, each of which is specifically incorporated by reference herein in their entirety: Beignon, et al., "The LTR72 Mutant of Heat-Labile Enterotoxin of Escherichia coli Enahnces the Ability of Peptide Antigens to Elicit CD4 T Cells and Secrete Gamma Interferon after Coapplication onto Bare Skin", Infection and Immunity (2002) 70(6):3012-3019; Pizza, et al., "Mucosal vaccines: non toxic derivatives of LT and CT as mucosal adjuvants", Vaccine (2001) 19:2534-2541; Pizza, et al., "LTK63 and LTR72, two mucosal adjuvants ready for clinical trials" Int. J. Med. Microbiol (2000) 290(4-5):455- 461; Scharton-Kersten et al., "Transcutaneous Imrnunization with Bacterial ADP-Ribosylating Exotoxins, Subunits and Unrelated Adjuvants", Infection and Immunity (2000) 68(9):5306-5313; Ryan et al., "Mutants of Escherichia coli Heat-Labile Toxin Act as Effective Mucosal Adjuvants for Nasal Delivery of an Acellular Pertussis Vaccine: Differential Effects of the

Nontoxic AB Complex and Enzyme Activity on Th1 and Th2 Cells" Infection and Immunity (1999) 67(12):6270-6280; Partidos et al., "Heat-labile enterotoxin of Escherichia coli and its site-directed mutant LTK63 enhance the proliferative and cytotoxic T-cell responses to intranasally co-immunized synthetic peptides", Immunol. Lett. (1999) 67(3):209-216; Peppoloni et al., "Mutants of the Escherichia coli heat-labile enterotoxin as safe and strong adjuvants for intranasal delivery of vaccines", Vaccines (2003) 2(2):285-293; and Pine et al., (2002) "Intranasal immunization with influenza vaccine and a detoxified mutant of heat labile enterotoxin from Escherichia coli (LTK63)" J. Control Release (2002) 85(1-3):263-270. Numerical reference for amino acid substitutions is preferably based on the alignments of the A and B subunits of ADP-ribosylating toxins set forth in Domenighini et al., Mol. Microbiol (1995) 15(6):1165-1167, specifically incorporated herein by reference in its entirety.

Thus, in the context of the invention, the word “toxin” is intended to mean toxins that have been detoxified such that they are no longer toxic to humans, or a toxin subunit or fragment thereof that is substantially devoid of toxic activity in humans.

Other detoxified toxins include the B subunit from *E. coli* labile toxin (LT), the amino terminal domain of the anthrax lethal factor (LF), *P. aeruginosa* exotoxin A, adenylate cyclase A from B. Pertussis, a derived or mutant from a toxin which is a family of the AB5 family, for example, the cholera toxin (CT), the Bordatella Pertussis toxin (PT) as well as the recently identified subtilase cytotoxins. (Paton et al, J Exp Med 2004, Vol 200 pp 35-46).

The labile toxin (LT) of *E. coli* consists of two subunits, a pentameric B subunit and a monomeric A subunit. The A subunit is responsible for toxicity, whilst the B subunit is responsible for transport into the cell. LT binds the G M1 ganglioside receptor.

A derivative of *E. coli* heat-labile toxin with equal or greater the 90% homology has greater than 90% homology at the amino acid level. In another embodiment the protein has equal or greater than 95% homology, for example 96, 97, 98 or 99%. For example, amino acid deletions may be made that do not affect function. In a further embodiment, a derivative is still able to bind the G M1 ganglioside receptor.

Thus, particular compositions of the invention include combinations of at least two, at least three, at least four or five *E. coli* antigens selected from the group consisting of orf405B, upec1232, orf3526, orf3515 and LTK63. Particular compositions of the invention include no more than two, no more than three, no more than four or no more than five antigens selected from the group consisting of orf405B, upec1232, orf3526, orf3515 and LTK63. Yet more particularly, compositions of the invention consist of, or consist essentially of, a combination of two, three, four or five antigens selected from the group consisting of orf405B, upec1232, orf3526, orf3515 and LTK63. Particular combinations include the following *E. coli* antigen(s)/immunogenic components:

Orf405B+orf3526

Orf405B+upec1232+orf3526

Orf405B+upec1232+orf3526+orf3515

Orf405B+orf3526+LTK63

Orf405B+upec1232+orf3526+LTK63

Orf405B+upec1232+orf3526+orf3515+LTK63

Antigen orf3526 comprises a zinc binding motif which encompasses amino acids at positions 1304-1308 (HEVGH underlined in SEQ ID 8) with reference to SEQ ID 8. Since this zinc binding motif may be associated with toxicity, orf3526 polypeptides which lack or have reduced zinc binding activity are particularly useful in combinations of the present invention. Preferably, zinc binding activity of a mutant orf3526 protein is either reduced by or reduced to at least 50%, at least 45%, at least 40%, at least 35%, at least 30%, at least 25%, at least 20%, at least 15%, at least 10% or at least 5% relative to or compared to wild-type orf3526. Zinc binding can be determined by atomic absorption and other assays will be known to one skilled in the art. Thus, mutations in the zinc binding motif are useful in reducing zinc binding and associated toxicity. For example, mutations in the zinc binding motif from wild-type HEVGH to AEVGH can reduce zinc binding to about 43% or more particularly, mutations from wild-type to AAVGA can reduce zinc binding to around 5%. Surprisingly an orf3526 mutant which comprises the AAVGA sequence has the added advantage that it co-elutes with native orf3526 and is present in only two isoforms (a monomer and truncated form) meaning that the efficiency of purification is simplified and significantly improved in comparison with other mutants tested. Compositions of the invention may comprise orf3526 mutants comprising a sequence motif selected from XEVGH, XXVGH, XEVGX, HXVGX or XXVGX, wherein in any such sequence motif X is not H or E; or X is not H, E or D; or X is not H, E, D, N, Q or C; or X is a non-polar amino acid, or X is selected from A or G, or X is preferably A.

Advantageously, vaccine combinations of the present invention may be used in combination with a Group B Streptococcus vaccine to prevent most cases of neonatal meningitis. Thus, in certain embodiments, the combinations of the invention may include: (i) one or more further, non *E. coli*, polypeptides that elicit antibody responses against Group B Streptococcal (GBS) proteins; (ii) a capsular saccharide from Group B Streptococcus; and/or (iii) one or more further immunogens that elicit antibody responses that recognise epitopes on non-GBS organisms. In other embodiments, the immunogenic combinations of the present invention are administered separately at substantially the same time as a GBS vaccine.

Particular GBS polypeptides include: 'GBS80' (SAG0645) a cell wall surface anchor family protein (see GI: 22533660); 'GBS1523' (SAN1518; SpbI), a cell wall surface anchor family protein (see GI: 77408651); 'GBS104' (SAG0649) (see GI: 22533664); 'GBS67' (SAG1408), a cell wall surface anchor family protein (see GI: 22534437); 'GBS59', a pilus backbone protein encoded by pathogenicity island 2a (BP-2a); 'GBS3' (SAG2603; BibA), a pathogenicity protein (see GI:22535109); 'SAN1485', a cell wall

surface anchor family protein ' (see GI: 77408233); 'GBS147' (SAG0416), a putative protease (see GI: GI:22533435); 'GBS328' (SAG1333) a 5'-nucleotidase family protein ' (see GI: 22534359).

Immunogenic compositions and medicaments

The Polypeptides described above are useful as active ingredients (immunogens) in immunogenic compositions of the invention, and such compositions may be useful as vaccines.
Immunogenic compositions will be pharmaceutically acceptable. They will usually include components in addition to the antigens *e.g.* they typically include one or more pharmaceutical carrier(s), excipient(s) and/or adjuvant(s). Thorough discussions of vaccine adjuvants are available in refs. 6 and 7.

Compositions will generally be administered to a mammal in aqueous form. Prior to administration, however, the composition may have been in a non-aqueous form. For instance, although some vaccines are manufactured in aqueous form, then filled and distributed and administered also in aqueous form, other vaccines are lyophilised during manufacture and are reconstituted into an aqueous form at the time of use. Thus a composition of the invention may be dried, such as a lyophilised formulation.

The composition may include preservatives such as thiomersal or 2-phenoxyethanol. It is preferred, however, that the vaccine should be substantially free from (*i.e.* less than 5μg/ml) mercurial material *e.g.* thiomersal-free. Vaccines containing no mercury are more preferred. Preservative-free vaccines are particularly preferred.

To improve thermal stability, a composition may include a temperature protective agent.
To control tonicity, it is preferred to include a physiological salt, such as a sodium salt. Sodium chloride (NaCl) is preferred, which may be present at between 1 and 20 mg/ml *e.g.* about 10±2mg/ml NaCl. Other salts that may be present include potassium chloride, potassium dihydrogen phosphate, disodium phosphate dehydrate, magnesium chloride, calcium chloride, *etc.*

Compositions will generally have an osmolality of between 200 mOsm/kg and 400 mOsm/kg, preferably between 240-360 mOsm/kg, and will more preferably fall within the range of 290-310 mOsm/kg.

Compositions may include one or more buffers. Typical buffers include: a phosphate buffer; a Tris buffer; a borate buffer; a succinate buffer; a histidine buffer (particularly with an aluminum hydroxide adjuvant); or a citrate buffer. Buffers will typically be included in the 5-20mM range.

The pH of a composition will generally be between 5.0 and 8.1, and more typically between 6.0 and 8.0 *e.g.* 6.5 and 7.5, or between 7.0 and 7.8.

The composition is preferably sterile. The composition is preferably non-pyrogenic *e.g.* containing <1 EU (endotoxin unit, a standard measure) per dose, and preferably <0.1 EU per dose. The composition is preferably gluten free.

The composition may include material for a single immunisation, or may include material for multiple immunisations (*i.e.* a 'multidose' kit). The inclusion of a preservative is preferred in multidose arrangements. As an alternative (or in addition) to including a preservative in multidose compositions, the compositions may be contained in a container having an aseptic adaptor for removal of material.

Human vaccines are typically administered in a dosage volume of about 0.5ml, although a half dose (*i.e.* about 0.25ml) may be administered to children.

In certain embodiments the vaccine composition will comprise one or more pharmaceutically acceptable carriers, diluents and/or adjuvants. Adjuvants which may be used in compositions of the invention include, but are not limited to:

- mineral salts, such as aluminium salts and calcium salts, including hydroxides (*e.g.* oxyhydroxides), phosphates (*e.g.* hydroxyphosphates, orthophosphates) and sulphates, *etc.* [*e.g.* see chapters 8 & 9 of ref. 8];
- oil-in-water emulsions, such as squalene-water emulsions, including MF59 (5% Squalene, 0.5% Tween 80, and 0.5% Span 85, formulated into submicron particles using a microfluidizer) [Chapter 10 of ref. 6, see also ref. 9-12, chapter 10 of ref. 13 and chapter 12 of ref. 14], complete Freund's adjuvant (CFA) and incomplete Freund's adjuvant (IFA);
- saponin formulations [chapter 22 of ref. 6], such as QS21 [15] and ISCOMs [chapter 23 of ref. 6];
- virosomes and virus-like particles (VLPs) [16-22];
- bacterial or microbial derivatives, such as non-toxic derivatives of enterobacterial lipopolysaccharide (LPS), Lipid A derivatives [23, 24], immunostimulatory oligonucleotides [25-30], such as IC-31™ [31] (deoxynucleotide comprising 26-mer sequence 5'-$(IC)_{13}$-3' (SEQ ID NO:56) and polycationic polymer peptide comprising 11-mer amino acid sequence KLKLLLLLKLK (SEQ ID NO:57)) and ADP-ribosylating toxins and detoxified derivatives thereof [32 - 41];
- human immunomodulators, including cytokines, such as interleukins (*e.g.* IL-1, IL-2, IL-4, IL-5, IL-6, IL-7, IL-12 [42, 43], interferons (*e.g.* interferon-γ), macrophage colony stimulating factor, and tumor necrosis factor;
- bioadhesives and mucoadhesives, such as chitosan and derivatives thereof, esterified hyaluronic acid microspheres [44] or mucoadhesives, such as cross-linked derivatives of poly(acrylic acid), polyvinyl alcohol, polyvinyl pyrollidone, polysaccharides and carboxymethylcellulos [45];

- microparticles (*i.e.* a particle of ~100nm to ~150μm in diameter, more preferably ~200nm to ~30μm in diameter, and most preferably ~500nm to ~10μm in diameter) formed from materials that are biodegradable and non-toxic (*e.g.* a poly(α-hydroxy acid), a polyhydroxybutyric acid, a polyorthoester, a polyanhydride, a polycaprolactone, *etc.*);
- liposomes [Chapters 13 & 14 of ref. 6, ref. 46-48];
- polyoxyethylene ethers and polyoxyethylene esters [49];
- PCPP formulations [50 and 51];
- muramyl peptides, including N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acetyl-normuramyl-L-alanyl-D-isoglutamine (nor-MDP), and N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-*sn*-glycero-3-hydroxyphosphoryloxy)-ethylamine MTP-PE); and
- imidazoquinolone compounds, including Imiquamod and its homologues (*e.g.* "Resiquimod 3M") [52 and 53].

The invention may also comprise combinations of one or more of the adjuvants identified above. For example, the following adjuvant compositions may be used in the invention: (1) a saponin and an oil-in-water emulsion [54]; (2) a saponin (*e.g.* QS21) + a non-toxic LPS derivative (*e.g.* 3dMPL) [55]; (3) a saponin (*e.g.* QS21) + a non-toxic LPS derivative (*e.g.* 3dMPL) + a cholesterol; (4) a saponin (*e.g.* QS21) + 3dMPL + IL-12 (optionally + a sterol) [56]; (5) combinations of 3dMPL with, for example, QS21 and/or oil-in-water emulsions [57]; (6) SAF, containing 10% squalane, 0.4% Tween 80™, 5% pluronic-block polymer L121, and thr-MDP, either microfluidized into a submicron emulsion or vortexed to generate a larger particle size emulsion. (7) Ribi™ adjuvant system (RAS), (Ribi Immunochem) containing 2% squalene, 0.2% Tween 80, and one or more bacterial cell wall components from the group consisting of monophosphorylipid A (MPL), trehalose dimycolate (TDM), and cell wall skeleton (CWS), preferably MPL + CWS (Detox™); and (8) one or more mineral salts (such as an aluminum salt) + a non-toxic derivative of LPS (such as 3dMPL).

Other substances that act as immunostimulating agents are disclosed in chapter 7 of ref. 6.

The use of an aluminium hydroxide and/or aluminium phosphate adjuvant is useful, particularly in children, and antigens are generally adsorbed to these salts. Squalene-in-water emulsions are also preferred, particularly in the elderly. Useful adjuvant combinations include combinations of Th1 and Th2 adjuvants such as CpG & alum or resiquimod & alum. A combination of aluminium phosphate and 3dMPL may be used.

The compositions of the invention may elicit both a cell mediated immune response as well as a humoral immune response.

Two types of T cells, CD4 and CD8 cells, are generally thought necessary to initiate and/or enhance cell mediated immunity and humoral immunity. CD8 T cells can express a CD8 co-receptor and are commonly referred to as Cytotoxic T lymphocytes (CTLs). CD8 T cells are able to recognized or interact with antigens displayed on MHC Class I molecules.

CD4 T cells can express a CD4 co-receptor and are commonly referred to as T helper cells. CD4 T cells are able to recognize antigenic peptides bound to MHC class II molecules. Upon interaction with a MHC class II molecule, the CD4 cells can secrete factors such as cytokines. These secreted cytokines can activate B cells, cytotoxic T cells, macrophages, and other cells that participate in an immune response. Helper T cells or CD4+ cells can be further divided into two functionally distinct subsets: TH1 phenotype and TH2 phenotypes which differ in their cytokine and effector function.

Activated TH1 cells enhance cellular immunity (including an increase in antigen-specific CTL production) and are therefore of particular value in responding to intracellular infections. Activated TH1 cells may secrete one or more of IL-2, IFN-γ, and TNF-β. A TH1 immune response may result in local inflammatory reactions by activating macrophages, NK (natural killer) cells, and CD8 cytotoxic T cells (CTLs). A TH1 immune response may also act to expand the immune response by stimulating growth of B and T cells with IL-12. TH1 stimulated B cells may secrete IgG2a.

Activated TH2 cells enhance antibody production and are therefore of value in responding to extracellular infections. Activated TH2 cells may secrete one or more of IL-4, IL-5, IL-6, and IL-10. A TH2 immune response may result in the production of IgG1, IgE, IgA and memory B cells for future protection.

An enhanced immune response may include one or more of an enhanced TH1 immune response and a TH2 immune response.

A TH1 immune response may include one or more of an increase in CTLs, an increase in one or more of the cytokines associated with a TH1 immune response (such as IL-2, IFN-γ, and TNF-β), an increase in activated macrophages, an increase in NK activity, or an increase in the production of IgG2a. Preferably, the enhanced TH1 immune response will include an increase in IgG2a production.

A TH1 immune response may be elicited using a TH1 adjuvant. A TH1 adjuvant will generally elicit increased levels of IgG2a production relative to immunization of the antigen without adjuvant. TH1 adjuvants suitable for use in the invention may include for example saponin formulations, virosomes and virus like particles, non-toxic derivatives of enterobacterial lipopolysaccharide (LPS), immunostimulatory oligonucleotides. Immunostimulatory oligonucleotides, such as oligonucleotides containing a CpG motif, are preferred TH1 adjuvants for use in the invention.

A TH2 immune response may include one or more of an increase in one or more of the cytokines associated with a TH2 immune response (such as IL-4, IL-5, IL-6 and IL-10), or an increase in the production of IgG1, IgE, IgA and memory B cells. Preferably, the enhanced TH2 immune response will include an increase in IgG1 production.

A TH2 immune response may be elicited using a TH2 adjuvant. A TH2 adjuvant will generally elicit increased levels of IgG1 production relative to immunization of the antigen without adjuvant. TH2 adjuvants suitable for use in the invention include, for example, mineral containing compositions, oil-emulsions, and ADP-ribosylating toxins and detoxified derivatives thereof. Mineral containing compositions, such as aluminium salts are preferred TH2 adjuvants for use in the invention.

A composition may include a combination of a TH1 adjuvant and a TH2 adjuvant. Preferably, such a composition elicits an enhanced TH1 and an enhanced TH2 response, i.e., an increase in the production of both IgG1 and IgG2a production relative to immunization without an adjuvant. Still more preferably, the composition comprising a combination of a TH1 and a TH2 adjuvant elicits an increased TH1 and/or an increased TH2 immune response relative to immunization with a single adjuvant (*i.e.*, relative to immunization with a TH1 adjuvant alone or immunization with a TH2 adjuvant alone).

The immune response may be one or both of a TH1 immune response and a TH2 response. Preferably, immune response provides for one or both of an enhanced TH1 response and an enhanced TH2 response.

The enhanced immune response may be one or both of a systemic and a mucosal immune response. Preferably, the immune response provides for one or both of an enhanced systemic and an enhanced mucosal immune response. Preferably the mucosal immune response is a TH2 immune response. Preferably, the mucosal immune response includes an increase in the production of IgA.

Infections can affect various areas of the body and so the compositions of the invention may be prepared in various forms. For example, the compositions may be prepared for parenteral administration as injectables, either as liquid solutions or suspensions. Solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection can also be prepared (*e.g.* a lyophilised composition or a spray-freeze dried composition). The composition may be prepared for topical administration *e.g.* as an ointment, cream or powder. The composition may be formulated for administration using a 'vaccine patch' or plaster. The composition may be prepared for oral administration *e.g.* as a tablet or capsule, as a spray, or as a syrup (optionally flavoured). The composition may be prepared for pulmonary administration *e.g.* as an inhaler, using a fine powder or a spray. The composition may be prepared as a suppository or pessary. The composition may be prepared for nasal, aural or ocular administration *e.g.* as drops. The composition may be in kit form, designed such that a combined composition is reconstituted just prior to

administration to a patient. Such kits may comprise one or more antigens in liquid form and one or more lyophilised antigens.

Where a composition is to be prepared extemporaneously prior to use (*e.g.* where a component is presented in lyophilised form) and is presented as a kit, the kit may comprise two vials, or it may comprise one ready-filled syringe and one vial, with the contents of the syringe being used to reactivate the contents of the vial prior to injection.

Delivery methods including parenteral injection (e.g., subcutaneous, intraperitoneal, intravenous, intramuscular, or interstitial injection) and rectal, oral (e.g., tablet, spray), vaginal, topical, transdermal are disclosed in WO 99/27961, transcutaneous methods in WO02/074244 and WO02/064162, intranasal in WO03/028760. Other routes of administration include ocular, aural, and pulmonary or other mucosal administration.

Particularly the compositions of the present invention may be administered via a systemic route or a mucosal route or a transdermal route or it may be administered directly into a specific tissue. As used herein, the term "transdermal delivery" includes intradermal (e.g., into the dermis or epidermis) and transdermal (e.g. "percutaneous") i.e., delivery by passage of an agent into or through at least a top layer of skin. As used herein, the term "systemic administration" includes but is not limited to any parenteral routes of administration. In particular, parenteral administration includes but is not limited to subcutaneous, intraperitoneal, intravenous, intraarterial, intramuscular, or intrasternal injection, intravenous, intraarterial, or kidney dialytic infusion techniques. Generally, the systemic, parenteral administration is intramuscular injection. As used herein, the term "mucosal administration" includes but is not limited to oral, intranasal, intravaginal, intrarectal, intratracheal, intestinal and ophthalmic administration. Novel direct delivery forms can also include transgenic expression of the combinations of polypeptides in foods, e.g., transgenic expression in a potato.

Immunogenic compositions used as vaccines comprise an immunologically effective amount of antigen(s), as well as any other components, as needed. By 'immunologically effective amount', it is meant that the administration of that amount to an individual, either in a single dose or as part of a series, is effective for treatment or prevention. This amount varies depending upon the health and physical condition of the individual to be treated, age, the taxonomic group of individual to be treated (*e.g.* non-human primate, primate, *etc.*), the capacity of the individual's immune system to synthesise antibodies, the degree of protection desired, the formulation of the vaccine, the treating doctor's assessment of the medical situation, and other relevant factors. It is expected that the amount will fall in a relatively broad range that can be determined through routine trials.

Pharmaceutically acceptable carriers

Compositions of the invention will typically, in addition to the components mentioned above, comprise one or more "pharmaceutically acceptable carriers." These include any carrier which does not itself induce the production of antibodies harmful to the individual receiving the composition. Suitable carriers typically are large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, and lipid aggregates (such as oil droplets or liposomes). Such carriers are well known to those of ordinary skill in the art. A composition may also contain a diluent, such as water, saline, glycerol, etc. Additionally, an auxiliary substance, such as a wetting or emulsifying agent, pH buffering substance, and the like, may be present. A thorough discussion of pharmaceutically acceptable components is available in Gennaro (2000) Remington: The Science and Practice of Pharmacy. 20th ed., ISBN: 0683306472. The compositions of the invention may be prepared in various forms (e.g., liquid, lyophilized), as is known in the art.

Methods of treatment, and administration of immunogenic or vaccine compositions of the invention

The invention also provides a method for raising an immune response in a subject, particularly a mammal, comprising the step of administering an effective amount of a composition of the invention. The immune response is preferably protective and preferably involves antibodies and/or cell-mediated immunity. The method may raise a booster response.

The invention also provides immunogenic combinations or compositions for use as a medicament *e.g.* for use in raising an immune response in a subject, such as a mammal.

The invention also provides the use of a combination of polypeptides or a composition of the invention in the manufacture of a medicament for raising an immune response in a subject, such as a mammal.

The invention also provides a delivery device pre-filled with an immunogenic composition of the invention.

By raising an immune response in the subject by these uses and methods, the subject, for example a mammal, can be protected against *E.coli* infection, e.g. more than one *E. coli* pathotype, including ExPEC and non-ExPEC strains. The invention is particularly useful for providing broad protection against pathogenic ExPEC *E.coli*, including intestinal pathotypes such as EPEC, EAEC, EIEC, ETEC and DAEC (Diffuse-adhering Escherichia coli) pathotypes. Thus the subject may be protected against diseases including, but not limited to peritonitis, pyelonephritis, cystitis, endocarditis, prostatitis, urinary tract infections (UTIs), meningitis (particularly neonatal meningitis), sepsis (or SIRS), dehydration, pneumonia, diarrhea (infantile, travellers', acute, persistent, *etc.*), bacillary dysentery, hemolytic uremic syndrome (HUS), pericarditis, bacteriuria, *etc.*

The subject is preferably a mammal, particularly a human, but by way of non-limiting example, may also be a cow, a pig, a sheep, a horse, a cat or a dog since *E.coli* disease is also problematic in these species. In certain embodiments the subject may be an avian subject such as, for example, a chicken, goose, turkey and the like.

Where the vaccine is for prophylactic use, the human is particularly a child (*e.g.* a toddler or infant) or a teenager; where the vaccine is for therapeutic use, the human is particularly a teenager or an adult. A vaccine intended for children may also be administered to adults *e.g.* to assess safety, dosage, immunogenicity, *etc.*

One way of checking efficacy of therapeutic treatment involves monitoring *E.coli* infection after administration of the compositions of the invention. One way of checking efficacy of prophylactic treatment involves monitoring immune responses, systemically (such as monitoring the level of IgG1 and IgG2a production) and/or mucosally (such as monitoring the level of IgA production), against the antigens in the compositions of the invention after administration of the composition. Typically, antigen-specific serum antibody responses are determined post-immunisation but pre-challenge whereas antigen-specific mucosal antibody responses are determined post-immunisation and post-challenge.

Another way of assessing the immunogenicity of the compositions of the present invention is to express the proteins recombinantly for screening patient sera or mucosal secretions by immunoblot and/or microarrays. A positive reaction between the protein and the patient sample indicates that the patient has mounted an immune response to the protein in question. This method may also be used to identify immunodominant antigens and/or epitopes within antigens.

The efficacy of compositions of the invention can also be determined *in vivo* by challenging animal models of *E.coli* infection, *e.g.*, guinea pigs or mice, with the vaccine compositions. A murine model of ExPEC and lethal sepsis is described in reference 58. A cotton rat model is disclosed in ref. 59

Dosage treatment can be a single dose schedule or a multiple dose schedule. In some embodiments, compositions of the invention are administered in combination with an antibiotic treatment regime. In one embodiment, the antibiotic is administered prior to administration of a composition of the invention. In another embodiment, the antibiotic is administered subsequent to the administration of a composition of the invention.

Multiple doses may be used in a primary immunisation schedule and/or in a booster immunisation schedule. In a multiple dose schedule the various doses may be given by the same or different routes *e.g.* a parenteral prime and mucosal boost, a mucosal prime and parenteral boost, *etc.* Multiple doses will

typically be administered at least 1 week apart (*e.g.* about 2 weeks, about 3 weeks, about 4 weeks, about 6 weeks, about 8 weeks, about 10 weeks, about 12 weeks, about 16 weeks, *etc.*).

Vaccines of the invention may be used to treat both children and adults. Thus a human patient may be less than 1 year old, 1-5 years old, 5-15 years old, 15-55 years old, or at least 55 years old. Particular patient groups for receiving the vaccines are the elderly (*e.g.* ≥50 years old, ≥60 years old, and preferably ≥65 years), the young (*e.g.* ≤5 years old), hospitalised patients, healthcare workers, armed service and military personnel, travellers, pregnant women, the chronically ill, or immunodeficient patients. The vaccines are not suitable solely for these groups, however, and may be used more generally in a population.

Vaccines of the invention are particularly useful for patients who are expecting a surgical operation, or other hospital in-patients. They are also useful in patients who will be catheterized. They are also useful in adolescent females (*e.g.* aged 11-18) and in patients with chronic urinary tract infections.

Vaccines of the invention may be administered to patients at substantially the same time as (*e.g.* during the same medical consultation or visit to a healthcare professional or vaccination centre) or in combination with other vaccines *e.g.* at substantially the same time as a measles vaccine, a mumps vaccine, a rubella vaccine, a MMR vaccine, a varicella vaccine, a MMRV vaccine, a diphtheria vaccine, a tetanus vaccine, a pertussis vaccine, a DTP vaccine, a conjugated *H.influenzae* type b vaccine, an inactivated poliovirus vaccine, a hepatitis B virus vaccine, a meningococcal conjugate vaccine (such as a tetravalent A-C-W135-Y vaccine), a respiratory syncytial virus vaccine, a *Streptococcal* vaccine such as a Group A *Streptococcal* vaccine or a Group B *Streptococcal* vaccine *etc.*

General

The practice of the present invention will employ, unless otherwise indicated, conventional methods of chemistry, biochemistry, molecular biology, immunology and pharmacology, within the skill of the art. Such techniques are explained fully in the literature. See, *e.g.*, references 60-61, *etc.*

In some implementations, the term "comprising" refers to the inclusion of the indicated active agent, such as recited polypeptides, as well as inclusion of other active agents, and pharmaceutically acceptable carriers, excipients, emollients, stabilizers, etc., as are known in the pharmaceutical industry. In some implementations, the term "consisting essentially of" refers to a composition, whose only active ingredient is the indicated active ingredient(s), however, other compounds may be included which are for stabilizing, preserving, etc. the formulation, but are not involved directly in the therapeutic effect of the indicated active ingredient. Use of the transitional phrase "consisting essentially" means that the scope of a claim is to be interpreted to encompass the specified materials or steps recited in the claim, and those that do not materially affect the basic and novel characteristic(s) of the claimed invention. See, In re Herz,

537 F.2d 549, 551-52, 190 USPQ 461, 463 (CCPA 1976) (emphasis in the original); see also MPEP § 2111.03. Thus, the term "consisting essentially of" when used in a claim of this invention is not intended to be interpreted to be equivalent to "comprising".

The term "about" in relation to a numerical value x means, for example, $x \pm 10\%$.

"GI" numbering is used herein. A GI number, or "GenInfo Identifier", is a series of digits assigned consecutively to each sequence record processed by NCBI when sequences are added to its databases. The GI number bears no resemblance to the accession number of the sequence record. When a sequence is updated (*e.g.* for correction, or to add more annotation or information) then it receives a new GI number. Thus the sequence associated with a given GI number is never changed.

References to a percentage sequence identity between two amino acid sequences means that, when aligned, that percentage of amino acids are the same (i.e. identical) in comparing the two sequences, relative to the longest of the sequences. This alignment and the percent homology or sequence identity can be determined using software programs known in the art. A preferred alignment is determined by the Smith-Waterman homology search algorithm using an affine gap search with a gap open penalty of 12 and a gap extension penalty of 2, BLOSUM matrix of 62.

While certain embodiments of the present invention have been described and specifically exemplified above, it is not intended that the invention be limited to such embodiments. Various modifications may be made thereto without departing from the scope and spirit of the present invention as set forth in the following claims.

MODES FOR CARRYING OUT THE INVENTION

Example 1

Bacterial Ig-like domain (group 1) protein (orf405), gspK (orf3515), upec-1232, and orf3526, each as more fully described herein, have been expressed, sequenced and purified. Sequences were obtained for the orthologs in various other *E.coli* strains. Distribution of each of the candidate antigens were determined in most pathogenic strains, specifically NMEC, APEC, UPEC, EHEC, EAEC, EIEC, EPEC, ETEC and AIEC. The presence of each of the antigens is shown in Figure 1A and B.

Example 2

Antigens were PCR amplified from the genomic DNA templates, cloned in pET-21b vectors (Novagen) and transformed in DH5α-T1 chemically competent cells for propagation (Invitrogen). BL21 (DE3) chemically competent cells were used for expression. All candidates were cloned and expressed without the signal sequence and as His-tag fusion proteins. Candidates were purified by affinity chromatography.

Antigen	Homology	Size (kDa)	Solubility	Yield (mg/L of growth)	% Purity
Orf3526	IHE3034/RS218/536	165	+	1,245 mg/L	90%
Orf3515	All	31	+	16,7 mg/L	95%
405B	All	46	+	5,3 mg/L	95%
Upec1232	CFT073		+		95%

Example 3

Protection was evaluated in a sepsis animal model. CD1 out bred female mice (5 weeks old) from Charles River Italia were immunized by subcutaneous injections at the 1st, 21st and 35th days with 20 µg of recombinant protein in Freund's adjuvant. Positive control was immunized with 10^8 heat-inactivated bacteria (65°C for 30 minutes) in 0.15 ml of physiological solution in Freund's adjuvant (Sigma), while negative control was immunized with physiologic solution in Freund's adjuvant. Challenge was done at the 49th day with a dose of 10^7 of fresh bacterial culture/mouse (LD_{80}) by intraperitoneal (for strains IHE3034 and CFT073) or intravenous (for strain 536) injection. Heparinised-blood samples were collected from survived mice at 24 hours after challenge to determine bacteremia levels and the mortality was observed for four days after challenge.

Candidate	Sepsis Animal Model		
	Survival with vaccination (%)	Survival without vaccination (%)	P value
upec-1232	15/30 (50)	3/36 (8)	0.0002
gspK (orf3515)	30/110 (27)	11/116 (9)	0.0005
405B (bacterial Ig-like domain (group 1) protein fragment)	17/63 (26)	9/66 (13)	0.07
Orf3526	8/8 (100)	2/8 (25)	-

Example 3A

Protection was evaluated in a sepsis animal model. CD1 mice were immunized by subcutaneous injections at day 0, 21 and 35 with 20 µg of recombinant protein in Freund's complete adjuvant or alum. Positive control was immunized with 10^8 heat-inactivated bacteria (65°C for 30 minutes) in 0.15 ml of physiological solution in Freund's complete adjuvant or alum, while negative control was immunized with physiologic solution in Freund's complete adjuvant or alum. Challenge was done at the 49th day with

a dose of 10^7 of fresh bacterial culture/mouse (LD_{80}) by intraperitoneal (for strains IHE3034 and CFT073) or intravenous (for strain 536) injection. Heparinised-blood samples were collected from survived mice at 24 hours after challenge to determine bacteremia levels and the mortality was observed for four days after challenge.

Protection using Freud's complete adjuvant:

Candidate	Sepsis Animal Model		
	Survival with vaccination (%)	Protection rate	P value
upec-1232	15/30 (50)	45%	0.0002
gspK (orf3515)	30/110 (27)	20%	0.0009
405B	25/81 (30)	18%	0.029
Orf3526	125/149 (84)	83%	<0.0001

Protection using alum - Homologous challenge:

Candidate	Sepsis Animal Model		
	Survival with vaccination (%)	Protection rate	P value
upec-1232-His	4/8 (50)	43%	0.28
405B-His	11/40 (27.5)	19%	0.046
Orf3526-native	7/8 (87.5)	87.5%	0.0014
Orf3526-His	74/102 (72.5)	70%	<0.0001

Protection using alum - Heterologous challenge:

Candidate	Sepsis Animal Model		
	Survival with vaccination (%)	Protection rate	P value
Orf3526-His	14/23 (61)	48%	0.018

Protection rate = ((% dead control - % dead immune) / (% dead control)) x 100

Example 3C

Cross-protection was evaluated in a sepsis animal model by active or passive immunisation. Mice were immunized with antigen 3526-his in alum before challenge (active immunization) or administered anti-3526-his antibodies after challenge (passive immunisation). The sequence of 3526-his is based on the sequence of the native 3526 protein from the NMEC strain IHE3034. Mice were challenged with strains IHE3034, B616, IN1S or 9855/93. % PE (protective efficacy) was calculated as: 1 – (% dead vaccinated / % dead control) x 100.

Challenge strain	Active immunisation		Passive immunisation	
	% PE (survival)	P value	% PE (survival)	P value
IHE3034 (NMEC)	70 (74/102)	<0.0001	100 (32/32)	<0.0001
B616 (NMEC)	85 (14/16)	0.0002	nd	nd
IN1S (SEPEC)	50 (23/40)	<0.0001	50 (20/32)	0.0026
9855/93 (SEPEC)	35 (11/24)	0.03	nd	nd

The results show that 3526 from ExPEC-NMEC strain IHE3034 confers protection in actively immunized mice against at least three additional ExPEC strains (one NMEC and two SEPEC). The passive immunization experiments confirm cross-protection against at least one additional ExPEC strain (SEPEC).

Example 4

Protection was evaluated in a sepsis animal model according to the following schedule:

Active immunization

- CD1 mice (4 weeks old) are immunized s.c.
- 20 µg antigen + Freund's adjuvant; 3 doses at 0, 21, 35 days
- 14 days after last immunization (11 weeks old mice) mice are infected at a lethal dose (LD_{80}) with pathogenic *E. coli* strains.

Passive immunization

- Mice or rabbit immune serum is administered i.v. to CD1 mice
- 24h after passive immunization mice are challenged i.p. with IHE3034 strain

- Blood is collected from the tail at 24 hours to evaluate bacteremia
- Mortality is monitored for 4 days after the infection and

Protection rate is calculated as = $\frac{(\%\ \text{dead ctrl.} - \%\ \text{dead immun.})}{\%\ \text{dead ctrl.}} \times 100$

Candidate	Survival (%) Vacc.	Survival (%) no Vacc.	P value*
Orf3515	21/64 (33)	5/67 (7)	0.0003
405B	17/55 (31)	7/58 (12)	0.01
Upec1232	8/23 (34)	2/28 (7)	0.03

Example 5

In order to study the gene distribution, genetic variability and protein expression of antigen orf3526, as well as to evaluate the effective vaccine coverage, we studied three different collections of human and animal isolates including different pathogenic (ExPEC, ETEC, EPEC) and faecal strains. Briefly,

Genomic DNA was prepared by culturing bacteria overnight at 37 °C in atmosphere humidified with 5% CO_2 in LB (Difco). Chromosomal DNA was prepared from 1.5 mL of culture using the GenElute Bacterial Genomic DNA Kit (Sigma) according to the manufacturer's instructions. DNA concentration was calculated by optical density determination at 260nm. About 100ng of chromosomal DNA was used as template for the amplification of antigen orf3526. The amplification enzyme used was the Phusion® DNA Polymerase (Finnzymes). All genes were amplified using primers external to the coding region. Primers were designed in conserved DNA region and the sequences are reported in Table 1. Antigen orf3526 was amplified using primers ECOK1_3385_1 and ECOK1_3385_22. PCR conditions were as follows: 35 cycles of denaturation at 98 °C for 10 s, annealing at 55 °C for 20 s, and elongation at 72 °C for 3 min. PCR products were purified with Agencourt® AMPure® protocol (Beckman Coulter) and sequenced on the capillary sequencer ABI3730xl DNA Analyzers (Applied Biosystems). Sequences were assembled with Sequencher 4.8 (Gene Codes) and aligned and analyzed using the Vector NTI Suite 10.

Table1. Primers list used for amplification and sequencing of *orf3526* (Orf03343_1 to Orf03343_22 correspond to SEQ ID NOs 32 to 53)

Orf03343_1	TGATGCCGTTTTCTTAAGAATGGAGGAA
Orf03343_2	GAGCCAGAACCTGTTCCTA
Orf03343_3	GTAAAGCCATCGCTGGATTCA
Orf03343_4	CCACCTCTTCATTGACCAGC
Orf03343_5	CGGAACAGTTTTATCAGTAT
Orf03343_6	CCCCGCGAACTTCATCAC
Orf03343_7	GCAAGGTCTTTGCCGAGTATC
Orf03343_8	TGATAAAAACTACTGGCTGGC
Orf03343_9	GGTTACCGATAACCATCAG
Orf03343_10	GCAGATACCAGCAAACCGA
Orf03343_11	TCATCACGTTTTCCATGATCAGC
Orf03343_12	GCGGATTTAGGCACCAACATTC
Orf03343_13	GAATGTTGGTGCCTAAATCCGC
Orf03343_14	GTGAACGTTTTAAAGCCCAGCTC
Orf03343_15	AACTATATGGAAAAACCGCTGAC
Orf03343_16	AAAGCGCCGTTCTATAAAGA
Orf03343_17	TTATAGAACGGCGCTTTTAC
Orf03343_18	CATCACCGATGGAGATCTGC
Orf03343_19	AAGATGAACCGTGTCGCTGAC
Orf03343_20	CTGGAACCTGTTCCAGTTGAT
Orf03343_21	ATCAACTGGAACAGGTTCCAG
Orf03343_22	TATTGCTGAAAAACATCAAAAAG

Elisa assay

orf3526 antigen detection and relative quantification in supernatants (SN) was performed by an antibody-sandwich ELISA targeting *orf3526* antigen with rabbit anti-*orf3526* antigen polyclonal antibody and revealed by alkaline phosphatase-conjugated anti-rabbit antibody. Briefly, the wells of microtiter plates (Nunc, Maxi Sorp) were coated overnight at 4°C with 0.22µm-filtered bacterial supernatant. Unbound SN was washed out twice with a solution of PBS-Tween (0.05%) (PBS-T), and non-specific binding sites were blocked with a PBST-BSA (1%) solution for 1 h at 37°C. The plates were further washed another three times with PBS-T before rabbit anti-*orf3526* polyclonal antibody serial dilutions were added to duplicate wells for 1 h 37 min at 37°C. After three washes with PBS-T, alkaline phosphatase-conjugated anti-rabbit polyclonal antibody was added. Subsequently, the microplates were incubated for 1 h at 37°C and washed three times with PBST, before revelation by adding the enzyme substrate. After a 30-min incubation in the dark at room temperature, the reaction was stopped by adding 50ul NaOH solution(3N). The plates were read at 405 nm in a microplate reader (TECAN).

Overall, ECOK1_3385 gene was present and expressed in more than 80% of the 417 strains analyzed, with an amino acid sequence identity never below 86%. In conclusion, the results presented here indicate that antigen orf3526 is well represented, conserved and expressed across pathogenic and faecal isolates indicating that this target may be a useful candidate for a broadly protective vaccine against E. coli (Figures 2(a), 2(b) and 2(c)).

Phylogenetic reconstruction

Phylogenetic tree of 217 amino acid sequences of *orf3526* antigen was computed using MEGA v.4 (ref Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24: 1596–1599.) using the Neighbor Joining algorithm from distance matrices between protein sequences computed using the Maximum Composite Likelihood (ref. Tamura K, Nei M, Kumar S (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl Acad Sci U S A 101: 11030–11035.) (Figure 3).

Example 6

orf3526 is protective in an avian model

Three day old chicks (Bar-On) were challenged i.p. with $5x10^6$ of E. coli O2 strain 1772, in two injections of 0.2 ml each, 3 hours apart. Antibodies (anti-orf3526), 0.15 ml, were applied s.c. (neck) 20 min after the first injection. Antibodies to antigen orf3526 were found to protect in passive immunization in a chicken model of sepsis (Figure 4).

Example 7

The purpose of this experiment was to evaluate the protective ability of conserved antigen orf3526 with and without LTK63 against diarrheal disease in piglets, caused by intestinal pathogenic K88 E.coli (ETEC).

Two studies were carried out:

1. 11NAHLW1019v: Production of Protein orf3526 Antiserum by Hyperimmunization (Phase I), Antiserum Oral Administration Pre-Challenge in Piglets (Phase II)
2. 11NAHLW1020v: ExPEC Antiserum Administration with Challenge Evaluation of orf3526 Protein

The first experiment has been designed to generate anti-sera for the orf3526 protein with and without LTK63 in CD/CD swine followed by a pilot administration with minimum and maximum doses as well as controls to immunologically naive piglets to assess preliminary efficacy of anti-sera type products in the challenge model.

The second experiment is a randomized, blinded trial designed to evaluate the orf3526 protein given in the form of an anti-sera to immunologically naive piglets followed by oral K88 E. coli challenge.
The following material was utilised:

- Purified recombinant orf3526 Protein
- Purified recombinant orf3526 Protein + LTK63
- Monoclonal Antibody to orf3526 for stability testing

Challenge

K88 challenge culture (lot no. TBD) was thawed at room temperature (~23 degrees C); pooled together, diluted 1:2 with sterile Peptone Buffer, 2.0 ml and re-dispensed into 3 ml cryovials and frozen < -60 degrees C. A Post-freeze Viability Count was performed in accordance to SO 6.001, to establish the amount of antigen being administered.

At the time of farrowing, each piglet had the date and time of farrowing recorded on the Farrowing/Challenge Form. The piglets were ear tagged prior to processing, due to the need for the piglets' identification number in recording time and date of birth. Piglets were allowed to suckle ad libitum. Within 6 hours (+/– 2 hours) of birth, piglets were weighed. Immediately following processing, piglets meeting the Post-inclusion Removal/Withdrawal criteria were utilized for challenge study.

The diluted and re-dispensed K88 Challenge Culture (lot no. TBD) was thawed at room temperature (~23 degrees C) and 2.0 ml administered orally to each piglet. Following challenge, piglets were placed back on the gilt.

The treatment outcome for Phase I was assessed by antibody titres to protein orf3526 in the sera collected from hyperimmunized pigs. The treatment outcome for Phase II was assessed by which group has the maximum protection from 2 doses of antisera orally pre-challenge; determined by the mortality/morbidity information gathered from clinical observations and necropsy results.

Example 8

Orf3526 protection in a murine model of intestinal tract colonization by intragastric infection with ETEC strain GL53-K88:

Mice received streptomycin (5g/liter) in their drinking water (enriched with 6,7% fructose) 48 to 24 h prior to infection to eradicate normal resident bacteria flora. Following this, mice were infected by oral lavage with 109 CFU of a suspension of GL53 strain in a final volume of 400 ul. To reduce the effect of stomach acidity on the bacterial organism, bicarbonate was aministered intragastrically 15 minutes prior to bacterial inoculation. 24 hours following infection, mice were euthanized and segments of ileum (2 cm) were harvested and homogenized. Serial dilutions of GL53, resistant to Kanamycin, are plated onto LB agar plates enriched with antibiotic. To confirm that the recovered bacteria were the inoculum strain, bacterial colonies are tested by PCR using primers encoding for LT holotoxin. To test the protective effect of antigen orf3526, mice were immunized intranasally on days 1, 7, 21, and 35 with antigen (20ug) used alone or in combination with LTR72 as mucosal adjuvant (ratio 1:10). On day 49, mice were infected by oral lavage with 109 CFU of GL53 strain.

As shown in Figure 5, intranasal immunization with antigen orf3526 reduced colonization in ileum tract following challenge with ETEC strain GL53. Thus, orf3526 is able to protect against challenge with the UPEC strain 536 in an UTI murine model.

Example 9

Antigen orf3526 was prepared and administered as previously described in combination with FCA, IC31, alum, MF59 or alone or in combination. Antigen orf3526 remained protective when administered with a variety of adjuvants:

Candidates	animal model		
	% survival immun.	% survival ctrl.	protection
pK1-3526 + FCA	8/8	2/8	100
pK1-3526 + IC31	7/8	1/8	86
pK1-3526 + alum	7/8	2/8	83
pK1-3526 + MF59	6/8	0/8	75
pK1-3526 + alum/IC31	8/8	2/8	100
pK1-3526 + MF59/IC31	7/8	2/8	83
pK1-3526ΔG + FCA	10/10	1/10	100
pK1-3526ΔGΔP + FCA	9/10	1/10	89

Example 10

The protective effect of antigens 405B and upec1232 were determined using a UTI model of infection in mice:

The bacteria used to infect the mice were grown in filter-sterilized human urine and were passaged three times. The bacteria were incubated at 37°C, shaken at 200 rounds/min overnight, and centrifuged at 6,500 × g for 10 min. The pellet was then suspended in phosphate-buffered saline (PBS) to a concentration of approximately 1010 CFU/ml.

Mice were anesthetized by intraperitoneal administration of 0.08 ml of a mixture of Hypnorm (fentanyl citrate, 0.315 mg/ml; fluanisone, 10 mg/ml) and Stesolid (diazepam, 5 mg/ml) at a ratio of 5:1.5. Anesthetized mice were inoculated transurethrally with the bacterial suspension (*E. coli* 536) by use of plastic catheters. 0.05 ml of bacterial suspension was injected in the bladder over 5 s in order to avoid vesicoureteral reflux (12, 18). The catheter was removed immediately after inoculation. Urine from each mouse was collected in Eppendorf tubes by gentle compression of the abdomen, and the mice were killed by cervical dislocation. The organs were removed aseptically, the bladders were cut off near the urethra, and the kidneys were removed by blunt dissection to avoid bleeding. The organs were placed in cryotubes (Nunc 363452) containing a 750-µl suspension of collagenase (500 U/ml; Sigma C9891) and were stored at −80°C. Prior to homogenization, the infected organs were incubated for 1.5 h at room temperature and were then homogenized manually with inoculating loops and a whirl mixer. Bacteria from the inoculum, bacteria that were recovered from the urine samples and bacteria from either the bladder or one of the kidneys were measured. The results, illustrated in Figure 6 and 7 respectively, demonstrate that antigens 405B and upec1232 prevent kidney colonisation in a UTI model of infection.

Example 11

Mutants/Variants of orf3526

Bacteria with one of each of three constructs expressing his-tagged variants of orf3526 were cultured in 30 ml of medium and induced to express the orf3526 variant at 25° C (orf3526 without the leader peptide (3526), orf3526 with the N-terminus removed through the gly-ser linker or gly-ser region (ΔG3526), and orf3526 with the N-terminus removed through the proline rich region (ΔG3526)). The bacteria were harvested and lysed by sonication. The soluble fractions were isolated and loaded on an IMAC column. The column was washed three times with 20mM imidazole buffer. The orf3526 variants were then eluted with three washes of 500 mM imidazole buffer. Removal of the N-terminus of orf3526 through the gly-ser linker or gly-ser region significantly increased solubility and yield of purified protein. The yield obtained was estimated by Bradford assay to be as follows: 0.18 mg of 3526 and 2.34 mg ΔG3526.

Example 12

Mutants/Variants of orf3526

Although the function of orf3526 is not known, analysis of the orf3526 sequence revealed several conserved motifs, most notably a zinc binding motif, possibly part of a metallo-protease function, and an imperfect GTP binding motif (Figure 8). Native orf3526 is a potential lipoprotein, also secreted into the culture supernatant. Sequence alignments studies show that the protein has homology to AcfD (accessory colonization factor) from Vibrio cholera. Native orf3526 is constitutively expressed and secreted by a Type 2 secretion system (T2SS). Seven mutants/variants of orf3526 were prepared as illustrated in Figure 10.

Zinc content of the various orf3526 derivatives was determined by atomic absorption spectroscopy. Results, illustrated in Figure 11, suggest the presence of a single zinc ion per protein molecule. The unexpected low zinc content of 3526 B his, actually containing the zinc binding motif, could be explained by misfolding of this truncated derivative, while the single amino acid exchange in the E1305 mutant apparently is not sufficient to completely abolish zinc binding (red boxes).

A purified triple mutant orf3526 (DG3526TL) protein was prepared wherein the amino acids at positions H1304, E1305 and H1308 according to SEQ ID 8 were substituted by another amino acid, specifically H1304A, E1305A and H1308A. Surprisingly, zinc affinity is almost completely lost in the DG3526TL triple mutant which comprised mutations in the zinc binding motif at amino acids 1304, 1305 and 1308 of SEQ ID 8 from the amino acid sequence HEVGH to AAVGA.

To allow functional and structural characterization, the soluble tagless recombinant protein DG3526TL (164 KDa) was purified by CaptoQ and butyl Sepharose chromatography. SE-HPLC/MALLS analysis revealed that DG3526TL exists in two isoforms (peaks A and B in Figure 9a) and one truncated form (peak C in Figure 9a). The isoforms can be separated on butyl Sepharose. However, without stabilization by glycerol, conversion of one form into the other was observed over time (Figure 9(a) and 9(b)). Thus, the triple mutant isoform is further advantageous in terms of purification.

The amino acid sequence of the triple mutant TL3M is:

>orf03526 Triple Mutant (SEQ ID 31)

MNKKFKYKKSLLAAILSATLLAGCDGGGSGSSSDTPSVDSGSGTLPEVKPDPTPTPEPTP
EPTPDPEPTPDPTPDPEPTPEPEPEPVPTKTGYLTLGGSQRVTGATCNGESSDGFTFTPG
NTVSCVVGSTTIATFNTQSEAARSLRAVDKVSFSLEDAQELANSENKKTNAISLVTSSDS
CPADAEQLCLTFSSVVDRARFEKLYKQIDLATDNFSKLVNEEVENNAATDKAPSTHTSTV
VPVTTEGTKPDLNASFVSANAEQFYQYQPTEIILSEGQLVDSLGNGVAGVDYYTNSGRGV
TDENGKFSFSWGETISFGIDTFELGSVRGNKSTIALTELGDEVRGANIDQLIHRYSTTGQ
NNTRVVPDDVRKVFAEYPNVINEIINLSLSNGATLDEGDQNVVLPNEFIEQFKTGQAKEI
DTAICAKTDGCNEARWFSLTTRNVNDGQIQGVINKLWGVDTNYQSVSKFHVFHDSTNFYG
STGNARGQAVVNISNSAFPILMARNDKNYWLAFGEKRAWDKNELAYITEAPSIVQPENVT
RDTATFNLPFISLGQVGEGKLMVIGNPHYNSILRCPNGYSWGGGVNSKGECTLSGDSDDM
KHFMQNVLRYLSNDIWQPNTKSIMTVGTNLENVYFKKAGQVLGNSAPFAFHEDFTGITVK
QLTSYGDLNPEEIPLLILNGFEYVTQWSGDPYAVPLRADTSKPKLTQQDVTDLIAYLNKG
GSVLIMENVMSNLKEESASSFVRLLDAAGLSMALNKSVVNNDPQGYPDRVRQRRATGIWV
YERYPAADGAQPPYTIDPNTGEVTWKYQQDNKPDDKPKLEVASWQEEVEGKQVTRYAFID
EAEYTTEESLEAAKAKIFEKFPGLQECKDSTYHYEINCLERRPGTDVPVTGGMYVPRYTQ
LNLDADTAKAMVQAADLGTNIQRLYQHELYFRTKGSKGERLNSVDLERLYQNMSVWLWND
TKYRYEEGKEDELGFKTFTEFLNCYANDAYAGGTKCSADLKKSLVDNNMIYGDGSSKAGM
MNPSYPLNYMEKPLTRLMLGRSWWDLNIKVDVEKYPGSVSAKGESVTENISLYSNPTKWF
AGNMQSTGLWAPAQQDVTIKSSASVPVTVTVALADDLTGREKHEVALNRPPRVTKTYTLE
ANGEVTFKVPYGGLIYIKGDSKDDVSANFTFTGVVKAPFYKDGEWKNDLDSPAPLGELES
ASFVYTTPKKNLEASNFTGGVAEFAKDLDTFASSMNDFYGRNDEDGKHRMFTYKNLTGHK
HRFTNDVQISIGDAHSGYPVMNSSFSTNSTTLPTTPLNDWLIW**AAVGA**NAAETPLNVPGA
TEVANNVLALYMQDRYLGKMNRVADDITVAPEYLDESNGQAWARGGAGDRLLMYAQLKEW
AEENFDIKQWYPDGELPKFYSDRKGMKGWNLFQLMHRKARGDDVGNSTFGGKNYCAESNG
NAADTLMLCASWVAQADLSEFFKKWNPGASAYQLPGATEMSFQGGVSSSAYSTLASLKLP
KPEKGPETINKVTEHKMSAE

Example 13

Antigen combinations

His-tagged antigens were combined and administered with alum. The protection rate was calculated:

Antigen(s)	Protection rate
Orf3526 (2µg)	61%
Orf3526 (2µg) + 405B (10µg) + upec-1232 (10µg)	87.5%

Example 14

Antigen combinations

Antigens were combined and administered with alum. Survival and protection rate were calculated following challenge with either IHE3034 (NMEC) or 9855/93 (SEPEC) ExPEC strains.

Antigen(s)	IHE3034 challenge		9855/93 challenge	
	Survival with vaccination	Protection rate	Survival with vaccination	Protection rate
Orf3526 (2μg)	57/88 (64%)	61%	n/a	n/a
Orf3526 (2μg) + 405B (10μg) + upec-1232 (10μg)	7/8 (87.5%)	87.5%	n/a	n/a
Orf3526 (10μg)	11/16 (68%)	63%	14/23 (61%)	48%
Orf3526 (10μg) + 405B (10μg) + upec-1232 (10μg)	7/8 (87.5%)	87.5%	n/a	n/a
Orf3526C (0.2μg)	7/16 (44%)	44%	n/a	n/a
Orf3526C (0.2μg) + 405B (20μg)	5/8 (62.5%)	62.5%	n/a	n/a

Example 15

Protective efficacy of 3625 in the intestinal colonization model.

Mice were immunized via the intramuscular route with the 3526 antigen with alum or MF59, on days 0, 21 and 35. Mice were challenged with GL53 (ETEC) on day 48 and bacterial titres were evaluated in the caecum. Figure 14(a) shows that bacterial titres were significantly reduced after immunisation with 3526 + alum, or after immunisation with 3526 + MF59, compared to adjuvant alone. The results are confirmed in an experiment were mice were immunised on days 0 and 21 only (Figure 14(b)).

Example 16

Protective efficacy of antigen combinations in the intestinal colonization model.

Mice were immunized with different combinations of antigens 405B, 1232 and 3526, with or without LTK63 (Figure 15), on days 1, 21 and 35. Mice were challenged with GL53 (ETEC) and bacterial titres were evaluated in the caecum. The results show that 405B+3526+LTK63 and 3526+1232+LTK63 significantly reduces intestinal colonization by GL53 in the caecum, compared to LTK63 alone.

Example 17

Protective efficacy of isoforms A, B, and C of 3526.

Mice were immunized with isoform A alone, isoform B alone, or a combination of isoforms A, B and C. Mice were challenged with IHE3034 and bacterial titres were evaluated. The results show that isoform B alone or combined with isoforms A and C confers greatest protective efficacy.

[peak A] pur_131	20 ug	44 (7/16)	36
[peak B] pur_131	20 ug	69 (11/16)	64,5
[peak A,B,C] pur_131	20 ug	62,5 (5/8)	75
[peak A,B,C] pur_131	5 ug	87,5 (7/8)	86
[peak A] pur_131	5 ug	37,5 (3/8)	28,5
[peak B] pur_131	5ug	87,5 (7/8)	86

ADDITIONAL REFERENCES

[1] Bodanszky (1993) *Principles of Peptide Synthesis* (ISBN: 0387564314).
[2] Fields et al. (1997) Meth Enzymol 289: Solid-Phase Peptide Synthesis. ISBN: 0121821900.
[3] Chan & White (2000) *Fmoc Solid Phase Peptide Synthesis.* ISBN: 0199637245.
[4] Kullmann (1987) *Enzymatic Peptide Synthesis.* ISBN: 0849368413.
[5] Ibba (1996) *Biotechnol Genet Eng Rev* 13:197-216.
[6] *Vaccine Design: The Subunit and Adjuvant Approach* (eds. Powell & Newman) Plenum Press 1995 (ISBN 0-306-44867-X).
[7] Vaccine Adjuvants: Preparation Methods and Research Protocols (Volume 42 of Methods in Molecular Medicine series). ISBN: 1-59259-083-7. Ed. O'Hagan.
[8] *Vaccine Design* (1995) eds. Powell & Newman. ISBN: 030644867X. Plenum.
[9] WO90/14837.
[10] WO90/14837.
[11] Podda & Del Giudice (2003) *Expert Rev Vaccines* 2:197-203.
[12] Podda (2001) *Vaccine* 19: 2673-2680.
[13] *Vaccine Design: The Subunit and Adjuvant Approach* (eds. Powell & Newman) Plenum Press 1995 (ISBN 0-306-44867-X).
[14] Vaccine Adjuvants: Preparation Methods and Research Protocols (Volume 42 of Methods in Molecular Medicine series). ISBN: 1-59259-083-7. Ed. O'Hagan.
[15] US 5,057,540.
[16] Niikura *et al.* (2002) *Virology* 293:273-280.
[17] Lenz *et al.* (2001) *J Immunol* 166:5346-5355.
[18] Pinto *et al.* (2003) *J Infect Dis* 188:327-338.
[19] Gerber *et al.* (2001) *J Virol* 75:4752-4760.
[20] WO03/024480.
[21] WO03/024481.

[22] Gluck *et al.* (2002) *Vaccine* 20:B10-B16.
[23] Meraldi *et al.* (2003) *Vaccine* 21:2485-2491.
[24] Pajak *et al.* (2003) *Vaccine* 21:836-842.
[25] Krieg (2003) *Nature Medicine* 9:831-835.
[26] McCluskie et al. (2002) FEMS Immunology and Medical Microbiology 32:179-185.
[27] WO98/40100.
[28] US 6,207,646.
[29] US 6,239,116.
[30] US 6,429,199.
[31] Schellack *et al.* (2006) *Vaccine* 24:5461-72.
[32] Johnson *et al.* (1999) *Bioorg Med Chem Lett* 9:2273-2278.
[33] Evans *et al.* (2003) *Expert Rev Vaccines* 2:219-229.
[34] Beignon *et al.* (2002) *Infect Immun* 70:3012-3019.
[35] Pizza *et al.* (2001) *Vaccine* 19:2534-2541.
[36] Pizza *et al.* (2000) *Int J Med Microbiol* 290:455-461.
[37] Scharton-Kersten *et al.* (2000) *Infect Immun* 68:5306-5313.
[38] Ryan *et al.* (1999) *Infect Immun* 67:6270-6280.
[39] Partidos *et al.* (1999) *Immunol Lett* 67:209-216.
[40] Peppoloni *et al.* (2003) *Expert Rev Vaccines* 2:285-293.
[41] Pine *et al.* (2002) *J Control Release* 85:263-270.
[42] WO99/40936.
[43] WO99/44636.
[44] Singh *et al]* (2001) *J Cont Release* 70:267-276.
[45] WO99/27960.
[46] US 6,090,406.
[47] US 5,916,588.
[48] EP-A-0626169.
[49] WO99/52549.
[50] Andrianov *et al.* (1998) *Biomaterials* 19:109-115.
[51] Payne et al. (1998) Adv Drug Delivery Review 31:185-196.
[52] Stanley (2002) *Clin Exp Dermatol* 27:571-577.
[53] Jones (2003) *Curr Opin Investig Drugs* 4:214-218.
[54] WO99/11241.
[55] WO94/00153.
[56] WO98/57659.
[57] European patent applications 0835318, 0735898 and 0761231.
[58] Durant *et al.* (2007) *Infect Immun* 75:1916-25.
[59] WO02/081653.
[60] Gennaro (2000) *Remington: The Science and Practice of Pharmacy.* 20th edition, ISBN: 0683306472.
[61] *PCR (Introduction to Biotechniques Series)*, 2nd ed. (Newton & Graham eds., 1997, Springer Verlag).

CLAIMS:

1. An immunogenic composition comprising a combination of (i) bacterial Ig-like domain protein fragment (orf405B) having the amino acid sequence set forth in SEQ ID NO:2 or a protein having at least 80% similarity thereto, and (ii) putative Lipoprotein (orf3526) having the amino acid sequence set forth in SEQ ID NO:8 or a protein having at least 80% similarity thereto.
2. The immunogenic composition of claim 1 which further comprises (iii) upec1232 having the amino acid sequence set forth in SEQ ID NO:4 or a protein having at least 80% similarity thereto.
3. The immunogenic composition of claim 1 or claim 2 which further comprises (iv) gspK (orf3515) having the amino acid sequence set forth in SEQ ID NO:30 or a protein having at least 80% similarity thereto.
4. The immunogenic composition of any of claims 1 to 3 which further comprises at least one bacterial toxin.
5. The immunogenic composition of claim 4 wherein the bacterial toxin is an Escherichia coli toxin.
6. The immunogenic composition of claim 5 wherein the bacterial toxin is modified heat-labile toxin of Escherichia coli (LTK63).
7. The immunogenic composition according to any one of claims 1 to 6 wherein the putative Lipoprotein (orf3526) is a mutant protein wherein at least one amino acid at positions 1304, 1305, 1306, 1307 and/or 1308 with reference to SEQ ID 8 is/are substituted by another amino acid and wherein the zinc binding activity of the mutant orf3526 protein is reduced by at least 50% relative to wild-type orf3526.
8. The immunogenic composition of claim 7 wherein the mutant orf3526 protein comprises the amino acid sequence of SEQ ID: 31
9. The immunogenic composition of any preceding claim which further comprises one or more pharmaceutically acceptable carriers, diluents and/or adjuvants.
10. The immunogenic composition of claim 9 which is a vaccine composition.
11. The immunogenic composition of any of claims 1 to 10 for use in medicine.
12. The immunogenic composition of any of claims 1 to 10 for use in treating or preventing infections by E. coli.
13. A method for treating or preventing E. coli infection in a mammal, which comprises administering to said mammal an effective amount of an immunogenic composition according to any one of claims 1 to 10.

14. The method of claim 13, or the immunogenic composition of claim 12 for the use of claim 12, wherein said infections by E. coli comprise infections by both extraintestinal and intraintestinal pathogenic E. coli.

15. The method of claim 13 or 14, or the immunogenic composition of claim 12 or 14, for the use of claim 12 or 14, wherein the immunogenic composition is administered to a mucosal surface.

16. The method of claim 15, or the immunogenic composition of claim 15, for the use of claim 15, wherein said mucosal surface comprises nasal epithelium.

17. The method of claim 15, or the immunogenic composition of claim 15, for the use of claim 15 wherein said mucosal surface comprises oral mucosa.

18. The method of claim 15, or the immunogenic composition of claim 15, for the use of claim 15, wherein said mucosal surface comprises a luminal surface of a gastrointestinal organ selected from the group consisting of: stomach, small intestine, large intestine, and rectum.

19. The method of claim 13 or 14, or the immunogenic composition of claim 12 or 14 for the use of claim 12 or 14 wherein the immunogenic composition is administered by parenteral administration.

20. A polypeptide comprising (i) amino acid sequence SEQ ID: 31, or (ii) an immunogenic fragment of SEQ ID: 31 which includes amino acid positions 1304, 1305 and 1308.

21. A polypeptide which binds to an antibody which antibody
(i) does bind to isoform B of the 3526 polypeptide obtainable by purification using size exclusion chromatography (e.g. CaptoQ and/or butyl sepharose chromatography) from a composition comprising recombinant 3526 polypeptide(s) of a fraction that elutes at around 17 mins, or that elutes after isoform A and/or before isoform C, but which antibody
(ii) does not bind to isoform A of the 3526 polypeptide obtainable by purification using size exclusion chromatography (e.g. CaptoQ and/or butyl sepharose chromatography) from a composition comprising recombinant 3526 polypeptide(s) of a fraction that elutes at around 16 mins, or of a fraction that elutes before isoform B and/or before isoform C,
or to isoform C of the 3526 polypeptide, obtainable by purification using size exclusion chromatography (e.g. CaptoQ and/or butyl sepharose chromatography) from a composition comprising recombinant 3526 polypeptide(s) of a fraction that elutes at around 19 mins, or that elutes after isoform A and/or after isoform

22. The immunogenic composition of any one of claims 1-12, wherein the putative lipoprotein (orf3526) is the polypeptide of claim 21.

Fig. 1A

Candidate	EXPEC	EHEC	EPEC	ETEC	AIEC
Orf3526	51/59	6/25	5/6	9/9	4/4
orf405B	54/59	24/25	6/6	8/9	3/4
Upec1232	46/59	4/25	2/4	8/9	1/4

Fig. 1B

	EPEC	ETEC	ExPEC	Fascal	EHEC	Other Pathovars
ECOK1_3385 distribution (PCR/Blast)	67/71 (94.3%)	32/34 (94.1%)	179/205 (87.3%)	132/156 (84.6%)	22/66 (39.4%)	58/71 (81.6%)
Expression (WB/Sumatant)	44/53 (75.5%)	23/24 (95.8%)	125/147 (85%)	79/112 (70.5%)	n.d	n.d

Fig. 2(a)

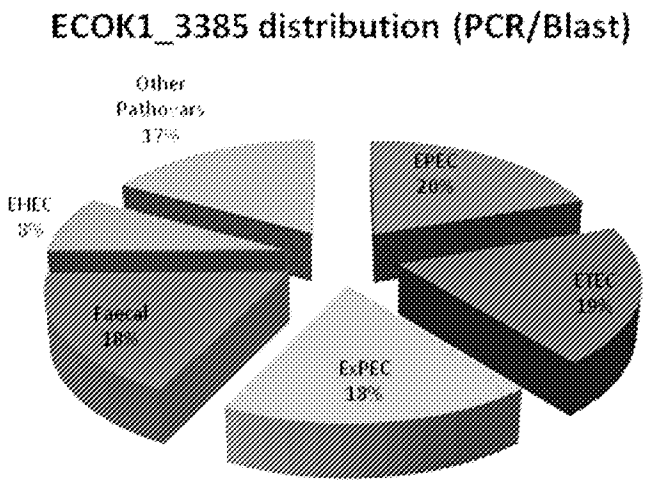


Fig. 2(b)

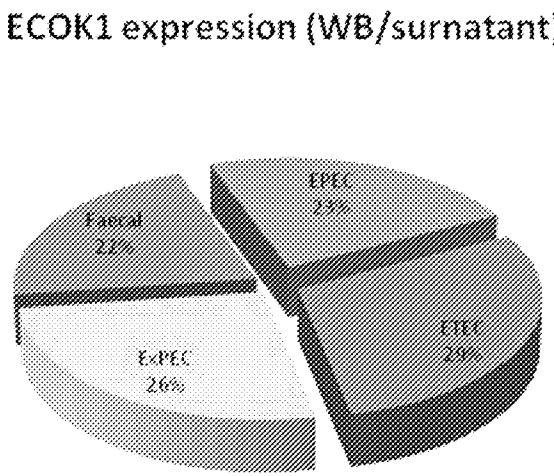
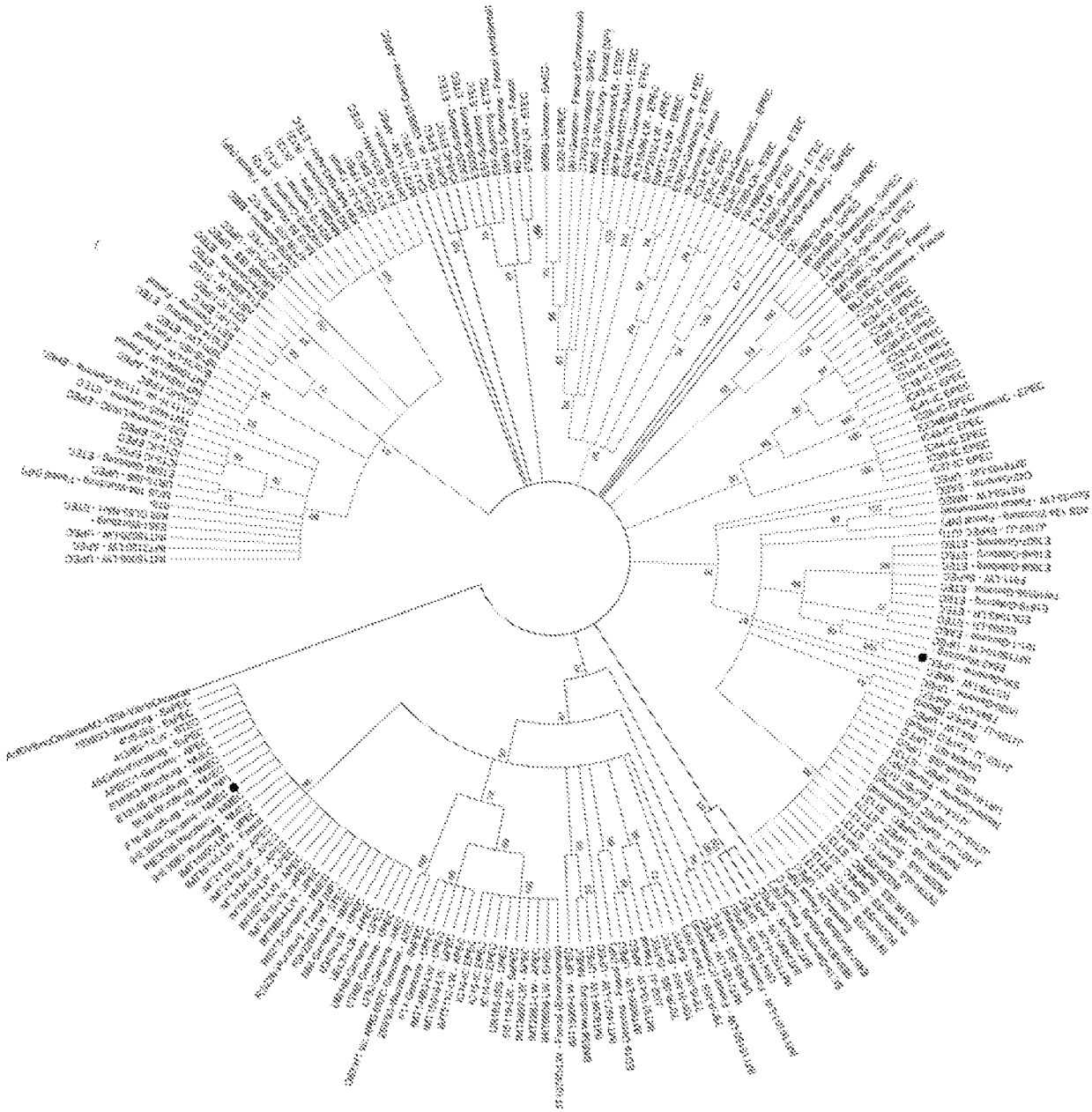


Fig. 2(c)

Fig. 3



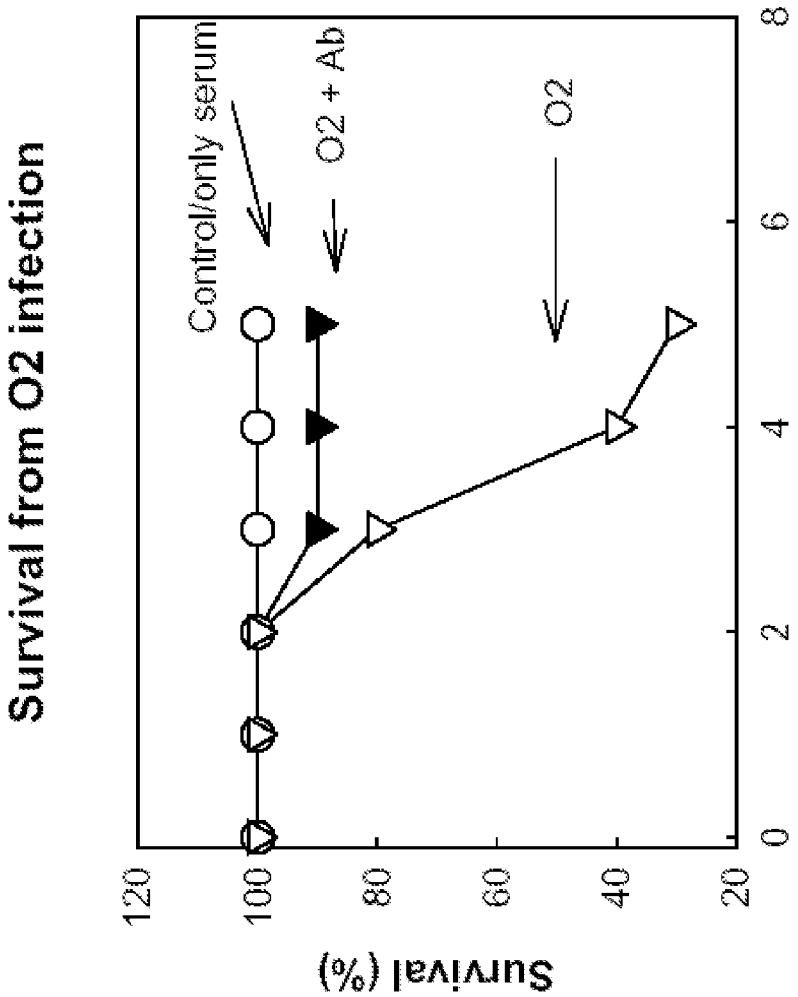

Survival from O2 infection
Control/only serum
O2 + Ab
O2
Survival (%)
120
100
80
60
40
20
0
2
4
6
8
Days

Fig. 4

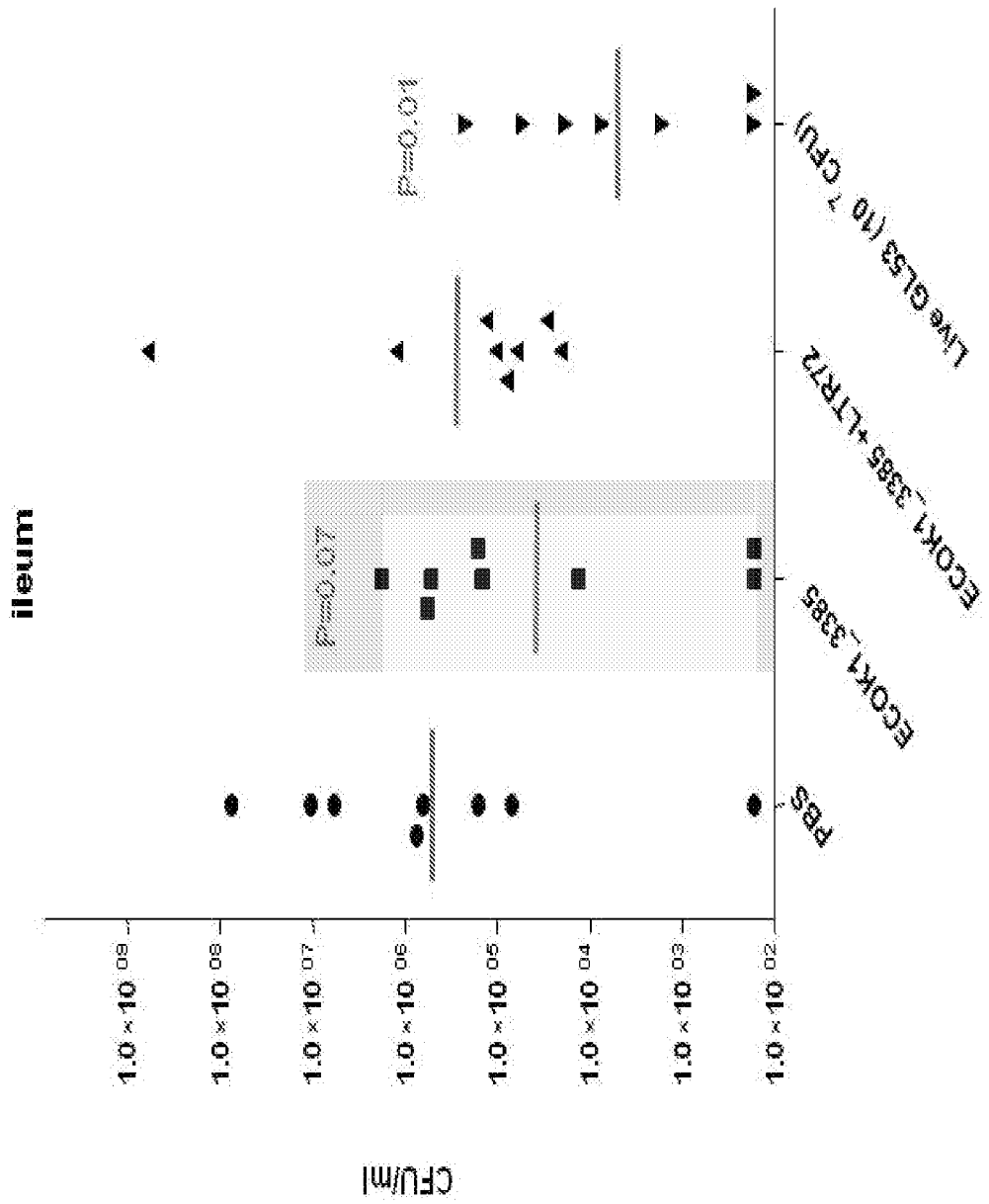

ileum
CFU/ml
1.0 × 10 09
1.0 × 10 08
1.0 × 10 07
1.0 × 10 06
1.0 × 10 05
1.0 × 10 04
1.0 × 10 03
1.0 × 10 02
P=0.07
P=0.01
PBS
ECOK1_3385
ECOK1_3385+LTR72
Live GL53 (10 7 CFU)

Fig. 5

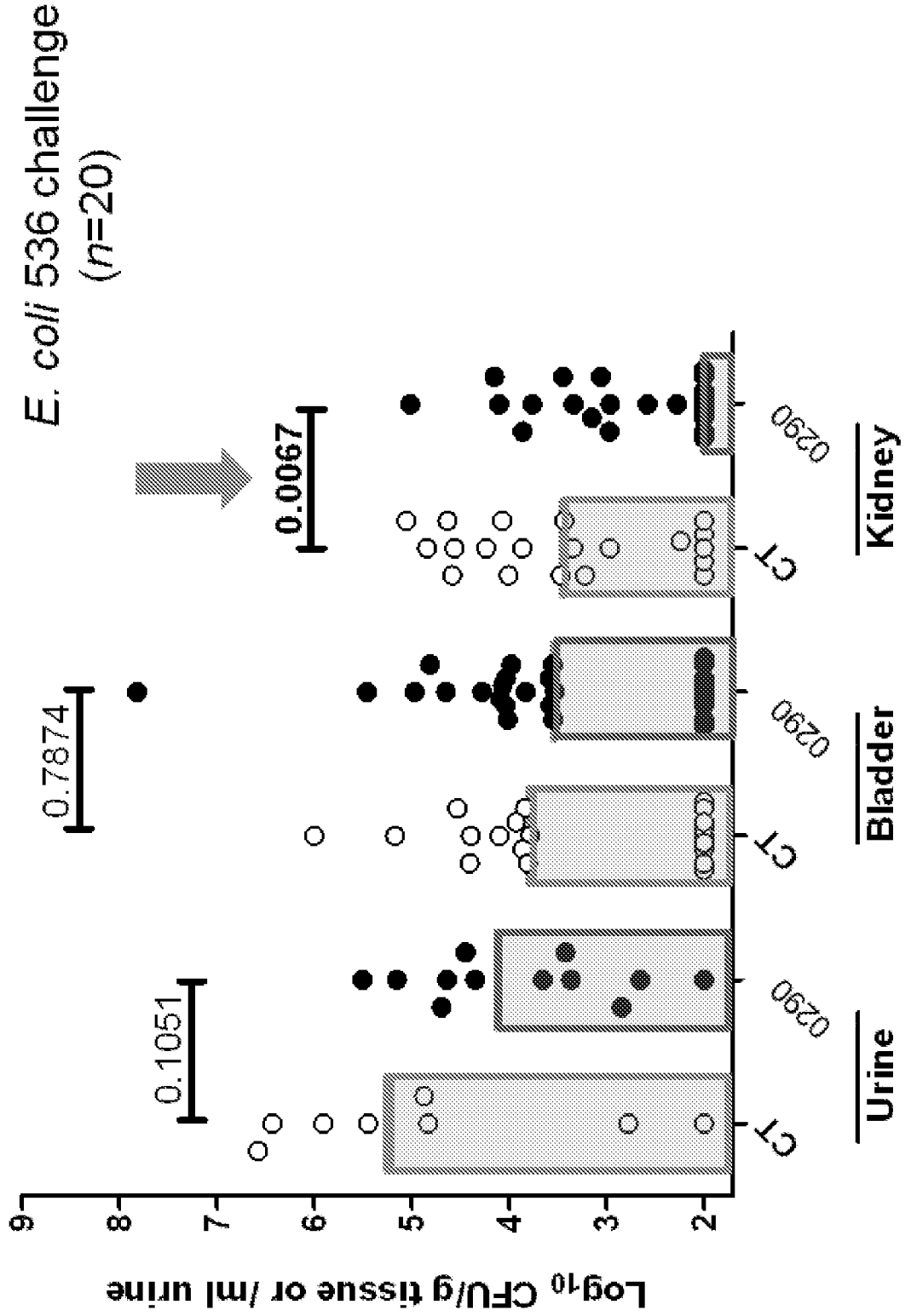

E. coli 536 challenge
(n=20)
0.1051
0.7874
0.0067
Log10 CFU/g tissue or /ml urine
9
8
7
6
5
4
3
2
CT
0290
CT
0290
CT
0290
Urine
Bladder
Kidney

Fig. 6

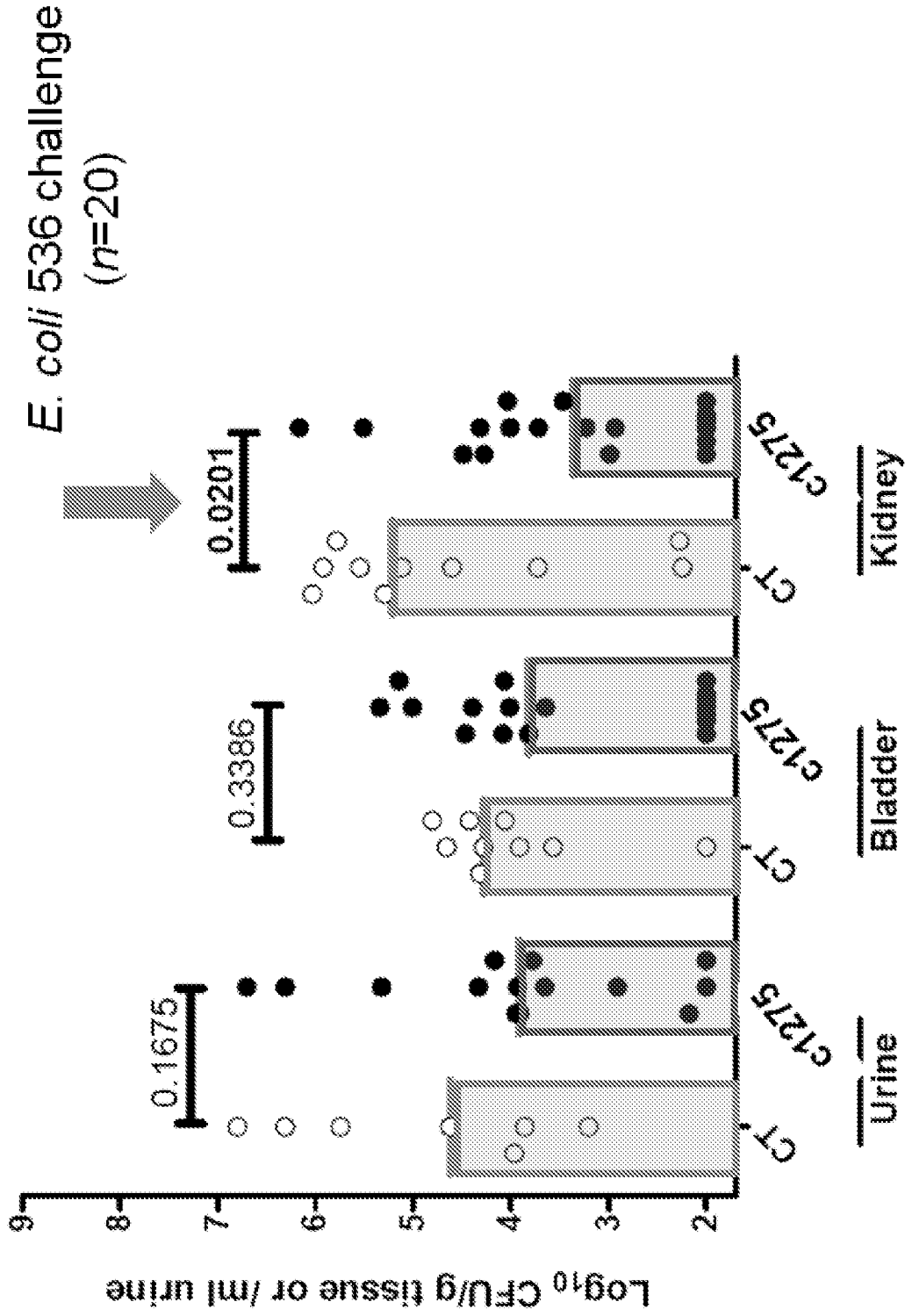

E. coli 536 challenge
(n=20)
Log10 CFU/g tissue or /ml urine
9
8
7
6
5
4
3
2
0.1675
0.3386
0.0201
CT
c1275
Urine
Bladder
Kidney

Fig. 7

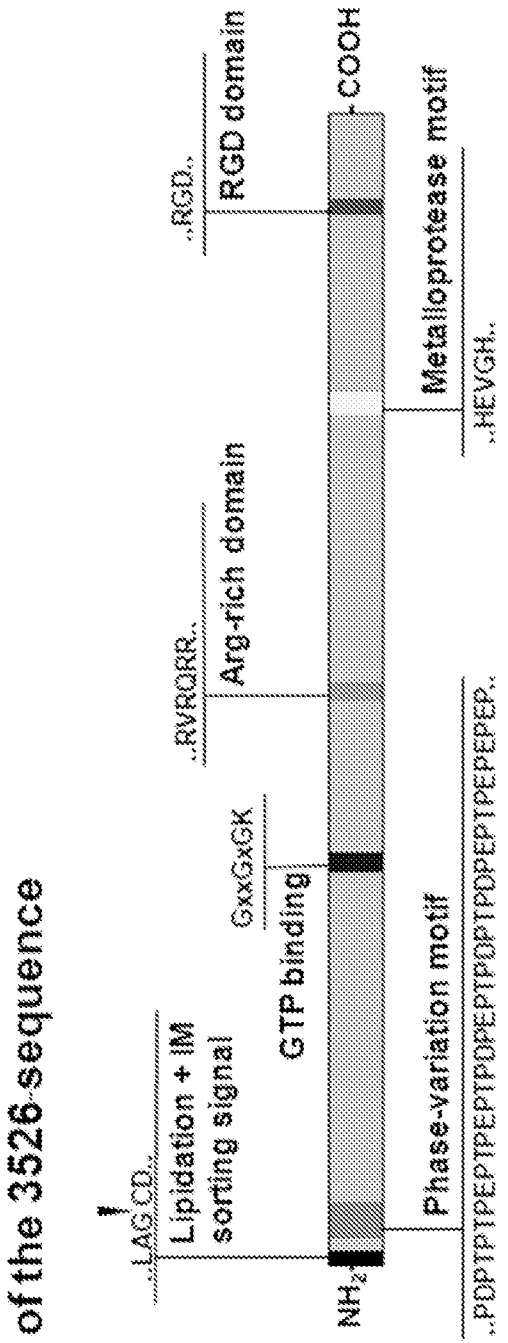

Features of the 3526 sequence
..LAGCD..
Lipidation + IM sorting signal
GxxGxGK
GTP binding
..RVRQRR..
Arg-rich domain
..RGD..
RGD domain
COOH
NH2
Phase-variation motif
..PDPTPEPTPEPTPDPEPTPDPTPDPEPTPEPEPEP..
Metalloprotease motif
..HEVGH..

Fig. 8

Isoforms of 3526

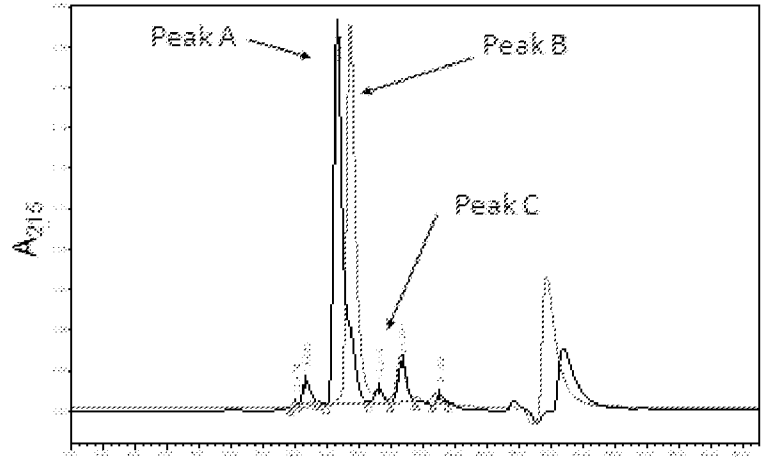


Fig. 9(a)

Peak	Ret. Time	MW by SEC	MW by MALLS	Status
A	16.6	380 KDa	165 KDa – 183 KDa	monomer
B	17.4	285 KDa – 290 KDa	168 KDa – 186 KDa	monomer
C	18.9 – 19.0	157 KDa – 162 KDa	121 KDa	truncated

Native 3526 coelutes with ΔG3526TL peak B

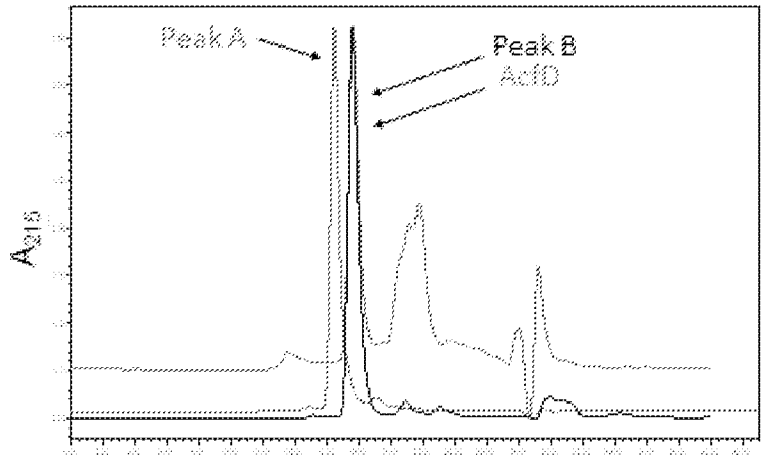


Fig. 9(b)

Peak	Ret. Time	MW by SEC	Status
A	16.5	380 KDa	monomer
B	17.6	285 KDa – 290 KDa	monomer
AcfD	17.6	285 KDa – 290 KDa	monomer

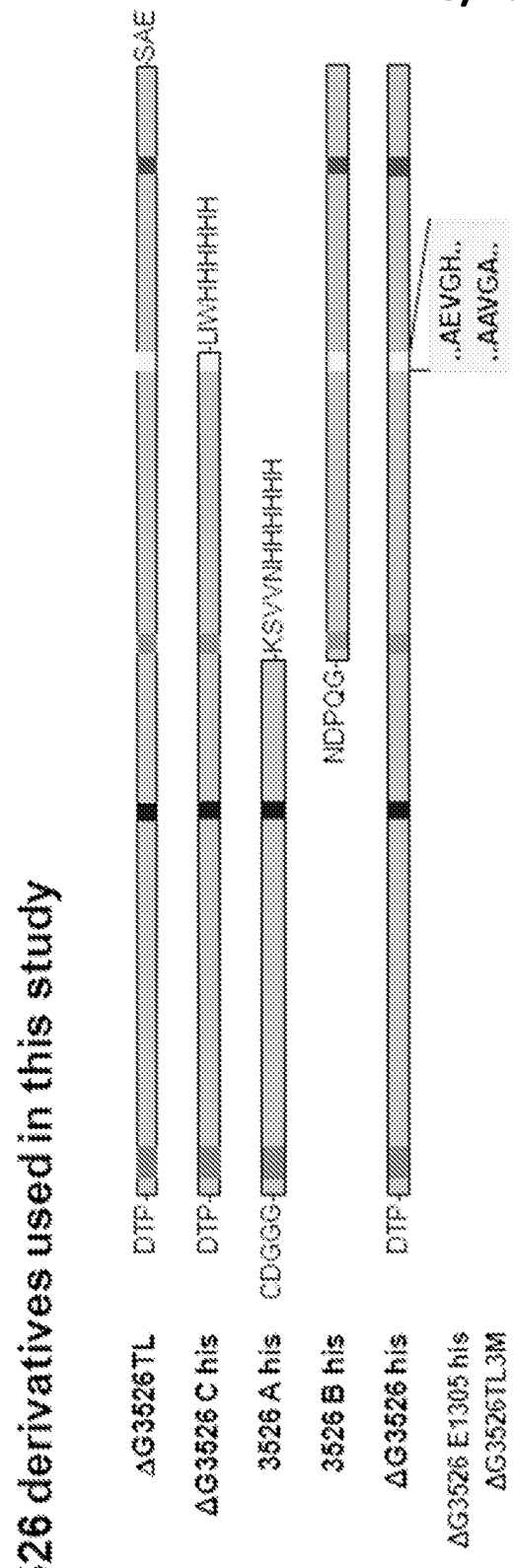

3526 derivatives used in this study
ΔG3526TL
ΔG3526 C his
3526 A his
3526 B his
ΔG3526 his
ΔG3526 E1305 his ΔG3526TL3M
DTP
SAE
LWHHHHHH
CDGSG
KSVVRHHHHHH
MDPQG
.AEVGH.
.AAVGA.

Fig. 10

Sample	MW (KDa)	Conc (µM)	Buffer	Batch	Zinc content (mol %)
ΔG3526TL	164	4.3	10mM KH_2PO_4, 150mM NaCl	ExPEC_102	**70**
ΔG3526 his	165	3	PBS+ 40% glycerol	ExPEC_115	**68**
3526 A his	81	4.9	PBS	ExPEC_092	**33**
3526 B his	86	4.4	PBS+ 40% glycerol	ExPEC_086	**5**
ΔG3526 C his	140	3.2	PBS	ExPEC_112	**18**
ΔG3526E1305-his	165	3.3	PBS+ 40% glycerol	ExPEC_099	**43**
ΔG3526TL3M	164	3.1	10mM KH_2PO_4, 150mM NaCl	ExPEC_117	**8**

Fig. 11

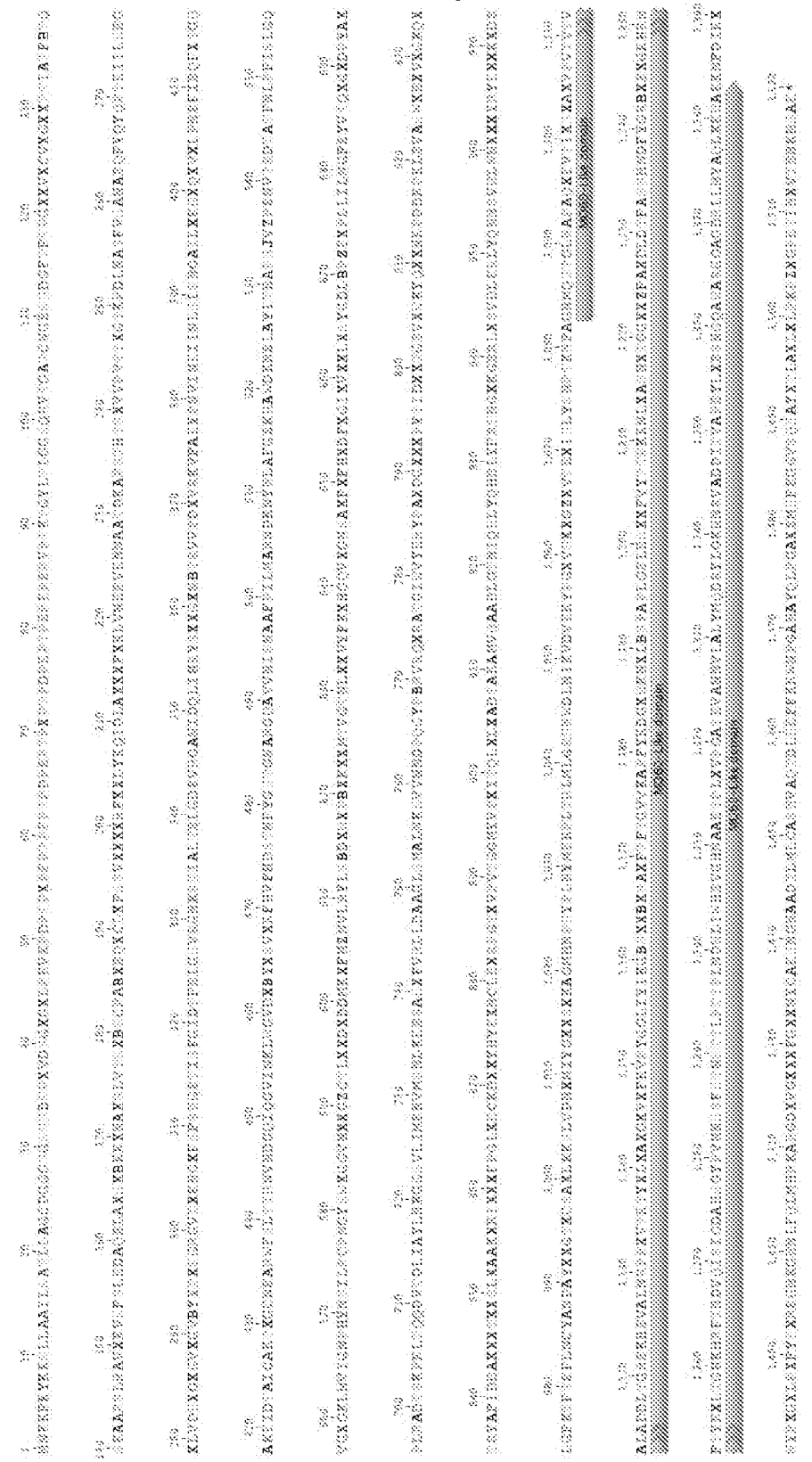


Fig. 12

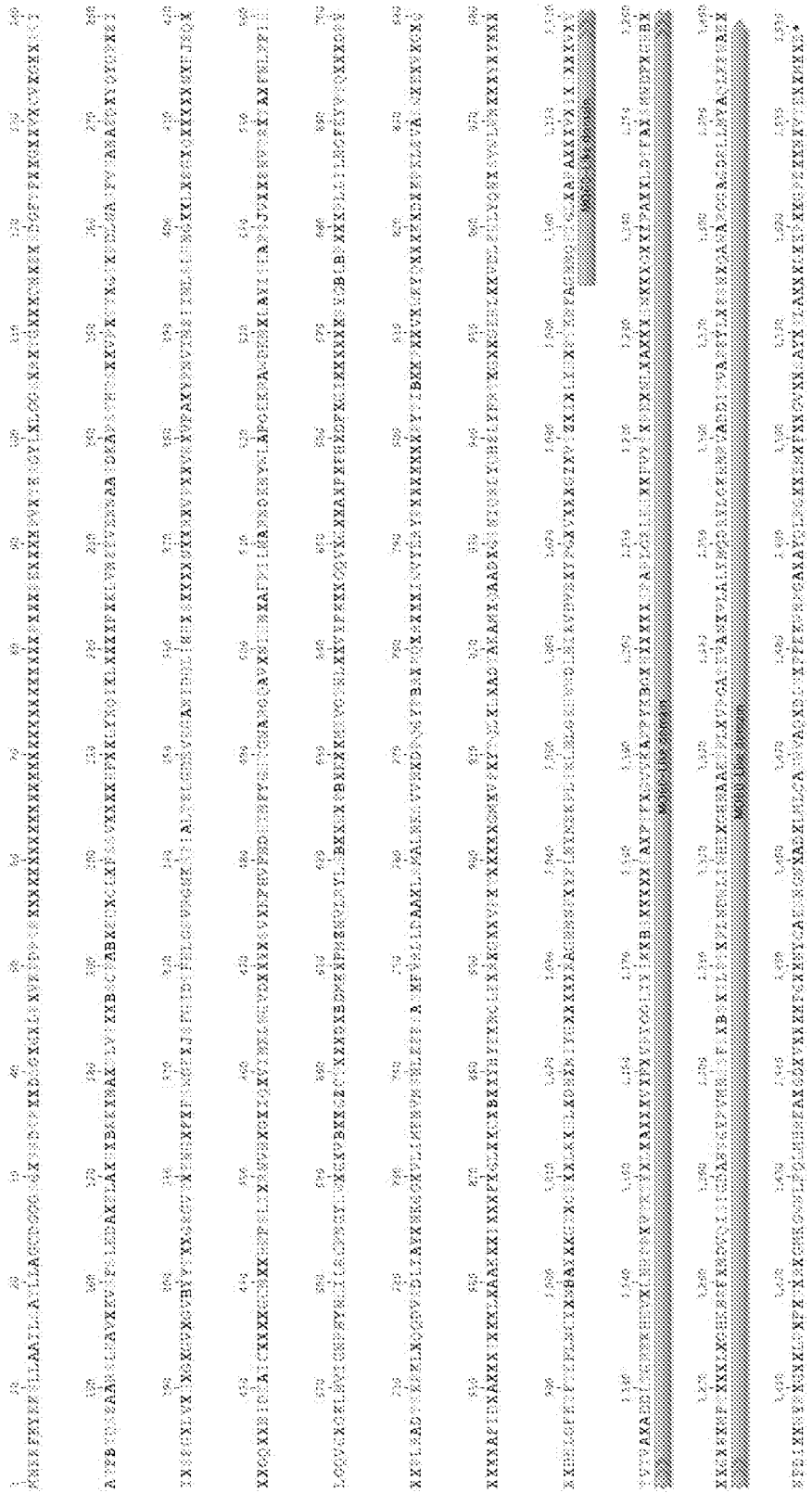


Fig. 13

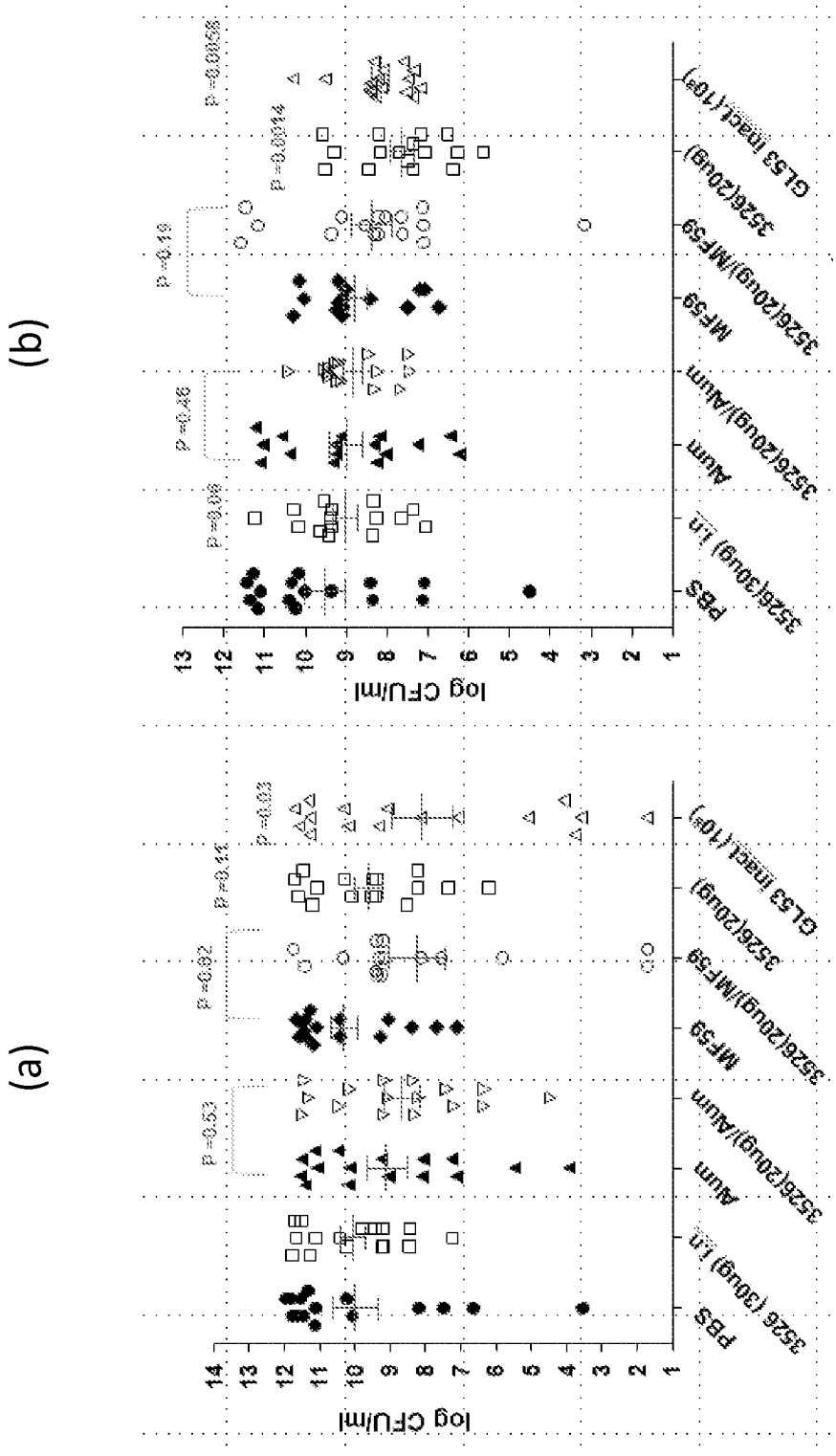

(a)
(b)
log CFU/ml
PBS
3526 (30ug) i.n
Alum
3526(20ug)/Alum
MF-59
3526(20ug)/MF-59
3526(20ug)
GLS3 Inact (10e)
p =0.53
p =0.92
p =0.11
p =0.03
p =0.16
p =0.46
p =0.06
p =0.0014
p =0.0058

Fig. 14

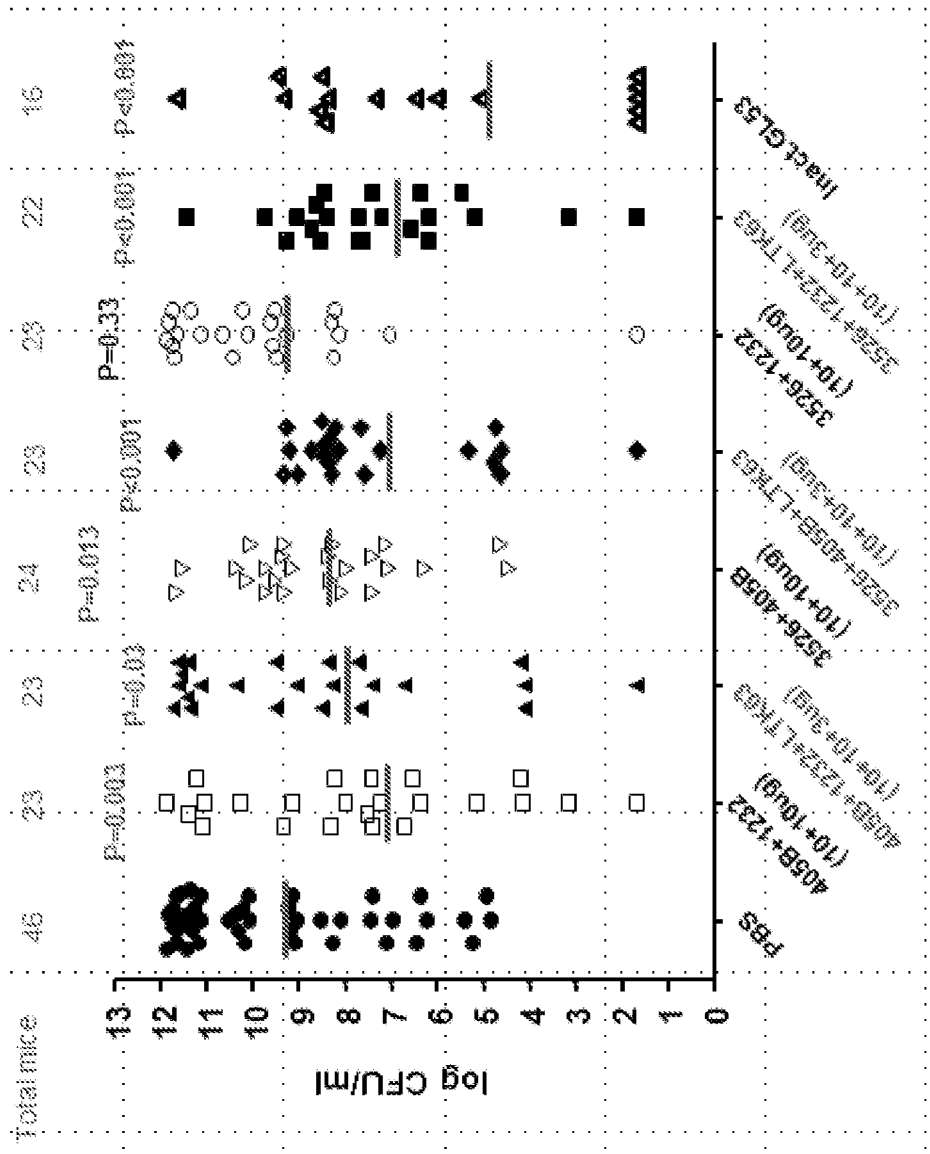

Total mice
46
23
23
24
23
23
22
16
P=0.003
P=0.03
P=0.013
P<0.001
P=0.33
P<0.001
P<0.001
log CFU/ml
13
12
11
10
9
8
7
6
5
4
3
2
1
0
PBS
405B+1232
(10+10ug)
1232+L TREG
(10+10+3ug)
3526+405B
(10+10ug)
405B+L TREG
(10+10+3ug)
3526+1232
(10+10ug)
1232+L TREG
(10+10+3ug)
Inact.GL53

Fig. 15