

Chimeric protein vaccine

Field of the invention

[0001] The invention relates to chimeric polypeptides that are useful for inducing an immune response to *P. gulae*, compositions comprising same and uses thereof for the prevention and treatment of *P. gulae* -related conditions and diseases.

Related application

[0002] This application claims priority from Australian provisional application AU 2023902373, the entire contents of which are hereby incorporated by reference.

Background of the invention

[0003] If dental plaque is left to accumulate around the tooth at the gingival (gum) margin this causes gingival inflammation (gingivitis). Chronic gingivitis can allow the emergence of a periodontal pathogen of the *Porphyromonas sp.* at the base of a periodontal pocket to result in a chronic infection and the development of severe disease. This severe form of periodontal disease is called periodontitis and can lead to tooth loss in an attempt by the immune system to eliminate the infection.

[0004] Periodontitis is an inflammatory disease of the supporting tissues of the teeth associated with a dysbiotic subgingival plaque which results in destruction of those tissues and loss of tooth attachment in humans and in companion animals. More than 80% of dogs show signs of periodontitis by age three and 70% of cats by the same age. Consequently there is a significant disease burden from periodontitis in companion animal populations.

[0005] The predominant periodontal pathogen in companion animals, particularly dogs, is *Porphyromonas gulae* (*P. gulae*).

[0006] There is currently no commercially approved treatment for use in preventing or reducing the incidence and/or severity of *P. gulae* infection or for treating *P. gulae* infection and disease in companion animals.

[0007] There is therefore a need for new and/or improved approaches for the design and manufacture of agents for treating, preventing or reducing severity of *P. gulae* infection.

[0008] Reference to any prior art in the specification is not an acknowledgment or suggestion that this prior art forms part of the common general knowledge in any jurisdiction or that this prior art could reasonably be expected to be understood, regarded as relevant, and/or combined with other pieces of prior art by a skilled person in the art.

Summary of the invention

[0009] The present invention provides a chimeric or fusion protein for inducing an immune response to *P. gulae*, the protein comprising a first polypeptide and a second polypeptide, wherein:

A) the first polypeptide comprises or consists of an amino acid sequence of the active site of an Arg- or Lys-gingipain homologue of *P. gulae*, or a sequence that is at least 80% identical thereto; and

B) the second polypeptide comprises or consists of:

the amino acid sequence of a DUF2436 domain of the Arg- or Lys-gingipain or hemagglutinin surface complexes of *P. gulae*, preferably as set forth in SEQ ID NO: 3 or 4, or a sequence at least 80% identical thereto; and

the amino acid sequence of an adhesin domain also found in surface complexes of the Arg- and Lys-gingipain homologues of *P. gulae*, preferably wherein the adhesin domain comprises the amino acid sequence of at least SEQ ID NO: 86 and/or SEQ ID NO: 85 (such as the sequence of SEQ ID NO: 88), or a sequence at least 80% identical thereto, more preferably wherein the adhesin domain comprises the amino acid sequence of SEQ ID NO: 20, or a sequence at least 80% identical thereto.

[0010] In further embodiments, the chimeric or fusion protein comprises one or more further polypeptides that comprise or consist of an amino acid sequence of the active site of an Arg- or Lys-gingipain homologue of *P. gulae*, or sequences that are at least 80% identical thereto. The one or more further polypeptides comprising or consisting of the active site of an Arg- or Lys-gingipain homologue of *P. gulae* may be located N-terminally to the first polypeptide, C-terminally to the first polypeptide, N-terminally to the second polypeptide or C-terminally to the second polypeptide. In certain embodiments, at least two further polypeptides that comprise or consist of an amino acid sequence of the active site of an Arg- or Lys-gingipain homologue of *P. gulae*, or sequences that are at least

80% identical thereto may be present. In such embodiments, the two further polypeptides may be located N-terminally to the second polypeptide, C-terminally to the second polypeptide, or N and C terminally to the second polypeptide.

[0011] The one or more further polypeptides may be linked to the first or second polypeptide of the chimeric or fusion protein, preferably via a linker of no more than 50 amino acids, or directly linked to the first or second polypeptide.

[0012] In any embodiment, the first polypeptide comprises or consists of an amino acid sequence selected from the group of: SEQ ID NOs: 1 or 2 sequences at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical thereto.

[0013] The one or more further polypeptides preferably comprise or consist of an amino acid sequence selected from the group of: SEQ ID NOs: 1 or 2, or sequences at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical thereto.

[0014] In any embodiment, the first polypeptide and the further polypeptide that comprises or consists of an amino acid sequence of the active site of an Arg- or Lys-gingipain homologue of *P. gulae*, comprises or consists of an amino acid sequence that is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to each other, or wherein the amino acid sequences are identical to each other. The first polypeptide and the further polypeptide may be derived from the active site of heterologous gingipains (e.g., from gingipains of different strains of *P. gulae*). The first polypeptide and the further polypeptide may have an amino acid sequence that is derived from different gingipains (e.g., wherein one of the polypeptides has an amino acid sequence of an active site from a Kgp and the other polypeptide has an amino acid sequence of an active site from an Rgp; or alternatively wherein one of the polypeptides has an amino acid sequence of an active site from a RgpA and the other polypeptide has an amino acid sequence of an active site from an RgpB).

[0015] In any embodiment, where the first polypeptide and/or the further polypeptide comprises or consists an amino acid sequence of the active site of an Arg- gingipain homologue of *P. gulae*, (R), such as exemplified in SEQ ID NO: 2, the amino acid

sequence may comprise substitution of the cysteine residue (located at position 5 in the sequence of SEQ ID NO: 2). Optionally, the amino acid substitution may be to a serine, valine or alanine residue.

[0016] In any embodiment, the amino acid sequence of the adhesin domain (found in a surface complex of the Arg- and Lys- gingipain homologues of *P. gulae*) further comprises one or more amino acid substitutions selected from:

- a) one or more cysteine amino acid substitutions, compared to the naturally occurring *P. gulae* sequences in corresponding regions;
- b) substitution of the proline and/or an asparagine residues in the sequence PxxN corresponding to, or at a position equivalent to, residues 6 to 9 of SEQ ID NO: 85 (equivalent to residues 68 to 71 of the sequence of SEQ ID NO: 20 or residues 68 to 71 of SEQ ID NO: 88);
- c) substitution of the motif NxFa to SxYQ in the sequence, corresponding to, or at a position equivalent to residues 2 to 5 of SEQ ID NO: 85 (equivalent to residues 64 to 67 of the sequence of SEQ ID NO: 20 or residues 64-67 of SEQ ID NO: 88)
- d) substitution of the tyrosine residue, corresponding to or at a position equivalent to residues at position 10 of SEQ ID NO: 86, and of the tryptophan residue, corresponding to or at a position equivalent to residue at position 23 of SEQ ID NO: 85, to alanine residues (equivalent to the tyrosine at residue position 10 and the tryptophan at residue position 85 of SEQ ID NO: 20).

[0017] The substitution of the cysteine amino acid residues in the adhesin domain (eg an adhesin domain as set forth in SEQ ID NO: 88 or 20), may be a substitution to a serine residue or a valine residue. Preferably, the one or more cysteine substitutions may comprise one or more substitutions to a serine residue.

[0018] Optionally, only one cysteine residue is substituted. In other embodiments, two cysteine residues are substituted. In certain embodiments, the cysteine residues are substituted to a combination of valine and serine residues. In other embodiments, all substituted cysteine residues are substituted to serine or all substitute cysteine residues are substituted to valine.

[0019] Preferably, the adhesin domain comprises or consists of the sequence set forth in any one of SEQ ID NOs: 20, 22, 23 or 24 or a sequence at least 80% identical thereto, wherein one or more cysteine residues are substituted to a serine or valine residue.

[0020] In any embodiment, the adhesin domain may comprise a substitution of the proline residue and/or a substitution of the asparagine, in the motif PxxN corresponding, or at a position equivalent to positions 68 to 71 of the sequence of SEQ ID NO: 20.

[0021] The proline amino acid substitution may be a substitution to an alanine residue.

[0022] The asparagine amino acid substitution may be a substitution to a proline residue or an alanine residue. Preferably the asparagine residue is substituted to a proline residue. In other embodiments, the asparagine residue is not substituted.

[0023] In particularly preferred embodiments, the motif PxxN in the adhesin domain, is substituted to AxxP (eg AVQP, SEQ ID NO: 64), for example, such that the adhesin domain comprises an amino acid sequence as set forth in SEQ ID NO: 21, or a sequence at least 80% identical thereto, wherein the sequence comprises AxxP at a position equivalent to positions 68 to 71 of the sequence of SEQ ID NO: 21.

[0024] In a particularly preferred embodiment, the amino acid sequence of the adhesin domain comprises one or both of:

- a) one or more cysteine amino acid substitutions, compared to the naturally occurring *P. gulae* Arg- or Lys-gingipain homologue sequences in corresponding regions;
- b) substitution of the proline and/or an asparagine residues in the sequence PxxN corresponding to, or at a position equivalent to, residues 68 to 71 of the sequence of SEQ ID NO: 20.

[0025] In especially preferred embodiments, the amino acid sequence of the adhesin domain comprises:

- a) substitution of the two cysteine amino acid residues of the adhesin A domain compared to the A domain found naturally occurring in the *P. gulae* Arg- and Lys-gingipain protein complex sequences, in corresponding regions;

b) substitution of the proline and/or an asparagine residues in the sequence PxxN corresponding to, or at a position equivalent to, residues 68 to 71 of the sequence of SEQ ID NO: 20.

[0026] Accordingly, the adhesin domain may comprise an amino acid sequence as set forth in any one of SEQ ID NOs: 24 to 27, or a sequence at least 80% identical thereto, wherein the sequence comprises AxxP at a position equivalent to positions 68 to 71 of the sequence of SEQ ID NO: 21, and wherein one or more cysteine residues are substituted to a serine or valine residue.

[0027] As noted above, the second polypeptide preferably comprises an amino acid sequence of a DUF2436 domain of the Arg- and Lys-gingipain surface complexes of *P. gulae* (such as defined in SEQ ID NOs: 3 and 4).

[0028] In any embodiment, the second polypeptide comprises an amino acid sequence of SEQ ID NO: 4, or a sequence at least 80% identical thereto.

[0029] In any embodiment, the second polypeptide comprises an amino acid sequence of a DUF2436 domain as set forth in SEQ ID NO: 3, optionally, further comprising one or more amino acid substitutions. Preferably the one or more amino acid substitutions are one or more substitutions of the cysteine residues in the domain.

[0030] In any embodiment, only one cysteine residue is substituted. Alternatively, any two, three or all four cysteine residues in the DUF domain are substituted.

[0031] The cysteine residues in the DUF domain may be substituted to any suitable amino acid residue for reducing the likelihood of disulphide bond formation between other cysteine residues in the chimeric or fusion protein. In preferred embodiments, the substitution may be to a valine residue, serine residue or an alanine residue. Optionally, the one or more cysteine residues are substituted to one or more serine residues. Optionally, the one or more cysteine residues are substituted to one or more alanine residues.

[0032] Exemplary amino acid sequences of DUF domains comprising one or more cysteine residue substitutions are set forth in SEQ ID NOs: 5 to 19. Accordingly, in any embodiment, the chimeric or fusion protein of the invention comprises a second

polypeptide comprising a DUF2436 amino acid sequence as set forth in any of SEQ ID NO: 5 to 19.

[0033] In accordance with the various embodiments disclosed above, it will be appreciated that the invention provides for various chimeric or fusion proteins for generating an immune response to *P. gulae*. Furthermore it will be appreciated that preferably the chimeric or fusion protein comprises a “core” structure comprising:

- an active site domain (eg as set for in either SEQ ID NO: 1 or 2, also referred to herein as a “K” domain or “R” domain, respectively to denote derivation from the active site from an Lys- or Arg-gingipain, respectively, of *P. gulae*)
- a DUF2436 domain (D) as herein defined; and
- an adhesin domain (A) as herein defined

wherein the domains may comprise further modifications as defined herein and optionally further comprising one or more additional active site domains.

[0034] Accordingly, in one embodiment, the chimeric or fusion protein comprises the amino acid sequence as set forth in any of SEQ ID NOs: 32 to 35, or a functional variant thereof having at least 80% identity thereto.

[0035] Alternatively, the chimeric or fusion protein may comprise the amino acid sequence as set forth in any of SEQ ID NOs: 77 to 82, or a functional variant thereof having at least 80% identity thereto.

[0036] In further embodiments, the chimeric or fusion protein (including one having an amino acid sequence of any one of SEQ ID NO: 32 to 35 or SEQ ID NOs: 77 to 82) further comprises one or more amino acid substitutions to the adhesin domain, as set forth herein (such as a) one or more cysteine amino acid substitutions, compared to the naturally occurring *P. gulae* Arg- or Lys-gingipain homologue sequences in corresponding regions; and/or b) substitution of the proline and/or an asparagine residues in the sequence PxxN corresponding to, or at a position equivalent to, residues 68 to 71 of the sequence of SEQ ID NO: 20). Accordingly, in certain embodiments, wherein the chimeric fusion protein comprises a DUF domain derived from SEQ ID NO: 4, the chimeric or fusion protein comprises an amino acid as set forth in any of SEQ ID NOs: 36 to 45 or a functional variant thereof having at least 80% identity thereto. In still further embodiments, wherein

the chimeric or fusion protein comprises a DUF domain derived from the amino acid sequence as set forth in SEQ ID NO: 3, the chimeric or fusion protein comprises an amino acid as set forth in any of SEQ ID NOs: 46 to 76 or a functional variant thereof having at least 80% identity thereto.

[0037] In especially preferred embodiments, the chimeric or fusion protein comprises or consists of the amino acid sequence of SEQ ID NO: 28 or 29 or functional variants thereof having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% identity thereto and preferably wherein the functional variant comprises the cysteine substitutions in the adhesin domain and the AVQP (SEQ ID NO: 64) substitutions as defined herein).

[0038] In any embodiment of the invention, the chimeric or fusion protein consists or consists essentially of the sequences of the first and second polypeptides as defined herein. It will be appreciated therefore that the chimeric or fusion proteins comprise an arrangement or configuration of domains that differs to the configuration of those domains in naturally occurring gingipain polypeptide sequences. In other words, the first and second polypeptides and domains therein have a differential spatial configuration to naturally occurring gingipain polypeptides.

[0039] In any embodiment of any aspect of the invention, the first and second polypeptides are linked. The first and second polypeptides may be linked directly, via a linker, or via a polypeptide sequence of no more than 100, preferably no more than 50 amino acids. Preferably the first and second polypeptides are directly linked, or linked by no more than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 amino acids. Most preferably, the first and second polypeptides are directly linked.

[0040] In any embodiment, the C-terminal residue of the first polypeptide may be linked to the N-terminal residue of the second polypeptide, directly, via a linker, or via a polypeptide sequence of no more than 50 amino acids. Alternatively, the N-terminal residue of the first polypeptide may be linked to the C-terminal residue of the second polypeptide, directly, via a linker, or via a polypeptide sequence of no more than 50 amino acids.

[0041] In any embodiment, the DUF2436 domain and adhesin domains derived from an Arg- or Lys-gingipain, (or from a hemagglutinin) and may be directly linked or joined via

a linker, or via a polypeptide sequence. Preferably, the DUF2436 and adhesin domains are linked via a short linker sequence comprising about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 amino acids. In a particularly preferred embodiment, the DUF2436 and adhesin domains are linked via a short linker sequence of no more than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15 amino acids.

[0042] In some embodiments, the DUF2436 and adhesin domains (with representative amino acid sequences of SEQ ID NOs: 3 and 4 and 20, respectively (or comprising various amino acid substitutions as described herein) may be linked via a peptide or polypeptide linker, as described elsewhere herein.

[0043] In any embodiment, the further polypeptide may be joined directly or via a linker to the chimeric or fusion protein comprising the first and second polypeptides. In embodiments where the further polypeptide is joined to the C terminal region of the fusion protein via the second polypeptide, preferably the C terminus of the second polypeptide is joined directly to the N terminus of the further polypeptide. (For example, the C terminus of the adhesin domain is preferably joined directly to the N terminus of the active site amino acid sequence). In instances where more than one further polypeptide is included, the copies of the further polypeptide may be directly joined to each other or joined via a linker sequence.

[0044] In any embodiment of any aspect of the invention, the linker region is an amino acid sequence of no more than 15 amino acids, and preferably greater than 2 amino acids. Suitable linkers for use in protein constructs, including those with minimal impact on solubility are known in the art. Useful linkers include glycine-serine (GlySer) and glycine-threonine (GlyThr) linkers, which are well-known in the art and comprise glycine and serine units combined in various orders. Examples include, but are not limited to, (GS), (GSGGS)_n (SEQ ID NO: 102), (GGGS)_n (SEQ ID NO: 103) and (GGGGS)_n (SEQ ID NO: 104), where *n* is an integer of at least one, typically an integer between 1 and about 10, for example, between 1 and about 8, between 1 and about 6, or between 1 and about 5. Other useful linkers include DSSG (SEQ ID NO: 105), DSSGAS (SEQ ID NO: 106), KLDSSG (SEQ ID NO: 107) or others described herein. In certain embodiments, the linker region may be derived from the native gingipain polyprotein sequence (such as the sequence PNGT, SEQ ID NO: 101).

[0045] The present invention also provides a nucleic acid encoding any chimeric or fusion protein as defined herein.

[0046] Preferably, the nucleic acid has a nucleotide sequence that encodes any one or more of the amino acid sequences defined in Table 1 herein.

[0047] Optionally, the nucleic acid has a nucleotide sequence as set forth in Table 2 herein (eg comprising the nucleic acid sequences of any one of SEQ ID NOs: 83, 84 and 89 to 96).

[0048] In any embodiment, such a nucleic acid is included in an expression construct in which the nucleic acid is operably linked to a promoter. Such an expression construct can be in a vector, e.g., a plasmid, or viral vector.

[0049] The present invention also provides a cell comprising a nucleic acid, or nucleic acid vector as herein described. Examples of cells of the present invention include bacterial cells, yeast cells, insect cells or mammalian cells. Preferably, the cell is isolated, substantially purified or recombinant.

[0050] The present invention also provides a composition comprising a chimeric or fusion protein as described herein, optionally in combination with a pharmaceutically acceptable carrier.

[0051] The composition may also comprise an adjuvant for potentiating an immune response to the chimeric or fusion protein.

[0052] The present invention accordingly further provides for a vaccine or immune stimulating composition for inducing an immune response to *P. gulae* in a subject, the composition comprising:

- an immunogen in the form of a chimeric or fusion protein as described herein, and
- an adjuvant, for potentiating the immune response to the immunogen in the subject.

[0053] Preferably, the sole immunogen provided in the compositions, vaccines or immune stimulating compositions of the invention, is a chimeric or fusion protein as herein described.

[0054] The present invention also provides a method for inducing an immune response in a subject to *P. gulae*, the method comprising administering to a subject in need thereof, a chimeric or fusion protein, vaccine or immune stimulating composition as described herein.

[0055] The present invention also provides a method of inducing a humoral immune response to *P. gulae* in a subject, the method comprising administering to the subject, a chimeric or fusion protein, composition, vaccine or immune stimulating composition as herein defined.

[0056] Preferably the immune response that is induced comprises a switch from a Th1 to a Th2 immune response.

[0057] The present invention also provides a method of inducing an immune response to *P. gulae* in a subject, the method comprising administering to the subject a first, priming dose of a chimeric or fusion protein, composition, vaccine or immune stimulating composition as herein defined, and further comprising administering a second, booster dose, of a chimeric or fusion protein, composition, vaccine or immune stimulating composition as herein defined.

[0058] It will be appreciated that in any embodiment, the immune response elicited by administration of a chimeric or fusion protein described herein, or vaccine or other composition comprising the same, is preferably antigen-specific. Accordingly, in preferred embodiments, the methods and compositions and chimeric proteins described herein, are for inducing an immune response, preferably a protective immune response, to *P. gulae* gingipain antigens.

[0059] In any embodiment, the compositions, chimeric proteins and methods of the invention may be for strengthening an immune response (such as a protective immune response) of a subject to *P. gulae*.

[0060] The present invention also provides for methods of immunising a subject against *P. gulae* infection, the method comprising administering to the subject, a chimeric or fusion protein, composition, vaccine or immune stimulating composition as herein defined.

[0061] In any embodiment, a subject who has received or has been administered a chimeric or fusion protein of the invention, or composition or vaccine including the same, has an increased level of protection against infection with *P. gulae*, or severity of one or more symptoms of occurring *P. gulae* infection, compared to a subject who has not received the protein, composition or vaccine.

[0062] Further still, the invention provides a method of treating a *P. gulae* infection in a subject, the method comprising administering to a subject in need thereof, a chimeric or fusion protein, composition, vaccine or immune stimulating composition as herein defined, thereby treating the occurring *P. gulae* infection in the subject.

[0063] The present invention provides a method for reducing or minimising the severity of a symptom associated with an infection with *P. gulae*, comprising administering to an individual in need thereof, a chimeric or fusion protein, composition, vaccine or immune stimulating composition as herein defined, wherein the symptoms are selected from the group consisting of swollen or puffy gums, gums that bleed easily, receding gums, periodontal pockets around the teeth, loss of tooth supporting tissues (periodontal ligament, cementum and/or alveolar bone) pus between gums and teeth, and gingivitis.

[0064] In any embodiment, administration of a chimeric protein as described herein may result in a reduction in the load of *P. gulae* bacteria (ie the number of *P. gulae* cells) in the saliva of a subject administered the chimera.

[0065] In any embodiment, administration of a chimeric protein as described herein may result in a reduction in the load of *P. gulae* bacteria (ie the number of *P. gulae* cells) in subgingival plaque of a subject administered the chimera.

[0066] The present invention also provides a method for treating *P. gulae* -related disease in a subject, the method comprising administering to an individual in need thereof, a chimeric or fusion protein, composition, vaccine or immune stimulating composition as herein defined. Preferably the *P. gulae* -related disease comprises periodontal disease.

[0067] The present invention provides for a method of treatment comprising administration of a chimeric or fusion protein, composition, vaccine or immune stimulating composition of the invention and may also comprise administration of one or more of: an antimicrobial compound, an anti-inflammatory agent.

[0068] It will also be appreciated that any method or use described herein may also include a method or use which comprises administration of one or more additional immunogens, therapeutic or prophylactic agents. In any embodiment, the chimeric or fusion protein of the invention may be a component of a combination vaccine or immune stimulating composition.

[0069] The invention also provides use of a chimeric or fusion protein as herein defined, in the manufacture of a medicament for:

- inducing an immune response (preferably a protective immune response) in a subject to *P. gulae*;
- immunising a subject against *P. gulae* infection;
- treating a *P. gulae* infection in a subject;
- minimising or reducing the severity of one or more symptoms of *P. gulae* infection;
or
- treating *P. gulae* -related disease in a subject.

[0070] The invention also provides a chimeric or fusion protein, composition, vaccine or immune stimulating composition as herein defined, for use in:

- inducing an immune response (preferably a protective immune response) in a subject to *P. gulae*;
- immunising a subject against *P. gulae* infection;
- treating a *P. gulae* infection in a subject;
- minimising or reducing the severity of one or more symptoms of *P. gulae* infection;
or
- treating *P. gulae* -related disease in a subject.

[0071] In any method, use or protein, composition or vaccine for use according to the invention, the subject may be any subject that has or is at risk of having an infection with *P. gulae*. In accordance with the present invention, the subject is preferably a veterinary

subject, such as a companion animal (such as a cat or a dog) that has, or is at risk of having an infection with *P. gulae*.

[0072] The present invention also provides a method for obtaining an antibody directed to *P. gulae*, the method comprising administering a chimeric or fusion protein, composition, vaccine or immune stimulating composition of the invention, to a non-human animal, thereby generating antibodies directed to *P. gulae* in the animal. Preferably the method further comprises isolating the antibody from the animal (eg, extracting from the blood of the animal) or from an egg thereof (in the case where the animal is an avian species, preferably chicken).

[0073] The present invention also provides an antibody preparation comprising an antibody directed to *P. gulae*, wherein the antibody preparation is obtained by administering a chimeric or fusion protein, composition, vaccine or immune stimulating composition of the invention, to a non-human animal, thereby generating antibodies directed to *P. gulae* in the animal, and isolating the antibodies from the animal or egg thereof.

[0074] The antibody directed to *P. gulae* may be used therapeutically to eliminate or reduce *P. gulae* infection or prophylactically, to prevent or reduce the severity of *P. gulae* infection.

[0075] The present invention also provides a kit comprising a composition comprising a chimeric or fusion protein as herein defined, wherein optionally the kit comprises one or more cytokines and/or adjuvants in sealed containers.

[0076] Preferably, the kit comprises a label or package insert indicating that the composition is used for immunising an individual, optionally wherein the label or package insert includes instructions for use.

[0077] Throughout this specification, unless the context requires otherwise, the words “comprise,” “comprises” and “comprising” will be understood to imply the inclusion of a stated step or element or group of steps or elements but not the exclusion of any other step or element or group of steps or elements. Thus, use of the term “comprising” and the like indicates that the listed elements are required or mandatory, but that other elements are optional and may or may not be present. By “consisting of” is meant including, and limited to, whatever follows the phrase “consisting of”. Thus, the phrase

“consisting of” indicates that the listed elements are required or mandatory, and that no other elements may be present. By “consisting essentially of” is meant including any elements listed after the phrase, and limited to other elements that do not interfere with or contribute to the activity or action specified in the disclosure for the listed elements. Thus, the phrase “consisting essentially of” indicates that the listed elements are required or mandatory, but that other elements (e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or more than 20 additional amino acid residues at the N-terminus or C-terminus of a polypeptide sequence) are optional and may or may not be present depending upon whether or not they affect the activity or action of the listed elements.

[0078] As used herein, except where the context requires otherwise, the term "comprise" and variations of the term, such as "comprising", "comprises" and "comprised", are not intended to exclude further additives, components, integers or steps.

[0079] Further aspects of the present invention and further embodiments of the aspects described in the preceding paragraphs will become apparent from the following description, given by way of example and with reference to the accompanying drawings.

Brief description of the drawings

[0080] Figure 1: Expression and Solubility of *P. gulae* chimeric proteins KD_FAK-2S-AVQP and KD_AAK-2S-AVQP. (A) Lysis fractions; (B) Small scale purification tests with Ni-NTA spin column: 1. clarified lysate before loaded to the column; 2. flow through; 3. column wash; 4. eluted protein fraction. Left side panel: KD_FAK-2S-AVQP and right side panel: KD_AAK-2S-AVQP in both A) and B).

[0081] Figure 2: SDS-PAGE analysis of expressed *P. gulae* antigens and the cell lysis fractions. (A) Expression of KD_FAK-2S-AVQP and KD_AAK-2S-AVQP in LB and TB media for 2-3 hours. (B) Lysis under non-reducing conditions. (C) (i & ii) Attempt to solubilise KD_FAK-2S-AVQP from the insoluble fraction of non-reducing lysis with reducing buffer (5 mM and 100 mM DTT); (iii) Lysis under reducing conditions (10 mM DTT) for proteins KD_FAK-2S-AVQP. M: protein standards; TC: total cell lysate; Sup: clarified lysate.

[0082] Figure 3: Ni-affinity chromatography. (A) Elution profile of KD_AAK-2S-AVQP under normal non-reducing conditions and reducing SDS-PAGE of the column fractions. (B) Elution profile of antigen-F under reducing conditions (R) and reducing SDS-PAGE

of the column fractions. (C) Elution profile of KD_FAK-2S-AVQP under normal non-reducing conditions (NR) and reducing SDS-PAGE of the column fractions. TC: total cell lysate; Sp: clarified supernatant before loaded to the column; FT: flow through; W: column wash. M: protein standards.

[0083] Figure 4: Purification with anion exchange chromatography (AIEX). (A) Elution profile of KD_AAK-2S-AVQP under normal non-reducing conditions and reducing SDS-PAGE of the column fractions. (B) Elution profile of KD_FAK-2S-AVQP under reducing conditions (R) and reducing SDS-PAGE of the column fractions. (C) Elution profile of KD_FAK-2S-AVQP under normal non-reducing conditions (NR) and reducing SDS-PAGE of the column fractions. BL: before loading; FT: flow through; M: protein standards.

[0084] Figure 5: Size exclusion chromatography under non-reducing conditions (NR). (A) Elution profile KD_AAK-2S-AVQP and reducing SDS-PAGE of the column fractions. (B) Elution profile of KD_FAK-2S-AVQP being purified under reducing conditions (R) before this step and reducing SDS-PAGE of the column fractions. (C) Elution profile of KD_FAK-2S-AVQP being purified under non-reducing conditions (NR) before this step and reducing SDS-PAGE of the column fractions. BL: before loading; M: protein standards.

[0085] Figure 6: PAGE analysis of the *P. gultae* antigen final products. (A) SDS-PAGE; (B) Native PAGE or on native gels. A-version: KD_AAK-2S-AVQP; F-R: KD_FAK-2S-AVQP purified under reducing conditions until the final size exclusion step; F-NR: KD_FAK-2S-AVQP purified under non-reducing conditions. R: reducing; NR: non-reducing; M: protein standards.

[0086] Figure 7: Mouse periodontitis model: therapeutic vaccination. Experimental timeline schematic.

[0087] Figure 8: *P. gultae* induced bone loss. Statistical analysis – One-way ANOVA and post-hoc Dunnet's T3. # (p<0.05, compared to Naïve control); ## (p<0.05, compared to infected control).

[0088] Figure 9: Anti-*P. gultae* antibody isotype response. Antibody titres in mouse sera (individual) towards heat-killed *P. gultae* whole cells. (A) Total IgG Titres; (B) IgG1 Subtype titres; (C) IgG2a Subtype titres.

[0089] Figure 10: Antibody IgG response against *P. gulae* protease complex.

Antibody titres in mouse sera (individual) towards purified *P. gulae* RgpA/Kgp protease complex. (A) Total IgG Titres; (B) IgG1 Subtype titres; (C) IgG2a Subtype titres.

[0090] Figure 11: Antibody IgG anti-vaccine antigen immune response. Antibody titres in mouse sera (Individual sera) towards vaccine antigen used to immunise mice.

[0091] Figure 12: Schematic of experimental protocol for canine serology studies.

[0092] Figure 13: IgG response to immunising antigen (KD_ΔAK-2S-AVQP-6His). A) pooled sera; B) individual sera; C) analysis of individual sera; D) titration of individual sera.

[0093] Figure 14: Serological endpoint titres (2xbaseline) vs *P. gulae* whole cells.

[0094] Figure 15: Comparison of titre responses to *P. gulae* KAS2 protein following immunisation with KD_ΔAK-2S-AVQP-6His.

[0095] Figure 16: Serological midpoint titres for each serum collection time point.

For the 10 dogs in each group, the midpoint titre values were plotted for each serum collection time point. Mean and standard error of the mean (SEM) are shown for each serum collection time point. The ordinate axis represents the serum dilution required to achieve the titration midpoint. Pairs of serum time points that showed a statistical significant difference ($p < 0.05$) between the means are represented by the bars above the graph.

[0096] Figure 17: Serological endpoint titres for each serum collection time point.

For the 10 dogs in each group, the endpoint titre values were plotted for each serum collection time point. Mean and standard error of the mean (SEM) are shown for each serum collection time point. The ordinate axis represents the serum dilution required to achieve the titration endpoint. Pairs of serum time points that showed a statistical significant difference ($p < 0.05$) between the means are represented by the bars above the graph.

Sequence information

[0097] Table 1: amino acid sequence information

Descriptor	SEQ ID No	Sequence
KAS (Kgp active site domain)	1	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDK
RAS (Rgp active site domain)	2	NSGICTTNYTGHGSDTAWGTSGFNTTHMKQLANYNQ
DUF2436 F variant	3	QKMRKEVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARII LEAHDVWGDGSGYQMLIDADHTAYGNEIPPTPLGQCLTEDC NIPATLYDPFEFKVPAAADPSCTPEYQVVDGVASIDLPAgiYDY VVVNPSPLGCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQN GTSFSTTGdYVTLTVTG
DUF2436 A variant	4	QRGPKTAPSVTHQTVQKGIRTSTAKDLRDIPIAGMARII DVGQDGTGYQMLWDDHDQYGASIPEDGFWFASSETIPDGLF DPFEYKVPVNADASFTPTNFVLDGTASADIPAGTYDYVIINPNP GLIYIVGQGVSKGNDYVVEAGKTYHFTVKRQGSgDAASVVVT G
DUF2436 F Cys1 sub	5	QKMRKEVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARII LEAHDVWGDGSGYQMLIDADHTAYGNEIPPTPLGQ <u>SL</u> TEDC NIPATLYDPFEFKVPAAADPSCTPEYQVVDGVASIDLPAgiYDY VVVNPSPLGCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQN GTSFSTTGdYVTLTVTG
DUF2436 F Cys2 sub	6	QKMRKEVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARII LEAHDVWGDGSGYQMLIDADHTAYGNEIPPTPLGQCLTED <u>S</u> NIPATLYDPFEFKVPAAADPSCTPEYQVVDGVASIDLPAgiYDY VVVNPSPLGCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQN GTSFSTTGdYVTLTVTG
DUF2436 F Cys3 sub	7	QKMRKEVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARII LEAHDVWGDGSGYQMLIDADHTAYGNEIPPTPLGQCLTEDC NIPATLYDPFEFKVPAAADPS <u>S</u> TPEYQVVDGVASIDLPAgiYDY

		VVNPSPGLCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQN GTSFSTTGDYVTLTVTG
DUF2436 F Cys4 sub	8	QKMRKEVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARII LEAHDVWGDGSGYQMLIDADHTAYGNEIPPTPLGQCLTEDC NIPATLYDPFEFKVPAAADPSCTPEYQVVDGVASIDLPAgiYDY VVNPSPGL <u>S</u> IVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQN GTSFSTTGDYVTLTVTG
DUF2436 F Cys1&2 sub	9	QKMRKEVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARII LEAHDVWGDGSGYQMLIDADHTAYGNEIPPTPLGQ <u>SL</u> TEDC NIPATLYDPFEFKVPAAADPSCTPEYQVVDGVASIDLPAgiYDY VVNPSPGLCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQN GTSFSTTGDYVTLTVTG
DUF2436 F Cys1&3 sub	10	QKMRKEVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARII LEAHDVWGDGSGYQMLIDADHTAYGNEIPPTPLGQ <u>SL</u> TEDC NIPATLYDPFEFKVPAAADPS <u>S</u> TPEYQVVDGVASIDLPAgiYDY VVNPSPGLCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQN GTSFSTTGDYVTLTVTG
DUF2436 F Cys1&4 sub	11	QKMRKEVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARII LEAHDVWGDGSGYQMLIDADHTAYGNEIPPTPLGQ <u>SL</u> TEDC NIPATLYDPFEFKVPAAADPSCTPEYQVVDGVASIDLPAgiYDY VVNPSPGL <u>S</u> IVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQN GTSFSTTGDYVTLTVTG
DUF2436 F Cys2&3 sub	12	QKMRKEVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARII LEAHDVWGDGSGYQMLIDADHTAYGNEIPPTPLGQCLTED <u>S</u> NIPATLYDPFEFKVPAAADPS <u>S</u> TPEYQVVDGVASIDLPAgiYDY VVNPSPGLCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQN GTSFSTTGDYVTLTVTG
DUF2436 F Cys2&4 sub	13	QKMRKEVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARII LEAHDVWGDGSGYQMLIDADHTAYGNEIPPTPLGQCLTED <u>S</u> NIPATLYDPFEFKVPAAADPSCTPEYQVVDGVASIDLPAgiYDY VVNPSPGL <u>S</u> IVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQN GTSFSTTGDYVTLTVTG
DUF2436 F Cys3&4 sub	14	QKMRKEVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARII LEAHDVWGDGSGYQMLIDADHTAYGNEIPPTPLGQCLTEDC NIPATLYDPFEFKVPAAADPS <u>S</u> TPEYQVVDGVASIDLPAgiYDY

		VVNPSPGLSIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQN GTSFSTTGDYVTLTVTG
DUF2436 F Cys1&2&3 sub	15	QKMRKEVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARII LEAHDVWGDGSGYQMLIDADHTAYGNEIPPTPLGQSLTEDS NIPATLYDPFEFKVPAAADPSSTPEYQVVDGVASIDLPAgiYDY VVNPSPGLCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQN GTSFSTTGDYVTLTVTG
DUF2436 F Cys1&2&4 sub	16	QKMRKEVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARII LEAHDVWGDGSGYQMLIDADHTAYGNEIPPTPLGQSLTEDS NIPATLYDPFEFKVPAAADPSCSTPEYQVVDGVASIDLPAgiYDY VVNPSPGLSIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQN GTSFSTTGDYVTLTVTG
DUF2436 F Cys1&3&4 sub	17	QKMRKEVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARII LEAHDVWGDGSGYQMLIDADHTAYGNEIPPTPLGQSLTEDC NIPATLYDPFEFKVPAAADPSSTPEYQVVDGVASIDLPAgiYDY VVNPSPGLSIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQN GTSFSTTGDYVTLTVTG
DUF2436 F Cys2&3&4 sub	18	QKMRKEVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARII LEAHDVWGDGSGYQMLIDADHTAYGNEIPPTPLGQCLTEDS NIPATLYDPFEFKVPAAADPSSTPEYQVVDGVASIDLPAgiYDY VVNPSPGLSIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQN GTSFSTTGDYVTLTVTG
DUF2436 F Cys1&2&3&4 sub	19	QKMRKEVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARII LEAHDVWGDGSGYQMLIDADHTAYGNEIPPTPLGQSLTEDS NIPATLYDPFEFKVPAAADPSSTPEYQVVDGVASIDLPAgiYDY VVNPSPGLSIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQN GTSFSTTGDYVTLTVTG
adhesin domain (ABM2+1+3)	20	DGGSSADYTYTVYRDGTKIKEGLTATTFEEEGVAPGNHEYCV EVKYQAGVSPKVCKGVTINGGNEFAPVQNLTGSA DGQKVTLK WDAPNGTNPNDPNPNPSL
Adhesin domain + AVQP	21	DGGSSADYTYTVYRDGTKIKEGLTATTFEEEGVAPGNHEYCV EVKYQAGVSPKVCKGVTINGGNEFAAVQPLTGSA DGQKVTLK WDAPNGTNPNDPNPNPSL
Adhesin domain + Cys1 substitution	22	DGGSSADYTYTVYRDGTKIKEGLTATTFEEEGVAPGNHEY[S/ V]VEVKYQAGVSPKVCKGVTINGGNEFAPVQNLTGSA DGQKV TLKWDAPNGTNPNDPNPNPSL

Adhesin domain + Cys2 substitution	23	DGGSSADYTYTVYRDGTKIKEGLTATTFFFFEGVAPGNHEYCV EVKYQAGVSPKV[S/V]KGVTINGGNEFAPVQNLTGSDGQKV TLKWDAPNGTNPNDPNPDNPNSL
Adhesin domain + Cys1&2 substitution	24	DGGSSADYTYTVYRDGTKIKEGLTATTFFFFEGVAPGNHEY[S/ V]VEVKYQAGVSPKV[S/V]KGVTINGGNEFAPVQNLTGSDGQ KVTLKWDAPNGTNPNDPNPDNPNSL
Adhesin domain + Cys1 substitution + AVQP	25	DGGSSADYTYTVYRDGTKIKEGLTATTFFFFEGVAPGNHEYSV EVKYQAGVSPKVCKGVTINGGNEFAAVQPLTGSDGQKVTLK WDAPNGTNPNDPNPDNPNSL
Adhesin domain + Cys2 substitution + AVQP	26	DGGSSADYTYTVYRDGTKIKEGLTATTFFFFEGVAPGNHEYCV EVKYQAGVSPKVSKGVTINGGNEFAAVQPLTGSDGQKVTLK WDAPNGTNPNDPNPDNPNSL
Adhesin domain + Cys1&2 substitution + AVQP	27	DGGSSADYTYTVYRDGTKIKEGLTATTFFFFEGVAPGNHEYSV EVKYQAGVSPKVSKGVTINGGNEFAAVQPLTGSDGQKVTLK WDAPNGTNPNDPNPDNPNSL
KD _F AK-2S- AVQP	28	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQKMRK EVSPARAKTLMKQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQCLTEDCNIPAT LYDPFEFKVPAAADPSCTPEYQVVDGVASIDLPAIYDYVVVN PSPGLCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF STTG DYVTLTVTGDGSSADYTYTVYRDGTKIKEGLTATTFFEE EGVAPGNHEYSEVEVKYQAGVSPKVSKGVTINGGNEFAAVQP LTGSADGQKVTLKWDAPNGTNPNDPNPDNPNSLNTGVSFV NYTAHGSETSWADPSLTTSQKALTNKDK
KD _A AK-2S- AVQP	29	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQRGPKT APSVTHQTVQKGIRTSTAKDLRDPAGMARIILEAHDVWQDG TGYQMLWDDHDQYGASIPEDGFWFASETIPDGLFDPFEYK VPVNADASFTPTNFVLDGTASADIPAGTYDYVIINPNPGLIYVG QGVSKGNDYVVEAGKTYHFTVKRQSGDAASVVVTGDGGS SADYTYTVYRDGTKIKEGLTATTFFFFEGVAPGNHEYSEVEKY QAGVSPKVSKGVTINGGNEFAAVQPLTGSDGQKVTLKWDA PNGTNPNDPNPDNPNSLNTGVSFVNYTAHGSETSWADPSL TTSQKALTNKDK

KD _F AK-2S-AVQP (expressed product)	30	MANTGVSVFN ^Y TAHGSETSWADPSLTTSQ ^L KALTNKDKQKM RKEVSPARAKTLM ^L KQKTEPSDLSKFRDEPIPGGMARIILEA HDVWGDGSGYQMLIDADHTAYGNEIPPTPLGQCLTEDCNIP ATLYDPFEFKVPAAADPSCTPEYQVVDGVASIDLPA ^G IYDYVV VNPSGLCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNG TSFSTTG ^D YVTLTVTGDGGSSADYTYTVYRDGT ^K IKIEGLTATT FEEEGVAPGNHEYS ^U VEVKYQAGVSPKV ^S KGVTINGGNEFA ^A <u>VQ</u> <u>PL</u> TGSADGQKVTLKWDAPNGTNPNDPNPDPNPSLNTGVS FVNYTAHGSETSWADPSLTTSQ ^L KALTNKDKLEHHHHHH
KD _A AK-2S-AVQP (expressed product)	31	MANTGVSVFN ^Y TAHGSETSWADPSLTTSQ ^L KALTNKDKQRG PKTAPSVTHQT ^V QKGIRTSTAKDLRDP ^I PAGMARIILEAHDVW QDGTGYQMLWDDDDHDQYGASIPEDGFWFAS ^E ETIPDGLFDPF EYKVPVNADASFTPTNFVLDGTASADIPAGTYDYVIINPNPGLI YIVGQGVSKGNDYVVEAGKTYHFTV ^K RQGS ^G DAASVVVTGD GGSSADYTYTVYRDGT ^K IKIEGLTATT ^F EEEGVAPGNHEYS ^U VE VKYQAGVSPKV ^S KGVTINGGNEFA ^A <u>VQ</u> <u>PL</u> TGSADGQKVTLK WDAPNGTNPNDPNPDPNPSLNTGVSFVNYTAHGSETSWA DPSLTTSQ ^L KALTNKDKLEHHHHHH
KD _A A	32	NTGVSVFN ^Y TAHGSETSWADPSLTTSQ ^L KALTNKDKQRGPKT APSVTHQT ^V QKGIRTSTAKDLRDP ^I PAGMARIILEAHDVWQDG TGYQMLWDDDDHDQYGASIPEDGFWFAS ^E ETIPDGLFDPFEYK VPVNADASFTPTNFVLDGTASADIPAGTYDYVIINPNPGLIYIVG QGVSKGNDYVVEAGKTYHFTV ^K RQGS ^G DAASVVVTGDGGS SADYTYTVYRDGT ^K IKIEGLTATT ^F EEEGVAPGNHEYC ^V EVKY QAGVSPKVCKGVTINGGNEFAPVQNL ^T GSADGQKVTLKWDA PNGTNPNDPNPDPNPSL
KD _A AK	33	NTGVSVFN ^Y TAHGSETSWADPSLTTSQ ^L KALTNKDKQRGPKT APSVTHQT ^V QKGIRTSTAKDLRDP ^I PAGMARIILEAHDVWQDG TGYQMLWDDDDHDQYGASIPEDGFWFAS ^E ETIPDGLFDPFEYK VPVNADASFTPTNFVLDGTASADIPAGTYDYVIINPNPGLIYIVG QGVSKGNDYVVEAGKTYHFTV ^K RQGS ^G DAASVVVTGDGGS SADYTYTVYRDGT ^K IKIEGLTATT ^F EEEGVAPGNHEYC ^V EVKY QAGVSPKVCKGVTINGGNEFAPVQNL ^T GSADGQKVTLKWDA PNGTNPNDPNPDPNPSLNTGVSVFN ^Y TAHGSETSWADPSL TTSQ ^L KALTNKDK

KD _F A	34	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQKMRK EVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQCLTEDCNIPAT LYDPFEFKVPAAADPSCTPEYQVVDGVASIDLPAgiYDYVVVN PSPGLCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF STTG DYVTLTVTGDGGSSADYTYTVYRDGTKIKEGLTATT FEE EGVAPGNHEYCVEVKYQAGVSPKVCKGVTINGGNEFAPVQN LTGSADGQKVTLKWDAPNGTNPNDPNPDPNPDSL
KD _F AK	35	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQKMRK EVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQCLTEDCNIPAT LYDPFEFKVPAAADPSCTPEYQVVDGVASIDLPAgiYDYVVVN PSPGLCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF STTG DYVTLTVTGDGGSSADYTYTVYRDGTKIKEGLTATT FEE EGVAPGNHEYCVEVKYQAGVSPKVCKGVTINGGNEFAPVQN LTGSADGQKVTLKWDAPNGTNPNDPNPDPNPDSLNTGVSFV NYTAHGSETSWADPSLTTSQKALTNKDK
KD _A A-1S	36	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQRGPKT APSVTHQTVQKGIRTSTAKDLRDIPIAGMARIILEAHDVWQDG TGYQMLWDDDDHDQYGASIPEDGFWFASETIPDGLFDPFEYK VPVNADASFTPTNFVLDGTASADIPAGTYDYVIINPNPGLIYIVG QGVSKGNDYVVEAGKTYHFTVKRQGS GDAASVVVTGDGGS SADYTYTVYRDGTKIKEGLTATT FEEEGVAPGNHEYSVEVKY QAGVSPKVCKGVTINGGNEFAPVQNLTGSADGQKVTLKWDA PNGTNPNDPNPDPNPDSL
KD _A A-1S*	37	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQRGPKT APSVTHQTVQKGIRTSTAKDLRDIPIAGMARIILEAHDVWQDG TGYQMLWDDDDHDQYGASIPEDGFWFASETIPDGLFDPFEYK VPVNADASFTPTNFVLDGTASADIPAGTYDYVIINPNPGLIYIVG QGVSKGNDYVVEAGKTYHFTVKRQGS GDAASVVVTGDGGS SADYTYTVYRDGTKIKEGLTATT FEEEGVAPGNHEYCVEVKY QAGVSPKVSKGVTINGGNEFAPVQNLTGSADGQKVTLKWDA PNGTNPNDPNPDPNPDSL

KD _A A-AVQP	38	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQRGPKT APSVTHQTVQKGIRTSTAKDLRDPAGMARIILEAHDVWQDG TGYQMLWDDDDHDQYGASIPEDGFWFASETIPDGLFDPFEYK VPVNADASFTPTNFVLDGTASADIPAGTYDYVIINPNPGLIYIVG QGVSKGNDYVVEAGKTYHFTVKRQSGDAASVVVTGDGGS SADYTYTVYRDGTKIKEGLTATTFEEEGVAPGNHEYCVVEVKY QAGVSPKVCKGVTINGGNEFA <u>A</u> <u>VQ</u> <u>PL</u> TGSADGQKVTLKWDA PNGTNPNDPNPDPNP
KD _A A-1S-AVQP	39	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQRGPKT APSVTHQTVQKGIRTSTAKDLRDPAGMARIILEAHDVWQDG TGYQMLWDDDDHDQYGASIPEDGFWFASETIPDGLFDPFEYK VPVNADASFTPTNFVLDGTASADIPAGTYDYVIINPNPGLIYIVG QGVSKGNDYVVEAGKTYHFTVKRQSGDAASVVVTGDGGS SADYTYTVYRDGTKIKEGLTATTFEEEGVAPGNHEYSVEVKY QAGVSPKVCKGVTINGGNEFA <u>A</u> <u>VQ</u> <u>PL</u> TGSADGQKVTLKWDA PNGTNPNDPNPDPNP
KD _A A-1S-AVQP*	40	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQRGPKT APSVTHQTVQKGIRTSTAKDLRDPAGMARIILEAHDVWQDG TGYQMLWDDDDHDQYGASIPEDGFWFASETIPDGLFDPFEYK VPVNADASFTPTNFVLDGTASADIPAGTYDYVIINPNPGLIYIVG QGVSKGNDYVVEAGKTYHFTVKRQSGDAASVVVTGDGGS SADYTYTVYRDGTKIKEGLTATTFEEEGVAPGNHEYCVVEVKY QAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> <u>VQ</u> <u>PL</u> TGSADGQKVTLKWDA PNGTNPNDPNPDPNP
KD _A AK-1S	41	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQRGPKT APSVTHQTVQKGIRTSTAKDLRDPAGMARIILEAHDVWQDG TGYQMLWDDDDHDQYGASIPEDGFWFASETIPDGLFDPFEYK VPVNADASFTPTNFVLDGTASADIPAGTYDYVIINPNPGLIYIVG QGVSKGNDYVVEAGKTYHFTVKRQSGDAASVVVTGDGGS SADYTYTVYRDGTKIKEGLTATTFEEEGVAPGNHEYS <u>S</u> VEVKY QAGVSPKVCKGVTINGGNEFAPVQNLTGSADGQKVTLKWDA PNGTNPNDPNPDPNP NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDK

KD _A AK-1S*	42	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQRGPKT APSVTHQTVQKGIRTSTAKDLRDPPIAGMARIILEAHDVWQDG TGYQMLWDDDDHDQYGASIPEDGFWFASETIPDGLFDPFEYK VPVNADASFTPTNFVLDGTASADIPAGTYDYVIINPNPGLIYIVG QGVSKGNDYVVEAGKTYHFTVKRQSGDAASVVVTGDGGS SADYTYTVYRDGTKIKEGLTATTFEEEGVAPGNHEYCVVEVKY QAGVSPKV <u>S</u> KGVTINGGNEFAPVQNLTGSDGQKVTLKWDA PNGTNPNDPNPDNPNSLNTGVSFVNYTAHGSETSWADPSL TTSQKALTNKDK
KD _A AK-AVQP	43	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQRGPKT APSVTHQTVQKGIRTSTAKDLRDPPIAGMARIILEAHDVWQDG TGYQMLWDDDDHDQYGASIPEDGFWFASETIPDGLFDPFEYK VPVNADASFTPTNFVLDGTASADIPAGTYDYVIINPNPGLIYIVG QGVSKGNDYVVEAGKTYHFTVKRQSGDAASVVVTGDGGS SADYTYTVYRDGTKIKEGLTATTFEEEGVAPGNHEYCVVEVKY QAGVSPKVCKGVTINGGNEFA <u>A</u> <u>VQ</u> <u>P</u> LTGSADGQKVTLKWDA PNGTNPNDPNPDNPNSLNTGVSFVNYTAHGSETSWADPSL TTSQKALTNKDK
KD _A AK-1S- AVQP	44	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQRGPKT APSVTHQTVQKGIRTSTAKDLRDPPIAGMARIILEAHDVWQDG TGYQMLWDDDDHDQYGASIPEDGFWFASETIPDGLFDPFEYK VPVNADASFTPTNFVLDGTASADIPAGTYDYVIINPNPGLIYIVG QGVSKGNDYVVEAGKTYHFTVKRQSGDAASVVVTGDGGS SADYTYTVYRDGTKIKEGLTATTFEEEGVAPGNHEY <u>S</u> VEVKY QAGVSPKVCKGVTINGGNEFA <u>A</u> <u>VQ</u> <u>P</u> LTGSADGQKVTLKWDA PNGTNPNDPNPDNPNSLNTGVSFVNYTAHGSETSWADPSL TTSQKALTNKDK
KD _A AK-1S- AVQP*	45	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQRGPKT APSVTHQTVQKGIRTSTAKDLRDPPIAGMARIILEAHDVWQDG TGYQMLWDDDDHDQYGASIPEDGFWFASETIPDGLFDPFEYK VPVNADASFTPTNFVLDGTASADIPAGTYDYVIINPNPGLIYIVG QGVSKGNDYVVEAGKTYHFTVKRQSGDAASVVVTGDGGS SADYTYTVYRDGTKIKEGLTATTFEEEGVAPGNHEYCVVEVKY QAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> <u>VQ</u> <u>P</u> LTGSADGQKVTLKWDA

		PNGTNPNDPNPDPNPNSLNTGVSFVNYTAHGSETSWADPSL TTSQKALTNKDK
KD _{F1S} A-2S- AVQP	46	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQKMRK EVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQ <u>SL</u> TEDCNIPAT LYDPFEFKVPAAADPSCTPEYQVVDGVASIDLPAgiYDYVVVN PSPGLCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF STTG DYVTLT VTGDGSSADYTYTVYRDGTKIKEGLTATT FEE EGVAPGNHEYS <u>VE</u> VKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> <u>AVQP</u> LTGSADGQKVTLKWDAPNGTNPNDPNPDPNPNSL
KD _{F1S} A-2S- AVQP*	47	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQKMRK EVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQCLTED <u>S</u> NIPAT LYDPFEFKVPAAADPSCTPEYQVVDGVASIDLPAgiYDYVVVN PSPGLCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF STTG DYVTLT VTGDGSSADYTYTVYRDGTKIKEGLTATT FEE EGVAPGNHEYSVEVKYQAGVSPKVSKGVTINGGNEFA <u>A</u> <u>AVQP</u> LTGSADGQKVTLKWDAPNGTNPNDPNPDPNPNSL
KD _{F1S} A-2S- AVQP**	48	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQKMRK EVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQCLTEDCNIPAT LYDPFEFKVPAAADPS <u>S</u> TPEYQVVDGVASIDLPAgiYDYVVVN PSPGLCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF STTG DYVTLT VTGDGSSADYTYTVYRDGTKIKEGLTATT FEE EGVAPGNHEYS <u>VE</u> VKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> <u>AVQP</u> LTGSADGQKVTLKWDAPNGTNPNDPNPDPNPNSL
KD _{F1S} A-2S- AVQP***	49	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQKMRK EVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQCLTEDCNIPAT LYDPFEFKVPAAADPSCTPEYQVVDGVASIDLPAgiYDYVVVN PSPGL <u>S</u> IVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF STTG DYVTLT VTGDGSSADYTYTVYRDGTKIKEGLTATT FEE EGVAPGNHEYS <u>VE</u> VKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> <u>AVQP</u> LTGSADGQKVTLKWDAPNGTNPNDPNPDPNPNSL

KD _{F2S} A-2S- AVQP	50	NTGVSVFVNYTAHGSETSWADPSLTTSQKALTNKDKQKMRK EVSPARAKTLMMLKQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQSLTEDSNIPATL YDPFEFKVPAAADPSCTPEYQVVDGVASIDLPAgiYDYVVVN SPGLCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSFS TTGDYVTLT VTGDGGSSADYTYTVYRDGTKIKEGLTATTFFEE GVAPGNHEY <u>S</u> VEVKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> VQ <u>P</u> LT GSADGQKVTLKWDAPNGTNPNDPNPDPNP
KD _{F2S} A-2S- AVQP*	51	NTGVSVFVNYTAHGSETSWADPSLTTSQKALTNKDKQKMRK EVSPARAKTLMMLKQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQSLTEDCNIPAT LYDPFEFKVPAAADPS <u>S</u> TPEYQVVDGVASIDLPAgiYDYVVVN PSPGLCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF STTG DYVTLT VTGDGGSSADYTYTVYRDGTKIKEGLTATTFEE EGVAPGNHEY <u>S</u> VEVKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> VQ <u>P</u> LTGSADGQKVTLKWDAPNGTNPNDPNPDPNP
KD _{F2S} A-2S- AVQP**	52	NTGVSVFVNYTAHGSETSWADPSLTTSQKALTNKDKQKMRK EVSPARAKTLMMLKQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQSLTEDCNIPAT LYDPFEFKVPAAADPSCTPEYQVVDGVASIDLPAgiYDYVVVN PSPGL <u>S</u> IVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF STTG DYVTLT VTGDGGSSADYTYTVYRDGTKIKEGLTATTFEE EGVAPGNHEY <u>S</u> VEVKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> VQ <u>P</u> LTGSADGQKVTLKWDAPNGTNPNDPNPDPNP
KD _{F2S} A-2S- AVQP***	53	NTGVSVFVNYTAHGSETSWADPSLTTSQKALTNKDKQKMRK EVSPARAKTLMMLKQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQCLTEDSNIPAT LYDPFEFKVPAAADPS <u>S</u> TPEYQVVDGVASIDLPAgiYDYVVVN PSPGLCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF STTG DYVTLT VTGDGGSSADYTYTVYRDGTKIKEGLTATTFEE EGVAPGNHEY <u>S</u> VEVKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> VQ <u>P</u> LTGSADGQKVTLKWDAPNGTNPNDPNPDPNP
KD _{F2S} A-2S- AVQP****	54	NTGVSVFVNYTAHGSETSWADPSLTTSQKALTNKDKQKMRK EVSPARAKTLMMLKQKTEPSDLSKFRDEPIPGGMARIILEAHD

		VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQCLTEDSNIPAT LYDPFEFKVPAAADPSCTPEYQVVDGVASIDLPAgiYDYVVVN PSPGL <u>S</u> IVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF STTG DYVTLT VTGDGGSSADYTYTVYRDG TKIKEGLTATT FEE EGVAPGNHEY <u>S</u> VEVKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> VQ <u>P</u> LTGSADGQKVTLKWDAPNGTNPNDPNPDPNPDSL
KD _{F2S} A-2S- AVQP*****	55	NTGVSVFNyTAHGSETSWADPSLTTSQlKALTNKDKQKMRK EVSPARAKTLMLKQQKTEPSDLSKFRDEIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQCLTEDCNIPAT LYDPFEFKVPAAADPS <u>S</u> TPEYQVVDGVASIDLPAgiYDYVVVN PSPGL <u>S</u> IVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF STTG DYVTLT VTGDGGSSADYTYTVYRDG TKIKEGLTATT FEE EGVAPGNHEY <u>S</u> VEVKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> VQ <u>P</u> LTGSADGQKVTLKWDAPNGTNPNDPNPDPNPDSL
KD _{F3S} A-2S- AVQP	56	NTGVSVFNyTAHGSETSWADPSLTTSQlKALTNKDKQKMRK EVSPARAKTLMLKQQKTEPSDLSKFRDEIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQ <u>S</u> LTEDSNIPATL YDPFEFKVPAAADPS <u>S</u> TPEYQVVDGVASIDLPAgiYDYVVVN SPGLCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSFS TTGDYVTLT VTGDGGSSADYTYTVYRDG TKIKEGLTATT FEE GVAPGNHEY <u>S</u> VEVKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> VQ <u>P</u> LT GSADGQKVTLKWDAPNGTNPNDPNPDPNPDSL
KD _{F3S} A-2S- AVQP*	57	NTGVSVFNyTAHGSETSWADPSLTTSQlKALTNKDKQKMRK EVSPARAKTLMLKQQKTEPSDLSKFRDEIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQ <u>S</u> LTEDSNIPATL YDPFEFKVPAAADPSCTPEYQVVDGVASIDLPAgiYDYVVVN SPGL <u>S</u> IVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSFS TTGDYVTLT VTGDGGSSADYTYTVYRDG TKIKEGLTATT FEE GVAPGNHEY <u>S</u> VEVKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> VQ <u>P</u> LT GSADGQKVTLKWDAPNGTNPNDPNPDPNPDSL
KD _{F3S} A-2S- AVQP**	58	NTGVSVFNyTAHGSETSWADPSLTTSQlKALTNKDKQKMRK EVSPARAKTLMLKQQKTEPSDLSKFRDEIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQ <u>S</u> LTEDCNIPAT LYDPFEFKVPAAADPS <u>S</u> TPEYQVVDGVASIDLPAgiYDYVVVN

		PSPGL <u>S</u> IVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF STTG DYVTLT VTGDGGSSADYTYTVYRDG TKIKEGLTATT FEE EGVAPGNHEYS <u>S</u> VEVKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> VQ <u>P</u> LTGSADGQKVTLKWDAPNGTNPNDPNPDPNPDSL
KD _{F3S} A-2S- AVQP***	59	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQKMRK EVSPARAKTLMMLKQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQCLTED <u>S</u> NIPAT LYDPFEFKVPAAADPS <u>S</u> TPEYQVVDGVASIDLPAgiYDYVVVN PSPGL <u>S</u> IVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF STTG DYVTLT VTGDGGSSADYTYTVYRDG TKIKEGLTATT FEE EGVAPGNHEYS <u>S</u> VEVKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> VQ <u>P</u> LTGSADGQKVTLKWDAPNGTNPNDPNPDPNPDSL
KD _{F4S} A-2S- AVQP	60	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQKMRK EVSPARAKTLMMLKQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQ <u>S</u> LTED <u>S</u> NIPATL YDPFEFKVPAAADPS <u>S</u> TPEYQVVDGVASIDLPAgiYDYVVVN SPGL <u>S</u> IVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSFS TTGDYVTLT VTGDGGSSADYTYTVYRDG TKIKEGLTATT FEE GVAPGNHEYS <u>S</u> VEVKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> VQ <u>P</u> LT GSADGQKVTLKWDAPNGTNPNDPNPDPNPDSL
KD _{F1S} AK-2S- AVQP	61	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQKMRK EVSPARAKTLMMLKQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQ <u>S</u> LTEDCNIPAT LYDPFEFKVPAAADPSCTPEYQVVDGVASIDLPAgiYDYVVVN PSPGLCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF STTG DYVTLT VTGDGGSSADYTYTVYRDG TKIKEGLTATT FEE EGVAPGNHEYS <u>S</u> VEVKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> VQ <u>P</u> LTGSADGQKVTLKWDAPNGTNPNDPNPDPNPDSLNTGVSFV NYTAHGSETSWADPSLTTSQKALTNKDK
KD _{F1S} AK-2S- AVQP*	62	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQKMRK EVSPARAKTLMMLKQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQCLTED <u>S</u> NIPAT LYDPFEFKVPAAADPSCTPEYQVVDGVASIDLPAgiYDYVVVN PSPGLCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF

		STTG DYVTLT VTGDGGSSADYTYTVYRDG TKIKEGLTATT FEE EGVAPGNHEY <u>S</u> VEVKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> VQ <u>P</u> LTGSADGQKVTLKWDAPNGTNPNDPNPDPNPSLNTGVSVF NYTAHGSETSWADPSLTTSQKALTNKDK
KD _{F1S} AK-2S- AVQP**	63	NTGVSVFNNTAHGSETSWADPSLTTSQKALTNKDKQKMRK EVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQCLTEDCNIPAT LYDPFEFKVPAAADPS <u>S</u> TPEYQVVDGVASIDLPA GIYDYVVVN PSPGLCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF STTG DYVTLT VTGDGGSSADYTYTVYRDG TKIKEGLTATT FEE EGVAPGNHEY <u>S</u> VEVKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> VQ <u>P</u> LTGSADGQKVTLKWDAPNGTNPNDPNPDPNPSLNTGVSVF NYTAHGSETSWADPSLTTSQKALTNKDK
KD _{F1S} AK-2S- AVQP***	65	NTGVSVFNNTAHGSETSWADPSLTTSQKALTNKDKQKMRK EVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQCLTEDCNIPAT LYDPFEFKVPAAADPSCTPEYQVVDGVASIDLPA GIYDYVVVN PSPGL <u>S</u> IVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF STTG DYVTLT VTGDGGSSADYTYTVYRDG TKIKEGLTATT FEE EGVAPGNHEY <u>S</u> VEVKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> VQ <u>P</u> LTGSADGQKVTLKWDAPNGTNPNDPNPDPNPSLNTGVSVF NYTAHGSETSWADPSLTTSQKALTNKDK
KD _{F2S} AK-2S- AVQP	66	NTGVSVFNNTAHGSETSWADPSLTTSQKALTNKDKQKMRK EVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQ <u>S</u> LTED <u>S</u> NIPATL YDPFEFKVPAAADPSCTPEYQVVDGVASIDLPA GIYDYVVVN SPGLCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSFS TTGDYVTLT VTGDGGSSADYTYTVYRDG TKIKEGLTATT FEE GVAPGNHEY <u>S</u> VEVKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> VQ <u>P</u> LT GSADGQKVTLKWDAPNGTNPNDPNPDPNPSLNTGVSVFN YTAHGSETSWADPSLTTSQKALTNKDK
KD _{F2S} AK-2S- AVQP*	67	NTGVSVFNNTAHGSETSWADPSLTTSQKALTNKDKQKMRK EVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQ <u>S</u> LTEDCNIPAT LYDPFEFKVPAAADPS <u>S</u> TPEYQVVDGVASIDLPA GIYDYVVVN PSPGLCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF

		STTG DYVTLT VTGDGGSSADYTYTVYRDG TKIKEGLTATT FEE EGVAPGNHEY <u>S</u> VEVKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> <u>VQP</u> LTGSADGQKVTLKWDAPNGTNPNDPNPDPNPSLNTGVSVFV NYTAHGSETSWADPSLTTSQ LKALT NKDK
KD _{F2S} AK-2S- AVQP**	68	NTGVSVFN NYTAHGSETSWADPSLTTSQ LKALT NKDKQKMRK EVSPARAKT LMLKQQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQ <u>S</u> LTEDCNIPAT LYDPFEFKVPAAADPSCTPEYQVVDGVASIDL PAGIYDYVVVN PSPGL <u>S</u> IVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF STTG DYVTLT VTGDGGSSADYTYTVYRDG TKIKEGLTATT FEE EGVAPGNHEY <u>S</u> VEVKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> <u>VQP</u> LTGSADGQKVTLKWDAPNGTNPNDPNPDPNPSLNTGVSVFV NYTAHGSETSWADPSLTTSQ LKALT NKDK
KD _{F2S} AK-2S- AVQP***	69	NTGVSVFN NYTAHGSETSWADPSLTTSQ LKALT NKDKQKMRK EVSPARAKT LMLKQQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQCLTED <u>S</u> NIPAT LYDPFEFKVPAAADPS <u>S</u> TPEYQVVDGVASIDL PAGIYDYVVVN PSPGLCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF STTG DYVTLT VTGDGGSSADYTYTVYRDG TKIKEGLTATT FEE EGVAPGNHEY <u>S</u> VEVKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> <u>VQP</u> LTGSADGQKVTLKWDAPNGTNPNDPNPDPNPSLNTGVSVFV NYTAHGSETSWADPSLTTSQ LKALT NKDK
KD _{F2S} AK-2S- AVQP****	70	NTGVSVFN NYTAHGSETSWADPSLTTSQ LKALT NKDKQKMRK EVSPARAKT LMLKQQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQCLTED <u>S</u> NIPAT LYDPFEFKVPAAADPSCTPEYQVVDGVASIDL PAGIYDYVVVN PSPGL <u>S</u> IVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF STTG DYVTLT VTGDGGSSADYTYTVYRDG TKIKEGLTATT FEE EGVAPGNHEY <u>S</u> VEVKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> <u>VQP</u> LTGSADGQKVTLKWDAPNGTNPNDPNPDPNPSLNTGVSVFV NYTAHGSETSWADPSLTTSQ LKALT NKDK
KD _{F2S} AK-2S- AVQP#	71	NTGVSVFN NYTAHGSETSWADPSLTTSQ LKALT NKDKQKMRK EVSPARAKT LMLKQQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQCLTEDCNIPAT

		LYDPFEFKVPAAADPS <u>S</u> TPEYQVVDGVASIDLPAgiYDYVVVN PSPGL <u>S</u> iVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF STTGdYVTLTVTGDGGSSADYTYTVYRDGTkIkeGLTATTfEE EGVAPGNHEy <u>S</u> VEVKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> VQ <u>P</u> LTGSADGQKVTLKWDAPNGTNPNDPNPDPNPSLNTGVSFV NYTAHGSETSWADPSLTTSQKALTNKDK
KD _{F3S} AK-2S- AVQP	72	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQKMRK EVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQ <u>S</u> LTED <u>S</u> NIPATL YDPFEFKVPAAADPS <u>S</u> TPEYQVVDGVASIDLPAgiYDYVVVN SPGLCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSFS TTGDYVTLTVTGDGGSSADYTYTVYRDGTkIkeGLTATTfEEE GVAPGNHEy <u>S</u> VEVKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> VQ <u>P</u> LT GSADGQKVTLKWDAPNGTNPNDPNPDPNPSLNTGVSFVNY TAHGSETSWADPSLTTSQKALTNKDK
KD _{F3S} AK-2S- AVQP*	73	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQKMRK EVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQ <u>S</u> LTED <u>S</u> NIPATL YDPFEFKVPAAADPSCTPEYQVVDGVASIDLPAgiYDYVVVN SPGL <u>S</u> iVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSFS TTGDYVTLTVTGDGGSSADYTYTVYRDGTkIkeGLTATTfEEE GVAPGNHEy <u>S</u> VEVKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> VQ <u>P</u> LT GSADGQKVTLKWDAPNGTNPNDPNPDPNPSLNTGVSFVNY TAHGSETSWADPSLTTSQKALTNKDK
KD _{F3S} AK-2S- AVQP**	74	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQKMRK EVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQ <u>S</u> LTEDCNIPAT LYDPFEFKVPAAADPS <u>S</u> TPEYQVVDGVASIDLPAgiYDYVVVN PSPGL <u>S</u> iVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF STTGdYVTLTVTGDGGSSADYTYTVYRDGTkIkeGLTATTfEE EGVAPGNHEy <u>S</u> VEVKYQAGVSPKV <u>S</u> KGVTINGGNEFA <u>A</u> VQ <u>P</u> LTGSADGQKVTLKWDAPNGTNPNDPNPDPNPSLNTGVSFV NYTAHGSETSWADPSLTTSQKALTNKDK

KD _{F3S} AK-2S- AVQP***	75	NTGVSVFN ^Y TAHGSETSWADPSLTTSQ ^L KALTNKDKQKMRK EVSPARAKTLMLKQQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQCLTEDS ^N IPAT LYDPFEFKVPAAADPS ^S TPEYQVVDGVASIDLPA ^G IYDYVVVN PSPGL ^S IVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF STTG ^D YVTLT ^V TG ^D GGSSADYTYTVYRDG ^T KIKEGLTATT ^F EE EGVAPGNHEY ^S VEVKYQAGVSPKV ^S KGVTINGGNEFA ^A AVQ ^P LTGSADGQKVTLKWDAPNGTNPNDPNPDPNPSLNTGVSVFV NYTAHGSETSWADPSLTTSQ ^L KALTNKDK
KD _{F4S} AK-2S- AVQP	76	NTGVSVFN ^Y TAHGSETSWADPSLTTSQ ^L KALTNKDKQKMRK EVSPARAKTLMLKQQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQ ^S LTEDS ^N IPATL YDPFEFKVPAAADPS ^S TPEYQVVDGVASIDLPA ^G IYDYVVVNP SPGL ^S IVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSFS TTG ^D YVTLT ^V TG ^D GGSSADYTYTVYRDG ^T KIKEGLTATT ^F EEEE GVAPGNHEY ^S VEVKYQAGVSPKV ^S KGVTINGGNEFA ^A AVQ ^P LT GSADGQKVTLKWDAPNGTNPNDPNPDPNPSLNTGVSVFN ^Y TAHGSETSWADPSLTTSQ ^L KALTNKDK
RD _A A	77	NSGICTTNYTGHGSDTAWGTSGFNTTHMKQLANYNQQRGPK TAPSVTHQTVQKGIRTSTAKDLRDPIAGMARIILEAHDVWQD GTGYQMLWDDDDHDQYGASIPEDGFWFASETIPDGLFDPFEY KVPVNADASFTPTNFVLDGTASADIPAGTYDYVIINPNPGLIYIV GQGVSKGNDYVVEAGKTYHFTVKRQSGDAASVVVTGDGG SSADYTYTVYRDG ^T KIKEGLTATT ^F EEEGVAPGNHEYC ^V EVK YQAGVSPKVCKGVTINGGNEFAPVQNLTGSADGQKVTLKWD APNGTNPNDPNPDPNPSL
RD _A AK	78	NSGICTTNYTGHGSDTAWGTSGFNTTHMKQLANYNQQRGPK TAPSVTHQTVQKGIRTSTAKDLRDPIAGMARIILEAHDVWQD GTGYQMLWDDDDHDQYGASIPEDGFWFASETIPDGLFDPFEY KVPVNADASFTPTNFVLDGTASADIPAGTYDYVIINPNPGLIYIV GQGVSKGNDYVVEAGKTYHFTVKRQSGDAASVVVTGDGG SSADYTYTVYRDG ^T KIKEGLTATT ^F EEEGVAPGNHEYC ^V EVK YQAGVSPKVCKGVTINGGNEFAPVQNLTGSADGQKVTLKWD

		APNGTNPNDPNPDPNPNSLNTGVSVNYTAHGSETSWADPS LTTSQKALTNKDK
KD _A AR	79	NTGVSVNYTAHGSETSWADPSLTTSQKALTNKDKQRGPKT APSVTHQTVQKGIRTSTAKDLRDPIAGMARIILEAHDVWQDG TGYQMLWDDHDQYGASIPEDGFWFASETIPDGLFDPFEYK VPVNADASFTPTNFVLDGTASADIPAGTYDYVIINPNPGLIYVG QGVSKGNDYVVEAGKTYHFTVKRQGSQDAASVVVTGDGGS SADYTYTVYRDGTKIKEGLTATTFEEGVAPGNHEYCVVKY QAGVSPKVCKGVTINGGNEFAPVQNLTGSDGQKVTLKWDA PNGTNPNDPNPDPNPNSLNSGICTTNYTGHGSDTAWGTSGF NTTHMKQLANYNQ
RD _F A	80	NSGICTTNYTGHGSDTAWGTSGFNTTHMKQLANYNQQKMRK EVSPARAKTLMKQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQCLTEDCNIPAT LYDPFEFKVPAAADPSCTPEYQVVDGVASIDLPAIYDYVVVN PSPGLCIVIPGIFGQTDITYGDDFRFEAGKIYHFLVAFQNGTSF STTGDIYVTLVTGDGSSADYTYTVYRDGTKIKEGLTATTFEE EGVAPGNHEYCVVKYQAGVSPKVCKGVTINGGNEFAPVQN LTGSADGQKVTLKWDAPNGTNPNDPNPDPNPNSL
KD _F AR	81	NTGVSVNYTAHGSETSWADPSLTTSQKALTNKDKQKMRK EVSPARAKTLMKQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQCLTEDCNIPAT LYDPFEFKVPAAADPSCTPEYQVVDGVASIDLPAIYDYVVVN PSPGLCIVIPGIFGQTDITYGDDFRFEAGKIYHFLVAFQNGTSF STTGDIYVTLVTGDGSSADYTYTVYRDGTKIKEGLTATTFEE EGVAPGNHEYCVVKYQAGVSPKVCKGVTINGGNEFAPVQN LTGSADGQKVTLKWDAPNGTNPNDPNPDPNPNSLNSGICTT NYTGHGSDTAWGTSGFNTTHMKQLANYNQ
RD _F AK	82	NSGICTTNYTGHGSDTAWGTSGFNTTHMKQLANYNQQKMRK EVSPARAKTLMKQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQCLTEDCNIPAT LYDPFEFKVPAAADPSCTPEYQVVDGVASIDLPAIYDYVVVN PSPGLCIVIPGIFGQTDITYGDDFRFEAGKIYHFLVAFQNGTSF STTGDIYVTLVTGDGSSADYTYTVYRDGTKIKEGLTATTFEE

		EGVAPGNHEYCVEVKYQAGVSPKVCKGVTINGGNEFAPVQN LTGSADGQKVTLKWDAPNGTNPNDPNPDPNPSLNTGVSFV NYTAHGSETSWADPSLTTSQKALTNKDK
ABM1 (PVQN and W target residues in bold)	85	GNEFAP PVQN LTGSADGQKVTLKWDAPNGT
ABM2 (Y target residue in bold)	86	DGGSSADYTYTVYRDGTKIKEGLTATTFEEEGVAPGNHEYCV EVKYQAGVSPKVCKGVTING
ABM3	87	NPNDPNPDPNPSL
ABM2+1	88	DGGSSADYTYTVYRDGTKIKEGLTATTFEEEGVAPGNHEYCV EVKYQAGVSPKVCKGVTINGGNEFAPVQNLTGSADGQKVTLK WDAPNGT
KD _(F) AΔABM3	97	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQKMRK EVSPARAKTLMKQQKTEPSDLSKFRDEPIPGGMARIILEAHD VWGDGSGYQMLIDADHTAYGNEIPPPTPLGQCLTEDCNIPAT LYDPFEFKVPAAADPSCTPEYQVVDGVASIDLPAIYDYVVVN PSPGLCIVIPGIFGQTDDTYGDDFRFEAGKIYHFLVAFQNGTSF STTG DYVTLVTGDDGGSSADYTYTVYRDGTKIKEGLTATTFEE EGVAPGNHEYCVEVKYQAGVSPKVCKGVTINGGNEFAPVQN LTGSADGQKVTLKWDAPNGT
KD _(A) AΔABM3	98	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQRGPKT APSVTHQTVQKGIRTSTAKDLRDPIPAGMARIILEAHDVWQDG TGYQMLWDDHDQYGASIPEDGFWFASETIPDGLFDPFEYK VPVNADASFTPTNFVLDGTASADIPAGTYDYVIINPNPGLIYVG QGVSKGNDYVVEAGKTYHFTVKRQGS GDAASVVVTGDGGS SADYTYTVYRDGTKIKEGLTATTFEEEGVAPGNHEYCVEVKY QAGVSPKVCKGVTINGGNEFAPVQNLTGSADGQKVTLKWD APNGT

KD _(A) ΔΔABM3 2S-AVQP	99	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQRGPKT APSVTHQTVQKGIRTSTAKDLRDPPIAGMARIILEAHDVWQDG TGYQMLWDDDDHDQYGASIPEDGFWFASETIPDGLFDPFEYK VPVNADASFTPTNFVLDGTASADIPAGTYDYVIINPNPGLIYIVG QGVSKGNDYVVEAGKTYHFTVKRQSGDAASVVVTGDGGS SADYTYTVYRDGTIKIEGLTATTFEEEGVAPGNHEYSVEVKY QAGVSPKVSKGVTINGGNEFAA <u>VQ</u> PLTGSADGQKVTLKWDA PNGT
KD _(A) AKΔABM3 2S-AVQP	100	NTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDKQRGPKT APSVTHQTVQKGIRTSTAKDLRDPPIAGMARIILEAHDVWQDG TGYQMLWDDDDHDQYGASIPEDGFWFASETIPDGLFDPFEYK VPVNADASFTPTNFVLDGTASADIPAGTYDYVIINPNPGLIYIVG QGVSKGNDYVVEAGKTYHFTVKRQSGDAASVVVTGDGGS SADYTYTVYRDGTIKIEGLTATTFEEEGVAPGNHEYSVEVKY QAGVSPKVSKGVTINGGNEFAA <u>VQ</u> PLTGSADGQKVTLKWDA PNGTNTGVSFVNYTAHGSETSWADPSLTTSQKALTNKDK
RD _A AK-2S- AVQP	114	NSGICTTNYTGHGSDTAWGTSGFNTTHMKQLANYNQQRGPK TAPSVTHQTVQKGIRTSTAKDLRDPPIAGMARIILEAHDVWQD GTGYQMLWDDDDHDQYGASIPEDGFWFASETIPDGLFDPFEY KVPVNADASFTPTNFVLDGTASADIPAGTYDYVIINPNPGLIYIV GQGVSKGNDYVVEAGKTYHFTVKRQSGDAASVVVTGDGG SSADYTYTVYRDGTIKIEGLTATTFEEEGVAPGNHEYSVEVK YQAGVSPKVSKGVTINGGNEFAA <u>VQ</u> PLTGSADGQKVTLKWD APNGTNPNDPNPDPNPSLNTGVSFVNYTAHGSETSWADPS LTTSQKALTNKDK
R(S) _{cys} D _A AK- 2S-AVQP	115	NSGI <u>S</u> TTNYTGHGSDTAWGTSGFNTTHMKQLANYNQQRGPK TAPSVTHQTVQKGIRTSTAKDLRDPPIAGMARIILEAHDVWQD GTGYQMLWDDDDHDQYGASIPEDGFWFASETIPDGLFDPFEY KVPVNADASFTPTNFVLDGTASADIPAGTYDYVIINPNPGLIYIV GQGVSKGNDYVVEAGKTYHFTVKRQSGDAASVVVTGDGG SSADYTYTVYRDGTIKIEGLTATTFEEEGVAPGNHEYSVEVK YQAGVSPKVSKGVTINGGNEFAA <u>VQ</u> PLTGSADGQKVTLKWD APNGTNPNDPNPDPNPSLNTGVSFVNYTAHGSETSWADPS LTTSQKALTNKDK

KD _A AK-2S	116	NTGVSFVNYTAHGSETSWADPSLTTSQLKALTNKDKQRGPKT APSVTHQTVQKGIRTSTAKDLRDIPIAGMARIILEAHDVWQDG TGYQMLWDDHDQYGASIPEDGFWFASETIPDGLFDPFEYK VPVNADASFTPTNFVLDGTASADIPAGTYDYVIINPNPGLIYIVG QGVSKGNDYVVEAGKTYHFTVKRQGSFDAASVVVTGDGGS SADYTYTVYRDGTKIKEGLTATTFEEEGVAPGNHEYSVEVKY QAGVSPKVSKGVTINGGNEFAPVQNLTGSDGQKVTLKWDA PNGTNPNDPNPDNPNSLNTGVSFVNYTAHGSETSWADPSL TTSQLKALTNKDK
KD _A A-2S-AVQP	117	NTGVSFVNYTAHGSETSWADPSLTTSQLKALTNKDKQRGPKT APSVTHQTVQKGIRTSTAKDLRDIPIAGMARIILEAHDVWQDG TGYQMLWDDHDQYGASIPEDGFWFASETIPDGLFDPFEYK VPVNADASFTPTNFVLDGTASADIPAGTYDYVIINPNPGLIYIVG QGVSKGNDYVVEAGKTYHFTVKRQGSFDAASVVVTGDGGS SADYTYTVYRDGTKIKEGLTATTFEEEGVAPGNHEYSVEVKY QAGVSPKVSKGVTINGGNEFAAVQPLTGSDGQKVTLKWDA PNGTNPNDPNPDNPNSL

[0098] Table 2: nucleic acid sequence information

Descriptor	SEQ ID No	Sequence
DNA Insert for gulae construct No 1 containing HAGF DUF PLUS AVQP 2xCYS>SER	83	CCATGGCAAATACCGGTGTCAGTTTTGTAACTATACGGCGCATGGATCT GAGACATCATGGGCTGATCCGTCCCTGACCACTAGCCAATTAAGGCTCT CACAAATAAGGACAAACAAAAATGAGAAAAGAGGTTTCTCCTGCGCGGG CAAAGACTCTGATGCTGAAACAGCAAAAGACAGAACCATCCGATCTTAGC AAATTCCGTGATGAGCCTATTCCGGGCGGTATGGCCCGTATTATCCTGGA GGCTCATGATGTGTGGGGAGACGGCTCAGGCTATCAAATGCTGATAGAC GCCGATCACACAGCTTATGGTAATGAAATCCCCCTCCACACCACTGGG GCAGTGTGACCGAAGACTGCAACATCCCTGCCACTTTGTATGATCCTT TCGAATTCAAAGTACCGGCAGCAGCCGATCCCTCCTGTACACCCGAGTAT CAGGTCGTAGACGGAGTGGCCAGTATCGACCTGCCGGCGGGGATATATG ATTATGTAGTGGTGAATCCGTGCCCCGGTCTGTGCATCGTTATTTCCCGGT ATTTTCGGGCAGACGGACGATACTTACGGGGACGATTTTCAAGTTCGAAG CGGGAAAGATTATCATTTCCTTGTGCGATTTCAAATGGAACCTCTTTCA GTACGACCGGTGACTATGTAACATTAACCGTCACCGGCGATGGAGGCTC CTCAGCCGATTATACCTATACGGTTTATCGCGATGGCACTAAGATCAAGG AGGGTCTGACGGCAACTACATTGGAAGAAGAGGGCGTAGCCCTGGCAA CCATGAGTATAGCGTGGAAGTTAAGTACCAGGCGGGCGTATCTCCGAAG GTAAGTAAGGGTGTTACGATAAATGGTGGCAACGAATTTGCTGCTGTACA GCCTCTGACGGGTAGTGCTGATGGACAGAAGGTAACGCTCAAGTGGGAT GCTCCCAACGGAACGAATCCTAATCCTGACCCTAATCCTGACCCTAATCC TAGTCTTAATACCGGTGTCAGTTTTGTAACTATACGGCGCATGGATCTGA GACATCATGGGCTGATCCGTCCCTGACCACTAGCCAATTAAGGCTCTCA CAAATAAGGACAAACTCGAG

DNA insert for gulae Construct No 2 containing HAGA DUF PLUS AVQP 2xCYS>SER	84	CCATGGCAAATACCGGTGTCAGTTTTGTAACTATACGGCGCATGGATCT GAGACATCATGGGCTGATCCGTCCCTGACCACTAGCCAATTAAGGCTCT CACAAATAAGGACAAACAGAGAGGGCCGAAGACGGCTCCATCTGTGACG CACCAGACGGTGCAGAAAGGTATTTCGAACATCCACAGCCAAGGATCTGC GAGATCCGATTCTGCCGGCATGGCACGAATTATCTTGAAGCTCATGAT GTGTGGCAAGACGGCACAGGCTATCAAATGCTTTGGGATGACGATCACG ATCAGTACGGCGCATCGATTCCCGAAGATGGTTTTGGTTTGCTAGCGAG ACAATTCCTGATGGCCTTTTCGATCCTTCGAGTATAAAGTCCGGTCAAT GCCGATGCTTCGTTTACTCCTACGAATTCGTGCTTGATGGAACGGCATC AGCCGATATTCTGCCGGCACTTATGACTATGTAATCATCAATCCGAATCC CGGCTTAATATATATTGTAGGGCAGGGTGTTTCCAAAGGCAATGATTATGT AGTAGAGGCCGGTAAGACTTATCACTTCACTGTCAAACGACAAGGTTCCG GCGATGCTGCATCCGTTGTGGTAACAGGCGATGGAGGCTCCTCAGCCGA TTATACCTATACGGTTTATCGCGATGGCACTAAGATCAAGGAGGGTCTGA CGGCAACTACATTTCGAAGAAGAGGGCGTAGCCCTGGCAACCATGAGTA TAGCGTGGAAGTTAAGTACCAGGCGGGCGTATCTCCGAAGGTAAAGTAAG GGTGTTACGATAAAATGGTGGCAACGAATTTGCTGCTGTACAGCCTTGAC GGGTAGTGCTGATGGACAGAAGGTAACGCTCAAGTGGGATGCTCCCAAC GGAACGAATCCTAATCCTGACCCTAATCCTGACCCTAATCCTAGTCTTAAT ACCGGTGTCAGTTTTGTAACTATACGGCGCATGGATCTGAGACATCATG GGCTGATCCGTCCCTGACCACTAGCCAATTAAGGCTCTCACAAATAAGG ACAACTCGAG
KD _(F) AK	89	AATACCGGTGTCAGTTTTGTAACTATACGGCGCATGGATCTGAGACATC ATGGGCTGATCCGTCCCTGACCACTAGCCAATTAAGGCTCTCACAAATA AGGACAAACAGAGAGGGCCGAAGACGGCTCCATCTGTGACGCACAGAC CTGATGCTGAAACAGCAAAAGACAGAACCATCCGATCTTAGCAATTCGG TGATGAGCCTATTCCGGGCGGTATGGCCCGTATTATCCTGGAGGCTCAT GATGTGTGGGGAGACGGCTCAGGCTATCAAATGCTGATAGACGCCGATC ACACAGCTTATGGTAATGAAATTCCTCCCTCCACACCACTGGGCGATGT TTGACCGAAGACTGCAACATCCCTGCCACTTTGTATGATCCTTTGCAATTC AAAGTACCGGCAGCAGCCGATCCCTCCTGTACACCCGAGTATCAGGTCG TAGACGGAGTGGCCAGTATCGACCTGCCGGCGGGGATATGATTATGT AGTGGTGAATCCGTGCCCCGGTCTGTGCATCGTTATCCCGGTATTTTCG GGCAGACGGACGATACTTACGGGGACGATTTCAAGTTTCAAGCGGGGAAA GATTTATCATTTCTTGTGCGATTTCAAATGGAACCTCTTTCACTACGAC CGGTGACTATGTAACATTAACCGTCACCGGCGATGGAGGCTCCTCAGCC GATTATACCTATACGGTTTATCGCGATGGCACTAAGGAGGCTCT GACGGCAACTACATTTCGAAGAAGAGGGCGTAGCCCTGGCAACCATGAG TATTGTGTGGAAGTTAAGTACCAGGCGGGCGTATCTCCGAAGGTATGTAA GGGTGTTACGATAAATGGTGGCAACGAATTTGCTCCTGTACAGAATCTGA CGGGTAGTGCTGATGGACAGAAGGTAACGCTCAAGTGGGATGCTCCCAA CGGAACGAATCCTAATCCTGACCCTAATCCTGACCCTAATCCTAGTCTTAA TACCGGTGTCAGTTTTGTAACTATACGGCGCATGGATCTGAGACATCAT GGGCTGATCCGTCCCTGACCACTAGCCAATTAAGGCTCTCACAAATAAG GACAAA
KD _(A) AK	90	AATACCGGTGTCAGTTTTGTAACTATACGGCGCATGGATCTGAGACATC ATGGGCTGATCCGTCCCTGACCACTAGCCAATTAAGGCTCTCACAAATA AGGACAAACAGAGAGGGCCGAAGACGGCTCCATCTGTGACGCACAGAC GGTGCAAGAAAGGTATTTCGAACATCCACAGCCAAGGATCTGCGAGATCCG ATTCTGCCGGCATGGCACGAATTATCTTGAAGCTCATGATGTGTGGCA AGACGGCACAGGCTATCAAATGCTTTGGGATGACGATCACGATCAGTACG GCGCATCGATTCCCGAAGATGGTTTTTGGTTTGCTAGCGAGACAATTCCT GATGGCCTTTTCGATCCTTTTCGAGTATAAAGTCCGGTCAATGCCGATGC TTCGTTTACTCCTACGAATTTTCGTGCTTGATGGAACGGCATCAGCCGATAT TCCTGCCGGCACTTATGACTATGTAATCATCAATCCGAATCCCGGCTTAAT ATATATTGTAGGGCAGGGTGTTTCCAAAGGCAATGATTATGTAGTAGAGG CCGGTAAGACTTATCACTTCACTGTCAAACGACAAGGTTCCGGCGATGCT GCATCCGTTGTGGTAACAGGCGATGGAGGCTCCTCAGCCGATTATACCTA TACGGTTTATCGCGATGGCACTAAGATCAAGGAGGGTCTGACGGCAACTA CATTTCGAAGAAGAGGGCGTAGCCCTGGCAACCATGAGTATTGTGTGGA AGTTAAGTACCAGGCGGGCGTATCTCCGAAGGTATGTAAGGGTGTACG ATAAATGGTGGCAACGAATTTGCTCCTGTACAGAATCTGACGGGTAGTGC TGATGGACAGAAGGTAACGCTCAAGTGGGATGCTCCCAACGGAACGAAT CCTAATCCTGACCCTAATCCTGACCCTAATCCTAGTCTTAATACCGGTGTC AGTTTTGTAACTATACGGCGCATGGATCTGAGACATCATGGGCTGATCC GTCCCTGACCACTAGCCAATTAAGGCTCTCACAAATAAGGACAAA

KD _(F) A	91	AATACCGGTGTCAGTTTTGTAACTATACGGCGCATGGATCTGAGACATC ATGGGCTGATCCGTCCCTGACCACTAGCCAATTAAGGCTCTCACAAATA AGGACAAACAAAAATGAGAAAAGAGGTTTCTCCTGCGCGGGCAAAGACT CTGATGCTGAAACAGCAAAAGACAGAACCATCCGATCTTAGCAAATTCGG TGATGAGCCTATTCCGGGCGGTATGGCCCGTATTATCCTGGAGGCTCAT GATGTGTGGGGAGACGGCTCAGGCTATCAAATGCTGATAGACGCCGATC ACACAGCTTATGGTAATGAAATCCCCCTCCACACCACTGGGGCAGTGT TTGACCGAAGACTGCAACATCCCTGCCACTTTGTATGATCCTTTCGAATTC AAAGTACCGGCAGCAGCCGATCCCTCCTGTACACCCGAGTATCAGGTGCG TAGACGGAGTGGCCAGTATCGACCTGCCGGCGGGGATATATGATTATGT AGTGGTGAATCCGTCGCCCGGTCTGTGCATCGTTATCCCGGTATTTTCG GGCAGACGGACGATACTTACGGGGACGATTTACAGTTTCAAGCGGGAAAA GATTTATCATTTCTTGTGCGATTTCAAATGGA ACCTCTTTCAGTACGACCGGTGACTATGTAACATTAACCGTCACCGGCGA TGGAGGCTCCTCAGCCGATTATACCTATACGGTTTATCGC GATGGCACTAAGATCAAGGAGGGTCTGACGGCACTACATTCGAAGAAG AGGGCGTAGCCCCCTGGCAACCATGAGTATTGTGTGGAAGTT AAGTACCAGGCGGGCGTATCTCCGAAGGTATGTAAGGGTGTACGATAAA TGGTGGCAACGAATTTGCTCCTGTACAGAATCTGACGGGT AGTGTGATGGACAGAAGGTAACGCTCAAGTGGGATGCTCCCAACGGAA CGAATCCTAATCCTGACCCTAATCCTGACCCTAATCCTAGT CTT
KD _(A) A	92	AATACCGGTGTCAGTTTTGTAACTATACGGCGCATGGATCTGAGACATC ATGGGCTGATCCGTCCCTGACCACTAGCCAATTAAGGCTCTCACAAATA AGGACAAACAGAGAGGGCCGAAGACGGCTCCATCTGTAGCAGCAGAC GGTGCAGAAAGGTATTCGAACATCCACAGCCAAGGATCTGCGAGATCCG ATTCTGCCGGCATGGCACGAATTATCTTGGAAGCTCATGATGTGTGGCA AGACGGCACAGGCTATCAAATGCTTTGGGATGACGATCAGCATCAGTACG GCGCATCGATTCCCGAAGATGGTTTTTGGTTTGTAGCGAGACAATTCTC GATGGCCTTTTCGATCCTTTTCGAGTATAAAGTTCGGTCAATGCCGATGC TTCGTTTACTCCTACGAATTTCTGTGCTTGATGGAACGGCATTACGCCGAT TCCTGCCGGCACTTATGACTATGTAATCATCAATCCGAATCCCGGCTTAAT ATATATTGTAGGGCAGGGTGTTCCAAAGGCAATGATTATGTAGTAGAGG CCGGTAAGACTTATCACTTCACTGTCAAACGACAAGGTTCCGGCGATGCT GCATCCGTTGTGTAACAGGCGATGGAGGCTCCTCAGCCGATTATACCTA TACGGTTTATCGCGATGGCACTAAGATCAAGGAGGGTCTGACGGCAACTA CATTGCAAGAAGAGGGCGTAGCCCCCTGGCAACCATGATGTTGTGGA AGTTAAGTACCAGGCGGGCGTATCTCCGAAGGTATGTAAGGGTGTACG ATAAATGGTGGCAACGAATTTGCTCCTGTACAGAATCTGACGGGTAGTGC TGATGGACAGAAGGTAACGCTCAAGTGGGATGCTCCCAACGGAACGAAT CCTAATCCTGACCCTAATCCTGACCCTAATCCTAGTCTT
KD _(F) AΔABM 3	93	AATACCGGTGTCAGTTTTGTAACTATACGGCGCATGGATCTGAGACATC ATGGGCTGATCCGTCCCTGACCACTAGCCAATTAAGGCTCTCACAAATA AGGACAAACAAAAATGAGAAAAGAGGTTTCTCCTGCGCGGGCAAAGACT CTGATGCTGAAACAGCAAAAGACAGAACCATCCGATCTTAGCAAATTCGG TGATGAGCCTATTCCGGGCGGTATGGCCCGTATTATCCTGGAGGCTCAT GATGTGTGGGGAGACGGCTCAGGCTATCAAATGCTGATAGACGCCGATC ACACAGCTTATGGTAATGAAATCCCCCTCCACACCACTGGGGCAGTGT TTGACCGAAGACTGCAACATCCCTGCCACTTTGTATGATCCTTTCGAATTC AAAGTACCGGCAGCAGCCGATCCCTCCTGTACACCCGAGTATCAGGTGCG TAGACGGAGTGGCCAGTATCGACCTGCCGGCGGGGATATATGATTATGT AGTGGTGAATCCGTCGCCCGGTCTGTGCATCGTTATCCCGGTATTTTCG GGCAGACGGACGATACTTACGGGGACGATTTACAGTTTCAAGCGGGAAAA GATTTATCATTTCTTGTGCGATTTCAAATGGAACCTCTTTCAGTACGAC CGGTGACTATGTAACATTAACCGTCACCGGCGATGGAGGCTCCTCAGCC GATTATACCTATACGGTTTATCGCGATGGCACTAAGATCAAGGAGGGTCT GACGGCAACTACATTGCAAGAAGAGGGCGTAGCCCCCTGGCAACCATGAG TATTGTGTGGAAGTTAAGTACCAGGCGGGCGTATCTCCGAAGGTATGTAA GGGTGTTACGATAAATGGTGGCAACGAATTTGCTCCTGTACAGAATCTGA CGGGTAGTGTGATGGACAGAAGGTAACGCTCAAGTGGGATGCTCCCAA CGGAACG
KD _(A) AΔABM 3	94	AATACCGGTGTCAGTTTTGTAACTATACGGCGCATGGATCTGAGACATC ATGGGCTGATCCGTCCCTGACCACTAGCCAATTAAGGCTCTCACAAATA AGGACAAACAGAGAGGGCCGAAGACGGCTCCATCTGTGACGCACCAGAC GGTGCAGAAAGGTATTCGAACATCCACAGCCAAGGATCTGCGAGATCCG ATTCTGCCGGCATGGCACGAATTATCTTGGAAGCTCATGATGTGTGGCA AGACGGCACAGGCTATCAAATGCTTTGGGATGACGATCAGCATCAGTACG

		GCGCATCGATTCCCGAAGATGGTTTTGGTTTGCTAGCGAGACAATTCCT GATGGCCTTTTCGATCCTTCGAGTATAAAGTTCCGGTCAATGCCGATGC TTCGTTTACTCCTACGAATTCGTGCTTGATGGAACGGCATCAGCCGATAT TCCTGCCGGCACTTATGACTATGTAATCATCAATCCGAATCCGGCTTAAT ATATATTGTAGGGCAGGGTGTTCGAAAGGCAATGATTATGTAGTAGAGG CCGGTAAGACTTATCACTTCACTGTCAAACGACAAGGTTCCGGCGATGCT GCATCCGTTGTGGTAACAGGGCGATGGAGGCTCCTCAGCCGATTATACCTA TACGGTTTATCGCGATGGCACTAAGATCAAGGAGGCTGACGGCAACTA CATTCGAAGAAGAGGGCGTAGCCCTGGCAACCATGAGTATTGTGTGGA AGTTAAGTACCAGGCGGGCGTATCTCCGAAGGTATGTAAGGGTGTACG ATAAATGGTGGCAACGAATTTGCTCCTGTACAGAATCTGACGGGTAGTGC TGATGGACAGAAGGTAACGCTCAAGTGGGATGCTCCCAACGGAACG
RD _(F) A	95	AATAGCGGTATATGTACCACAACTATACGGGACATGGAAGCGACACTGC TTGGGGTACTTCGGGCTTCAATACTACTCATATGAAACAGCTTGCCAACTA CAACCAACAAAAAATGAGAAAAGAGGTTTCTCCTGCGCGGGGCAAGACTC TGATGCTGAAACAGCAAAAGACAGAACCATCCGATCTTAGCAAATTCGGT GATGAGCCTATTCGGGGCGGTATGGCCCGTATTATCCGTGAGGCTCATG ATGTGTGGGGAGACGGCTCAGGCTATCAAATGCTGATAGACGCCGATCA CACAGCTTATGGTAATGAAATTCCTCCCTCCACACCACTGGGGCAGTGTT TGACCGAAGACTGCAACATCCCTGCCACTTTGTATGATCCTTTCGAATTCA AAGTACCGGCAGCAGCCGATCCCTCCTGTACACCCGAGTATCAGGTCAGT AGACGGAGTGGCCAGTATCGACCTGCCGGCGGGGATATATGATTATGTA GTGGTGAATCCGTCGCCCCGGTCTGTGCATCGTTATTCGGGTATTTTCGG GCAGACGGACGATACTTACGGGGACGATTCAGGTTGGAAGCGGGAAAG ATTTATCATTTTCTTGTGCGATTTCAAAATGGAACCTCTTTTCAGTACAG GGTGACTATGTAACATTAACCGTCACCGGCGATGGAGGCTCCTCAGCCG ATTATACCTATACGTTTATCGCGATGGCACTAAGATCAAGGAGGGTCTG ACGGCAACTACATTCGAAGAAGAGGGCGTAGCCCTGGCAACCATGAGT ATTGTGTGGAAGTTAAGTACCAGGCGGGCGTATCTCCGAAGGTATGTAAG GGTGTTACGATAAATGGTGGCAACGAATTTGCTCCTGTACAGAATCTGAC GGGTAGTGCTGATGGACAGAAGGTAACGCTCAAGTGGGATGCTCCCAAC GGAACGAATCCTAATCCTGACCCTAATCCTGACCCTAATCCTAGTCTT
RD _(A) A	96	AATAGCGGTATATGTACCACAACTATACGGGACATGGAAGCGACACTGC TTGGGGTACTTCGGGCTTCAATACTACTCATATGAAACAGCTTGCCAACTA CAACCAACAGAGAGGGCCGAAGACGGCTCCATCTGTGACGCACCAGACG GTGCAGAAAGGTATTCGAACATCCACAGCCAAGGATCTGCGAGATCCGAT TCCTGCCGGCATGGCAGCAATTATCTTGAAGCTCATGATGTGTGGCAAG ACGGCACAGGCTATCAAATGCTTGGGATGACGATACGATCAGTACGG CGCATCGATTCCCGAAGATGGTTTTGGTTTGCTAGCGAGACAATTCCTG ATGGCCTTTTCGATCCTTTCGAGTATAAAGTTCCGGTCAATGCCGATGCTT CGTTTACTCCTACGAATTCGTGCTTGATGGAACGGCATCAGCCGATATT CCTGCCGGCACTTATGACTATGTAATCATCAATCCGAATCCCGGCTTAATA TATATTGTAGGGCAGGGTGTTCGAAAGGCAATGATTATGTAGTAGAGGC CGGTAAGACTTATCACTTCACTGTCAAACGACAAGGTTCCGGCGATGCTG CATCCGTTGTGGTAACAGGCGATGGAGGCTCCTCAGCCGATTATACCTAT ACGGTTTATCGCGATGGCACTAAGATCAAGGAGGGTCTGACGGCAACTA CATTCGAAGAAGAGGGCGTAGCCCTGGCAACCATGAGTATTGTGTGGA AGTTAAGTACCAGGCGGGCGTATCTCCGAAGGTATGTAAGGGTGTACG ATAAATGGTGGCAACGAATTTGCTCCTGTACAGAATCTGACGGGTAGTGC TGATGGACAGAAGGTAACGCTCAAGTGGGATGCTCCCAACGGAACGAAT CCTAATCCTGACCCTAATCCTGACCCTAATCCTAGTCTT

Detailed description of the embodiments

[0099] It will be understood that the invention disclosed and defined in this specification extends to all alternative combinations of two or more of the individual features mentioned or evident from the text or drawings. All of these different combinations constitute various alternative aspects of the invention.

[0100] Reference will now be made in detail to certain embodiments of the invention. While the invention will be described in conjunction with the embodiments, it will be understood that the intention is not to limit the invention to those embodiments. On the contrary, the invention is intended to cover all alternatives, modifications, and equivalents, which may be included within the scope of the present invention as defined by the claims.

[0101] One skilled in the art will recognize many methods and materials similar or equivalent to those described herein, which could be used in the practice of the present invention. The present invention is in no way limited to the methods and materials described. It will be understood that the invention disclosed and defined in this specification extends to all alternative combinations of two or more of the individual features mentioned or evident from the text or drawings. All of these different combinations constitute various alternative aspects of the invention.

[0102] All of the patents and publications referred to herein are incorporated by reference in their entirety.

[0103] For purposes of interpreting this specification, terms used in the singular will also include the plural and vice versa.

[0104] In work leading to the present invention, the inventors investigated various chimeric or fusion proteins for use in inducing immune responses to *P. gulae* and methods for large-scale production of such chimeras for use as vaccine candidates.

[0105] Recent studies in dogs and cats suggest that companion animals host similar periodontal pathogens with a species *Porphyromonas gulae* closely related to the human keystone pathogen *P. gingivalis* commonly found in cats and dogs with severe disease. Although *P. gulae* is a known causative agent for periodontitis in companion animals, no preventative or therapeutic treatment currently exists.

[0106] *P. gingivalis* and *P. gulae*, although different species, possess similar virulence domains, albeit encoded by different genes, but proteins secreted and assembled at the cell surface in a similar fashion. This led the inventors to develop chimeric or fusion proteins comprising these virulence genes, for use in inducing an immune response to *P. gulae*, and methods and uses comprising the same.

Gingipains

[0107] The pathogenicity of *P. gulae* is attributed to a number of surface-associated virulence factors that include cysteine proteinases (gingipain-homologues), fimbriae, haem-binding proteins, and outer membrane transport proteins amongst others. In particular, the extracellular Arg- and Lys-specific proteinases 'gingipain-homologues' (RgpA/B and Kgp) of *P. gingivalis*, a related pathogen, have been implicated as major virulence factors that are critical for colonisation, penetration into host tissue, dysregulation of the immune response, dysbiosis and disease.

[0108] The gingipains, in particular the Lys-specific proteinase Kgp are essential for the ability of a related pathogen, *P. gingivalis*, to induce alveolar bone resorption in the mouse periodontitis model. The gingipains have also been found in gingival tissue at sites of severe periodontitis at high concentrations proximal to the subgingival plaque and at lower concentrations at distal sites deeper into the gingival tissue. Lys-specific and Arg-specific proteinases have been shown to degrade a variety of host proteins *in vitro*, e.g., fibrinogen, fibronectin, and laminin. Plasma host defence and regulatory proteinase inhibitors α -trypsin, α 2-macroglobulin, anti-chymotrypsin, antithrombin III and antiplasmin are also degraded by Lys- and Arg- proteinases from *P. gingivalis*. This has led to the development of a cogent mechanism to explain the keystone role played by *P. gingivalis* in the development of chronic periodontitis.

[0109] The RgpA, RgpB and Kgp genes of *P. gingivalis* all encode an N-terminal signal peptide of ~22 amino acids in length, an unusually long propeptide of ~200 amino acids in length, and a catalytic domain of ~480 amino acids. C-terminal to the catalytic domain is a large hemagglutinin-adhesin (HA) domain which is comprised of a "Domain of Unknown Function" (termed DUF2436 which is defined as conserved Pfam Domain of Unknown Function; IPR018832) and an adhesin domain (comprising adhesin binding domains). The particular arrangement of the adhesin domains and DUF varies between Kgp and RgpA/B and in particular between Kgp and Rgp gingipains from different Porphyromonas species.

[0110] *In vivo*, the RgpA and Kgp precursor proteins are cleaved into multiple domains that remain non-covalently associated forming large outer membrane protein complexes. *In vivo*, Arg- and Lys-specific proteinases are therefore found in a cell-associated complex of non-covalently associated proteinases and adhesins. One such complex has been designated the RgpA-Kgp proteinase-adhesin complex (previously referred to as the PrtR-PrtK proteinase-adhesin complex). The complex is composed of a 45kDa Arg-

specific calcium-stabilised cysteine proteinase and seven sequence-related adhesin domains.

[0111] As used herein a Lys-gingipain catalytic domain sequence around the active site Histidine may also be referred to a KAS or K domain. As used herein an Arg-gingipain catalytic domain sequence around the active site Histidine may also be referred to as a RAS or R domain. Typically, the catalytic domain of the Lys-gingipain or Arg-gingipains is located in the N-terminal region of the protein. Exemplary Histidine active site peptides, as found within the catalytic domains are set out in Table 1 as SEQ ID NOs: 1 and 2.

[0112] As used herein an HA domain of an Arg- or Lys-gingipain of *P. gulae* will be understood to typically either refer to the region of an Arg- or Lys-gingipain homologue that is C-terminal to the catalytic or active site domain or to an homologous HA domain sequence encoded by a separate polyadhesin gene such as Hag. The HA domain typically comprise a Domain of Unknown Function (DUF) domain (especially DUF 2436 conserved Pfam Domain of Unknown Function; IPR018832) and an adhesin domain comprising adhesin binding domains (ABMs). Once these domains are expressed and secreted to the cell surface they form non-covalent complexes of the Arg- and Lys-specific proteinases catalytic domains together with the DUF and ABM domains to form a virulence coat.

First polypeptide

[0113] The chimeric or fusion proteins of the present invention comprise a first polypeptide that comprises or consists of an amino acid sequence of the active site of an Arg-X or Lys-X proteinase homologue (also referred to herein as Arg- or Lys-gingipain, respectively) of *P. gulae*, or a sequence that is at least 80% identical thereto.

[0114] In any embodiment, the first polypeptide comprises or consists of an amino acid sequence selected from the group of: SEQ ID NOs: 1 or 2, or sequences at 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical thereto.

[0115] In embodiments where the first polypeptide comprises or consists of the amino acid sequence of SEQ ID NO: 2, optionally the sequence may comprise a substitution of the cysteine residue at position 5. The substitution may be to a serine, valine or alanine

residue, or any other residue suitable for reducing the likelihood that the cysteine in the peptide sequence could form disulphide bonds.

[0116] Preferably, the chimeric or fusion protein comprises one or more further polypeptides that comprise or consist of an amino acid sequence of the active site of an Arg- or Lys-gingipain homologue of *P. gulae*, or sequences that are at least 80% identical thereto. The one or more further polypeptides comprising or consisting of the active site of an Arg- or Lys-gingipain of *P. gulae* may be located N-terminally to the first polypeptide, C-terminally to the first polypeptide, N-terminally to the second polypeptide or C-terminally to the second polypeptide. The one or more further polypeptides may be linked to the first or second polypeptide of the chimeric or fusion protein, preferably via a linker of no more than 50 amino acids, or directly linked to the first polypeptide.

[0117] The one or more further polypeptides preferably comprise or consist of an amino acid sequence selected from the group of: SEQ ID NOs: 1 or 2, or sequences at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical thereto.

[0118] In embodiments where the one or more further polypeptides comprises or consists of the amino acid sequence of SEQ ID NO: 2, optionally the sequence may comprise a substitution of the cysteine residue at position 5. The substitution may be to a serine, valine or alanine residue, or any other residue suitable for reducing the likelihood that the cysteine in the peptide sequence could form disulphide bonds.

[0119] In any embodiment, the first polypeptide and the further polypeptide that comprise or consist of an amino acid sequence of the active site of an Arg- or Lys-gingipain of *P. gulae*, comprise or consist of an identical amino acid sequence, or sequences that are at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to each other.

[0120] In preferred embodiments, the chimeric or fusion proteins of the invention comprise no more than 2 or no more than 3 or no more than 4 polypeptides that comprise or consist of an amino acid sequence of the active site of an Arg- or Lys-gingipain of *P. gulae*, or sequences that are at least 80% identical thereto. Preferably, the chimeric or fusion proteins of the invention have fewer than 5, more preferably fewer than 4, most preferably fewer than 3 polypeptides that comprise or consist of an amino acid sequence

of the active site of an Arg- or Lys-gingipain of *P. gulae*, or sequences that are at least 80% identical thereto.

Second polypeptide

[0121] The chimeric or fusion proteins of the present invention comprise a second polypeptide that comprises or consists of:

i) the amino acid sequence of a DUF2436 domain of the Arg- and Lys-gingipain surface complexes of *P. gulae*, preferably wherein the amino acid sequence of the DUF2436 domain is as set forth in SEQ ID NO: 3 or 4, or a sequence at least 80% identical thereto; and

ii) the amino acid sequence of an adhesin domain of the Arg- and Lys-gingipain homologue surface complexes of *P. gulae*, preferably wherein the adhesin domain comprises the amino acid sequence of at least SEQ ID NO: 86 and/or SEQ ID NO: 85 (such as the sequence of SEQ ID NO: 88), or a sequence at least 80% identical thereto, more preferably wherein the adhesin domain comprises the amino acid sequence of SEQ ID NO: 20, or a sequence at least 80% identical thereto.

DUF domain

[0122] In *P. gulae*, there are two different forms of the DUF2436 domain: the “F” variant (eg as set out in SEQ ID NO: 3) and the “A” variant (eg as set out in SEQ ID NO: 4). The inventors have found that both variants are useful for inclusion in chimeric or fusion proteins of the invention and for inducing an immune response to *P. gulae*.

[0123] In certain embodiments, the DUF2436 domain comprises the amino acid sequence of the “A” variant, as set forth in SEQ ID NO: 4, or functional variants thereof having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% identity thereto.

[0124] In further embodiments, the inventors contemplate the inclusion of various substitutions to cysteine residues found in the F variant of the DUF 2436 domain. The inventors believe that reducing the number of cysteine residues in the domain will improve solubility and ease of manufacture of the chimeric or fusion proteins of the invention.

[0125] Accordingly, the invention contemplates chimeric or fusion proteins as defined herein, comprising a second polypeptide where the second polypeptide comprises a DUF2436 domain with one or more cysteine substitutions.

[0126] The one or more cysteine residues may be substituted to any amino acid selected from: serine, alanine, valine.

[0127] Preferably the one or more cysteine residues are substituted to one or more serine residues.

[0128] In preferred embodiments, the sequence of the DUF2436 domain comprises the amino acid sequence of the "A" variant, as set forth in SEQ ID NO: 2, or functional variants thereof having, such as those having amino acid sequences set forth in SEQ ID NOs: 5 to 19, or sequences at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% identity thereto.

Adhesin domain

[0129] The inventors have also established that various domains in gingipains have a propensity to multimerise, contributing to reduced solubility of the chimeric or fusion proteins and reduced immunogenicity when multimerised or aggregated. Thus, the formation of multimers has implications for ease of large-scale production of chimeric proteins derived from gingipain sequences and for use as vaccines, but may also affect the generation of protective responses.

[0130] In particular, the inventors have identified two mechanisms for multimer formation: the formation of disulphide bridges between cysteine residues and beta-strand exchange between ABM domains within the adhesin domain.

[0131] Accordingly, in order to reduce multimerisation of the chimeric proteins proposed for use according to the present invention, the inventors targeted the cysteine residues of the adhesin domain which may be responsible for beta sheet formation (and thereby contributing to multimerisation).

[0132] The inventors identified that one or more of following modifications contributes to reducing multimerisation of chimeric fusion proteins derived from Kgp and Rgp polypeptides:

- a) one or more cysteine amino acid substitutions, compared to the naturally occurring *P. gulae* Arg- or Lys-gingipain sequences in corresponding regions;
- b) substitution of the proline and/or an asparagine residues in the sequence PxxN corresponding to, or at a position equivalent to, residues 6 to 9 of SEQ ID NO: 85 (equivalent to residues 68 to 71 of the sequence of SEQ ID NO: 20 or residues 68 to 71 of SEQ ID NO: 88);
- c) substitution of the motif NxFA to SxYQ in the sequence, corresponding to, or at a position equivalent to residues 2 to 5 of SEQ ID NO: 85 (equivalent to residues 64 to 67 of the sequence of SEQ ID NO: 20 or residues 64-67 of SEQ ID NO: 88);
- d) substitution of the tyrosine residue, corresponding to or at a position equivalent to residues at position 10 of SEQ ID NO: 86, and of the tryptophan residue, corresponding to or at a position equivalent to residue at position 23 of SEQ ID NO: 85, to alanine residues (equivalent to the tyrosine at residue position 10 and the tryptophan at residue position 85 of SEQ ID NO: 20).

[0133] In particularly preferred embodiments of the invention, the second polypeptide comprises one or more cysteine amino acid substitutions compared to the naturally occurring adhesin domain sequence of a *P. gulae* gingipain sequence.

[0134] The cysteine amino acid substitution may be a substitution to a serine residue or a valine residue. Preferably, the one or more cysteine substitutions comprise one or more substitutions to a serine residue.

[0135] One or both cysteine residues in the adhesin domain may be substituted to a serine or valine residue, preferably a serine residue. In preferred embodiments of the invention, one or both of the cysteine residues at positions 41 and 55, or positions equivalent thereto of SEQ ID NO: 20, are substituted, optionally to a serine or valine residue, preferably to a serine residue. An exemplary sequence of the adhesin domain, comprising one or two amino acid substitutions is set forth in SEQ ID NOs: 22 to 24.

[0136] Accordingly, in a particularly preferred embodiment, the chimeric or fusion protein of the invention comprises an amino acid sequence corresponding to the sequence as set forth in any one of SEQ ID NOs: 22 to 24, or sequences at least 80% identical thereto and comprising the cysteine substitutions.

[0137] The inventors have further established that multimerisation may also be reduced by the mutation of a conserved motifs present in the adhesin domain.

[0138] In one example, the inventors considered modification of the motif PxxN (eg PVQN (SEQ ID NO: 108) in *P. gulae* Kgp, as set forth in SEQ ID NO: 20), was found to substantially contribute to reduced multimerisation and reduction in beta-strand exchange between ABM domains.

[0139] In preferred embodiments, the chimeric or fusion proteins of the invention therefore comprise a modification of the PxxN motif in the region of the chimeric or fusion protein corresponding to the adhesin domain of *P. gulae* gingipain. Accordingly, the second polypeptide preferably comprises a proline substitution and an asparagine substitution, in the sequence PxxN corresponding, or at a position equivalent to positions 68 to 71 of the sequence of SEQ ID NO: 20.

[0140] The proline amino acid substitution is preferably a substitution to an alanine residue.

[0141] The asparagine amino acid substitution may be a substitution to a proline residue or an alanine residue. Preferably the asparagine residue is substituted to a proline residue. In other embodiments, the asparagine residue is not substituted.

[0142] In further examples, the inventors considered the motif NEFA (SEQ ID NO: 109) in the sequence in the adhesin domain of the Kgp/Rgp polyprotein. This sequence is defined at residues 64 to 67 of SEQ ID NO: 20 herein. The inventors believe that substitution of the motif NEFA (SEQ ID NO: 109) to SEQY (SEQ ID NO: 110), through substitution of the asparagine, phenylalanine and alanine residues to serine, glutamine and tyrosine, respectively, significantly reduces multimerisation.

[0143] In another example, the inventors determined that substitution of the tyrosine residue in the adhesin domain corresponding or at a position equivalent to residues at position 10 of SEQ ID NO: 86, and of the tryptophan residue corresponding or at a position equivalent to residue at position 23 of SEQ ID NO: 85 to alanine residues, also significantly reduced multimerisation.

[0144] Finally, the inventors found that multimerisation was practically eliminated through the combination of cysteine modifications and one or more of the substitutions:

- i) substitution of the proline and/or an asparagine residues in the sequence PxxN corresponding to, or at a position equivalent to, residues 68 to 71 of the sequence of SEQ ID NO: 20;
- ii) substitution of the motif NxFA to SxYQ in the sequence, corresponding to, or at a position equivalent to residues 64 to 67 of the sequence of SEQ ID NO: 20; or
- iii) substitution of the second tyrosine residue, corresponding to or at a position equivalent to residues at position 10 of SEQ ID NO: 86, and of the tryptophan residue corresponding or at a position equivalent to residue at position 23 of SEQ ID NO: 85.

[0145] In particularly preferred embodiments, the inventors found that the combination of one or more cysteine modifications, preferably 2 cysteine substitutions to serine, and modification of the PXXN motif of ABM1 to AXXP, eliminated multimerisation of the resulting recombinant chimeric protein.

[0146] Taken together, the cysteine modifications and the modification of the PXXN motif, the present invention therefore provides for chimeric or fusion proteins as described herein, wherein the second polypeptide of the chimeric protein comprises an amino acid sequence corresponding to a region of an adhesin domain of a *P. gulae* Arg or Lys gingipain, wherein the sequence is as set forth in any of SEQ ID NOs: 21 to 27, preferably comprising a sequence as set forth in SEQ ID NO: 27, or a sequence at least 80% identical thereto, wherein or both cysteine residues are substituted to serine residues and/or wherein the proline and asparagine residues in the sequence PXXN at positions 72 to 75 of SEQ ID NO: 23 or positions equivalent to, are substituted.

Linking of the first and second polypeptides

[0147] In the chimeric or fusion proteins of the present invention, the C-terminal residue of the first polypeptide may be covalently linked to the N-terminal residue of the second polypeptide (corresponding to an adhesin domain polypeptide) or the N-terminal residue of the first peptide may be covalently linked to the C-terminal residue of the second polypeptide (corresponding to an adhesin domain polypeptide). In this arrangement, the first peptide and adhesin domain polypeptide, are said to be "directly linked" or "adjacent".

[0148] In other embodiments, the chimeric or fusion protein includes a linker for linking the first peptide to an adhesin domain polypeptide. The linker may be any linker able to join a peptide to a polypeptide, including both amino acid and non-amino acid linkers.

[0149] Preferably, the linker is non-immunogenic. Typically, the linker is comprised of amino acids, and may therefore be termed a peptide linker.

[0150] A linker is usually a peptide having a length of up to 20 amino acids, although may be longer. The term “linked to” or “fused to” refers to a covalent bond, e.g., a peptide bond, formed between two moieties. Accordingly, in the context of the present invention the linker may have a length of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 or 22 or more amino acids. For example, the herein provided chimeric or fusion proteins may comprise a linker between the first polypeptide comprising or consisting of an amino acid sequence of a *P. gulae* gingipain active domain, and the second polypeptide corresponding to the adhesin domain of a *P. gulae* gingipain, such as between the N-terminus of the second polypeptide and the C-terminus of the first polypeptide. Such linkers have the advantage that they can make it more likely that the different polypeptides of the fusion protein fold independently and behave as expected. Suitable linkers may be up to 50 amino acids in length, although less than 20, less than 15 or less than five amino acids is preferred. The linker may function to bring the first peptide and adhesin domain polypeptide into a closer spatial arrangement than normally observed in a *P. gulae* trypsin-like enzyme. Alternatively, it may space the first polypeptide and the second polypeptide (corresponding to an adhesin domain polypeptide) apart.

[0151] Suitable linkers for use in protein constructs, including those with minimal impact on solubility are known in the art. The linker may be any linker known in the art to the skilled person and may be a flexible linker (such as those comprising repeats of glycine and serine residues), a rigid linker (such as those comprising glutamic acid and lysine residues, flanking alanine repeats) and/or a cleavable linker (such as sequences that are susceptible by protease cleavage). Examples of such linkers are known to the skilled person and are described for example, in Chen et al., (2013) Advanced Drug Delivery Reviews, 65: 1357-1369.

[0152] Useful linkers include glycine-serine (GlySer) linkers, which are well-known in the art and comprise glycine and serine units combined in various orders. Examples include,

but are not limited to, (GS), (GSGGS)_n (SEQ ID NO: 102), (GGGS)_n (SEQ ID NO: 103) and (GGGGS)_n (SEQ ID NO: 104), where *n* is an integer of at least one, typically an integer between 1 and about 10, for example, between 1 and about 8, between 1 and about 6, or between 1 and about 5.

[0153] In some embodiments, the peptide linker may include the amino acids glycine and serine in various lengths and combinations. In some aspects, the peptide linker can include the sequence Gly-Gly-Ser (GGS), Gly-Gly-Gly-Ser (GGGS, SEQ ID NO: 103) or Gly-Gly-Gly-Gly-Ser (GGGGS, SEQ ID NO: 104) and variations or repeats thereof. In some aspects, the peptide linker can include the amino acid sequence GGGGS (a linker of 6 amino acids in length, SEQ ID NO: 104) or even longer. The linker may comprise a series of repeating glycine and serine residues (GS) of different lengths, i.e., (GS)_n where *n* is any number from 1 to 15 or more. For example, the linker may be (GS)₃ (i.e., GSGSGS, SEQ ID NO: 111) or longer (GS)₁₁ or longer. It will be appreciated that *n* can be any number including 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or more. Fusion proteins having linkers of such length are included within the scope of the present invention. Similarly, the linker may be a series of repeating glycine residues separated by serine residues. For example (GGGGS)₃ (i.e., the linker may comprise the amino acid sequence GGGGSGGGGSGGGGS, (G4S)₃, SEQ ID NO: 112) and variations thereof.

[0154] In one embodiment, the peptide linker can include the amino acid sequence GGGGS (a linker of 6 amino acids in length, SEQ ID NO: 104) or even longer. The linker may a series of repeating glycine and serine residues (GS) of different lengths, i.e., (GS)_n where *n* is any number from 1 to 15 or more. For example, the linker may be (GS)₃ (i.e., GSGSGS, SEQ ID NO: 111) or longer (GS)₁₁ or longer. It will be appreciated that *n* can be any number including 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or more.

[0155] Other useful linkers include DSSG (SEQ ID NO: 105), DSSGAS (SEQ ID NO: 106), KLDSSG (SEQ ID NO: 107) and variations thereof. Examples of other suitable linkers are described in Chen et al., (2013) *Advanced Drug Delivery Reviews*, 65: 1357-1369.

Chimeric or fusion proteins and recombinant proteins

[0156] The chimeric or fusion proteins of the invention can be prepared by any of a number of conventional techniques although typically, the polypeptides are made using recombinant technology.

[0157] In the case of recombinant polypeptides, a DNA fragment encoding a desired peptide can be subcloned into an appropriate vector using well-known molecular genetic techniques (see, e.g., Maniatis et al., *Molecular Cloning: A Laboratory Manual*, 2nd ed. (Cold Spring Harbor Laboratory, 1982); Sambrook et al., *Molecular Cloning A Laboratory Manual*, 2nd ed. (Cold Spring Harbor Laboratory, 1989). The fragment can be transcribed and the polypeptide subsequently translated *in vitro*. Commercially available kits also can be employed (e.g., such as manufactured by Clontech, Palo Alto, Calif.; Amersham Pharmacia Biotech Inc., Piscataway, N.J.; InVitrogen, Carlsbad, Calif., and the like). The polymerase chain reaction optionally can be employed in the manipulation of nucleic acids.

[0158] A "fragment" is a portion of a polypeptide of the present invention that retains substantially similar functional activity or substantially the same biological function or activity as the polypeptide, which can be determined using assays described herein.

[0159] "Percent (%) amino acid sequence identity" or "percent (%) identical" with respect to a polypeptide sequence, i.e. a polypeptide of the invention defined herein, is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in the specific polypeptide of the invention, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Reference herein to variants having "at least x% sequence identity" to a recited sequence, means that the variant is at least x% identical to the recited sequence.

[0160] In various aspects and embodiments of the invention, the defined polypeptides are described by reference to variants having at least 80% homology to a reference sequence or more. Percentage (%) homology generally refers to a polypeptide of the invention defined herein, defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in the specific polypeptide of the invention, after aligning the sequences and introducing gaps, if necessary, to achieve the

maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity.

[0161] The amino acids glycine, alanine, valine, leucine and isoleucine can often be substituted for one another (amino acids having aliphatic side chains). Of these possible substitutions it is preferred that glycine and alanine are used to substitute for one another (since they have relatively short side chains) and that valine, leucine and isoleucine are used to substitute for one another (since they have larger aliphatic side chains which are hydrophobic). Other amino acids which can often be substituted for one another include: phenylalanine, tyrosine and tryptophan (amino acids having aromatic side chains); lysine, arginine and histidine (amino acids having basic side chains); aspartate and glutamate (amino acids having acidic side chains); asparagine and glutamine (amino acids having amide side chains); and cysteine and methionine (amino acids having sulphur containing side chains).

[0162] Substitutions of this nature are often referred to as "conservative" or "semi-conservative" amino acid substitutions.

[0163] Amino acid deletions or insertions can also be made relative to the native sequence of the *P. gulae* protein. Thus, for example, amino acids which do not have a substantial effect on the activity of the polypeptide, or at least which do not eliminate such activity, can be deleted. Such deletions can be advantageous, particularly with longer polypeptides since the overall length and the molecular weight of a polypeptide can be reduced whilst still retaining activity. This can enable the amount of polypeptide required for a particular purpose to be reduced - for example, dosage levels can be reduced.

[0164] Amino acid insertions relative to the sequence of the native polypeptide can also be made. This can be done to alter the properties of a polypeptide for use in the present invention (e.g. to enhance antigenicity).

[0165] Amino acid changes can be made using any suitable technique e.g. by using site-directed mutagenesis or solid-state synthesis.

[0166] It should be appreciated that amino acid substitutions or insertions within the scope of the present invention can be made using naturally occurring or non-naturally occurring amino acids..

[0167] Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms (non-limiting examples described below) needed to achieve maximal alignment over the full-length of the sequences being compared. When amino acid sequences are aligned, the percent amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain percent amino acid sequence identity to, with, or against a given amino acid sequence B) can be calculated as: percent amino acid sequence identity = $X/Y100$, where X is the number of amino acid residues scored as identical matches by the sequence alignment program's or algorithm's alignment of A and B and Y is the total number of amino acid residues in B. If the length of amino acid sequence A is not equal to the length of amino acid sequence B, the percent amino acid sequence identity of A to B will not equal the percent amino acid sequence identity of B to A.

[0168] In calculating percent identity, typically exact matches are counted. The determination of percent identity between two sequences can be accomplished using a mathematical algorithm. A non-limiting example of a mathematical algorithm utilized for the comparison of two sequences is the algorithm of Karlin and Altschul (1990) Proc. Natl. Acad. Sci. USA 87:2264, modified as in Karlin and Altschul (1993) Proc. Natl. Acad. Sci. USA 90:5873-5877. Such an algorithm is incorporated into the BLASTN and BLASTX programs of Altschul et al. (1990) J. Mol. Biol. 215:403. To obtain gapped alignments for comparison purposes, Gapped BLAST (in BLAST 2.0) can be utilized as described in Altschul et al. (1997) Nucleic Acids Res. 25:3389. Alternatively, PSI-Blast can be used to perform an iterated search that detects distant relationships between molecules. See Altschul et al. (1997) supra. When utilizing BLAST, Gapped BLAST, and PSI-Blast programs, the default parameters of the respective programs (e.g., BLASTX and BLASTN) can be used. Alignment may also be performed manually by inspection. Another non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the ClustalW algorithm (Higgins et al. (1994) Nucleic Acids Res. 22:4673-4680). ClustalW compares sequences and aligns the entirety of the amino acid or DNA sequence, and thus can provide data about the sequence conservation of the entire amino acid sequence. The ClustalW algorithm is used in several commercially available DNA/amino acid analysis software packages, such as the ALIGNX module of the Vector NTI Program Suite (Invitrogen Corporation, Carlsbad, CA). After alignment of amino acid sequences with ClustalW, the percent amino acid identity can be assessed. A non-limiting

examples of a software program useful for analysis of ClustalW alignments is GENEDOC™ or JalView (<http://www.jalview.org/>). GENEDOC™ allows assessment of amino acid (or DNA) similarity and identity between multiple proteins. Another non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Myers and Miller (1988) CABIOS 4:11-17. Such an algorithm is incorporated into the ALIGN program (version 2.0), which is part of the GCG Wisconsin Genetics Software Package, Version 10 (available from Accelrys, Inc., 9685 Scranton Rd., San Diego, CA, USA). When utilizing the ALIGN program for comparing amino acid sequences, a PAM 120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used.

[0169] The polypeptide desirably comprises an amino end and a carboxyl end. The polypeptide can comprise D-amino acids, L-amino acids or a mixture of D- and L-amino acids.

[0170] The term "conservative substitution" as used herein, refers to the replacement of an amino acid present in the native sequence in the peptide with a naturally or non-naturally occurring amino acid or a peptidomimetic having similar steric properties. Where the side-chain of the native amino acid to be replaced is either polar or hydrophobic, the conservative substitution should be with a naturally occurring amino acid, a non-naturally occurring amino acid or with a peptidomimetic moiety which is also polar or hydrophobic (in addition to having the same steric properties as the side-chain of the replaced amino acid).

[0171] Conservative amino acid substitution tables providing functionally similar amino acids are well known to one of ordinary skill in the art. The following six groups are examples of amino acids that may be considered to be conservative substitutions for one another:

- 1) Alanine (A), Serine (S), Threonine (T);
- 2) Aspartic acid (D), Glutamic acid (E);
- 3) Asparagine (N), Glutamine (Q);
- 4) Arginine (R), Lysine (K);
- 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); and

6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W).

[0172] As naturally occurring amino acids are typically grouped according to their properties, conservative substitutions by naturally occurring amino acids can be determined bearing in mind the fact that replacement of charged amino acids by sterically similar non-charged amino acids are considered as conservative substitutions. For producing conservative substitutions by non-naturally occurring amino acids it is also possible to use amino acid analogues (synthetic amino acids) well known in the art. A peptidomimetic of the naturally occurring amino acid is well documented in the literature known to the skilled person and non-natural or unnatural amino acids are described further below. When affecting conservative substitutions the substituting amino acid should have the same or a similar functional group in the side chain as the original amino acid.

[0173] The phrase "non-conservative substitution" or a "non-conservative residue" as used herein refers to replacement of the amino acid as present in the parent sequence by another naturally or non-naturally occurring amino acid, having different electrochemical and/or steric properties. Thus, the side chain of the substituting amino acid can be significantly larger (or smaller) than the side chain of the native amino acid being substituted and/or can have functional groups with significantly different electronic properties than the amino acid being substituted. Examples of non-conservative substitutions of this type include the substitution of phenylalanine or cyclohexylmethyl glycine for alanine, isoleucine for glycine, or $\text{-NH-CH[(-CH}_2\text{)}_5\text{-COOH]-CO-}$ for aspartic acid. Non-conservative substitution includes any mutation that is not considered conservative.

[0174] A non-conservative amino acid substitution can result from changes in: (a) the structure of the amino acid backbone in the area of the substitution; (b) the charge or hydrophobicity of the amino acid; or (c) the bulk of an amino acid side chain. Substitutions generally expected to produce the greatest changes in protein properties are those in which: (a) a hydrophilic residue is substituted for (or by) a hydrophobic residue; (b) a proline is substituted for (or by) any other residue; (c) a residue having a bulky side chain, e.g., phenylalanine, is substituted for (or by) one not having a side chain, e.g., glycine; or (d) a residue having an electropositive side chain, e.g., lysyl, arginyl, or histadyl, is substituted for (or by) an electronegative residue, e.g., glutamyl or aspartyl.

[0175] Alterations of the native amino acid sequence to produce mutant polypeptides, such as by insertion, deletion and/or substitution, can be done by a variety of means known to those skilled in the art. For instance, site-specific mutations can be introduced by ligating into an expression vector a synthesized oligonucleotide comprising the modified site. Alternately, oligonucleotide-directed site-specific mutagenesis procedures can be used, such as disclosed in Walder et al., Gene 42: 133 (1986); Bauer et al., Gene 37: 73 (1985); Craik, Biotechniques, 12-19 (January 1995); and U.S. Pat. Nos. 4,518,584 and 4,737,462. A preferred means for introducing mutations is the QuikChange Site-Directed Mutagenesis Kit (Stratagene, LaJolla, Calif.).

[0176] Any appropriate expression vector (e.g., as described in Pouwels et al., Cloning Vectors: A Laboratory Manual (Elsevier, N.Y.: 1985)) and corresponding suitable host can be employed for production of recombinant polypeptides. Expression hosts include, but are not limited to, bacterial species within the genera *Escherichia*, *Bacillus*, *Pseudomonas*, *Salmonella*, mammalian or insect host cell systems including baculovirus systems (e.g., as described by Luckow et al., Bio/Technology 6: 47 (1988)), and established cell lines such as the COS-7, C127, 3T3, CHO, HeLa, and BHK cell lines, and the like. The skilled person is aware that the choice of expression host has ramifications for the type of polypeptide produced. For instance, the glycosylation of polypeptides produced in yeast or mammalian cells (e.g., COS-7 cells) will differ from that of polypeptides produced in bacterial cells, such as *Escherichia coli*.

[0177] Alternately, a polypeptide of the invention can be synthesized using standard peptide synthesizing techniques well-known to those of ordinary skill in the art (e.g., as summarized in Bodanszky, Principles of Peptide Synthesis (Springer-Verlag, Heidelberg: 1984)). In particular, the polypeptide can be synthesized using the procedure of solid-phase synthesis (see, e.g., Merrifield, J. Am. Chem. Soc. 85: 2149-54 (1963); Barany et al., Int. J. Peptide Protein Res. 30: 705-739 (1987); and U.S. Pat. No. 5,424,398). If desired, this can be done using an automated peptide synthesizer. Removal of the t-butyloxycarbonyl (t-BOC) or 9-fluorenylmethyloxycarbonyl (Fmoc) amino acid blocking groups and separation of the polypeptide from the resin can be accomplished by, for example, acid treatment at reduced temperature. The polypeptide-containing mixture can then be extracted, for instance, with dimethyl ether, to remove non-peptidic organic compounds, and the synthesized polypeptide can be extracted from the resin powder (e.g., with about 25% w/v acetic acid). Following the synthesis of the polypeptide, further

purification (e.g., using high performance liquid chromatography (HPLC)) optionally can be done in order to eliminate any incomplete polypeptides or free amino acids. Amino acid and/or HPLC analysis can be performed on the synthesized polypeptide to validate its identity. For other applications according to the invention, it may be preferable to produce the polypeptide as part of a larger fusion protein, such as by the methods described herein or other genetic means, or as part of a larger conjugate, such as through physical or chemical conjugation, as known to those of ordinary skill in the art and described herein.

[0178] In any embodiment of the invention, the chimeric or fusion protein of the invention may comprise additional amino acid residues to facilitate expression in a recombinant expression system and/or to facilitate purification of the protein. Thus, the proteins defined herein may include additional amino acids such as one, two, three, four, or five amino acids in the N-terminal region. Typically the additional amino acids will include an N-terminal methionine for facilitating expression in recombinant expression systems although it will be appreciated that typically such N-terminal residues are cleaved following translation of the protein. As such it will be appreciated that for some of the proteins described herein (eg SEQ ID NO: 30 and 31), the N terminal methionine will be processed by the expression organism and the final protein product will not comprise the N terminal methionine). In certain embodiments, the N-terminal amino acids include at least methionine and an alanine residue.

[0179] Further, a chimeric or fusion protein according to the invention may include additional amino acids such as one, two, three, four, or five amino acids in the N or C-terminal region, preferably to facilitate purification. It will be understood that such amino acid residues may facilitate the inclusion of a purification tag in the protein (such as histidine tags and the like). Such residues may not be included where untagged versions of the protein are produced.

[0180] A polypeptide of the invention may also be modified by, conjugation or fusion to another moiety to facilitate purification, or increasing the *in vivo* half-life of the polypeptides, or for use in immunoassays using methods known in the art. For example, a polypeptide of the invention may be modified by glycosylation, acetylation, pegylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to a cellular ligand or other protein, etc.

Nucleic acids

[0181] Nucleic acid molecules that encode any of the chimeric or fusion proteins or polypeptides of the invention are also within the scope of the invention. The nucleic acids are useful, for example, in making the polypeptides of the present invention and as therapeutic agents. They may be administered to cells in culture or *in vivo* and may include a secretory signal that directs or facilitates secretion of the polypeptide of the invention from the cell. Also within the scope of the invention are expression vectors and host cells that contain or include nucleic acids of the invention (described further below). While the nucleic acids of the invention may be referred to as "isolated," by definition, the polypeptides of the invention are not wild-type polypeptides and, as such, would not be encoded by naturally occurring nucleic acids. Thus, while the polypeptides and nucleic acids of the present invention may be "purified," "substantially purified," "isolated," "recombinant" or "synthetic" they need not be so in order to be distinguished from naturally occurring materials.

[0182] An "isolated" nucleic acid molecule is a nucleic acid molecule that is identified and separated from at least one contaminant nucleic acid molecule with which it is ordinarily associated in the natural source of the nucleic acid. An isolated nucleic acid molecule is other than in the form or setting in which it is found in nature. Isolated nucleic acid molecules therefore are distinguished from the nucleic acid molecule as it exists in natural cells. However, an isolated nucleic acid molecule includes nucleic acid molecules contained in cells that ordinarily express Kgp where, for example, the nucleic acid molecule is in a chromosomal location different from that of natural cells.

[0183] The terms "nucleic acid molecule" and "polynucleotide" are used interchangeably herein and refer to a polymeric form of nucleotides of any length, either deoxyribonucleotides or ribonucleotides, or analogues thereof. Non-limiting examples of polynucleotides include a gene, a gene fragment, messenger RNA (mRNA), cDNA, recombinant polynucleotides, plasmids, vectors, isolated DNA of any sequence, isolated RNA of any sequence, nucleic acid probes, and primers. A polynucleotide of the invention may be provided in isolated or purified form. A nucleic acid sequence which "encodes" a selected polypeptide is a nucleic acid molecule which is transcribed (in the case of DNA) and translated (in the case of mRNA) into a polypeptide *in vivo* when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a start codon at the 5' (amino) terminus and a translation stop codon at

the 3' (carboxy) terminus. For the purposes of the invention, such nucleic acid sequences can include, but are not limited to, cDNA from viral, prokaryotic or eukaryotic mRNA, genomic sequences from viral or prokaryotic DNA or RNA, and even synthetic DNA sequences. A transcription termination sequence may be located 3' to the coding sequence.

[0184] Polynucleotides of the invention can be synthesised according to methods well known in the art, as described by way of example in Sambrook et al (1989, *Molecular Cloning—a laboratory manual*; Cold Spring Harbor Press).

[0185] The polynucleotide molecules of the present invention may be provided in the form of an expression cassette which includes control sequences operably linked to the inserted sequence, thus allowing for expression of the polypeptide of the invention in vivo in a targeted subject. These expression cassettes, in turn, are typically provided within vectors (e.g., plasmids or recombinant viral vectors) which are suitable for use as reagents for nucleic acid immunization. Such an expression cassette may be administered directly to a host subject. Alternatively, a vector comprising a polynucleotide of the invention may be administered to a host subject. Preferably the polynucleotide is prepared and/or administered using a genetic vector. A suitable vector may be any vector which is capable of carrying a sufficient amount of genetic information and allowing expression of a polypeptide of the invention.

[0186] The present invention thus includes expression vectors that comprise such polynucleotide sequences. Thus, the present invention provides a vector for use in preventing or treating an inflammatory disease or condition comprising a polynucleotide sequence which encodes a polypeptide of the invention and optionally one or more further polynucleotide sequences which encode different polypeptides as defined herein.

[0187] Furthermore, it will be appreciated that the compositions and products of the invention may comprise a mixture of polypeptides and polynucleotides. Accordingly, the invention provides a composition or product as defined herein, wherein in place of any one of the polypeptide is a polynucleotide capable of expressing said polypeptide.

[0188] Expression vectors are routinely constructed in the art of molecular biology and may for example involve the use of plasmid DNA and appropriate initiators, promoters, enhancers and other elements, such as for example polyadenylation signals which may

be necessary, and which are positioned in the correct orientation, in order to allow for expression of a peptide of the invention. Other suitable vectors would be apparent to persons skilled in the art. By way of further example in this regard we refer to Sambrook et al.

[0189] Thus, a polypeptide of the invention may be provided by delivering such a vector to a cell and allowing transcription from the vector to occur. Preferably, a polynucleotide of the invention or for use in the invention in a vector is operably linked to a control sequence which is capable of providing for the expression of the coding sequence by the host cell, i.e. the vector is an expression vector.

[0190] “Operably linked” refers to an arrangement of elements wherein the components so described are configured so as to perform their usual function. Thus, a given regulatory sequence, such as a promoter, operably linked to a nucleic acid sequence is capable of effecting the expression of that sequence when the proper enzymes are present. The promoter need not be contiguous with the sequence, so long as it functions to direct the expression thereof. Thus, for example, intervening untranslated yet transcribed sequences can be present between the promoter sequence and the nucleic acid sequence and the promoter sequence can still be considered “operably linked” to the coding sequence.

[0191] A number of expression systems have been described in the art, each of which typically consists of a vector containing a gene or nucleotide sequence of interest operably linked to expression control sequences. These control sequences include transcriptional promoter sequences and transcriptional start and termination sequences. The vectors of the invention may be for example, plasmid, virus or phage vectors provided with an origin of replication, optionally a promoter for the expression of the said polynucleotide and optionally a regulator of the promoter. A “plasmid” is a vector in the form of an extra-chromosomal genetic element. The vectors may contain one or more selectable marker genes, for example an ampicillin resistance gene in the case of a bacterial plasmid or a resistance gene for a fungal vector. Vectors may be used *in vitro*, for example for the production of DNA or RNA or used to transfect or transform a host cell, for example, a mammalian host cell. The vectors may also be adapted to be used *in vivo*, for example to allow *in vivo* expression of the polypeptide.

[0192] A “promoter” is a nucleotide sequence which initiates and regulates transcription of a polypeptide-encoding polynucleotide. Promoters can include inducible promoters (where expression of a polynucleotide sequence operably linked to the promoter is induced by an analyte, cofactor, regulatory protein, etc.), repressible promoters (where expression of a polynucleotide sequence operably linked to the promoter is repressed by an analyte, cofactor, regulatory protein, etc.), and constitutive promoters. It is intended that the term “promoter” or “control element” includes full-length promoter regions and functional (e.g., controls transcription or translation) segments of these regions.

[0193] A polynucleotide, expression cassette or vector according to the present invention may additionally comprise a signal peptide sequence. The signal peptide sequence is generally inserted in operable linkage with the promoter such that the signal peptide is expressed and facilitates secretion of a polypeptide encoded by coding sequence also in operable linkage with the promoter.

[0194] Typically a signal peptide sequence encodes a peptide of 10 to 30 amino acids for example 15 to 20 amino acids. Often the amino acids are predominantly hydrophobic. In a typical situation, a signal peptide targets a growing polypeptide chain bearing the signal peptide to the endoplasmic reticulum of the expressing cell. The signal peptide is cleaved off in the endoplasmic reticulum, allowing for secretion of the polypeptide via the Golgi apparatus.

Immunogenic and vaccine compositions

[0195] The invention further provides compositions comprising the chimeric or fusion proteins defined herein, and the use of such chimeric or fusion proteins in immunogenic or vaccine compositions in the treatment or prevention of *P. gulae* infection.

[0196] The term “vaccine composition” used herein is defined as a composition used to elicit an immune response against an antigen (immunogen) within the composition in order to protect or treat an organism against disease.

[0197] As used herein, the terms “immunostimulating composition”, “vaccine composition” and “immunogenic composition” may generally be used interchangeably.

[0198] The immunostimulating compositions or vaccines of the invention may suitably include a pharmaceutically acceptable carrier, excipient, diluent, adjuvant, vehicle, buffer

or stabiliser in addition to one or more peptides of the invention as the therapeutically or prophylactically active ingredient. Such carriers include, but are not limited to, saline, buffered saline, dextrose, liposomes, water, glycerol, polyethylene glycol, ethanol and combinations thereof.

[0199] The immunostimulating compositions or vaccine compositions can be adapted for administration by any appropriate route, for example by the parenteral (including subcutaneous, intramuscular, intravenous or intradermal or by injection into the cerebrospinal fluid), oral (including buccal or sublingual), nasal, topical (including buccal, sublingual or transdermal), vaginal or rectal route. In certain embodiments, the vaccine or immune stimulating composition may be delivered as a bolus injection or via slow release composition. Such compositions can be prepared by any method known in the art of pharmacy, for example by admixing peptides with the carrier(s) or excipient(s) under sterile conditions. Typically, the vaccine composition is adapted for administration by the subcutaneous, intramuscular, intravenous or intradermal route, typically by injection. Alternatively, the vaccine composition may be adapted for oral or nasal administration.

[0200] An immunostimulating composition or vaccine composition adapted for parenteral administration may be an aqueous and non-aqueous sterile injection solution which can contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation substantially isotonic with the blood of the intended recipient; and aqueous and non-aqueous sterile suspensions which can include suspending agents and thickening agents. Excipients which can be used for injectable solutions include water, alcohols, polyols, glycerine and vegetable oils, for example. The composition can be presented in unit-dose or multi-dose containers, for example sealed ampoules and vials, and can be stored in a freeze-dried (lyophilized) condition requiring only the addition of the sterile liquid carrier, for example water for injections, immediately prior to use. Extemporaneous injection solutions and suspensions can be prepared from sterile powders, granules and tablets.

[0201] An immunostimulating or vaccine composition adapted for oral administration, can be presented as discrete units such as capsules or tablets or lozenges; as powders or granules; as solutions, syrups or suspensions (in aqueous or non-aqueous liquids; or as edible foams or whips; or as emulsions).

[0202] Suitable excipients for tablets or hard gelatine capsules include lactose, maize starch or derivatives thereof, stearic acid or salts thereof. Suitable excipients for use with soft gelatine capsules include for example vegetable oils, waxes, fats, semi-solid, or liquid polyols etc.

[0203] For the preparation of solutions and syrups, excipients which can be used include for example water, polyols and sugars. For the preparation of suspensions, oils (e.g. vegetable oils) can be used to provide oil-in-water or water in oil suspensions.

[0204] An immunostimulating or vaccine composition adapted for nasal administration wherein the carrier is a solid include a coarse powder having a particle size for example in the range 20 to 500 microns which is administered in the manner in which snuff is taken, i.e. by rapid inhalation through the nasal passage from a container of the powder held close up to the nose. A suitable composition wherein the carrier is a liquid, for administration as a nasal spray or as nasal drops, may comprise an aqueous or oil solution of the active ingredient.

[0205] Compositions adapted for administration by inhalation include fine particle dusts or mists that can be generated by means of various types of metered dose pressurised aerosols, nebulizers or insufflators.

[0206] An immunostimulating or vaccine composition adapted for transdermal administration may be presented as a discrete patch intended to remain in intimate contact with the epidermis of the recipient for a prolonged period of time. For example, the active ingredient can be delivered from the patch by iontophoresis as generally described in *Pharmaceutical Research*. 3(6):318 (1986).

[0207] A composition adapted for topical administration may be formulated as an ointment, cream, suspension, lotion, powder, solution (eg mouth wash) paste, gel, spray, aerosol or oil. For infections of the eye or other external tissues, for example mouth and skin, the composition may be applied as a topical ointment or cream. When formulated in an ointment, the active ingredient can be employed with either a paraffinic or a water-miscible ointment base. Alternatively, the active ingredient can be formulated in a cream with an oil-in-water cream base or a water-in-oil base. A pharmaceutical composition adapted for topical administration to the eye may comprise eye drops wherein the active ingredient is dissolved or suspended in a suitable carrier, especially an aqueous solvent.

A pharmaceutical composition adapted for topical administration in the mouth may comprise lozenges, pastilles or mouth washes.

[0208] The immunostimulating or vaccine composition may contain preserving agents, solubilising agents, stabilising agents, wetting agents, emulsifiers, sweeteners, colourants, odourants, salts (substances of the present invention can themselves be provided in the form of a pharmaceutically acceptable salt), buffers, coating agents or antioxidants.

[0209] The vaccine composition of the invention may also contain one or more other prophylactically or therapeutically active agents in addition to the chimeric or fusion protein as defined herein.

[0210] A chimeric or fusion protein for use in the vaccine compositions of the invention may or may not be lyophilised.

[0211] The vaccine compositions of the invention may also include a pharmaceutically acceptable adjuvant in addition to the peptide(s) as defined herein. Adjuvants are added in order to enhance the immunogenicity of the vaccine composition.

[0212] Suitable adjuvants for inclusion in a vaccine composition are known in the art and include incomplete Freund's adjuvant, complete Freund's adjuvant, Freund's adjuvant with MDP (muramyl dipeptide), alum (aluminium hydroxide), alum plus Bordetella pertussis and immune stimulatory complexes (ISCOMs, typically a matrix of Quil A containing viral proteins), QS- 21, Detox-PC, MPL-SE, MoGM-CSF, TitreMax-G, CRL-1005, GERBU, TERamide, PSC97B, Adjuver, PG-026, GSK-I, GcMAF, B-aletine, MPC-026, Adjuvax, CpG ODN, Betafectin, and MF59.

[0213] The vaccine compositions of the invention may also include or be co-administered with one or more co-stimulatory molecules.

[0214] Dosages of the vaccine composition of the present invention can vary between wide limits, depending upon the age and condition of the individual to be treated, etc. and a physician will ultimately determine appropriate dosages to be used.

[0215] This dosage can be repeated as often as appropriate. For example, an initial dose of the vaccine may be administered and then a booster administered at a later date.

[0216] For administration to mammals, and particularly humans, it is expected that the daily dosage of the active agent will be from 1 µg/kg to 10 mg/kg body weight, typically around 10 µg /kg to 1 mg/kg body weight. The physician in any event will determine the actual dosage which will be most suitable for an individual which will be dependent on factors including the age, weight, sex and response of the individual. The above dosages are exemplary of the average case. There can, of course, be instances where higher or lower dosages are merited, and such are within the scope of this invention.

[0217] The vaccine composition of the invention can be administered by any convenient route as described herein, such as via the intramuscular, intravenous, by inhalation, intraperitoneal or oral routes or by injection into the cerebrospinal fluid.

[0218] The vaccine composition of the invention can be provided in unit dosage form, will generally be provided in a sealed container and may be provided as part of a kit. Such a kit would normally (although not necessarily) include instructions for use. It can include a plurality of said unit dosage forms.

[0219] Accordingly, in yet another aspect, the present invention provides a kit of parts comprising a vaccine composition of the invention and one or more cytokines and/or adjuvants in sealed containers.

[0220] Methods for immunising a subject using the subject chimeric or fusion proteins

[0221] The present invention provides methods and compositions for treating or preventing infection or minimising the likelihood of infection with *P. gulae*, in an individual in need thereof, the methods comprising administering a fusion or chimeric protein of the invention.

[0222] The present invention also provides for methods and compositions for inducing a humoral immune response in a subject to *P. gulae*. The humoral response may be for the purposes of obtaining protective/therapeutic anti-*P. gulae* antibodies directly in the individual requiring protection/therapy. Alternatively, the humoral response may be for the purpose of generating antibodies which are then isolated from the subject (or egg thereof), such that the antibodies can then be directly administered to an individual requiring treatment/protection with the antibodies.

[0223] As such, the present invention includes methods and compositions for preventing infection with *P. gulae*, minimising the likelihood of infection and/or reducing the severity and duration of *P. gulae* infection in a subject.

[0224] The present invention also provides a method for obtaining an antibody directed to *P. gulae*, the method comprising administering a chimeric or fusion protein, composition, vaccine or immune stimulating composition of the invention, to a non-human animal, thereby generating antibodies directed to *P. gulae* in the animal. Preferably the method further comprises isolating the antibody from the animal (eg from the blood of the animal) or from an egg of the animal (eg in the case of generating IgY antibodies from chickens).

[0225] The present invention also provides an antibody preparation comprising an antibody directed to *P. gulae*, wherein the antibody preparation is obtained by administering a chimeric or fusion protein, composition, vaccine or immune stimulating composition of the invention, to a non-human animal, thereby generating antibodies directed to *P. gulae* in the animal, and isolating the antibodies from the animal or egg thereof.

[0226] The antibody directed to *P. gulae* may be used therapeutically to eliminate or reduce *P. gulae* infection or prophylactically, to prevent or reduce the severity of *P. gulae* infection.

[0227] As used herein, the terms "treatment" or "treating" of a subject includes the application or administration of a composition of the invention to a subject (or application or administration of a compound of the invention to a cell or tissue from a subject) with the purpose of delaying, slowing, stabilizing, curing, healing, alleviating, relieving, altering, remedying, less worsening, ameliorating, improving, or affecting the disease or condition, the symptom of the disease or condition, or the risk of (or susceptibility to) the disease or condition. The term "treating" refers to any indication of success in the treatment or amelioration of an injury, pathology or condition, including any objective or subjective parameter such as abatement; remission; lessening of the rate of worsening; lessening severity of the disease; stabilization, diminishing of symptoms or making the injury, pathology or condition more tolerable to the subject; slowing in the rate of degeneration or decline; making the final point of degeneration less debilitating; or improving a subject's physical or mental well-being.

[0228] As used herein, "preventing" or "prevention" is intended to refer to at least the reduction of likelihood of the risk of (or susceptibility to) acquiring a disease or disorder (i.e., causing at least one of the clinical symptoms of the disease not to develop in a subject that may be exposed to or predisposed to the disease but does not yet experience or display symptoms of the disease). Biological and physiological parameters for identifying such subjects are provided herein and are also well known by physicians.

[0229] The vaccine compositions of the invention can be administered to subjects felt to be in greatest need thereof. The vaccine compositions of the invention can also be administered to subjects suspected of having or diagnosed with having infection with *P. gulae*.

[0230] The compositions and methods of the present invention extend equally to uses in both human and/or veterinary medicine, generation of diagnostic agents or the generation of other treatment reagents.

[0231] As used herein, the term "subject" shall be taken to mean any animal, preferably a non-human animal. Exemplary subjects include but are not limited companion animals (cat, dog, guinea pig, and the like).

[0232] As used herein, the terms "subject", "individual" and "patient" may be used interchangeably.

[0233] The skilled person will be familiar with methods for determining successful vaccination/immunisation with a chimeric or fusion protein or composition as described herein. For example, the skilled person will be familiar with methods for quantifying the antibodies generated following immunisation and/or for quantifying the extent of the humoral (Th2) response induced following immunisation.

Kits

[0234] In another embodiment there is provided a kit or article of manufacture including one or more proteins, polypeptides or polynucleotides of the invention and/or immunogenic composition as described above.

[0235] In yet another aspect, the present invention provides a kit of parts comprising a vaccine composition of the invention and one or more adjuvants for separate, subsequent or simultaneous administration to a subject.

[0236] In other embodiments there is provided a kit for use in a therapeutic or prophylactic application mentioned above, the kit including:

- a container holding a protein, polypeptide, polynucleotide or immunogenic composition of the invention;
- a label or package insert with instructions for use.

[0237] In any embodiment the kit may contain one or more further active principles or ingredients for eliciting an immune response to *P. gulae* in a subject.

[0238] The kit or “article of manufacture” may comprise a container and a label or package insert on or associated with the container. Suitable containers include, for example, bottles, vials, syringes, blister pack, etc. The containers may be formed from a variety of materials such as glass or plastic. The container holds a therapeutic composition which is effective for treating the condition and may have a sterile access port (for example the container may be an intravenous solution bag or a vial having a stopper pierceable by a hypodermic injection needle). The label or package insert indicates that the therapeutic composition is used for treating the condition of choice. In one embodiment, the label or package insert includes instructions for use and indicates that the therapeutic or prophylactic composition can be used to treat an inflammatory disease or condition described herein.

[0239] The kit may comprise (a) a therapeutic or prophylactic composition; and (b) a second container with a second active principle or ingredient contained therein. The kit in this embodiment of the invention may further comprise a package insert indicating the composition and other active principle can be used to treat a disorder or prevent a complication stemming from an inflammatory disease or condition described herein. Alternatively, or additionally, the kit may further comprise a second (or third) container comprising a pharmaceutically-acceptable buffer, such as bacteriostatic water for injection (BWFI), phosphate-buffered saline, Ringer's solution and dextrose solution. It may further include other materials desirable from a commercial and user standpoint, including other buffers, diluents, filters, needles, and syringes.

[0240] In any embodiment the therapeutic composition may be provided in the form of a device, disposable or reusable, including a receptacle for holding the therapeutic, prophylactic or immunogenic composition. In one embodiment, the device is a syringe,

autoinjector or nanopatch. The device may hold between 0.1 to 2 mL of the therapeutic or immunogenic composition. The therapeutic or prophylactic composition may be provided in the device in a state that is ready for use or in a state requiring mixing, dissolution or resuspension or addition of further components.

[0241] It will be understood that the invention disclosed and defined in this specification extends to all alternative combinations of two or more of the individual features mentioned or evident from the text or drawings. All of these different combinations constitute various alternative aspects of the invention.

[0242] The following examples are presented in order to more fully illustrate some embodiments of the invention. They should, in no way be construed, however, as limiting the broad scope of the invention. One skilled in the art can readily devise many variations and modifications of the principles disclosed herein without departing from the scope of the invention.

Examples

[0243] The following examples describe a series of *in vitro* and *in vivo* studies relating to the chimeric or fusion proteins of the invention. Example 1 describes *in vitro* materials and methods. Example 2 describes the results of *in vitro* studies. Example 3 describes materials and methods for *in vivo* studies and Example 4 describes the results of those *in vivo* studies.

Example 1: *In vitro* materials and methods

[0244] Cloning of *Porphyromonas gulae* constructs

[0245] DNA fragments encoding KD_FAK-2S-AVQP (DUF F variant; SEQ ID NO: 28) and KD_AAK-2S-AVQP (DUF A variant, SEQ ID NO: 29) containing NcoI and XhoI restriction sequences at the 5' and 3' ends were synthesized and ligated into transport vector, pBHA, by Bioneer Pacific (Aus.) to generate pBHA-gulA and pBHA-gulF. The cloned inserts were verified by DNA sequencing (Bioneer Pacific). pBHA-F and pBHA-A vectors were introduced into *E. coli* ABLE-K (Agilent Technologies) chemically competent cells following manufacturer's protocol and recombinant strains were selected on LB containing Amp (100 µg/mL). Plasmid DNA was purified from positive clones and digested with NcoI and XhoI and ligated to NcoI/XhoI digested pET28b plasmid vector

(Novagen). Ligation products were introduced into *E. coli* ABLE-K chemically competent cells and recombinant strains were selected on LB containing Kanamycin (30 µg /mL). Plasmid DNA from positive clones was purified, digested with NcoI/XhoI and plasmid digests were subject to agarose gel electrophoresis to verify the presence of correctly sized insert. Confirmed recombinant pET-gulA and pET-gulF variant constructs were then introduced into the *E. coli* expression host, BL21 (DE3) (Invitrogen) for recombinant protein expression. Select positive clones were grown in LB-Kan (30 µg/mL) to approximately mid log phase and glycerol stocks (25 % glycerol) were snap frozen for storage at -70°C.

[0246] Small scale expression test of expression constructs

[0247] Small scale expression was performed to check for soluble expression levels as follows. A single bacterial colony from freshly transformed LB-Kan plates was transferred to 5 mL LB-Kan media and grow O/N at 37°C with shaking at 200 rpm. These starter cultures were then used to inoculate fresh LB-Kan media at a 1:100 ratio of 5 mL culture in a 50 mL falcon tube. Cell cultures were grown at RT (approximately 24-26°C) until the cell density reached OD600 ≈ 0.6 - 0.8, then induced with 0.5 mM IPTG at RT for 16 hours. Two mL of the cell culture was centrifuged in an Eppendorf tube and resuspended in 0.5 mL of lysis buffer [20 mM Na-phosphate, 500 mM NaCl, 0.5% (v/v) TritonX-100, 20 mM Imidazole, 1x proteinase inhibitor, pH7.8] followed by brief sonication. A sample of total cell lysate was collected for electrophoretic analysis and the remainder was centrifuged at full speed using a bench top Eppendorf centrifuge. The soluble fraction was collected and the insoluble fraction was resuspended in Lysis Buffer. Total cells, soluble cell fraction and Insoluble cell fraction equivalent to 4 µL of induced cell culture were then subjected to SDS-PAGE to assess expression and solubility.

[0248] Mini purification of soluble recombinant proteins by Ni-NTA spin column

[0249] A mini-scale purification was performed on the two recombinant *P. gulate* constructs to assess for His-Tag integrity/availability for Nickel chromatography purification as follows. Soluble protein fraction (0.5 mL taken from the 2 mL induced culture above) was loaded onto Ni-NTA spin column according to manufacturers' instructions and the eluate representing a crude Nickel purification was collected, the concentration measured and approximately 3 to 5 µg of purified r-protein (alongside

samples of the flow through and washed column fractions) were resolved on SDS-PAGE analysis.

[0250] Protein expression and cell lysis of His-tagged candidates

[0251] The His-tagged vaccine candidates for the animal model were expressed as recombinant C-terminally His-tagged fusion proteins (SEQ ID NO: 30 and 31, plus His tag) in *E. coli* BL21(DE3) as previously described. The cells were grown at 37°C in LB medium or Terrific broth supplemented with 50 µg/mL kanamycin. At culture OD600 = 1.0-1.2, protein expression was induced with 0.2 mM IPTG at 32°C for 2-3 hours. The cells were harvested at 8000 g by centrifugation at 4°C and stored at -80°C for later use. The cells were lysed on ice by sonication for 15 min with 35% power output after thawed and resuspended in the lysis buffer [10 mM imidazole, 50 mM Tris, 300 mM NaCl, pH 7.5, 1x EDTA free protease inhibitor cocktail (Sigma) (for the reducing lysis buffer for the four Cys antigen-F: supplemented with 10 mM DTT)]. Cell lysate was clarified by centrifugation for 40 min at 25,000 g at 4°C.

[0252] Purification of a His-tagged candidate from inclusion bodies

[0253] Following lysis of cells, the insoluble pellet was washed twice with buffer PBS500 (20 mM NaPi, 500 mM NaCl, pH 7.4). The protein expressed as inclusion bodies was solubilised with 8 M urea at room temperature on a rolling platform for 1 hour in PBS500. The protein extract was centrifuged at 20,000 g. The supernatant was further filtered through a 0.22 µm filter unit and then mixed with the nickel affinity resin (ThermoFisher) with gentle stirring for 2 hours in the urea containing buffer PBSU (8 M urea, 20 mM NaPi pH 7.8, 500 mM NaCl, plus 20 mM imidazole). After extensive wash with PBSU (pH 7.8 and then pH 6.5), the bound target protein was eluted off the resin with 500 mM imidazole in the same buffer. The eluted protein was stepwise dialysed into 6 M, 4 M and then 2 M urea phosphate buffers in a dialysis tube with molecular weight cutoff of 3.5 kDa (Fisher Biotech, Australia). Target protein in the 2 M urea buffer was buffer exchanged using a PD10 gel filtration column to remove urea before animal model experiments.

[0254] Purification of His-tagged candidates from soluble fractions

[0255] *Ni-affinity chromatography*

[0256] This is the first chromatographic step for purification of His-tagged proteins. The clarified cell lysates were filtered through a 0.22 µm filter unit and loaded to a HisTrap or HisPrep Ni-affinity column (GE Healthcare) in the loading buffer TBS300 (50 mM Tris-Cl, 300 mM NaCl, pH 7.5), plus 10 mM imidazole. For reducing purification, cell lysate was diluted 4 fold with the imidazole containing TBS300 and then filtered before loaded to the column in the reducing loading buffer with inclusion of 2 mM DTT. Columns were washed extensively with appropriate loading buffer and then 20 mM imidazole in TBS300 or plus 2 mM DTT for reducing conditions. Bound proteins were eluted with a 20-350 mM imidazole gradient in TBS300 with the absorbance being monitored at 280 nm. Peak fractions were analysed with SDS-PAGE. Eluted target proteins were concentrated using the Amicon filter units with a 10 kDa molecular weight cut-off. The resulting protein solutions were stored on ice for further purification.

[0257] *Anion exchange chromatography*

[0258] The online ExPASy ProtParam tool (<https://web.expasy.org/protparam/>) predicted that the antigens had an acidic isoelectric point (pI), 5.07 for antigen-A and 4.87 for antigen-F. Thus, anion exchange chromatography was applied following the Ni-affinity chromatography step. Briefly, the concentrated proteins from Ni-affinity purification were diluted 10 folds into Buffer A (50 mM Tris.Cl, 20 mM NaCl, pH 7.5; 2 mM DTT for reducing buffer) to reduce ionic strength and then loaded to an anion exchange HiTrap Q column (GE Healthcare) in Buffer A. The proteins were eluted with a NaCl gradient from 50 to 350 mM and then from 350 to 700 mM. The absorbance at 280 nm was monitored in the process. Target proteins in the peak fractions were verified with SDS-PAGE and concentrated using the Amicon filter units. The protein solutions were stored on ice for further purification and buffer exchange using size exclusion chromatography.

[0259] Size exclusion chromatography

[0260] Size exclusion chromatography was performed on a HiLoad Superdex 200 column (GE Healthcare). The target protein solutions from anion exchange purification were concentrated and loaded to the size exclusion column, eluting in the buffer of TBS150 (50 mM Tris, 150 mM NaCl, pH 7.5) with absorbance at 280 nm being monitored. Target proteins in the peak fractions were verified with SDS-PAGE. The fractions of the best protein purity were pooled and concentrated. After quantitated by determination of

absorbance at 280 nm on a Cary UV-vis spectrometer and calculation with theoretical extinction coefficients, target proteins were stored at -80°C in aliquots for future use.

[0261] ESI-LC-MS intact protein analysis

[0262] Intact protein mass spectrometric analysis was performed on an electrospray ionisation time-of-flight mass spectrometer (ESI-TOF) coupled with liquid chromatography. With programmed automation, protein samples passed through a C4 HPLC column (Phenomenex) with 0.1% formic acid aqueous solution as buffer A and acetonitrile + 0.1% formic acid as buffer B. Acetonitrile gradient for elution was set to be 5–60% and then 60-95% buffer B for a total of 15 minutes. MS data were collected from 400 to 3200 m/z on the Agilent 6520 QTOF mass spectrometer operated in positive mode, with in-situ internal mass reference standards. Mass spectra were deconvoluted to obtain the intact protein molar masses using Agilent Mass Hunter Qualitative Analysis software (B.05) with maximum entropy algorithm.

[0263] SEC-MALS

[0264] SEC-MALS were run in buffer TBS150 at room temperature on a SEC-MALS system (Wyatt Technology). Species were resolved on a HPLC size exclusion column (Shim-Pack Bio Diol-300) and passed to the Wyatt 18-angle light scattering detector and Wyatt refractive index monitor. The concentrations of the samples were 2 mg/mL for the antigens and 5 mg/mL for the control protein BSA which was used to set up the method. Measurements were started following the stepwise instructions on screen of the control station with auto-injected volume set to be 10 µL for each run. Data was processed with Wyatt Technology ASTRA software for the information about sample homogeneity, protein aggregation and molar masses of the species in the solution.

[0265] Determination of endotoxin

[0266] The protein samples were diluted to 0.5 mg/mL using freshly prepared storage buffer TBS150. Endotoxin content was determined using the commercial Pierce Chromogenic Endotoxin Quantitation kit (Thermo Scientific) as per manufacturer instructions.

[0267] Assessment of nucleic acid contamination

[0268] A260/A280 ratio was determined on a NanoDrop lite UV spectrophotometer (Thermo Scientific) to assess nucleic acid contamination.

[0269] Culture of Bacteria for Mouse Model of Periodontitis

[0270] *P. gulae* was obtained from the culture collection of the Oral Health Cooperative Research Centre, The Melbourne Dental School, University of Melbourne, Australia. *P. gulae* was cultivated (5% CO₂, 10% H₂, and 85% N₂) in an anaerobic chamber (Whitley MG500 anaerobic workstation) at 37°C in brain–heart infusion (BHI) broth (BD Bacto Laboratories, USA) supplemented with cysteine (1 g/L; Sigma-Aldrich, Australia), tryptic soy broth (5% w/v, BD Bacto Laboratories, USA), hemin (5 µg/mL, Calbiochem, Netherlands), and Vitamin K (10 µg/mL, Sigma-Aldrich, USA).

[0271] Absorbance of batch cultures were monitored at OD_{650nm} using a spectrophotometer (model 295E, Perkin-Elmer, Germany). Bacterial cells were harvested during late exponential growth by centrifugation (7,000 g, 20 min, 4 °C). Bacterial purity was routinely confirmed by Gram stain (Slots 1982).

[0272] Preparation of Heat Killed Bacteria

[0273] *P. gulae* culture was harvested (6,500 g, 4°C), washed once with phosphate buffered saline (PBS) (0.01 M Na₂HPO₄, 1.5 mM KH₂PO₄ and 0.15 M NaCl, pH 7.4) then pelleted by centrifugation (7,000 g, 20 min 4°C). Bacterial cells were resuspended in PBS and heated to 65°C for 15 minutes. The suspension was centrifuged (7,000 g, 20 min 4 °C) and resuspended in sterile PBS and this was repeated once. After the second wash, the supernatant was discarded and the cell pellet was resuspended in sterile PBS to obtain a cell density of 2 x 10¹⁰ cells/mL, and protein concentration determined using Biorad Protein Assay Dye Reagent Concentrate (Life Science, NSW, Australia).

[0274] Animal Ethics

[0275] All animal experimental procedures were carried out in strict accordance with the recommendations in the Australian Code of Practice for the Care and Use of Animals for Scientific Purposes.

Example 2: Results of *in vitro* studies

[0276] Antigen expression and solubility- small scale

[0277] Small scale expression tests of recombinant proteins comprising either DUF variants A or F, induced with 0.5 mM IPTG at RT, showed that both variants expressed relatively high levels of soluble recombinant protein. Fig. 1 shows recombinant proteins of the expected size in the soluble fractions that are very prominent relative to the background of *E. coli* soluble proteins, suggestive of high levels of soluble expression. At RT, the chimeric protein comprising DUF variant A exhibited very high levels of overall solubility with little or no recombinant protein present in the insoluble fraction that represents inclusion bodies. By comparison the chimeric protein comprising the DUF variant F exhibited a high level of total recombinant protein expression in the insoluble fraction. Nevertheless, despite the relatively high level of KD_FAK-2S-AVQP recombinant protein as insoluble inclusion bodies, there was still a relatively high level of soluble protein expressed in the soluble fraction due to the extremely high overall level of total recombinant protein expressed by this strain under the conditions tested.

[0278] Antigen expression and solubility- large scale

[0279] Both KD_FAK-2S-AVQP and KD_AAK-2S-AVQP proteins were expressed well at a similar level in either LB or TB medium (Fig. 4). Under non-reducing conditions, antigen-A had a high solubility as evidenced by SDS-PAGE analysis of the lysis fractions (Fig. 2). In contrast, the KD_FAK-2S-AVQP protein was much less soluble with its major portion being found in the precipitate after lysis under the same non-reducing conditions (Fig. 2). In addition, reducing buffer containing either low (5 mM) or high (100 mM) concentration of DTT failed to extract antigen-F from the precipitate fraction (Fig. 2). However, when the cells were lysed under reducing conditions, KD_FAK-2S-AVQP was highly soluble with most protein being in the soluble fraction (Fig. 2). The low solubility of KD_FAK-2S-AVQP was likely due to the presence of disulfide bonds formed by wrongly paired Cys residues which caused the protein to have a misfolded structure. Once this abnormal structure formed, it was irreversible and its disulfide bonds were not accessible to the reductant if without assistance of unfolding force such as SDS and heating. Thus, purification of KD_FAK-2S-AVQP from soluble fractions under non-reducing conditions was not an ideal approach although low temperature expression may improve soluble expression levels for this protein in the beginning as seen from small scale trial. In the end, the samples of this protein for the animal model experiments were prepared from lysis and purification under reducing conditions until the last step for removal of the included reductant DTT.

[0280] Purification of the His-tagged control antigen from inclusion bodies

[0281] This procedure was only applied to purification of a prior art *P. gingivalis* chimera (termed KDcAK1n) which was expressed in the inclusion bodies as described previously (O'Brien-Simpson et al. 2016, NPJ Vaccines 1:16022). KDcAK1n was stable in 2M urea after stepwise dialysed into 2 M urea phosphate buffer and thus stored at -80°C after quantitated for further use. No precipitation was observed after buffer exchange with gel filtration before the animal model experiments.

[0282] *P. gultae* antigen purification

[0283] For the animal model experiments described in examples 3 and 4, KDAK-2S-AVQP was purified under non-reducing conditions and KDFAK-2S-AVQP was purified under reducing conditions until the final buffer exchange and size exclusion purification step (Figs. 3 to 5). Non-reducing conditions for the purification was also trialed for KDFAK-2S-AVQP. KDAK-2S-AVQP expressed in LB medium was used for purification under non-reducing conditions. KDFAK-2S-AVQP expressed in TB medium was used for purification under reducing conditions and the protein expressed in LB was used for non-reducing purification. KDAK-2S-AVQP had the highest yield (78 mg/L culture, 18.6 mg/g wet cells) (Table 1). KDFAK-2S-AVQP also had a high yield of its final product when purified from lysis supernatant under reducing conditions (34 mg/L culture, 7.6 mg/g wet cells), however when purified under non-reducing conditions its yield was much lower (3.5 mg/L culture, 1.1 mg/g wet cells) (Table 1). Identities of both proteins were confirmed with the first Met residue missing by intact protein MS spectrometry (Table 3).

[0284] Table 3: Final products of *P. gultae* antigens purified from 1 L culture in LB or TB Broth.

Antigens	aa ¹	Expected MW (Da) (Met 1 off)	Measured MW (Da) ²	£280	Concentration (mg/mL) ³	Yield (mg)/L culture
<i>P. gultae</i> _cHis-KDAK-2S-avqp-A	358 (-1)	38274.88	38275.69 38458.98 38641.78	55350	20.470	78 (18.6 mg/g wet cells)
<i>P. gultae</i> _cHis-KDAK-2S-avqp-F_R	377 (-1)	40480.72	40478.85 40661.63	45840	17.951	34 (7.6 mg/g wet cells)
<i>P. gultae</i> _cHis-KDAK-2S-avqp-F_NR	377 (-1)	40480.72	40477.92 40660.87 40841.57	45960	2.139	3.5 (1.1 mg/g wet cells)

¹ (-1) denotes the missing first Met residue.

² The MWs in red were detected as minor peaks with additional 1 or 2x 183 Da to the measured target monomeric molar masses.

³ The listed concentrations were determined based on the absorbance and theoretical extinction coefficients at 280 nm.

NB: All the final products of these antigens were stored in TBS buffer (50 mM Tris, 150 mM NaCl, pH 7.5).

[0285] KD_FAK-2S-AVQP and KD_AAK-2S-AVQP were concentrated up to 20 and 18 mg/mL, respectively, in TBS150 buffer and these concentrations were not the maximum that could be achieved (Table 1). The majority of expressed KD_FAK-2S-AVQP precipitated into the insoluble fraction from non-reducing lysis. Once it was purified under reducing conditions to a high quality, it stayed in solution without reversion to precipitation even when the reducing factor was removed. This solution stability was also resistant against freeze-thaw cycles in air. Thus, it was possible to prepare enough material from one round of purification for the animal model.

[0286] Removal of DnaK

[0287] The host molecular chaperone protein DnaK of ~70 kDa appeared to contaminate the purification of the *P. gulyae* antigens. Due to its interactions with target proteins, DnaK eluted into the fractions with most overlapping the target protein fractions in the Ni affinity purification step. This contamination, including other impurities, was well separated from target proteins by anion exchange chromatography (Fig. 4). The interactions between DnaK and the target proteins did not seem to be dependent on reducing conditions. Nevertheless, under reducing conditions, KD_FAK-2S-AVQP eluted into two peaks possibly due to the presence of two forms (Fig. 4). Essentially, almost all the fractions in the second peak had DnaK contamination. This was possibly due to a small portion of misfolded antigen-F that had higher interactions with reduced DnaK. Therefore, the second peak was excluded and the best fractions in the early major peak were used for further purification with size exclusion chromatography (Fig. 5).

[0288] Degradation and site prediction

[0289] Some minor bands could be seen below each major full-length target protein on the SDS gels of the anion exchange fractions including the doublet bands in some fractions for KD_FAK-2S-AVQP, suggesting the occurrence of degradation (Fig. 6). Most degradants were essentially removed by size exclusion chromatography (Fig. 5) and a

high quality of final product for each protein was thus achieved despite existence of minor degraded species (Fig. 6). The high purity and high homogeneity of the final products were also supported by SEC-MALS analysis, which showed the predominant existence of monomer in solution at 2 mg/mL with high stability for both proteins (Table 4).

[0290] Table 4: Peak area percentages in SEC-MALS chromatograms of the *P. gulae* antigens

Peak#	KD _A AK-2S-AVQP		KD _F AK-2S-AVQP _R ^a		KD _F AK-2S-AVQP _NR ^b	
	Ret. Time	Area%	Ret. Time	Area%	Ret. Time	Area%
1	8.133	0.78	8.083	0.73	7.4	0.599
2	8.917	98.948	8.85	98.954	8.858	99.14
3	12.883	0.272	12.908	0.316	12.925	0.26

^a purified under reducing conditions before the final size exclusion step;

^b purified under non-reducing conditions.

[0291] Intramolecular disulfide bonds in antigen-F

[0292] Slight band shift exhibited in the SDS-PAGE profile of the F-version antigen (KD_FAK-2S-AVQP) in response to the reducing conditions, suggesting the presence of intramolecular disulfide bonds (Fig. 9A). Even for KD_FAK-2S-AVQP purified involving reducing purification procedures (F-R), it had a similar band shift to the protein purified under non-reducing conditions (F-NR). F-R also formed intramolecular disulfide bonds upon removal of reducing factors at the size exclusion purification step. This intramolecular disulfide bridge may play a role in maintaining the protein structure and make the protein more compact, resulting in a faster mobility in the SDS gel under non-reducing denaturing conditions (Fig. 6A). There was no such band shift for KD_AAK-2S-AVQP (Fig. 6A). The slight reduction dependent band shift for KD_FAK-2S-AVQP was not resolved in native PAGE of unheated samples (Fig. 6B). In contrast to the results from SDS-PAGE, for heated samples on native gel without SDS, the negative charge on thiol ions of reduced KD_FAK-2S-AVQP appeared to take effects on the protein mobility and make the protein move faster than the unreduced sample (Fig. 6B).

[0293] F-NR seemed to have a band shift slightly more than F-R on the SDS gel, whether the samples were heated or not (Fig. 6A). It was hypothesized that F-NR formed intramolecular disulfide bonds as early as at the stage of lysis and F-R formed the

disulfide bonds only after removal of the reducing agent at the final purification stage. This difference may have a subtle impact on the protein structure. However, F-R purified from the reducing agent was found to have a similar overall structure to F-NR with minor difference in compactness.

[0294] Oligomeric state

[0295] A faint band at around 90 kDa appearing in the SDS gel lanes of heated non-reduced samples of F-R and F-NR suggested the presence of a small amount of dimer due to intermolecular disulfide formation (Fig. 6A). It is possible that this was an artifact from heating denaturation to temporarily expose the Cys residues for intermolecular disulfide formation as this was not the case for the samples unheated (Fig. 6A).

[0296] Since no effects of reducing conditions were seen on the PAGE profiles of KD_FAK-2S-AVQP in native gels, whether the samples were either heated or not (Fig. 6B), a possibility of progressive formation of multimers bridged by intermolecular disulfide bonds was excluded for this containing protein.

[0297] SEC-MALS data showed some signs of minor or minimal multimerisation of the two antigens. The species with estimated molar masses of 71.2 kDa for antigen-A and 84.7 kDa for F-R (reduction involved purification) may be assigned to the dimeric forms of the two proteins. As for F-NR (purified under non-reducing conditions), a peak of 394.2 kDa may be due to the presence of a small amount of decamer of this protein, although this species was not detected by SDS-PAGE. Cys residues in KD_FAK-2S-AVQP may have made this protein prone to form a higher level of multimerisation when it was in non-reducing conditions from the stage of lysis. Therefore, the trace amount of decamer may be intermolecular disulfide linked. Importantly, these multimerised species were less than 1% of the total protein amount as estimated by integration of the SEC-MALS liquid chromatograms (Table 4).

[0298] Assessment of endotoxin and nucleic acid contaminations

[0299] Endotoxin was determined to have similarly low levels in the two antigens. At a protein concentration of 0.5 mg/mL, both proteins had endotoxin contamination at 2.24 and 2.36 EU/mL, respectively (Table 5). In addition, the A260/A280 value was lower than 0.6 for both antigens, indicating minimal contamination with nucleic acids (Table 5).

[0300] Table 5: Assessment of endotoxin and nucleic acid contamination in *P. gulae* antigens.

Antigens	Buffer ^a	A260/A280 (antigen 0.5 mg/mL)	Endotoxin (EU/mL, at antigen 0.5 mg/mL)
cHis-KDAK-2S-avqp_F	TBS150	0.56	2.36
cHis-KDAK-2S-avqp_A	TBS150	0.57	2.24

^a TBS150: 50 mM Tris, 150 mM NaCl, pH 7.5.

Example 3: materials and methods for mouse *in vivo* studies

[0301] Mouse Periodontitis Model

[0302] Mice (female BALB/c; 6-8 weeks old, 10 mice/group), on Day 0 were intra-orally inoculated with *P. gulae* consisting of four doses of *P. gulae* [1×10^{10} viable *P. gulae* cells per dose suspended in 20 μ L PG buffer (50 mM Tris-HCL, 150 mM NaCl, 10 mM MgSO₄ and 14.3 mM mercaptoethanol, pH 7.4) containing 2% w/v carboxymethylcellulose (CMC, Sigma, New South Wales, Australia)], with each dose given two days apart. The inoculum was prepared anaerobically and then immediately applied to the gingival margin of the maxillary molar teeth. The number of viable bacteria in each inoculum was verified by flow cytometry and CFU counts on blood agar. Groups of animals consisted of: *P. gulae* orally inoculated (infected control), a non-bacterial inoculated control, and immunised groups. For the therapeutic vaccination periodontitis model 21 mice were immunised on day 19 after the first oral inoculation with 100 μ g of vaccine candidate in saline/alum (Alhydrogel; 2% aluminium hydroxide wet gel suspension; Invivogen) via the intraperitoneal route. Mice received a second immunisation (100 μ g in saline/alum) on day 40 via the subcutaneous route. On Day 62, mice were bled by cardiac puncture and killed. Maxillae were removed and halved through the midline, with 10 halves used to determine alveolar bone loss. Sera were used to determine the antibody profile using ELISA.

[0303] Measurement of alveolar bone loss in mouse maxillae

[0304] Maxillae to be examined for bone loss were boiled (1 min) in deionised water, mechanically defleshed, and immersed in 2% w/v potassium hydroxide (16 h, 25°C). Maxillae were washed twice with deionised water (25 °C), dried (1 h, 37 °C) and stained

with 0.5% w/v aqueous methylene blue. Coded digital images of the buccal aspect of the maxillae were captured with an Olympus DP12 digital camera mounted on a dissecting microscope, using ImageJ imaging software for analysis (<https://imagej.nih.gov/ij/index.html>) to assess horizontal bone loss. Maxillae were oriented so that the buccal and lingual molar cusps were superimposed. Images were captured with a micrometre in frame, so that measurements could be normalised for each image. Horizontal bone loss was defined as the loss occurring in a horizontal plane, perpendicular to the alveolar bone crest that resulted in a reduction of the crest height. The visible area from the cemento-enamel junction (CEJ) to the alveolar bone crest (ABC) for each molar was measured using ImageJ version 1.3k imaging software to give the total visible CEJ-ABC area in mm². *P. gulae*-induced alveolar bone loss in mm² was calculated by subtracting the total visible CEJ-ABC area of the uninoculated (N-C) group from the total visible CEJ-ABC area of each experimental group. Alveolar bone loss measurements were determined twice in a random and blinded protocol. Data are expressed as the mean +/- standard deviation in mm² and were analysed using a one-way ANOVA and Dunnetts T3 post-hoc test.

[0305] Enzyme-Linked Immunosorbent Assay (ELISA)

[0306] ELISAs were performed to evaluate subclass antibody in sera using a solution (1 µg/mL) of either heat-killed (HK) *P. gulae* cells, recombinant domain subunits or biotinylated peptide epitopes in 0.1M PBS (pH 7.4) to coat wells (16 h, 4 °C) of flat-bottom polyvinyl microtitre plates (Microtitre; Dynatech Laboratories, McLean, VA, US).

[0307] In these experiments the following antibody dilutions were used; a dilution of 1/4000 dilution of goat anti-mouse; IgG (M8642), IgG1 (M8770), IgG2a (M4434) antibodies (Sigma, New South Wales, Australia). A 1/4000 dilution of a horseradish peroxidase-conjugated swine anti-goat IgG antibody (M5420; Sigma, New South Wales, Australia) was used to develop ELISA experiment. For the epitope ELISAs biotinylated peptides were bound to pre-blocked streptavidin coated flat bottom plates (Pierce; Thermo-Fisher) at 10 µg/mL. Following incubation with sera, the ELISA was developed with 1/4000 goat anti-mouse IgG and 1/4000 horseradish peroxidase-conjugated swine anti-goat IgG antibody. All optical density measurements were conducted on a Wallac VICTOR3 1420 Multilabel counter (Perkin Elmer) at 405nm.

Example 4: results of mouse *in vivo* studies

[0308] *P. gulae*-induced Alveolar bone loss in mouse maxillae.

[0309] The animal model used was the therapeutic vaccine mouse periodontitis model (Fig. 7) developed by O'Brien-Simpson et al (2016, *supra*). Compared to the naïve control, the infected control animals developed significant levels ($P < 0.001$) of alveolar bone loss (Fig. 8). The original *P. gingivalis* chimera (KDcAK1n) was used as a positive vaccine control in this study. New vaccines of the present invention and based on the *P. gulae* virulence domain sequence, KD_FAK-2S-AVQP and KD_AAK-2S-AVQP provided significant levels of protection compared to the infected control animals ($P < 0.01$) (Fig. 8).

[0310] Antibody response

[0311] Serum antibody subclass responses of immunised mice in the periodontitis model were examined by ELISA. Antisera were used to probe heat killed *P. gulae* as the adsorbed antigen (Fig. 9). Antibody responses are expressed as the ELISA titre obtained minus double the background level, with each titre representing the mean $\pm$ s.d. of the 10 individual mice. All antigens tested generated IgG responses towards *P. gulae*, with the strongest response observed with the *P. gingivalis* KDcAK1n vaccine. The two *P. gulae* vaccines, KD_FAK-2S-AVQP and KD_AAK-2S-AVQP, generated significant levels of antibodies against whole cells compared to the negative control. (Fig. 9).

[0312] Antibody titres against *P. gulae* purified RgpA/Kgp protease complex (Fig. 10) were also measured.

[0313] Both the KD_FAK-2S-AVQP and KD_AAK-2S-AVQP proteins generated significant antibodies against the RgpA/Kgp protease complex, with KD_FAK-2S-AVQP generating higher titres against the RgpA/Kgp protease complex. Finally, Fig. 11 shows that the two *P. gulae* vaccines generated high titre antibody responses against themselves.

[0314] In conclusion the inventors have generated *P. gulae* vaccines using *P. gulae* specific sequences derived from a complex polyprotein gingipain. Both proteins tested protected as vaccines in the mouse animal model.

Example 5: materials and methods for canine *in vivo* studies

[0315] A series of experiments were planned for determining the ability of a chimeric fusion protein to raise an immune response in dogs.

[0316] 9 dogs were divided into 3 groups as follows:

- Group 1 – low dose: (Jake 3962, Phoebe 0346 and Rachael 8562) were immunised with a low dose (100 µg) of *Porphyromonas gulae* KDAK-2S-avqp-6His recombinant protein.
- Group 2 - mid-dose: (Forest 5635, Milan 3711 and Moo 5498) were immunised with a medium dose (200 µg) of *Porphyromonas gulae* KDAK-2S-avqp-6His recombinant protein.
- Group 3 – high dose: (Bruno 8086, Kale 5636 and Lachie 5634) were immunised with a high dose (400 µg) of *Porphyromonas gulae* KDAK-2S-avqp-6His recombinant protein.

[0317] Dogs were immunised twice, first on Day 0 (“prime”) and again on Day 21 (“boost”), according to the protocol in Fig. 12. For each dog, serum samples were collected on:

- day 0 (= pre-immune serum, aka pre-treatment, i.e., immediately prior to receiving the prime dose).
- day 21 (= primed serum, aka pre-second treatment, i.e., 21 days after receiving the prime dose and immediately prior to receiving the booster).
- day 35 (= boosted serum, i.e., 14 days after receiving the booster).

[0318] Optimization ELISAs

[0319] *Purified Porphyromonas gulae KDAK-2S-avqp-6His recombinant protein*

[0320] ELISA plate wells were coated with 100 µL/well of *Porphyromonas gulae* KDAK-2S-avqp-6His recombinant protein at concentrations of 10, 2.5, 1, 0.5 and 0.1 µg/mL (dilutions made in 1×PBS). A PBS-only control was also included. After overnight incubation at 4 °C, the recombinant protein solution was discarded, wells were washed twice with distilled water, and free protein binding sites were blocked by the addition of 200 µL/well of Block Solution. After incubation at 4 °C for 6 hours, the Block Solution was discarded, wells were rinsed twice with Wash Solution, and serum dilutions were applied as described below.

[0321] *Biotinylated peptides*

[0322] ELISA plate wells were coated with 100 μ L/well of 10 μ g/mL streptavidin solution (dilution made in 1 $\times$ PBS). After overnight incubation at 4 $^{\circ}$ C, the streptavidin solution was discarded, wells were washed twice with distilled water, and free protein binding sites were blocked by the addition of 200 μ L/well of Block Solution (10% (w/v) non-fat milk powder in 1 $\times$ PBS). After incubation at 4 $^{\circ}$ C for 6 hours, the Block Solution was discarded, wells were rinsed twice with Wash Solution, and aliquots of 100 μ L of biotinylated synthetic peptides (Pgul_KAS2 (SEQ ID NO: 1) and P. gul_KAS2_scrambled: Biotin-KYKGWTNNSSTVLLQTNATLGVETFTHSPDSASDAK, SEQ ID NO: 113) prepared in Dilution Buffer at 10, 2.5, 1, 0.5 and 0.1 μ g/mL were added to the blocked streptavidin-coated wells. A Dilution Buffer-only control was also included. After overnight incubation at 4 $^{\circ}$ C, the peptide solutions were discarded, wells were washed six times with Wash Solution, and serum dilutions applied as described below.

[0323] *Whole heat-killed Porphyromonas gulae cells*

[0324] ELISA plate wells were coated with 100 μ L/well of whole heat-killed *Porphyromonas gulae* cells at concentrations of 10, 2.5, 1, 0.5 and 0.1 μ g/mL (dilutions made in 1 $\times$ PBS). A PBS-only control was also included. After overnight incubation at 4 $^{\circ}$ C, the *P. gulae* cells were discarded, wells were washed twice with distilled water, and free protein binding sites were blocked by the addition of 200 μ L/well of Block Solution. After incubation at 4 $^{\circ}$ C for 6 hours, the Block Solution was discarded, wells were rinsed twice with Wash Solution, and serum dilutions were applied as described below.

[0325] Optimization ELISAs: Serum application

[0326] Aliquots of each of the Day 0 serum samples were pooled and diluted 1/10 in Dilution Buffer. Similarly, the day 35 serum samples were pooled and diluted 1/10 in Dilution Buffer. Aliquots (100 μ L) of the 1/10 diluted serum samples were added to the blocked ELISA plates and 4-fold dilution series of the 1/10 serum samples were constructed by serially transferring 25 μ L into 75 μ L of Dilution Buffer. Control wells that received only Dilution Buffer (i.e., no serum) were also included. Plates were incubated overnight at 4 $^{\circ}$ C and detection antibody was applied as described below.

[0327] Detection Antibody

[0328] After overnight incubation, serum samples were discarded, and the wells were rinsed six times with Wash Solution. Aliquots (100 μ L) of horseradish peroxidase (HRP)-conjugated Goat anti-Dog IgG (Fc specific) antiserum (10 mg/mL), diluted 1/5,000 in Dilution Buffer, were added to each well. After incubation at room temperature (22 °C) for 2 hours, the labelled antibody was discarded, wells were rinsed six times with Wash Solution and HRP substrate added as described below.

[0329] HRP substrate

[0330] TMB Substrate Buffer Solution (90 μ L) was added to the washed ELISA plates. For optimization ELISAs, colour development was followed spectrophotometrically at 370 nm in kinetic mode using a SpectraMax iD5 spectrophotometer. For titring of individual sera, the colour was allowed to develop for 20 minutes, the reaction stopped by the addition of 50 μ L of 1 M H₂SO₄ to each well, and the plate read in stopped ELISA mode (at 450 nm) in the SpectraMax iD5 spectrophotometer. Data was analysed as described below.

[0331] ELISAs for titring of individual dog serum**[0332]** *Purified Porphyromonas gulae KDAK-2S-avqp-6His recombinant protein*

[0333] ELISA plate wells were coated with 100 μ L/well of 0.1 μ g/mL *Porphyromonas gulae* KDAK-2S-avqp-6His recombinant protein (dilution made in 1 $\times$ PBS). After overnight incubation at 4 °C, the recombinant protein solution was discarded, the wells were washed twice with distilled water and free protein binding sites were blocked by the addition of 200 μ L/well of Block Solution. After incubation at 4 °C for 6 hours, the Block Solution was discarded, and the wells were rinsed twice with Wash Solution. Aliquots (100 μ L) of pre-immune, primed and boosted sera from each dog were diluted 1/10, 1/10 and 1/500, respectively, in Dilution Buffer, added to the ELISA plate, and a 1/5 dilution series was constructed by serially transferring 20 μ L into 80 μ L of Dilution Buffer. Dilution Buffer-only control wells (i.e., no serum) were also included. After overnight incubation at 4 °C, detection antibody was applied as described above.

[0334] *Biotinylated P_gul_KAS2 peptide*

[0335] ELISA plate wells were coated with 100 µL/well of 10 µg/mL streptavidin solution (dilution made in 1×PBS). After overnight incubation at 4 °C, the streptavidin solution was discarded, the wells were washed twice with distilled water and free protein binding sites were blocked by the addition of 200 µL/well of Block Solution. After incubation at 4 °C for 6 hours, the Block Solution was discarded, wells were rinsed twice with Wash Solution, and 100 µL aliquots of 0.1 µg/mL biotinylated P_gul_KAS2 peptide (dilution made in Dilution Buffer) were added to each well of the streptavidin-coated plates. After overnight incubation at 4 °C, the peptide solution was discarded, and wells were rinsed six times with Wash Solution. Aliquots (100 µL) of pre-immune, primed and boosted sera from each dog were diluted 1/10 in Dilution Buffer, added to the ELISA plate, and a 1/5 dilution series was constructed by serially transferring 20 µL into 80 µL of Dilution Buffer. Dilution Buffer-only control wells (i.e., no serum) were also included. After overnight incubation at 4 °C, detection antibody was applied as described above.

[0336] *Whole heat-killed Porphyromonas gulae cells*

[0337] ELISA plate wells were coated with 100 µL/well of 1 µg/mL whole heat-killed Porphyromonas gulae cells (dilution made in 1×PBS). After overnight incubation at 4 °C, the P. gulae cells were discarded, the wells were washed twice with distilled water and free protein binding sites were blocked by the addition of 200 µL/well of Block Solution. After incubation at 4 °C for 6 hours, the Block Solution was discarded, and wells were rinsed twice with Wash Solution. Aliquots (100 µL) of pre-immune, primed and boosted sera from each dog were diluted 1/10 in Dilution Buffer, added to the ELISA plate, and a 1/5 dilution series was constructed by serially transferring 20 µL into 80 µL of Dilution Buffer. Dilution Buffer-only control wells (i.e., no serum) were also included. After overnight incubation at 4 °C, detection antibody was applied as described above.

[0338] Data Analysis: Determination of Antibody Titres

[0339] ELISA assays based on peroxidase-TMB systems can be monitored in two ways; continuously recording absorbance at 370 nm; or acid-stopping the reaction following a set time and measuring absorbance at 450 nm. Continuous assays tend to be more accurate with a wider dynamic range than stopped assays but become impractical with multiple plates. Both assay types were deployed during this trial.

[0340] All analyses were performed using scripts written for R-4.2.1 (R Foundation). Initial slopes of continuously read ELISA assays were calculated by quadratic regression of the absorbance response to account for curvature. The linear parameter of the quadratic equation corresponds to the slope at time = 0 (i.e., initial slope). Dose response curves for both continuous and acid-stopped assays were calculated using the drc-library of functions for R. A 4-parameter log-logistic equation (LL.4) was used for all samples (Equation 1, below). To cope with samples with low responses which did not cover the full range, the maximum and minimum parameters were determined at a plate level, the slope parameter was set at 1.0, while the inflection point parameter was unique for each serum sample.

[0341] Two antibody titre values were calculated and reported: midpoint titres are equivalent to the inflection point parameter in the LL.4 model. Endpoint titres were calculated as the predicted point on the LL.4 curve corresponding to twice the minimum response.

[0342] Equation 1:

$$f(x) = a + \frac{b - a}{1 + \left(\frac{x}{d}\right)^c}$$

[0343] where : x = concentration of serum; f(x) = absorbance response; a = maximum response; b = minimum response; c = slope (Hill coefficient); d = Inflection point (midpoint)

[0344] Example 6: results from initial canine serology studies

[0345] Figures 13 to 15 show preliminary results from the experimental protocol outlined in Example 5. Briefly, following immunisation, sera were collected at the time points indicated and ELISA was used to determine immune response to the immunising antigen (KDAK-2S-avqp-6His), *P. gulae* KAS2 peptide or *P. gulae* whole cells.

[0346] As shown in Figure 13, all dogs showed a robust immune response to antigen following prime and boost immunisation.

[0347] The results shown in Figure 14 indicate that the response to *P. gulae* whole cells correlates strongly with the response to immunising antigen.

[0348] The results shown in Figure 15 indicate that the response to *P. gulae* KAS2 peptide correlates strongly with the response to immunising antigen.

[0349] Overall, the results indicate that the vaccines are well tolerated by the animals and generated a strong immune response in the canine study.

Example 7: canine study proof of concept study

[0350] A further canine study, similar to the one discussed at Examples 5 and 6, was conducted. Briefly, 20 dogs were organised into “treatment” or “control” groups (10 dogs in each group). Dogs in the treatment group were immunised 3 times with chimeric protein antigen KDAK-2S-avqp as herein described, beginning on day 0 (prime), followed by a booster immunisation on day 21 (boost 1) and again on day 42 (boost 2). The control group received no treatment.

[0351] Sera were collected at the following time points:

- Day 0 = pre-immune response (immediately prior to receiving the prime dose)
- Day 21 = prime response (21 days after receiving the prime dose and immediately prior to receiving the boost-1 dose)
- Day 42 = boost-1 response (21 days after receiving the boost-1 dose and immediately prior to receiving the boost-2 dose)
- Day 56 = boost-2 response (i.e., 14 days after receiving the boost-2 dose).

[0352] Serum IgG responses against the *P. gulae* KAS2 peptide (SEQ ID NO: 1) were measured using ELISA, similarly to the methods described in Example 5.

[0353] The results, summarised in Figures 16 and 17, show a robust immune response was generated following the immunisation protocol.

[0354] Briefly, when either mean midpoint or mean endpoint titres were used as a measure of immune responses to the KAS2 peptide, serum collected from the “control” group dogs at day 21, day 42 or day 56 showed less than a 1.3-fold increase relative to the mean pre-immune (day 0) titre. This finding is consistent with this group being comprised of control animals that received no treatment.

[0355] In contrast, all dogs in the treatment group produced titratable boosted (day 42 and day 56) peptide-specific IgG responses that were greater than their pre-immune (day 0) responses across most of the serum dilutions. A side-by-side comparison of the midpoint and endpoint titres for all treatment group dogs showed a progressive increase in titre following administration of successive doses, which appears to reach a maximum in the day 42 serum samples.

[0356] When the peptide-specific serum IgG responses of all ten dogs in the treatment group were analysed collectively, serum harvested at days 21, 42 and 56 showed statistically significant ($p < 0.05$) higher mean midpoint and mean endpoint titre values compared to the mean values for the pre-immune (day 0) serum samples. In addition, titres measured at days 42 and 56 were statistically significantly higher than day 21 titres ($p < 0.05$). When using mean midpoint titres as a measure of response, day 21, day 42 and day 56 serum samples were, respectively, 5×, 94× and 96× greater than the pre-immune titres.

[0357] Example 8: reduction of bacterial load

[0358] In parallel with the experiments described in Example 7, samples of saliva and subgingival plaque are collected from the canine subjects, both at the beginning and conclusion of the study. Samples are assessed for oral *P. gulae* load, using the qRT-PCR method described in Maruyama et al., (2018) Polish Journal of Veterinary Sciences, 21: 127-132.

[0359] The results will indicate a reduction in oral *P. gulae* load in saliva of subjects inoculated with the protein antigen KDAK-2S-avqp, compared to controls.

[0360] The results will indicate a reduction in oral *P. gulae* load in subgingival plaque of subjects inoculated with the protein antigen KDAK-2S-avqp, compared to controls.

[0361] In addition, plasma samples are obtained from subjects and assessed for the levels of inflammatory cytokines present therein. The results will indicate a reduction in inflammatory cytokines present in the animals inoculated with the protein antigen KDAK-2S-avqp, compared to controls. The cytokines measured are one or more of: GM-CSF, IFN- γ , IL-2, IL-6, IL-7, IL-8, IL-10, IL-15, IL-18, IP-10, KC-like MCP-1 and TNF- α . To measure all of these cytokines, the Milliplex Canine cytokine/chemokine magnetic bead panel is used.

[0362] Example 9: assessment of alternative protein antigen

[0363] A similar set of experiments to those describes in Example 7 and 8 are performed using one of the following alternative protein antigens:

- K_DF_AK-2S-AVQP (eg SEQ ID NO: 28)
- R_DA_AK-2S-AVQP (eg SEQ ID NO: 114)
- R(S)_DA_AK-2S-AVQP (cysteine in R domain substituted to serine) (eg SEQ ID NO: 115)
- K_DF(S)_AK-2S-AVQP (all 4 cysteines in D domain substituted to serine) (eg SEQ ID NO: 76)
- K_DA_AK-1S-AVQP (eg SEQ ID NO: 40)
- K_DA_AK-2S (eg SEQ ID NO: 116)
- K_DA_A-2S-AVQP (eg SEQ ID NO: 117)
- K_DA_AK (eg SEQ ID NO: 31)
- K_DA_A (eg SEQ ID NO: 32)

[0364] Chimeras that do not comprise amino acid substitutions in the D or A domains are anticipated to be less soluble than chimeras having such substitutions. Nonetheless, the results will indicate that all tested chimeras are capable of inducing a robust immune response, similarly to the results shown in Example 7.

[0365] Chimeras that comprise only a single “K” domain are also anticipated to induce a slightly less significant immune response compared to chimeras having two copies of the K or R domain. Nonetheless, the results will indicate that all tested chimeras are capable of inducing a robust immune response, similarly to the results shown in Example 7.

[0366] It will be understood that the invention disclosed and defined in this specification extends to all alternative combinations of two or more of the individual features mentioned or evident from the text or drawings. All of these different combinations constitute various alternative aspects of the invention.

CLAIMS

1. A chimeric or fusion protein for inducing an immune response to *P. gulae*, the protein comprising a first polypeptide and a second polypeptide, wherein:

A) the first polypeptide comprises or consists of an amino acid sequence of the active site of an Arg- or Lys-gingipain homologue of *P. gulae*, or a sequence that is at least 80% identical thereto; and

B) the second polypeptide comprises or consists of:

the amino acid sequence of a DUF2436 domain of the Arg- or Lys-gingipain surface complexes of *P. gulae*; and

the amino acid sequence of an adhesin domain of surface complexes of the Arg- and Lys-gingipain homologues of *P. gulae*, preferably wherein the adhesin domain comprises the amino acid sequence of at least SEQ ID NO: 86 and/or SEQ ID NO: 85, or a sequence at least 80% identical thereto.

2. The chimeric or fusion protein of claim 1, wherein the amino acid sequence of the DUF2436 domain of the gingipain homologue surface complexes comprises or consists of an amino acid sequence as set forth in SEQ ID NO: 3 or 4, or a sequence at least 80% identical thereto.

3. The chimeric or fusion protein of claim 1, wherein the amino acid sequence of the DUF2436 domain of the gingipain homologue surface complexes comprises or consists of an amino acid sequence as set forth in SEQ ID NO: 4, or a sequence at least 80% identical thereto.

4. The chimeric or fusion protein of any one of claims 1 to 3, wherein the amino acid sequence of the adhesin domain of the gingipain homologue surface complexes comprises or consists of the amino acid sequence of SEQ ID NO: 88 or 20, or a sequence at least 80% identical thereto.

5. The chimeric or fusion protein of any one of claims 1 to 4, wherein the chimeric or fusion protein comprises one or more further polypeptides that comprise or consist of an amino acid sequence of the active site of an Arg- or Lys-gingipain homologue of *P. gulae*, or a sequence at least 80% identical thereto.

6. The chimeric or fusion protein of claim 5, wherein the one or more further polypeptides comprising or consisting of the active site of an Arg- or Lys-gingipain homologue of *P. gulae* is located N-terminally to the first polypeptide.
7. The chimeric or fusion protein of claim 5, wherein the one or more further polypeptides comprising or consisting of the active site of an Arg- or Lys-gingipain homologue of *P. gulae* is located C-terminally to the first polypeptide,
8. The chimeric or fusion protein of claim 5, wherein the one or more further polypeptides comprising or consisting of the active site of an Arg- or Lys-gingipain homologue of *P. gulae* is located N-terminally to the second polypeptide.
9. The chimeric or fusion protein of claim 5, wherein the one or more further polypeptides comprising or consisting of the active site of an Arg- or Lys-gingipain homologue of *P. gulae* is located C-terminally to the second polypeptide.
10. The chimeric or fusion protein of any one of claims 1 to 9, wherein the protein comprises at least two further polypeptides that comprise or consist of an amino acid sequence of the active site of an Arg- or Lys-gingipain homologue of *P. gulae*, or sequences that are at least 80% identical thereto may be present.
11. The chimeric or fusion protein of claim 10, wherein the two or more further polypeptides comprising or consisting of the active site of an Arg- or Lys-gingipain homologue of *P. gulae* are located N-terminally to the second polypeptide, C-terminally to the second polypeptide, or N and C terminally to the second polypeptide.
12. The chimeric or fusion protein of any one of claims 5 to 11, wherein the one or more further polypeptides are linked to the first or second polypeptide of the chimeric or fusion protein via a linker of no more than 50 amino acids, or directly linked to the first or second polypeptide.
13. The chimeric or fusion protein of any one of the preceding claims wherein the first polypeptide comprises or consists of an amino acid sequence selected from the group of: SEQ ID NOs: 1 or 2 sequences at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical thereto.
14. The chimeric or fusion protein of any one of claims 5 to 13, wherein the one or more further polypeptides comprise or consist of an amino acid sequence selected from

the group of: SEQ ID NOs: 1 or 2, or sequences at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical thereto.

15. The chimeric or fusion protein of any one of claims 5 to 14, wherein the first polypeptide and the one or more further polypeptide comprise or consist of an amino acid sequence that is at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to each other.

16. The chimeric or fusion protein of any one of claims 5 to 14, wherein the first polypeptide and the one or more further polypeptide comprise or consist of identical amino acid sequences.

17. The chimeric or fusion protein of any one of claims 5 to 15, wherein the first polypeptide comprises or consists of the amino acid sequence of an active site of a Kgp gingipain homologue (eg as set forth in SEQ ID NO: 1) and the further polypeptide comprises or consists of the amino acid sequence of an active site of an Rgp gingipain homologue (eg as set forth in SEQ ID NO: 2).

18. The chimeric or fusion protein of any one of claims 5 to 15, wherein the first polypeptide comprises or consists of the amino acid sequence of an active site of an Rgp gingipain homologue (eg as set forth in SEQ ID NO: 2) and the further polypeptide comprises or consists of the amino acid sequence of an active site of a Kgp gingipain homologue (eg as set forth in SEQ ID NO: 1).

19. The chimeric or fusion protein of any one of claims 5 to 15, wherein the first polypeptide and the further polypeptides comprise or consist of the amino acid sequence of an active site of a Kgp gingipain homologue (eg as set forth in SEQ ID NO: 1).

20. The chimeric or fusion protein of any one of claims 5 to 15, wherein the first polypeptide and the further polypeptide comprise or consist of the amino acid sequence of an active site of an Rgp gingipain homologue (eg as set forth in SEQ ID NO: 2).

21. The chimeric or fusion protein of any one of the preceding claims wherein the amino acid sequence of the adhesin domain of the surface complex of *P. gulae* further comprises one or more amino acid substitutions selected from:

- a) one or more cysteine amino acid substitutions, compared to the naturally occurring *P. gulae* sequences in corresponding regions;
- b) substitution of the proline and/or an asparagine residues in the sequence PxxN corresponding to, or at a position equivalent to, residues 6 to 9 of SEQ ID NO: 85 (equivalent to residues 68 to 71 of the sequence of SEQ ID NO: 20 or 88);
- c) substitution of the motif NxFA to SxYQ in the sequence, corresponding to, or at a position equivalent to residues 2 to 5 of SEQ ID NO: 85 (equivalent to residues 64 to 67 of the sequence of SEQ ID NO: 20 or residues 64-67 of SEQ ID NO: 88);
- d) substitution of the tyrosine residue, corresponding to or at a position equivalent to residues at position **10** of SEQ ID NO: 86, and of the tryptophan residue, corresponding to or at a position equivalent to residue at position **23** of SEQ ID NO: 85, to alanine residues (equivalent to the tyrosine at residue position 10 and the tryptophan at residue position 85 of SEQ ID NO: 20).

22. The chimeric or fusion protein of claim 21, wherein the amino acid sequence of the adhesin domain further comprises:

- a) one or more cysteine amino acid substitutions, compared to the naturally occurring *P. gulae* sequences in corresponding regions.

23. The chimeric or fusion protein of claim 21, wherein the amino acid sequence of the adhesin domain further comprises:

- b) substitution of the proline and/or an asparagine residues in the sequence PxxN corresponding to, or at a position equivalent to, residues 68 to 71 of the sequence of SEQ ID NO: 20.

24. The chimeric or fusion protein of claim 21, wherein the amino acid sequence of the adhesin domain further comprises:

- a) one or more cysteine amino acid substitutions, compared to the naturally occurring *P. gulae* Arg- or Lys-gingipain sequences in corresponding regions; and
- b) substitution of the proline and/or an asparagine residues in the sequence PxxN corresponding to, or at a position equivalent to, residues 68 to 71 of the sequence of SEQ ID NO: 20;

25. The chimeric or fusion protein of any one of claims 21, 22 and 24, wherein the substitution of the cysteine amino acid residues in the adhesin domain is a substitution to a serine residue or a valine residue.
26. The chimeric or fusion protein of claim 25, wherein the one or more cysteine substitutions comprise one or more substitutions to a serine residue.
27. The chimeric or fusion protein of any one of claims 21, 22 and 24 to 26, wherein only one cysteine residue in the adhesin domain is substituted.
28. The chimeric or fusion protein of any one of claims 21, 22, and 24 to 26, wherein two cysteine residues are substituted.
29. The chimeric or fusion protein of any one of claims 21, 22 and 24 to 27, wherein the adhesin domain comprises or consists of the sequence set forth in any one of SEQ ID NOs: 20 or 23 or a sequence at least 80% identical thereto, wherein one or more cysteine residues are substituted to a serine or valine residue.
30. The chimeric or fusion protein of any one of claims 21, 22 and 24 to 27, wherein the adhesin domain comprises or consists of the sequence set forth in any one of SEQ ID NO: 23 or a sequence at least 80% identical thereto, wherein two cysteine residues are substituted to a serine or valine residue.
31. The chimeric or fusion protein of any one of claims 24 to 30, wherein the adhesin domain comprises a substitution of the proline residue and/or a substitution of the asparagine, in the motif PxxN corresponding, or at a position equivalent to positions 68 to 71 of the sequence of SEQ ID NO: 20.
32. The chimeric or fusion protein of claim 31, wherein the proline amino acid substitution is a substitution to an alanine residue.
33. The chimeric or fusion protein of claim 31, wherein the asparagine amino acid substitution is a substitution to a proline residue or an alanine residue, preferably a proline residue.
34. The chimeric or fusion protein of claim 31, wherein the motif PxxN in the adhesin domain, is substituted to AxxP (eg AVQP, SEQ ID NO: 64).

35. The chimeric or fusion protein of claim 23, wherein the adhesin domain comprises or consists of an amino acid sequence as set forth in SEQ ID NO: 21, or a sequence at least 80% identical thereto, wherein the sequence comprises AxxP at a position equivalent to positions 68 to 71 of the sequence of SEQ ID NO: 21.
36. The chimeric or fusion protein of any one of claims 34 to 35, wherein the adhesin domain comprises or consists of an amino acid sequence as set forth in any one of SEQ ID NOs: 24 to 27, or a sequence at least 80% identical thereto, wherein the sequence comprises AxxP at a position equivalent to positions 68 to 71 of the sequence of SEQ ID NO: 21, and wherein one or more cysteine residues are substituted to a serine or valine residue.
37. The chimeric or fusion protein of any one of claims 4 to 36, wherein the second polypeptide comprises an amino acid sequence of a DUF2436 domain as set forth in SEQ ID NO: 3, further comprising one or more amino acid substitutions.
38. The chimeric or fusion protein of claim 37, wherein the one or more amino acid substitutions are one or more substitutions of the cysteine residues in the domain.
39. The chimeric or fusion protein of claim 38, wherein only one cysteine residue is substituted.
40. The chimeric or fusion protein of claim 38, wherein any two, three or all four cysteine residues in the DUF domain are substituted.
41. The chimeric or fusion protein of any one of claims 38 to 40, wherein the cysteine residues in the DUF domain are substituted to a valine residue, serine residue or an alanine residue.
42. The chimeric or fusion protein of claim 41, wherein the one or more cysteine residues are substituted to one or more serine residues.
43. The chimeric or fusion protein of any one of claims wherein the DUF domain comprising one or more cysteine residue substitutions comprises an amino acid sequence as set forth in any one of SEQ ID NOs: 5 to 19.
44. The chimeric or fusion protein of any one of claims 1 to 43, wherein the chimeric or fusion protein comprises or consists of the amino acid sequence as set forth in any of

SEQ ID NOs: 32 to 35 97 or 98, or a functional variant thereof having at least 80% identity thereto.

45. The chimeric or fusion protein of any one of claims 1 to 43, wherein the chimeric or fusion protein comprises or consists of the amino acid sequence as set forth in any of SEQ ID NOs: 77 to 82, or a functional variant thereof having at least 80% identity thereto.

46. The chimeric or fusion protein of any one of claims 1 to 43, wherein the chimeric or fusion protein comprises or consists of an amino acid as set forth in any of SEQ ID NOs: 36 to 45 or 99, or a functional variant thereof having at least 80% identity thereto.

47. The chimeric or fusion protein of any one of claims 1 to 43, wherein the chimeric or fusion protein comprises or consists of an amino acid as set forth in any of SEQ ID NOs: 46 to 76 or a functional variant thereof having at least 80% identity thereto.

48. The chimeric or fusion protein of any one of claims 1 to 43, wherein the chimeric or fusion protein comprises or consists of the amino acid sequence of SEQ ID NO: 28 29, or 100, or functional variants thereof having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% identity thereto and preferably wherein the functional variant comprises cysteine substitutions in the adhesin domain and the AVQP (SEQ ID NO: 64) substitutions to the motif PVQN (SEQ ID NO: 108).

49. The chimeric or fusion protein of any one of claims 1 to 43, wherein the chimeric or fusion protein comprises or consists of the amino acid sequence of SEQ ID NO: 29 or a functional variant thereof having at least 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% identity thereto.

50. A nucleic acid encoding any chimeric or fusion protein as defined herein in any one of the preceding claims, optionally, wherein the nucleic acid has a nucleotide sequence that encodes any one or more of the amino acid sequences defined in Table 1 herein.

51. The nucleic acid of claim 50, wherein the nucleic acid has a nucleotide sequence as set forth in Table 2 herein.

52. A vector comprising a nucleic acid of claim 50 or 51.

53. A cell comprising a vector of claim 52, or a nucleic acid of claim 50 or 51.

54. A composition comprising a chimeric or fusion protein of any one of claims 1 to 49, optionally in combination with a pharmaceutically acceptable carrier.
55. The composition of claim 54, further comprising an adjuvant for potentiating an immune response to the chimeric or fusion protein.
56. A vaccine or immune stimulating composition for inducing an immune response to *P. gulae* in a subject, the composition comprising:
- an immunogen in the form of a chimeric or fusion protein of any one of claims 1 to 49, and
 - an adjuvant, for potentiating the immune response to the immunogen in the subject.
57. The vaccine of claim 56, wherein the chimeric or fusion protein is the sole immunogen in the composition.
58. A method for inducing an immune response in a subject to *P. gulae*, the method comprising administering to a subject in need thereof, a chimeric or fusion protein of any one of claims 1 to 49, composition of claim 54 or 55 or vaccine or immune stimulating composition of claim 56.
59. A method of inducing a humoural immune response to *P. gulae* in a subject, the method comprising administering to the subject, a chimeric or fusion protein, of any one of claims 1 to 49, composition of claim 54 or 55 or vaccine or immune stimulating composition of claim 56.
60. The method of claim 58 or 59, wherein the method is for strengthening an immune response (such as a protective immune response) of a subject to *P. gulae*.
61. A methods of immunising a subject against *P. gulae* infection, the method comprising administering to the subject, a chimeric or fusion protein, of any one of claims 1 to 49, composition of claim 54 or 55 or vaccine or immune stimulating composition of claim 56.
62. The method of any one of claims 58 to 61, wherein the subject who has received or has been administered the chimeric or fusion protein, or composition or vaccine, has an increased level of protection against infection with *P. gulae*, or severity of one or more

symptoms of *P. gulae* infection, compared to a subject who has not received the protein, composition or vaccine.

63. A method of treating a *P. gulae* infection in a subject, the method comprising administering to a subject in need thereof, a chimeric or fusion protein, of any one of claims 1 to 49, composition of claim 54 or 55 or vaccine or immune stimulating composition of claim 56, thereby treating the *P. gulae* infection in the subject.

64. A method for reducing or minimising the severity of a symptom associated with an infection with *P. gulae*, comprising administering to an individual in need thereof, a chimeric or fusion protein, of any one of claims 1 to 49, composition of claim 54 or 55 or vaccine or immune stimulating composition of claim 56, optionally wherein the symptoms are selected from the group consisting of swollen or puffy gums, gums that bleed easily, receding gums, periodontal pockets around the teeth, loss of tooth supporting tissues (periodontal ligament, cementum and/or alveolar bone) pus between gums and teeth, and gingivitis.

65. A method for treating occurring *P. gulae* -related disease in a subject, the method comprising administering to an individual in need thereof, a chimeric or fusion protein, of any one of claims 1 to 49, composition of claim 54 or 55 or vaccine or immune stimulating composition of claim 56.

66. The method of claim 65, wherein the *P. gulae*-related disease comprises periodontal disease.

67. The method of any one of claims 58 to 66, further comprising administering an antimicrobial compound, and/or an anti-inflammatory agent or further comprising administering one or more further immunogens or stimulating compositions.

68. Use of a chimeric or fusion protein of any one of claims 1 to 49, in the manufacture of a medicament for:

- inducing an immune response (preferably a protective immune response) in a subject to *P. gulae*;
- immunising a subject against *P. gulae* infection;
- treating a *P. gulae* infection in a subject;

- minimising or reducing the severity of one or more symptoms of *P. gulae* infection;
or

treating *P. gulae* -related disease in a subject.