



(51) International Patent Classification:

C07K 14/31 (2006.01) *C12N 15/11* (2006.01)
A61K 39/085 (2006.01) *G01N 33/569* (2006.01)
C07K 16/12 (2006.01) *G01N 33/68* (2006.01)

BioScience, 76 Molecular Sciences Building, University of Queensland, Brisbane (AU).

(21) International Application Number:

PCT/GB2010/001916

(74) Agent: **CHAPMAN, Paul**; Marks & Clerk LLP, Aurora, 120 Bothwell Street, Glasgow G2 7JS (GB).

(22) International Filing Date:

15 October 2010 (15.10.2010)

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

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/252,026 15 October 2009 (15.10.2009) US

(71) Applicant (for all designated States except US): **UNIVERSITY COURT OF THE UNIVERSITY OF EDINBURGH** [GB/GB]; Old College, South Bridge, Edinburgh EH8 9YL (GB).

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

(72) Inventor; and

(71) Applicant : **BANNOEHR, Jeanette** [DE/GB]; Laboratory for Bacterial Evolution and Pathogenesis, The Chancellor's Building, University of Edinburgh, Edinburgh EH16 4SB (GB).

(72) Inventors; and

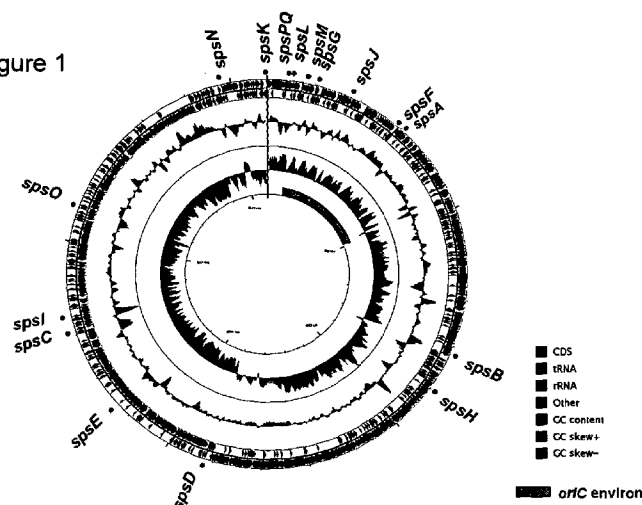
(75) Inventors/Applicants (for US only): **FITZGERALD, J., Ross** [IE/GB]; Laboratory for Bacterial Evolution and Pathogenesis, The Chancellor's Building, University of Edinburgh, Edinburgh EH16 4SB (GB). **BEN ZAKOUR, Nouri, L.** [FR/AU]; School of Chemistry and Molecular

Published:

— without international search report and to be republished upon receipt of that report (Rule 48.2(g))

(54) Title: STAPHYLOCOCCAL ANTIGENS

Figure 1



(57) Abstract: The present invention provides novel sequences encoding *Staphylococcus pseudintermedius* proteins/nucleic acids potentially useful in the treatment and/or prevention of canine disorders. In particular, the various protein and/or nucleic acid sequences described herein may find application as vaccines for use in treating and/or preventing a variety of canine diseases and/or conditions caused or contributed to by *Staphylococcus pseudintermedius*.

STAPHYLOCOCCAL ANTIGENS

FIELD OF THE INVENTION

The present invention provides novel staphylococcal cell wall associated proteins, genes encoding the same and vaccines for use in treating/preventing Staphylococcal infections.

BACKGROUND OF THE INVENTION

Skin diseases are a major cause of morbidity in dogs and an important animal welfare issue (Hill et al, 2006). In particular, superficial bacterial folliculitis (pyoderma) caused by *Staphylococcus pseudintermedius* (formerly known as *Staphylococcus intermedius*) is one of the most common diseases seen in small animal veterinary practice, worldwide (Hill et al, 2006). Superficial pyoderma is characterized by the formation of follicular pustules and is often associated with pruritus, alopecia, erythema and swelling. This may develop into deep pyoderma which typically includes pain, crusting, odor, and exudation of blood and pus. The disease often occurs as a secondary infection in dogs with atopic dermatitis (AD) resulting from a type I hypersensitivity reaction (IgE antibody-associated) to environmental allergens (Hill et al, 2006). Treatment of canine pyoderma is often difficult without resorting to aggressive, medium-term administration of systemic antibacterial agents to prevent relapse of infection, and such therapy predisposes to the development of bacterial resistance that may be transferred to bacteria infecting humans (Guardabassi et al, 2004). Worryingly, methicillin-resistant *S. pseudintermedius* has recently emerged as a major problem in veterinary clinics worldwide (Bannoehr et al, 2007). Although rare, several episodes of life-threatening infections of humans by *S. pseudintermedius* have been reported with the typical route of transmission being through dog bite wounds (Bannoehr et al, 2007). Previously, crude vaccine preparations based on *Staphylococcus aureus* phage lysate or *S. pseudintermedius* autogenous bacterin have shown promise as adjunctive therapies for treatment of pyoderma (Curtis et al, 2006), and a rationally-designed effective vaccine would be a highly desirable means to reducing or eliminating the suffering associated with the disease.

Accordingly, the present invention aims to obviate one or more of the problems associated with the prior art.

SUMMARY OF THE INVENTION

The present invention is based upon the identification of novel gene sequences encoding proteins potentially useful in the treatment and/or prevention of canine disorders. In particular, the proteins encoded by the genes described herein, may find application in the treatment and/or prevention of diseases caused or contributed to by the bacterial pathogen *Staphylococcus pseudintermedius*.

The inventors have identified a number of *Staphylococcus pseudintermedius* genes encoding proteins which may broadly be classed as members of the cell-wall anchored (CWA) family of proteins. In certain embodiments, these CWA proteins may be further grouped as surface proteins known as Microbial Surface Components Recognising Adhesive Matrix Molecules (MSCRAMM). It should be understood that while a number microbial organisms may be known to express MSCRAMM type proteins, the term "MSCRAMM" describes the phenotypic function of a wide range of diverse surface-associated proteins of Gram-positive bacteria. As such, while MSCRAMM proteins may all possess cell-wall anchor motifs and signal sequences for cell wall transportation, the proteins belonging to this group may otherwise be structurally diverse. Furthermore, bacterial species within a particular genus, for example the genus *Staphylococcus*, may possess unique MSCRAMM profiles.

In view of the above, the present invention relates to a group of surface expressed proteins derived from *Staphylococcus pseudintermedius* that may be referred to either as CWA or MSCRAMM proteins.

As such, a first aspect of this invention provides an isolated and/or substantially purified *Staphylococcus pseudintermedius* CWA or MSCRAMM nucleic acid or protein sequence comprising a nucleic acid or amino acid sequence homologous or identical to any one of the nucleic acid or amino acid sequences provided as SEQ ID NOS: 1-36 below.

SEQ ID NO: 1

```
atggaaaacaaaaacttttttagtattcgtaaactatctattggtgtaggttcttgctta
atcgcgagttctttacttgtaaacacgccaaagttttgctgaagaaacagataatgcgaac
attaatgacgcacacaaaaacgccttttatgaaattttacatttgccaaacttaactgaa
gagcaacaaaatggattcatccaaagtcttaagatgatccaagtgtgagcaacgacatt
ttagtagaagctaagaaattaaatgacactcaagctaaacctgattacagtgaagcaca
caaatgcatttttatgaaattttacatttgtcaaaacttaactgaagagcaacaaaatgga
ttcatccaaagtcttaagatgatccaagtgtgagcaacgacattttagtagaagctaag
aagttaaatgacactcaagctaaacctgattacagtgaagcacaacaaaatgcattttat
```

gaaattttacatttgtcaaacttaactgaagagcaacaaaatgggttcatccaaagcctt
 aaagatgatccaagtgttaagtaaaagaaattttagcagaagctaagaagttaaatgatagt
 caagcacctaaagttgataaagctaaaaaaactgacaaagctgaagcgaaagcagatgat
 aaagctaaaggtgaagaagccaaaaaatctgaagacaaaaaagatagcaaagcagataag
 gcaaaatcgaaaaacgctacacatgttggttaaacctgggtgaaacttttagataatattgct
 aaagatcatcatacaactgttgataaaaattgctaaagataacaaaataaaagataaaaaat
 gtgattaaactaggtcaaaaacttggttggtgataaaacaaaagcaactcaaggaaaacaa
 gaagctgtagcgaaaaatgaagtgaaggctttacctaatactgggtgaaaatgatgatatc
 gcattattcagcacacagttgcgggtggcgtaagtatcgcttttaggttcattattatta
 ggaagaaacagaaaaactagctaa

The protein sequence translated from SEQ ID NO 1 is designated SEQ ID NO: 2 and is shown below:

SEQ ID NO: 2

MENKNFFSIRKLSIGVGSCLIASLLVNTPSFAETDNANINDAQQNAFYEILHLPNLT
 EEQQNGFIQSLKDDPSVSNLILVEAKKLNDTQAKPDYSEAQQNAFYEILHLSNLTEE
 QQNGFIQSLKDDPSVSNLILVEAKKLNDTQAKPDYSEAQQNAFYEILHLSNLTEEQQ
 NGFIQSLKDDPSVSKEILAEAKKLND SQAPKVDKAKKTDKAEAKADDKAKGEEAKKS
 EDKKDSKADKAKSKNATHVVKPGETLDNIAKDHHTTVDKIAKDNKIKDKNVIKLGQKL
 VVDKQKATQKGQEAVAKNEVKALPNTGENDDIALFSTTVAGGVSIALGSLLLGRNRK
 TS

SEQ ID NO: 3

atggaaaaacaaaaactttttcagcattcgtaaattatcaattgggggtgggttcatgttta
 atcgcgagctctttacttgtgaatacaccaagtttcgcagaagaaggagataataacgca
 gaagcgcaacaaaaacgctttctctgaggtagtaaaattacctaaccttagcgaagaacaa
 cgtaatggtttcattcaaagccttaaagatgatccaagtacaagtcaagatgtgcttaat
 gaagctaaaaaattaaatgatagtcaagagggatctcaacctgctcctgattacagtgat
 gaacaacaaaatgcattttatgaaattttacaccttccaaacttaactgaagaacaacgc
 aatggctatatattcaaagctttaaagatgaccaagtgtgaagcgctaattcttctgttgaa
 gctaaaaatatgaatgttaaccaaacacctacacaacctgcgccaagtttcgatgaagcg
 caacaaaatgcattctatgagattgtaaacttaccaaactcttactgaagagcaacgtaac
 ggtttcatccaaagccttaaagacgatccaagtgtgaagtaagatatccttgttgaaagct
 aaaaagttaaatgacagccaagcaaaacctgattacagtgaagcgcaacaaaatgcattt
 tatgaaattttacaccttccaaacttaactgaagaacaacgtaacggtttcatccaaagc
 cttaaagacgatccgagtgtaagtagtgatattcttgctgaagctaagaaattaaatgac
 agccaagcgctaaagaagacaacaacgtaaaagacaataattcaggtgaaaacaaagct
 gaagacaaaggcaacaaagaaaacaaagctgaagataaaggcagcaagaagacaagct

gaagataaaggcagcaaagaagacaaagctgaagataaaggcagcaaagaagacaaagct
 gaagataaaggcagcaaagaagacaaagctgaagataaaggcagcatagaagataaagct
 aaagacaaagacacaaagaaggcaaagctgcagacaaaggtatggacaaagcgaaagat
 gcaatgcatgtcggttcaacctggtgaaacagtagaaaaaattgctaaagctaataacaca
 actgtagaacaatcgctaaagataatcatttagaagataaaaaacatgattttaccaggt
 caaaaacttggtgttgacaacaaaaagcaatgaaagacagccaagaagctaagcaaac
 cacgaaatgaaagctttacctgaaacaggtgaagaaaacgatatggcattattcgggtaca
 tcacttacaggtggtccttagcttagcattaggtttatacatcttaggacgtggcagaaaa
 acaaaactaa

The protein sequence translated from SEQ ID NO 3 is designated SEQ ID NO: 4 and is shown below:

SEQ ID NO: 4

MENKNFFSIRKLSIGVGSCLIASLLVNTPSFAEEGDNNAEAQQNAFSEVVKLPNLSEEQRNGFIQSLKDDPS
 TSQDVLNEAKKLNDSEQSQPAPDYSDEQQNAFYIEILHLPNLTEEQRNGYIQSLKDDPSVSANILVEAKNMNV
 NQTPTQPAPSFDEAQQNAFYIEIVNLPNLTEEQRNGFIQSLKDDPSVSKDILVEAKKLNDSEQAKPDYSEQQNA
 FYIEILHLPNLTEEQRNGFIQSLKDDPSVSSDILAEAKKLNDSEQPKEDNNVKDNNNGENKAEDKGNKENKAED
 KGSKEDKAEDKGSKEDKAEDKGSKEDKAEDKGSKEDKAEDKGSIEDKAKDKDNKEGKAADKGMKAKDAMHV
 VQPGETVEKIAKANNTTVEQIAKDNHLEDKNMILPGQKLVDNQKAMKDSQEAKANHEMKALPETGEENDMALF
 GTSLTGGLSLALGLYILGRGRKTN

SEQ ID NO: 5

gtgtacaaaaatgaagaagaaaagcattcaataagaaagttatctataggagccgcatct
 gtcattgttgggggactcatgtatggtgttttgggaaatgatgaagctcaagcgaatgaa
 gatgtcactgaaacaactgggagaaaattcagtgacaacgcaagcttctgagcaacatttg
 caagtggagcagtacctcaagaaggcaataatgtaaatgtatcctctgtaaaagtacct
 acgaatacggcaacgcaagcacaagaagatgttgcaagtgtatccgatgttaaagcacat
 gctgatgatgcattacaagtacaagaagtagtcatactgatgggtgttcttcagaattc
 aagcaggagacagcttatgcgaatcctcaaacagctgagacagttaaacctaatagtga
 gcagtgcattcagctctgaatacaggagataagcaaaaaccgcatcatctagccgcaaagaa
 gatgagactatgcttcagcagcaacaagttgaagccaaaatgttgtagtgagtgaggaggaa
 gtgtctaaagaagaaaatactcaagtgtatgcaatcccctcaagacgttgacaacatgta
 ggtggtaaagatatctctaattgaggttgtagtgataggagtgatataaaggatttaac
 agcgaaactactattcgacctcatcagggacaaggtggttaggtgaattatcaattaaag
 tttcctagcaatgtaaagccaggcgatcagtttactataaaattatctgacaatatcaat
 acacatgggtgttctgttgaaagaaccgcaccgagaatcatggctaaaaatactgaaggt
 gcgacggatgtaattgctgaaggtctagtgttggaagatggtaaaaccatcgatatataca
 tttaaagactatgtaaatggcaagcaaaatttgactgctgagttatcagtgagctatttc
 gtaagtcgggaaaaagtcttgactactgggacacaaacattcacgacgatgatcggtaat
 cattcaacgcaatccaatattgacgtttattatgataatagtcattatgtagatggacgt
 atttcgcaagtgaacaaaaagaagctaaatttcaacaaatagcatacattaacccta
 atggctatttaaattggcagggggacaattgcagtttaattggtgaagtggtcagtggtacgact

aaagacttaatgcaacctacagtgcgtgtatatcaatataaaggacaaggtgttcctcct
 gaaagtattactatagaccctaatatgtgggaagaaatcagcataaacgatactatggta
 agaaaatatgatgggtggctatagcttgaatctggataccagcaagaatcaaaaatatgcc
 atctattatgaaggggcatatgatgcgcaagctgacacactgttgtatagaacatatata
 cagtcattaaacagttactatccgttcagttaccaaaaaatgaacggtgtgaagttttac
 gaaaacagtgcgagtggaagcggtaggttgaaaccgaaaccacctgaacaacccaaaacca
 gaacctgaaattcaagctgatgtagtagatattattgaagatagccatgtgattgatata
 ggatggaatacagcagttggagaagaaagtgaggcaaaccaaggccctcaagaagaaatc
 acggaaaatcacgacatcgaagtcattgaggaaaacaacttggtggaaatgacagaagat
 acagcagttggagaagaaagtgaggcaaaccaaggccctcaagaagaaatcacggaaaat
 cagcagatcgaagtcattgaagaaaacaacttagtggaatgacagaagatacagcgttg
 gaagaagaaagtgaggcaaatcaaggtcctcaagaagagatcacagaaaaccacgatatc
 gaagtcattgaagaaaacaacttggtggaaatgacagaagatacagcgttggaagaagaa
 agtgaggcaaatcaaggtcctcaagaagagatcacagaaaaccacgacatcgaagtcatt
 gaagaaaataacttagtagaaatgacagaagatacagcagttggagaagaaagtgaggca
 aatcaaggtcctcaagaagagatcacagaaaaccacgatatcgaagtcattgaggaaaac
 aacttagtggaatgacagaagatacagcagttggagaagaaagtgaggcaaaccaaggt
 cctcaagaagaaatcacggaaaatcacgacatcgaagtcattgaagaaaacaacttggtg
 gaaatgacagaagatacagcgttggaagaagaaagtgaggcaaatcaaggtcctcaagaa
 gagatcacagaaaaccacaacatcgaagtcattgaagaaaacaacttggtggaaatgaca
 gaagatacagcagttggagaagaaagtgaggcaaacccaggacctcaagaagaagtaaca
 gagaatcaacctcagcaagaagaaatcatggaaaaccaagaagtcgaaaagaaaggcgat
 agtaacttggtagaaagtacaaaaactccaaaggccgaagaatcagttagcgttcagcca
 actttagaagacaaaaacacaaagaaccacgttaacacagtagtagtgaatacgaaggta
 tctgaagttaaagaaaaggatccccaccatacaaaaagcactaccagatacggggacaacc
 tctcgaagtcattccatgatgattcctctccttcttgttgctgggtcagtagtggtgta
 cgtcgaaagaaaaagcatagtaaggtgaattaa

The protein sequence translated from SEQ ID NO 5 is designated SEQ ID NO: 6 and is shown below:

SEQ ID NO: 6

VYKNEEEKHSIRKLSIGAASVIVGGLMYGVLGNDEAQANEDVTETTGRNSVTTQASEQHLQVEAVPQEGNNVN
 VSSVKVPTNTATQAQEDVASVSDVKAHADDALQVQESSHTDGVSSSEFKQETAYANPQTAETVKPNSEAVHQSE
 YEDKQKPVSSSRKEDETMLQQQQVEAKNVVSAEEVSKEENTQVMQSPQDVEQHVGGKDISNEVVVDRSDIKGF
 NSETTIRPHQGGRLNYQLKFPSNVKPGDQFTIKLSDNINTHGVSVERTAPRIMAKNTEGATDVIAEGLVLE
 DGKTIVYTFKDYVNGKQNLTAELSVSYFVSPEKVLTTGTQTFTTMIGNHSTQSNIDVYYDNSHYVDGRISQVN
 KKEAKFQQIAYINPNGYLNGRGTIAVNGEVVSGTTKDLMQPTVRVYQYKGQGVPPESITIDPNMWEEISINDT
 MVRKYDGGYSLNLDTSKNQKYAIYYEGAYDAQADTLLYRTYIQSLNSYYPFSYQKMNGVKFYENSASGSGELK
 PKPPEQPKPEPEIQADVVDIIEDSHVIDIGWNTAVGEESGANQGPQEEITENHDIEVIEENNLVEMTEDTAVG
 EESGANQGPQEEITENHDIEVIEENNLVEMTEDTALEESGANQGPQEEITENHDIEVIEENNLVEMTEDTAL
 EEESGANQGPQEEITENHDIEVIEENNLVEMTEDTAVGEESGANQGPQEEITENHDIEVIEENNLVEMTEDTA

VGEEESGANQGPQEEITENHDIEVIEENNVLVEMTEDTALEEESGANQGPQEEITENHNIEVIEENNVLVEMTEDT
AVGEEESGANPGPQEEVTENQPPQEEIMENQVEVEKKGDSNLVESTKTPKAEESVSVQPTLEDKNTKNHVNTVVV
NTKVSEVKEKDPHHTKALPDTGTTSRSHSMIPLLLVAGSVVLLRRKKKHSKVN

SEQ ID NO: 7

atgaataaatcaagaactaaacatttttaattttttatcaaaacgtcagaatcggtatgct
attcgccacttttcagctggtactgtgtcagtgctttagtaggagcagctttcttgctaggt
gtccatacgagtgatgcatctgctgcagaacaagatcaaacatctgaagcaaagcaaaac
ctctttgatgcttccgctatttttggcgctttaacagagacgaacgaaaaggtagcaciaa
gtgacgccaaacagaaaaaatctttcatcagttgaagaaatgagagataaaaggcgcaact
ggaaatggaccatcaataacatcactacaaactgtagaacaaaaataatgcagtacaacct
acagcaacacctattaatgacacagaaaattcaaccgaagcccctatgaaagaacaatcg
aatgatgcacaaacgactgacgaaagtaacaatgccactcagaaaaataatactgaaccc
caagcaaaacaatgaaatatcagcgcgtaatgcaaaaacaacagcatatttaacaagtga
acctttacaacagcaacgtctacaactgatatgcctacacagaaacaagaatatccatct
ttagaaaatccaacaaatcaatcgcaaacgaacagagcacaaccaccaacaatggaagca
cccaaactggcagaaggattagacaatctattaaaaaaatcaactttcgaaagtatgtac
gtgacaaaaagaaatcaatttgacaaagagacggcttctaaaacaaaagcatggccgagt
gatgttggtccagaaaatcaagtagagatacttgctgatgcaattcaaaatggctatatc
aaatctgtaaatgatgtgaccaataaagcacatacgttatctggacgtgcatggatgta
gaccacggaacaccaacgacaatagctaattggtttaacacctgttccagagggcactaaa
gtttattttgcggtggatagatcaagatggtgccacttcgccaatgtatacagcaaaaacg
acaagtagattaagcgctgcggttggaatcaagtgggtccaggtgcttatgctttcgat
ttacgcacaggttggtatagatgctaaaggaaaacaccacgtatatagagcagtaaaagggt
caatattataaaatatggatcaatgattttagaactaaagacggtaatatcgctacaatg
ttacgtggtgcaggaggatattgttccgggaacgtacgtggattctgtgacatacaacaat
atggggcaattttccattaattggtacaaaatagcaacgtacaggtatctttatgacaacg
ataccttcagaaaaatatattaatcaaaaacattacgtgaaagatacaaaagggtgctgca
gcaaatccagccgtcacgataattgaaaataactttgtgagcggcaagtttggtatagaa
acaggtgctggagattatgtgaactcagcgacaggtccaaaccacaatgcaagagatgtc
gttgccctctggatacaaaagtgggtcatgtcatcattaacagatcaaggtgctaaagcctac
gatgcgcaagtcaatcgcttgccgaagaaagatcgagcagaagcagcacgtcaattatta
ataaaacatccagaatatatcgagcaactgtagaagggtataacgaatgagtgggggaga
tatacattgcgtttccctaaaggcacattcaacaaagaccatctttacggttacgtattg
gattttgatggtgaaattgtaaaaacttattcaggttttacttcaccagagttccggaga
ccgaattataatttgaccgttacaccgcaaacagctccctattatagaccggttcgacgt
gcatgggtcaatgttaattttgcggttattgaagcaccacaatctcaaatcgaaataaaa
gaatttgatgcaacctctaaccctgcgcacgtgggcaaacagcaactattgatatacata
gggtatgcctaaaacttcattacttacacgtgtacaatggaaagattcatcgggcagttat
gttgaggatagtggtcctgtttttacggaagaagaggctgaacatatagcggaatttgta
ataccgtctagcgcaaaatcaggcggaagtgtatactgtacaactcggtggtaggtaatcat
atcgtagcttcggactctcttattgtacatgtcaatgaagaagcggcgacatatcatccg

atatacccatcgacaacagtagaatcaggtcaaagagtaacgattccagcacctaagaat
atggatggcaaacctttactagatggcacaacttttgaaaaagggtcatcacgtaccaact
tgggcttttagtgaatggtgatggctcgattacagtaaaacctggagaaaaagtagcagag
ggtgagtatgatattccagtgattgtgacatatccagatgggttctaaaaacacaatcttt
gcacctgtgaccgttgaagaaaaacaaccaatggcatcgcaatatgagccaataacaact
ggagtatcgaaaccattttgaaacccagtaatgccaaactgatgtaacagattcaattcaa
gtaccgaactatccattggaagggcaacaaccgacagtaacagtggtgatgaaagtcaa
ttaccagatggaacaacagaagggttacaaggatatagatgtaacagtgacataccagac
ggaacaaaggatcgtgtcaaagttccagtcgtaacggaacaacaattagatagtgataaa
tatgatccggtcgcaacaggtatcttgaaaccgtttggtactccaacaacagaggaagac
gttataaaattagtgagataccgaaatatccaacagacttaacacaacccaaaagtaaca
gtgacggttccaaatactttaccggatgggcaaacgccaggtaaagtagacgttgatgtg
acagtaacgtatccagatgggtccacagatcacatttcagttccagtttgacaaaacaag
catctggataaagacaaatataacccaataacgactgggggtatcgaaaccatttggaatc
ccagtaacgccaaactgatgtaacagattcaattcaagtaccgaactatccattggaaggg
caacaactgacagtaacagtggtgatgaaacacaattaccagatggaacaacagaagggt
cacaaggatatagatgtaacagtgacataccagacggaacaaaggatcatatcaaagtt
ccagtcgtaacggaacaaacaatcagataatgaaaaatatgagccaacaactaacggaatc
acgaaaaagtagcgtatccctacgacagaggatgaagtgatagatatagttcgaaattcca
tattttccagtagatggcgtgcaacctattgtaacggtaaattgatcctagactattgcca
aatggtcaaaaagaagggtcaaataatggtccagtcacagtgacgtatccggatggcaca
aaagatctcatgacaggttccggttattacaggttaagcaagcagaaaaatgaaaaatacgat
ccaatcacattaggagtaactaaagattatggtgatcctacaactgcaaacgatgtgaca
aagtcaatccaataaccaatatccagcaggtggcgaacaaccaatcgcaacagcggat
gatgaaagtcaattaccggatggcacagtagaaggtaaagtggatattccagtcacagtg
acgtatccggatgggtactcaggatcatatcactgtcccagtatcccaatcaacaacga
gataatcaaaaagccagtaaaagctgtgacgaaaatacatgggtatatcggtaacaggcact
gatatgacagatactaagaaaaatcataactatccagcaggtgggtgaacaacctaaagtt
actgtgaaagatgacgatcaattatcagagggttaaagtcgattcaacagtggggtgcggat
aatgtgacaactacagatgatttatcaagcgttaactgcggtatctcatggtcatcaaca
agtgtacaaacaacaaaagagaaccaatcagtgcatgatgaagaggtgacgatcccaaca
gttgacatgtgtctacaataatgacaggtgtggttaaagggtgagcaagaagcgacggat
atcgtggctagaccacatgttgaaacaactcaactcccatcaatttcagctcaagcaaca
gttaaaaaactaccagaaacgggtgaaaacaatgaacaatcaggtgttttattaggtgga
tttattgcgttcatgggtagcttacttttattcggcagacgtcgcaaaccaagaaagat
taa

The protein sequence translated from SEQ ID NO 7 is designated SEQ ID NO: 8 and is shown below:

SEQ ID NO:8

MNKSRTKHFN FLSKRQNRYA IRHFSAGTVS VLVGAAFLLG VHTSDASAAE
 QDQTSEAKQN LFDASAIFGA LTETNEKVAQ VTPTEKNLSS VEEMRDKGAT
 GNGPSITSLQ TVEQNNAVQP TATPINDTEN STEAPMKEQS NDAQTTDES
 NATQKNNTEP QANNEISARN AKTTAYLTSE TFTTATSTTD MPTQKQEYPS
 LENFTNQSQT NRAQPPTMEA PKLAEGLDNL LKKSTFESMY VTKRNQFDKE
 TASKTKAWPS DVVPENQVEI LADAIQNGYI KSVNDVTNKA HTLSGRAWML
 DHGTPTTIAN GLTPVPEGTK VYLRWIDQDG ATSPMYTAKT TSRLSAADGN
 QVGPGAYAFD LRTGWIDAKG KHHVYRAVKG QYYKIWINDF RTKDGNIATM
 LRVAGGYVPG TYVDSVTYNN MGQFPLIGTN MQRTGIFMTT IPSEKYLISK
 HYVKDTKGAA ANPAVTIEN NFVSGKVWIE TGAGDYVNSA TGPNNHAKDV
 VASGYKVVMS SLTDQGAKEY DAQVNRLPKK DRAEAAARQLL IKHPEYIAAT
 VEGITNEWGR YTLRFPKGTF NKDHLGYGYL DFDGEIVKTY SGFTSPEFRR
 PNYNLTVTPQ TAPYYRPVRR AWWNVNFAVI EAPQSQIEIK EFDATSNPAH
 RGQTATIDII GMPKTSLLTR VQWKDSSGSI VEDSGPFVTE EEAEHIAEFV
 IPSSAKSGEV YTVQLVVGNH IVASDSLIVH VNEEAATYHP IYPSTTVESG
 QRV TIPAPKN MDGKPLLDGT TFEKGHHVPT WALVNGDGS I TVKPGKVAE
 GEYDIPVIVT YPDGSKNTIF APVTVEEKQP MASQYEPITT GVSKPFGNPV
 MPTDVTDSIQ VPNYPLEGQQ PTVTVDDSEQ LPDGTTEGYK DIDVTVTYPD
 GTKDRVKVPV VTEQQLDSDK YDPVATGILK PFGTPTTEED VIKLVEIPKY
 PTDLTQPKVT VTVPNTLPDG QTPGKVDVDV TVTYPDGSTD HISVPVWTKN
 HLDKDKYNPI TTGVSKPFGI PVTPTDVTDS IQVPNYPLEG QQLTVTVDD
 TQLPDGTTEG HKDIDVTVTY PDGTDKHIKV PVVTEKQSDN EKYEPTTNGI
 TKKYGIPTTE DEVIDIVRIP YFPVDGVQPI VTVNDPRLLP NGQKEGQINV
 PVTVTYPDGT KDLMTVPVIT GKQAENEKYD PITLGVT KDY GDPTTANDVT
 KSIQIPTYP A GGEQPIATAD DESQLPDGTV EGKVDIPVTV TYPDGTQDHI
 TVPVFTNQQR DNQKASKAVT KIHGISVTGT DMTDTKKNHN YPAGGEQPKV
 TVKDDQLSE GKVDSTVGAD NVT TTDLLSS VTAVSHGHQT SVQTTKENQS
 VHDEEVTIPT VAHVSTIMTG VVKGEQEATD IVARPHVETT QLPSISAQAT
 VKKLPETGEN NEQSGVLLGG FIAFMGSLLL FGRRRKPKKD

SEQ ID NO: 9

atgtttaatcaacaaaaacaactatggtatccggaatatgcaatcgggacttcatca
 gtattattagcatgacattatattatcacacatgacgcaactgcatctgcagctgaaaac
 aatacaactgcaaagacagagacaaatcaagcagcaacaatttcttctcgcaacttcgcca
 accgacgtcgctcaacctaattgcagacacgaatgctacaacggcgactaaagagacaaca
 ccacaatcagattcaacagcattaccgcaagcagcagcgcaacctcaaacgggccaaca
 gcatcgaaagacacagtagatacaataaaacgcaaacagcagattccacaaccgctcct
 cctgtgacagacgcgcaaaaagctaattgacgacacaacacagccagaagctgcgactgta
 gccaaaaaagaagatgctcagacaccatcgactgcagaccctacaccacaagcgcaaca

ccgcctcagtcaaaagcacctcaagaaacgcaacaacaatcaacagttgaagatacaacg
ccacaacaaaacgcatcaactgaagcacaccctaaaaatgtagataccgcttcaacaaaa
caacaacaaaacgcatcaaccgcaccgacaccttacacacaacaagcagacgaagca
atgacagatgtcacaacaaccagtgctgcagacgaacgtacagccgttagccctgcagaa
gatcaacctaaaaatacgaacacagctgacaaagcaaccgttgcgacaccaccacgtgac
aatgctaagactgctgatccgaacaaaaagatgacacgtgcagcaacgacacaacaagat
gatgccgtcgatacattgaagtcaaaagaaatgacagcaacgatcgataaaagttttcca
gccgttaaataattacacgttgaaaaatggtaaaaaagtcgatgcacaactgacggatgca
cgtcaaatacatcgtcaatggtgaagtcattacaccaacagtc aaatacaacaaaattgat
gatcatacggctgaatatgacttaacagcacaaaatgattcacgttcgattgatgccaat
tttaaatttcgtttatcagttgaaggtaagaccgttgatttacaatgacagattacacg
aacaacaacacagatccacaaaacgtcattcgcaacttttagctttgtaagtcaatcgctc
gtatctgtaaacaatcaacagaaaaatgccaaactgcaaacatcgaaactgtctacaaat
acaatgaaaagcggcgataaatcatatcatatcgatgaaaatttcaaaaacgacttcaac
gactttatgatgtacggtttcgtgtcaaatgatgattacagtgaggattgtggagtaac
gcacaaattggcgtcggcattggtgaacaagacttcttacgtgtctacgcacagctctata
caaacagatatcggggcgtcgtgtcgggttaggctcaatgccatgggtttatccaaaagac
gctgcacatccagatgcgaaaaatcaaggactactccacatgtcaaagttgcaattgcg
gaagatgaaaatcaagatggtgaaattaactggcaagacggtgcaattgcttatcgtagc
attatgaacaatccatatggtgccgaagaagtacctgaccttggtgggtaccgtatcgcg
atgaactttggttctcaagcgcaaaaccatttttaagacggttagatggtgtgaaaaaa
ttctatctcaatacagatgggttagggcaatccattttattaaaaggttataacagtgaa
ggccacgactctggctatttagattacgcgaatattggtcaacgtataggtggcgtgaaa
gactttaaaacgttacttcaaaaagggcagattatggcgacggttcgggtcttcatgtg
aatgcatctgaaacatatccagagctctcaagcattcaatcctgccctcttacgtaaagat
gcgaatggaaactatatgtatggctggaactggctcgatcaaggctttaacatcgatgca
gattacgatttaatacacgggcgtaaagaacgcttcgaagcactcaacaaaattgtcgggt
gatgacctcgactttatttatgtcgatgtatgggggaatggacaatccggcgacaataca
gcttgccatcacatcaattagccaaagaaatcaacgacttaggatggcgcgctcggtgtc
gaatggggtcacgggtatggaatatgactccacggtccaacattgggcagccgacttaacg
tatggatcgtacaaaaataaagggttaactcagaggtagcacgcttcttaacgcaaccat
caaaaagattcatgggtcggtaactatccaaaataactcaggtgcagctgacttccattg
ctcggcgggttatgacatgaaagattttgaagggtggcaaggctcgtaacgattactctgct
tacattaaaaatatatttcaatggtgatgtaccaacaaagtttttacaacattataaagt
atgcgtattgtcgatggtgagcctgttaaaatgactgccaatggtcaaacgattgactgg
acaccagaaatgcaagtcgatttcaaaaatgaagccgggtgatcaagtcactgttaaacgt
aaatctaacgactatgaaaacgacactgacaactaccgctcacgtacaatcgaattgaat
ggtcgcacagtaactcgatggcgattcataccttttaccatggaattgggatgcgaacggc
caaccattaaactggcgataacgaaaaattatatcactggaataaaaaaggcgggttcaacg
acttggaactgcctgaatcatgggatacagaccaagtcgtgctatacgaattatctgaa
acgggtcgtaagtcaccacgtacagtggcagtgaaagaccatcaagtgacactcgataat

attaaagcagacacaccgtatgtcggtttataaagtcgcacaaccagacaacacagatgtg
 aactggagcgaagacatgcacgtgaaagatgccggcttcaactcacaacaactgacacct
 tggacaatcgaaggcaatcgagataaagtgagcatcgaaaagtcgacaacatcaaataaa
 atgctaaaaatcgatagtcacaacaaaaaacgcaattaacgcaacaattgacagggttta
 gtgccaggacaacggttacgctgtctatgttggcatcgataaaccgcagtgatgcagcggcg
 catattgcagtgacacataacggtaaaaacgctcgcaagtaacgaaacagggtcaatcgatc
 gcgaaaaactatgtgaaagcagatgcacatagtaacaatgctgcgacggtttaaaaatggc
 ggtagttacttccaaaacatgtacgtgtacttgcgttgccgagaagatggtaaaagcagac
 ttgacgattcaacgcgacccagggtgaaggggccacttatttcgatgatattcgtgtgtta
 gaaaataacgcgaatctccttcaaaaacggcacattcaaccaagacttcgaaaatgtacca
 caagggttattcccggttcgtcgtgtcagaagttgaaggcgttgaagataatcgcgttcac
 ttatctgaaaagcagcgacccgtatacacaacgcggatggaataataaacgtgtcgatgat
 gtcattgatggcaaatggtcacttaagtaaacgggtcaaacagggtaaagataaaatggtc
 atccaaacgattccgcaaaacttctacttcgaaccaggaaaaacgtatgaagtgtcattt
 gattatgaagcaggttctgatgatacgtatgcatttgcgacaggtagtggggacatttct
 aaaaatcgtaactttgaaaagacaccattgaaaaatacagtcgatgggtggcaaaagcgaaa
 cgggtgacattttaaagtgacgggtgatgaaaatgggtcaaacttggatcggtatttactca
 acgaaaacaccaatgatccacgaggcgtgaaaaatggcaatcaaactcaacttcgaaggg
 acgaaagatttcatctagacaaccttctatccgtgaaattgacgcaccgaaagcctgat
 gccacacaagaaagcgggtgatagcgacccaatgaatgaaacagatgagcgtaacgtcaat
 tcaaacgggtacattagccgatcatagtgagacaactgatgtcaatgtcagtgcaacggca
 gatgatacagcagtcaaaaggcgaaatgacgacaaacagaacagatgcaccaactgttaca
 ctgcctgaagcaacgatagtagatgaaggcacgtcaaactcctgtcactacaacaccaacg
 aatacaacacaagctatgacaaataaggctgatgagatgccacaaacgatgaacaatggt
 cctttaactagcatcgctaccgatatgatgcagttctcatgcggtggattccatggcagca
 acactagcagctacaaatcaagtggcggcacctgtgcgtcaaacagcaggacctatgcaa
 catggtatggacagtgcttcaacgcaacacgcacccatacaagttgacaatgtcacagca
 ccaccattaccagatgaacagtttgccgaattacctaaaactggggatacgactccaaat
 acacgtggacctttaatggcgatgatagttggcgagtccttaacagcattcggattcaga
 cgccaacgtaaaagaaaaatag

The protein sequence translated from SEQ ID NO 9 is designated SEQ ID NO: 10 and is shown below:

SEQ ID NO: 10

MFNQKQHYGIRKYAIGTSSVLLGMTLFITHDATASAAENNTAKTETNQAATISSRTSPTDVAQPNADTNAT
 TATKETTPQSDSTALPQAAAQFQTGQTASKDVTDNKTQTADSTTAPPVTDAPKANDDTTQPEAATVAKKEDA
 QTPSTADPTPQAQQPPQSKAPQETQQQSTVEDTTPQQNASTEAHPKNVDASTKQQQTTPSTAPTPYTQQADE
 AMTDVTTTSSVDSNVQPLAPAEDQPKNTNTADKATVATPFRDNAKTADPNKKMTRAATTQQDDAVDTLKSKEMT
 ATIDKSFPVAVKYYTLKNGKKVDAQLTARQIIVNGEVITPTVKYNKIDDHAEYDLTAQNDSRSIDANFKFRL

SVEGKTVDLQMTDYTNNTDPQNVIRNFSFVSQSLVSVNNQQKNAKLQTSKLSTNTMKSGDKSYHIDENFKND
 FNDFFMYGFVSNDAYSAGLWSNAQIGVGIGEQDFLRVYAQSIQTDIGVAVGLGSMFWFIQKDAHPDAKNQGL
 LPHVKVAIAEDENQDGEINWQDGAIAYSIMNNPYGAEEVPDLVGRIAMNFGSQAQNPFLKTLTLDGVKKFYLN
 TDGLGQSILLKGYNSEGHDSGHLDYANIGQRIIGGVKDFKTLTLLQKGADYGARFGLHVNASETYPESQAFNPALL
 RKDANGNYMYGWNWLDQGFNIDADYDLIHGRKERFEALKQIVGDDLDIFIYVDVWNGQSGDNTAWPSHQLAKE
 INDLGWRVGVWGHGMEYDSTFQHWAADLTYSYQNKGINSEVARFLRNHQKDSWVGNYPKYSGAADFPLLGG
 YDMKDFEGWQGRNDYSAYIKNIFNVDVPTKFLQHYKVMRIVDGEVPVKMTANGQTDWTPMQVDLQNEAGDQV
 TVKRKSN DYENDTDNYSRSTIELNGRTVLGD SYLLPWNWDANGQPLTGDNEKLYHWNKKGGSTTWTLPESWD
 TDQVVLIELSETGRKSPRTVAVKDHQVTLDNIAKADTPYVVYKVAQPDNTDVNWSEDMHVKDAGFNSQQLT PWT
 IEGNRDKVSI EKSTTSNEMLKIDSPTKTTQLTQQLTGLVPGQRYAVYVGIDNRS DAAAHIAVTHNGKTLASNE
 TGQSI AKNYVKADAHSNNAATFKNGGSYFQNMVYFVAPEDGKADLTIQRPGE GATYFDDIRVLENNANLLQ
 NGTFNQDFENVPQGLFPFVSEVEGVEDNRVHLSEKHAPYTQRGWNNKRVDDVIDGKWSLKVNGQTGKDKMVI
 QTIPQNFYFEPGKTYEVSFDYEAGSDDTYAFATGSGDISKNRNFEKTP LKNTVDGGKAKRVTFKVTGDENGQT
 WIGIYSTKTPNDPRGVKNGNQINFEGTKDFILDNLSIREIDAPKPDATQESGDSAPMNETDERNVNSNGTLAD
 HSETTDVNVSATADDTAVKGEMTTNRTDAPT VTLPEATIVDEGTSNPVTTTTPTNTTQAMTNKADEMPQTMNNV
 PLTSIATDMMQSHAVDSMAATLAATNQVAAPVRQTAGPMQHGMDSASTQHAPIQVDNVTAPPLPDEQFAELPK
 TGDTPNTRGPLMAMIVGAVLTAFGFRRQRKEK

SEQ ID NO: 11

atgacaagaaaatttagggaatttaagaaaagttaagtgaagaaaaagcaagagt gaaa
 ctttacaagtcaggtaaaaaactgggttaaagctggaattaaagaatttcagttattaaaa
 gcattaggtctatcttttttaagccatgacattgttaaaggatgaaaatggagaagtaacg
 acacaatttggggaacagttgaagaaaaatgcattaagaacaactgcttttgcggtgga
 atgttcacagttaatatgttgcatgaccaacaagcatttgcggtgctcggtgacactata
 acttctgaactggcaacccaaaagtcaaactattggcgatcaaacatcaattgttattgaa
 aagtctacatcgctcagatcaatcaacgaacccaataacagaaagtgaagtaaacacgat
 tctgaaagtatctcattatctgagcatcaaacatcagagtcacaagctctttcaacgtca
 acttccaaatcaatatcaacttcagtagaggaatcagaatcaacatcaaaagattctcat
 actaaaactcaagatagtcattcagatagtcattcagtcacaagtcagaggtaaatggc
 tcttccaaccacgagcaatcaacaccacacactgcacaaagtcttacgagcctattctatt
 gagagccaaacgtcgacttcaaatacatcattgaaggaaactaaagaaggggaattgtca
 aaaaacctttcgaagttatctcaaaatcaaaacatcaaacttcattgaagaacatacgtg
 cggtcagcagatttgagctcaggttatatcaggatttagagcggttactatgtaccaaga
 tcaagaacaacaccaacgacaaaagtctacacagggcaaggaagcttcagaggttagaggt
 agaattaaatataatattttctacaaagttgtcggttacaagtaatggcaaagaaatgaag
 atccgctatacattgagtcagatgatccaaacacgtctaattgttgaaaaacctaggtgg
 gcaggacagaaacgatttggtattcataatacttgggatgaaggtcctggtcgcgggcaa
 ttaaagttagggtcggttcggcaaaccaacagttatatacaggagaaactagaccgaat
 tatggtagctgggttgccacacctataacgaaatattgttcaggcgatcgtaaaatgggt
 ttttactggcaagctgctgtacttgcaccgagacatggagagaagggagaaggaatcaca
 gcagaaattacagttcctattgttaacccttctggaagatttaattgggaattccatcct

gtcgggtcaacaggacggaggttggtggcaaaactgactactttgaaaatgtatggattcga
gactatgaccatattacaaatatattcaaactaaggaaggcagagcctcagtttcgcac
tctattttctcaggtgaaagcaagtgaatcgagatcgacatcgctcatacaatcggagtct
attagaagatcacagtcctatatctgagagtgaatctattgtagccgcaagtcattcggca
agtgtagcaaaatcgcaatccatctcgagaagtcaatctgtggcgaaatcacaatcgatc
tcaagaagtcagtcgatcgacacagccgatcagcaagtggtggcaaaatcgcaatccatc
tcaagaagtcagtcgatcgacacagccgatcagcaagtggtggcaaaatcacaatcgatt
tcaagaagtcagtcgatcgacacagccgatcagcaagtggtggcgaaatctcaatcgatt
tcaagaagtcagtcgaattgcgagagccaatcagcaagtggtggcaaaatcacagtcgatt
tcaagaagtcagtcgaattgcgagagccaatcagcaagtggtggcgaaatcgcaatcgatt
tcaagaagtcagtcgattgcacatagccgatcagcaagtgtagcggaatcacagtcgatt
tcaagaagtcagtcgattgcgaatagccaatctgtagcagcgagtgaaatcagagagtgta
tcaatatcattgtctaaaaagcagtcgaatcgatgagtaattctgaaagtcagcaaaa
tcacactcgctttcgggtgaaaaggtctaactggattaaaaagtcaaaagcggcttcagta
agaaagtcacattcactttcggtaagaaaatctaattcggcgaaaaggtcacatgctatt
tcggtaagaaagtcaaagtcattatcagttaaaaagtcgaatttcgcagagccaatcagca
agtggtggcgaaatcgcaatcgatttcaagaagtcgaatcagtagcagcgagtgagtcggca
tcgctaagtaagtcgaagagcacatcgctcagtaactcagtgagtgagagaaatcgacg
tcattaagtcggttcagcaagtgtagcaaaatcgcaatcgatttcaagaagccaatcagta
gtagcgagcgaatcggcatcggttaagtaagtcgaagagcacatcgctcagtaactcagtg
agtgagagaaatcgacgtcattaagtcgatcagcaagtgtagcaaaatcgcaatcgatt
tcaagaagccaatcgggtggcagcgagcgaatcggcatcggttaagtaagtcgaagagcaca
tcgctcagtaactcagtgagtgagagaaatcgacgtcattaagtcgatcagcaagtgta
gcaaaatcgcaatcgatttcaagaagccaatcgggtggcagcgagcgaatcggcatcggtta
agtaagtcgaagagcacatcgctcagtaactcagtgagtgagagaaatcgacgtcatta
agtcgatcagcaagtggtggcaaaatcgcaatcgatttcaagaagccaatcagtagtagcg
agcgaatcggcatcggttaagtaagtcgaagagcacatcgctcagtaactcagtgagtgca
gagaaatcgacgtcattaagtcgatcagcaagtgtagcaaaatcgcaatcgatttcaaga
agccaatcgggtggcagcgagcgaatcggcatcggttaagtaagtcgaagagcacatcgctc
agtaactcagtgagtgagagaaatcgacgtcattaagtcgatcagcaagtggtggcaaaa
tcgcaatcgatttcaagaagccagtcagtagcagcaagtgagtcggcatcattaagtaag
tcgaagagcacatctttaagcaactcagtgagtgtagagaaatcgacgtcattaagtcga
tcagcaagtggtggcgaaatcgcaatcgatttcaagaagtcgaatcagtagcagcgagtgag
tcggcatcgctaagtaagtcgaagagcacatcgctcagtaactcagtgagtgagagaaa
tcgacgtcattaagtcggttcagcaagtgtagcaaaatcgcaatcgatttcaagaagccag
tcagtagcagcaagtgagtcggcatcattgagtaaatcaacaagtcagtcacaagtgac
tcagatagcgcgtaacatcaacatctgtatcagatagcgattcagcttcattgagtaag
tcgactagtacatcaacaagcgattcagacagcgcgtcagcatcattgagcaagtcaca
agtacatcaacgagcgactcagatagcgcatcgacatcaacatcagtatcagatagcgac
tcggcatcggttgagtaaatcgacaagcacgtcaacaagtgattcagacagcacgtctact
tcattgagtaagtcgacaagtcacatcgacaagtgattcagatagtcggtcaaaatcaacg

[illegible]

tcaatcagtgcgagtcfaatcaagtagtttatctgcttcaacatccacatcgattgaaaag
tctgagttacatcaacaagtggtcaaattcaactaatgctgcttaagtagctcatct
tcacttagtacatcagcaagtaacttctgtaagcgaagtgcacatctgtcacacattctgaa
aatgatttaagtgcatctaacgatagagatacatccggatcagtaagtcatttgcttct
gaaaatacatcattaagtgattctgcatcaattagtggaagtttctagtagtacgtcc
gcgtcaacttcgaaatcatcatcactttcagcaagcgcgttacatgataagcatgtatca
gaaagcacttctgcatcattaagtagtgagattcaagtcgtgcttcggcatcagtgta
acgtcattatcagaatcagatagtggttaataagactctgaatcaattagcgtttccgag
cacacatcaacattacaatcaggtagtcattcactatcacaacaacaatcagcagaatta
tcacaatcagagcaaacatcacaatcacaacgcatttcaacaagtgctcagtatccggt
atgaaatcagaaagtgctgctaaggatctgaatcgctatctacgtctcaatcaaaagta
gatagtcfaatcacaatccggtatctgaatcagcgcagcaactcacgagtgtaagagattca
aaatcaacaagcgttcaatgcatcgatcattgtcagagtcagtatctcaaagtatgtca
cttattgatcagtcagaaagtgattcaacatctatatcgatttcgacgtcaatcagtgat
gaagactctatgctgtattctatgagtgattccgcatcgatcagtactaaggcatcaagt
agtatgtctacttcgacaagcgaagagcatgccaacagtcattctcagtcctgaatcgaca
gcacggttgaaagtatctcaagaaatgagtgcatcggcttcaacaagcaaatctgagtc
caatcagagtcagtatcagtaagtaacgaagaatcaaatatctcatctatgcaagagtc
tttgtagagagtgcaaaagcatcgcgtagtgcatctatgagcgttgcaaaatctgaagcc
tctgaatcacagctattaagtgagtcctaatgcttcggtgaagccaatcagcaagcacaagt
agtaaagcatcagcaagtcagtcagaatctatttcaacgtcactcagcgtatctgaagca
actcatggaaaaccgagaaatcattcagaaagtgcatcagcaagtcattattagaagaa
aatgagtcattaagcgaattcagcatcaacaagtggtgaagattcagaaagtgcatcagca
tctctgtcgggtgatcaatcacaatcagcaagtgcatgaaatcaacacatgcacagaa
aaagcttcagtgaaatacagtgcaaacgcacgaagcgtgcacatcgacatctatc
tctaactcgaaatctaaagtcattgcgagtgaaatcgaagtcacaagcatatcaacatat
gaatcgttgtaatatcgactagtaagaacaatcaacgcgtgatcagtgagtgagtcg
acatcaacgtctaaagtgaaagtcagaaagcgactcggcatcaacgtcgacatctgaatca
atctcaattagcgcgaatcgttcagggttacacatcgtctaaacgttcggtacaaatgag
gaagcacaatcaacgagcgaattcattatcagtaatgcaatctgaagggttcagtaagtga
tcgcaatctttaagtatatcagataagacatcacagtccttatcggaatcaatatcgcat
tcagaaagtgactctgatagtaactcagtgcttattagtcagagacatctgaacaacat
tcgggtgtcagacagtgactcgatgtcaatttcggaagcgaatctattgcatatagtcaa
tcagcagtgaaatcagaatcaacaagtatcgcaaaatctgatagtatttcgaactcatta
tctgtttcattaagtgaaatcagaaagtgaaagcaagcacatcagcttcagtgagtcacat
gaaagtcagtcgtgaaagggttctctatcaacaagtatcttgaaacagtcattcagcatct
actcatcaatcaacagaagcttctcaaagtacatcaacttcaaaagttgaggaagcatca
ttgagtgactctgcttctgtatcagattcacaatcactttcaatgagtcagtgagaaatca
caaagtgcatcgacttcaaaatctacgagtcgtgtcaaaaactatttctgagtcagagtc
gtgagtgcatcaacatcaacaagtgaaagtgtaagtcagaaagcaagcgaatttgatca
gcagtagactcattgagtcagtaacttctaacggaagcacaacgaaagaagatgcgag

acatttgtatccacagtagattcattgaaagacaaagcatcaaataatggtacaccatca
gagtttgcgtcagcagtgaaatcaacacacgcacagtcagtgagtggtcagcatcagaaagt
acgtcagcatcaacatcaacaagtgaagctgtaagtagagaagcaagcgaatttgtatca
gcagtgaaattcgttgagtgaaagcgaacttctaacggaagcacaacgaaagaagatgagcag
acatttgtatccacagtagattcattgaaagacaaagcatcaaataatggtacaccatca
gagtttgcgtcagcagtgaaatcaacacacgcacagtcagtgagtggtcagcatcagaaagt
acatcagcatcaacatcaacaagtgaagctgtaagtagagaagcaagcgaatttgtatca
gcagtagactcattgagtgcaagtaacttctaacggaagcacaacgaaagaagatgcaagc
acatttgtatccacagtagattcattgaaagataaagcatcaaacaatggtacaccatca
gaatttgaatcagttgtgaaatcagtagacagcatcaatgagtgcatcagcaagtgcgtca
acatcagcatctacatcagcatctacatctacaagtgaagctgcaagtgagaagcaagc
gaattagaatcagtaaggaaatcattatccaatggagcatcaaacggtagcacagcaaga
gaaggtgcaagcacatttgtatcaacggtagattcattgaaagataaagcatcaaacaat
ggtacagcatcagaatttgaatcagttgtgaagtcagtagacagcatcaacaagtgcatca
gcaagtgagtcacagtcagcatcaacatcagcaagtgatcagcaagtagagaagcaag
gaatttgtatcagcagtgagcatcattaagcagttcagcatggaacggaagcactacagga
gaaggtgcaagcacatttgtatcaacagttgattcatcgaaagattcagcgtcagacaaa
gcttcaccatcagaatcagaatcagttgtgaagtcagtagacagcatcaacgagtagcatca
gcaagtggtcagcgtcggcaagtagcatcgacatcaacaagtgaaagctgtaagc
acagaagcaagtgagtttgtatcagcagtgaaactcattaagcagtgaaagcatcgaaacggc
agcacaacaagagaaggtgcaagcacatttgtatcaacagtagattcattgaaagacaaa
gcatcaaacaatggtacagcatcagaatttgaatcagttgtgaagtcagtagacagcatca
atgagtagcatcagcaagtggtcagcatcagaaagtagtcggcatcgacatcgacaagt
gaagctgtaagtagagaagcaagcagtcagcatcgataagtgtagatcaatgtcagtgagc
gcatcaacaagtgcttcaatgagcgtatcagtggtcaaacagtggtcagtgagtgactct
atttcagtaagtgcatcaacaagtgaaaccttaactcggtaagcacttctatgagtagttct
ctttcaacatcggcatcaacgccatcagaaattacttcaagttcgtcatcaagcgattca
gagacagttcaaaaagtagtttctaagatgaacagcacgctacaaataaagttgaaaaa
ttacctgacacaggtcaatcaacgacacaaactgggtttattgggtggagtaggtgcttta
cttacaggccttgggtttactcaaaaaatcaagaaaacaaaaagatgaagaaacatcatca
catgaataa

The protein sequence translated from SEQ ID NO 11 is designated SEQ ID NO: 12
and is shown below:

SEQ ID NO: 12

MTRKFREFKK SLSEKARVK LYKSGKNWVK AGIKEFQLLK ALGLSFLSHD
IVKDENGVT TQFGEQLKKN ALRTTAFAGG MFTVNMLHDQ QAFAASDAPI
TSELATKSQT IGDQTSIVIE KSTSSDQSTN PITESSEKHD SESISLSEHQ
TSESTSLSTS TSKSISTSVE ESESTSKDSH TKTQDSQSDS HQSTSQEVNG

SSNHEQSTPH TAQSLTSLSI ESQTSTSNST LKETKEGELS KNLSKLSQNO
 NIKLHEEHTM RSADLSSGYT GFRAAYYVPR SRTTPTTKVY TGQGSFRGRG
 RIKYNIFYKV VVTSNGKEMK IRYTSLQDDP NTSNVEKPRW AGQKRFGIHN
 TWDEGPGRGQ LKLGSAFGKP TVIQGETRPN YGSWVGTPIT KYVSGDRTNG
 FYWQAAVLAP RHGEKGEGIT AEITVPIVNP SGRFNWEFHP VGQQDGVGGK
 TDYFENVWIR DYDPYKYIQ TKEGRASVSH SISQVKASES RSTSLIQSES
 IRRSQSISES ESIVAASHSA SVAKSQSISR SQSVAKSQSI SRSQSIAHSR
 SASVAKSQSI SRSQSIAHSR SASVAKSQSI SRSQSIAHSR SASVAKSQSI
 SRSQSIAQSQ SASVAKSQSI SRSQSIAQSQ SASVAKSQSI SRSQSIAHSR
 SASVAESQSI SRSQSIAHSR SVAASESESL SISLSKKQSI SMSNSESAAK
 SHSLSVKRSN WIKKSKAASV RKSHSLSVRK SNSAKRSHAI SVRKSLSLSV
 KKSISQSQSA SVAKSQSISR SQSVAASESA SLSKSKSTSL SNSVSAEKST
 SLRSASVAK QSISRQSV VASESASLSK SKSTSLNSV SAEKSTSLSR
 SASVAKSQSI SRSQSVAASE SASLSKSKST SLSNSVSAEK STSLRSASV
 AKSQSISRQ SVAASESASL SKSKSTSLSN SVSAEKSTSL SRSASVAKSQ
 SISRSQSVVA SESASLSKSK STSLNSVSA EKSTSLRSA SVAKSQSISR
 SQSVAASESA SLSKSKSTSL SNSVSAEKST SLRSASVAK QSISRQSV
 AASESASLSK SKSTSLNSV SVEKSTSLSR SASVAKSQSI SRSQSVAASE
 SASLSKSKST SLSNSVSAEK STSLRSASV AKSQSISRQ SVAASESASL
 SKSTSTSTSD SDSASTSTSV SDSASLSK STSTSTSDSD SASASLSKST
 STSTSDSDSA STSTSVSDSD SASLSKSTST STSDSDSTST SLSKSTSTST
 SDSASASKST SVSDSTASL SKSTSTSTSD SDSASKSTSV SDSTASLRK
 SASTSTSDSD STSTSLKST STSTSDSDSA SKSTSVSDSD SASLSKSTST
 STSDSDSASK STSVSDSDSA SLSKSTSTST SDSASASKST SVSDSTSTSL
 SKSTSTSTSD SDSASTSTSV SDSTASLSK STSTSTSDSD SASTSVSDSD
 SASLSKSTST STSDSDSVST STSVSDSDSA SLSKSTSTST SDSASASKST
 SVSDSTSTSL SKSTSTSTSD SDSASTSVSD STSASLSKST STSTSDSDSA
 SKSTSVSDSD SASLSKSTST STSDSDAST STSVSDSDSA SLSKSTSTST
 SDSASTSL SKSTSTSTSD SDSTSTSLSK STSTSTSDSD ASKSTSVSD
 SDSASLSKST STSTSDSDSA SKSTSVSDSD SASLSKSTST STSDSDSASK
 STSVSDSESA SLSKSTSTST SDSASTST SVSDSTSVSL SKSTSTSTSD
 SDSTSTSLSK STSTSTSDSD SDSASLSKST STSTSDSDSV STSTSVSDSD
 SASLSKSTST STSDSDAST STSVSDSGSA SLSKSTSTST SDSASASKST
 SVSDSDSASL SKSTSTSTSD SDSASTSTSV SDSTASLSK STSTSTSDSD
 SASTSTSVSD SDSASLSKST STSTSDSDSA SKSTSVSDSD SASLSKSTST
 STSESDSASK STSVSDSDSA SLSKSTSTST SDSASTST SVSDSDSASL
 SKSTSTSTSD SDSASKSTSV SDSASLSK STSTSTSESD SASTSTLVSD
 STSVLSQST SVDKOSTAKG STELVNVASL SISASQSSSL SASTSTIEK
 SESTSTSGSN STNASLSSSS SLSTASTSV SEVTSVTHSE NDLSASNDRD
 TSGSVSQFAS ENTSLSDSAS ISGEVSSSTS ASTSKSSSL ASALHDKHVS
 ESTSASLSSG DSSRASASVS TSLSESDSAL IDSESISVSE HTSTLQSGSH

SLSQQQSAEL SQSEQTSQSQ RISTSASVSA MKSESAAKVS ESLSTSQSKV
 DSQSQSVSES ASNSRVSRDS KSTSASMHRS LSESVSQSMS LIDQSEDST
 SISISTSISD EDSMLYSMSD SASISTKASS SMSTSTSEEH ANSHSQSEST
 ASVEVSQEMS ASASTSKSES QSESVSVSNE ESNISSMQES FVESAKASRS
 ASMSVAKSEA SESQLLSESN ASVSQSASTS SKASASTSES ISTSLSVSEA
 THGKPRNHSE SASASQLLEE NESLSDSAST SVEDSESASA SLSVYQSQA
 SALKSTHASE KASVNTSANA SKRASASTSI SNSKSKVIAS ESKSTSISTY
 ELSISTSKE QSTRVSVSES TSTSKVKSES DSASTSTSES ISISANRSGY
 TSSKRSVQMS EAQSTSDSLV VMQSEGSVSV SQSLSIDKT SQSLSESISH
 SESDSDSNSV SISQETSEQH SVSDSDSMSI SESESIAYSQ SASESESTSI
 AKSDSISNSL SVSLSESESE ASTSASVSTS ESTSVKGSLV TSILNSQSAS
 THQSTEASQS TSTSKVEEAS LSDSASVSDS QSLSMSHEKS QSASTSKSTS
 LSKTISESES VSASTSTSEA VSTEASEFVS AVDSLQVTS NGSTTKEDAS
 TFVSTVDSLK DKASNNGTPS EFASAVKSTH ASVSVSASES TSASTSTSEA
 VSTEASEFVS AVNSLSEATS NGSTTKEDAS TFVSTVDSLK DKASNNGTPS
 EFASAVKSTH ASVSVSASES TSASTSTSEA VSTEASEFVS AVDSLQVTS
 NGSTTKEDAS TFVSTVDSLK DKASNNGTPS EFESVVKSVH GSMSASASAS
 TSASTSASTS TSEAASAEAS ELESVRKSLV NGASNGSTAR EGASTFVSTV
 DSLKDKASNN GTASEFESVV KSVHGSTSAS ASASTSASTS ASESASTEAS
 EFVSAVASLS SSAWNGSTTG EGASTFVSTV DSSKDSASDK ASPSESESVV
 KSVHGSTSTS ASVSASASTS ASTSTSEAVS TEASEFVSAV NSLSSEASNG
 STTREGASTF VSTVDSLKDK ASNNGTASEF ESVVKSVHGS MSTASVSAS
 ESTSASTSTS EAVSTEASES ASISVSMSVS ASTSASMSVS VNSVSVSDS
 ISVSASTSEP NSVSTSMSSS LSTSASTPSE ITSSSSSDS ATVQKVSKD
 EQHATNKVEK LPDTGQSTTQ TGLLGGVGAL LTGLGLLKKS RKQKDEETSS
 HE

SEQ ID NO: 13

atgaaaaagtctagaaaaaagcgtatcgattttttacctaaccgtcaaaatcgatatgcg
 atacgtcgtttttcagtaggcactgcgtcaattctcgttgagcaacattaatttttggg
 attcattcaaatgatgcatcggcagcagtagaagacgcaacatctcaagaagcaggaaca
 actaacgaaaattcaaatagtagagaagaagcaacaacaaacgaaagtacaactgttgaa
 gcaccaacaagtgaagaagcaacaacggaagagcaatcagtagaggcgccaacaagtga
 gaagtaacaacggaagagcaatcagtagaggcaccaacaagtgaagaagtaacaacggaa
 gagcaatcagtagaagcgccaacaagtgaagaagtaacaacggaagagcaatcagtagaa
 gcgccaacaagtgaagaagtaacaacggaagagcaatcagtagaggcaccaacaagtga
 gaagtaacaacggaagagcaatcagtagaggcaccaactagtgaagaagtaactacggaa
 gagcaatcagtagaagcaccaacaagtgaagaagcaacaacggaagagcaatcagtagaa
 gcaccaacaagtgaagaagcaactacaaaaactcctgtaaaagaagaacatcctcaaca
 caagaaaattcacccacgactacactagaagaacaattttcaaatgaattcaatcagtta
 acatctacagaagataaaacaaactacacacgtgaatatttaactcaaaacacaaatctt

tcggcagaacaagtggaagcaacagttgaacgcttgaatttaagtcaagaaaatgtaaca
gccaagatatctatttcgcattacttaaagatttagctgatcaacaagatgccttatta
ccacgtgtaacacttttggcgcgtagagattctgagctcacaaacgaagcgtctatcgct
ttaactgaaaatagtccaatgttcgcgcgagcattagcgaatagtccttctggcaatgat
gtggtgtcagaagaagataatattattgtggtgatgcactcgcaaatggatacatcaat
tcacaaacagatgcaacaaatgcggaacaaatcattgtctggtcgtgcatgggttgtggat
acagggacaccagcgacaatgtcaaacggcttaacagctgttcagaaggcacaacaaagtc
tacatgcaatggattgatacagatggcgcggtttcaccagtgatcaagcaagcacaaca
aataaattgagttcaagtggtagccaagtaggtccaggtgcatatgcatttgattta
cgtgaagcatggatagactcaaatggcaaagcgacagatatgaagcgtcaagtggccaa
tattatcggttatggattgatgactacaaaacagtagatgggaatacggcaaccatgtta
cgccaagcaggtgggtttctccctgggttcattgttaattcggtgacaggtacaatatt
gggtcaattcccacttatcggaacgaacatgcaacgtacaggtatctttatgggtgtgata
ccaacgaacgattacatgactacagatacaagcaattggattcaagataatgaaggacct
atctcaaaccagcagtaacgagcacaagtgaaattgtcagtggttaaagtatgggtctgag
acaggttcaggtgactatgcgaactctgcgacaggtccaaactttaactcaggtgatatt
gcacgtgaagggttatcaagttgtcatgtcttcattaacaagtgtggtgccaagcgtat
aaagcacaagtcgaatcggtgccaacagaccaacaagcggcagcagcacaccaattattc
aaagaccaccagaattttatttctgcgacagtgacgggtaaaactgatgcaaacgggtgcg
tatacattacgtttcccttcaggctcattgagtaaagattatctttatgggttatgtgatg
gataataagggcaacttggttaagggtatctcatcattcacgtcacctttattccggttcg
cctaacagtaacttatctttcgcgccacaaacagcgccatatcatagaccagccaaaaat
gcttgggtgaatgtgaactttgcgcttgtagaacaattgaaacaagtatagacatcacg
aactttgatgtgacagccaaccagcgcaacgtggtgatacggctatcattgatgtgact
tctacagcattgtcaccattacctacgcatgttgagtgagagattcaaaagggaatgtc
gttcaaaaaagtggagatgtcactacggtagaagaagctgaaacggcaggcacatttact
attcctgatgatgcgaaaacaggtgaaatctatacagtttatattgtttcaggaggcaat
gaagttgcagcagactcactgattgtccaagtgaagaaaatgcggcaacctatgaacct
gtatatccaacaacaacagttgaacaagaccaaaactgtaacaattcctacacctacaaat
gaagatgggttagcattaccagacggaacaaagttcgaaggtggcaacaatgtacctgaa
tgggcaactgtgaatgaagatgggttctattttcaatttcaccaaatacaagatgtggaaaaa
ggtaactataatgtgcctgttgctgcacatatccagatgggtcaaaagaaacagtatatt
gcaccagttttagttcaagaagctgttccaactgcagaacaatacgatccaacaattgaa
acaattaataaggaatatgggtactactgcaacagaagatgaaattaaaggcgcaatcaca
attccggattaccaacagatggagatcaaccaacaatcacgattgacgacccaactcaa
attccaaatggaacagaagaaggcacagtgaaatgtaggtgtcactgtcacttatccagat
ggttcaacagacaaaattaacagtagcagtcgttacaggttaagcaagcggataacgataag
tacacaccagaaacaacaccaattacgaaagacttcgggtacaggtgtaacagaagacgaa
gtgaaaggtgcagtcactgttccggattaccaacagatggagaccaaccaacaattacg
attgacgacccaagtcagttgcctgatgggttcaaaagaaggaaacaacggatgtcgacgta
acagtggaatatccagacggcacaacagatcacatcacagttccagtgactgttggaag

caagcggataatgataagtacacaccagaaacaacaccaattacgaaagacttcggtaca
ggtgtaacagaagacgaagtgaaggtgcagtcactgttccggattacccaacagacggt
gaccaaccaacaattacaattgatgatccaaatcaattaccggacggttcacaagaaggt
acgactgatgtaaagtgaacagtggaatatccagatggcacaacagatcacatcacagtt
ccagtgactgttggaaagcaagcggataatgataagtacacaccagaaacaacaccaatt
acgaaagacttcggtacaggtgtaacagaagacgaagtgaaggtgcagtcactgttccg
gattacccaacagatggagatcaaccaacggttacaattgatgatccaaatcaattaccg
gacggttcacaagaaggtacgactgatgtaaagtgaacagtggaatatccagacggcaca
acagatcacatcacagttccagtgactgttggaaagcaagcggataatgataagtacaca
ccagaaacaacaccaattacgaaagacttcggtacaggtgtaacagaagacgaagtgaag
ggtgcagtcactgttccggattacccaacagacggtgaccaaccaacggttacaattgat
gatccaaatcaattaccggacggttcacaagaaggtacgactgatgtaaagtgaacagtg
gaatatccagatggcacaacagatcacatcacagttccagtgactgttggaaagcaagcg
gataacgataagtacacaccagaaacaacaccaattacgaaagacttcggtacaggtgta
acagaagacgaagtgaaggtgcagtcactgttccggattacccaacagatggagatcaa
ccaacggttacaattgacgatccgagtcagttaccagatggctcacaagaaggcacaaca
gatgtgaatgtaacagtggaatatccagatggcacaacagaccacatcacagttccagt
actgttggtaagcaagcagataacgataagtacacgccagaaacaacaccaattacgaaa
gacttcggtacaggtgtaacagaagacgaagtgaaggtgcagtcactgttccggattac
ccaacagatggagaccaaccaacaattacaattgacgatccgagtcagttaccagacggt
tcacaagaaggtacgactgatgtaaagtgaacagtggaatatccagatggcacaacagat
cacatcacagttccagtgactgttggtaagcaagcagataacgataagtacacaccagaa
acaacaccaattacgaaagacttcggtacaggtgtaacagaagacgaagtgaaggtgca
gtcactgttccggattacccaacagatggagaccaaccaacaattacaattgacgatccg
agtcagttaccagacggttcacaagaaggtacgactgatgtaaagtgaacagtggaatat
ccagatggcacaacagatcacatcacagttccagtgactgttggaaagcaagcagataac
gataagtacacaccagaaacaacaccaattacgaaagacttcggtacaggtgtaacagaa
ggcgaagtgaagattcaatcacattcccggttacccaacagatggagaccaaccaaca
attacaattgacgacccaagtcagttaccagatgggttcacaagaaggtacgactgatgtc
gatgtaacagtggaatatccagacggcacaacagatcacattacagttccagtgactgtt
ggaaagcaagcagataacgataagtacacaccagaaacagaaggtgtcaacaaagatcat
ggtacgtcagtaacagaagatgaagtgaaggtgcagtcactgttccgggatacccaaca
gatggagatcaaccaacggttacaattgatgatccaagtcaattgccggacggttcacaa
gaaggtacgactgatgtaaagtgaacagtggaatatccagacggcacaacagaccacatt
acagtccagtaactgttggtaaacacctaactaaagataacggggctacagataatgat
ggcgacatgaatcaaggcacagatgaaggaaatagtgtactgatcatggcgacaatgta
aaacaagattcaaacggaaactatacgccggttgaacaacgtgacaatcatgcgacttca
cctgcaacagatatggatccaatgccaagcaatagccaaacaacttttgatggcataaat
gcaaaaggttcaacttcagagaaagcaaacataaacaacagtcctgagcaattaccagac
acaggtgaagcaatacacaaaatggtgcacttttaggcggattatttgcagcacttggg
ggcttattcttaatcggcagacgtcgtaaagaaaaagaaggcaataa

The protein sequence translated from SEQ ID NO 13 is designated SEQ ID NO: 14 and is shown below:

SEQ ID NO: 14

```

MKKSRKKRID FLPNRQNRYA IRRFSVGTAS ILVGATLIFG IHSNDASAAV
EDATSQEAGT TNENSNSTEE ATTNESTTVE APTSEEATTE EQSVEAPTSE
EVTTEEQSVE APTSEEVTTT EQSVEAPTSE EVTTEEQSVE APTSEEVTTT
EQSVEAPTSE EVTTEEQSVE APTSEEVTTT EQSVEAPTSE EATTEEQSVE
APTSEEATTK TPVKEETSST QENSPTTTLE EQFSNEFNQL TSTEDKNTYT
REYLTQNTNL SAEQVEATVE RLNLSQENVV AQDIYFALLK DLADQQDALL
PRVTLAARD SELTNEASIA LTENSPMFRA ALANSPSGND VVSEEDNIIV
ADALANGYIN SQTDATNAAN TLSGRAVVVD TGTPATMSNG LTAVPEGTKV
YMQWIDTDGA VSPVYQASTT NKLSSSSGGSQ VGPGAYAFDL REAWIDSNGK
AHRYEASSGQ YYRLWIDDYK TVDGNATATML RQAGGFFPGS YVNSVTGNNI
GQFPLIGTNM QRTGIFMGVI PTNDYMTTDT SNWIQDNEGP ISNPAVTSTS
EFVSGKVWSE TGSGDYANSA TGPNFNSGDI AREGYQVMS SLTSAGAQAY
KAQVESLPTD QQAAAAHQLF KDHPEFISAT VTGKTDANGA YTLRFPSGSL
SKDYLYGYVM DNKGNLVKGY SSFTSPLFRS PNSNLSFAPQ TAPYHRPAKN
AWNVNMFALV ETIETSIDIT NFDVTANPAQ RGDATIDVT STALSPLPTH
VEWRDSKGNV VQKSGDVTTV EEAETAGTFT IPDDAKTGEI YTVYIVSGGN
EVAADSLIVQ VQENAAATYEP VYPTTTVEQD QVTIPTPTN EDGLALPDGT
KFEGGNNVPE WATVNEDGSI SISPNQDVEK GNYNVPVVVT YPDGSKETVF
APVLVQEAVP TAEQYDPTIE TINKEYGTTA TEDEIKGAIT IPDYPTDGDQ
PTITIDDPDQ IPNGTEEGTV NVGVTVTPD GSTDKLTPV VTGKQADNDK
YTPETTPITK DFGTGVTEDE VKGAVTVPDY PTDGDQPTIT IDDPSQLPDG
SKEGTTDQDV TVEYPDGTTD HITVPVTVGK QADNDKYTPE TTPITKDFGT
GVTEDEVKGA VTVPDYPTDG DQPTITIDDP NQLPDGSQEG TTDVNVTVY
PDGTTDHITV PVTVGKQADN DKYTPETTPI TKDFGTGVTE DEVKAVTVP
DYPTDGDQPT VTIDDPNQLP DGSQEGTTDV NVTVEYPDGT TDHITVPVTV
GKQADNDKYT PETTPITKDF GTGVTEDEVK GAVTVPDYPT DGDQPTVTID
DPNQLPDGSQ EGTDDVNVTV EYPDGTTHI TVPVTVGKQA DNDKYTPETT
PITKDFGTGV TEDEVKAVT VPDYPTDGDQ PTVTIDDPDQ LPDGSQEGTT
DVNVTVYEPD GTTDHITVPV TVGKQADNDK YTPETTPITK DFGTGVTEDE
VKGAVTVPDY PTDGDQPTIT IDDPSQLPDG SQEGTTDVNV TVEYPDGTTD
HITVPVTVGK QADNDKYTPE TTPITKDFGT GVTEDEVKGA VTVPDYPTDG
DQPTITIDDP SQLPDGSQEG TTDVNVTVY PDGTTDHITV PVTVGKQADN
DKYTPETTPI TKDFGTGVTE GEVKDSITIP GYPTDGDQPT ITIDDPSQLP
DGSQEGTTDV DVTVEYPDGT TDHITVPVTV GKQADNDKYT PETEGVNKDH
GTSVTEDEVK GAVTVPGYPT DGDQPTVTID DPSQLPDGSQ EGTDDVNVTV

```

EYPDGTDDHI TVPVTVGKQP TKDNGATDND GDMNQGTDEG NSATDHGDNV
KQDSNGNYTP VEQRDNHATS PATDMDPMPs NSQTTFDGIN AKGSTSEKAN
HKQQSEQLPD TGESNTQNGA LLGGLFAALG GLFLIGRRRK EKEGK

SEQ ID NO: 15

atgacagaacgaaaatccccttcattctcaaaacatgcgatcatcgtttagtcaaagctggt
actgtcctttttatttggttggttagtggaactgcaaatgccttcaacattgtcacacgaaatg
acagcgatagctcagacagatgcgactgatgatttgaaaacattacgtgaaaatgcagat
aaaaaagtgaagcggttacaatatTTTaaatacggattataaaaaatgaatttcttgcggtta
attcgtgaatatgatacgtcgtcaaaaaatattgaagtgggtgttgacgaagcagaagca
gccaatcgtctagctcatgacgctcaatcggaacgatgaaatacaaacctgaattagatgcc
attgatgaaaaaattagcgcggttaaaggcaaggttgatgaagggtcaacgagaatcaact
gaagcgcggtcaagatgtaacgtcaacagagacaaagagtgctgaatcagaaggaagagag
ccatccactgaaggcgagagcaaaagtaaaggagtcattcttcagcacaacgattgttagca
cctcatcatggtcaacaagatgtgagcgcaactgaaagaccatattaagaacgatgtcgat
acacttaacaagactatgcaacgcaagacaagcaagtgaaccactccagggcattgac
agtgaatcacacgcattgaccatttctggttcagaaagcggtggatcacaagtctgacaat
tattttgaagaaaaacgtcaacattttacaaaactttgaacaagacattaaaaaacgtacg
gacatttctgggactgagaaggcgactttgcttgatgatgcgaaaacggttagccaaccaa
ctgaacgcgcaaaatgatacgattttaactgaacttcaacagcatgacgataaacgtgca
gcagttgaatcgatattaggtgagatttttaatgcacaagaagcggtgaacgtgcgaaa
cagatagatgttaaagggtaaaacagatcaacaattggcaaacgaaattcatcaacaagcg
gacggacttatcaaacgtcgagtgatgatttattggttaggaatgttggaataaattca
aatacacaaggtctagtggaaagcatttttacgaacacgctttgacaaacaagaagcgac
aaaattgccggcgaaatcatgcaaggcaagccttcaaatacagcgatactcgaccgcttg
aaagaccatttttaaagcggaatggttaaggcgagtgagatgatattttaaatgcgttaatt
aataatacggatgcagatgctgaagtgattgaatcaattctagggggccgtcttaatgca
gaaaatgcaaaattgattgccgatcgtgtacagcaagataaaaaagaagacacatcaaaac
ttaaaggcgattgaagacgaacttagtgcgcaagcgaatcgattgttaacgttacggaag
caattgcaacaaatccgtcataatacgaacagatatgaatgacttggttgaccactg
cgctgtattgcaaatattctcggtggtggtttaaatcgtgacgacattcactcttcaggt
cgtacgaatgacaaattgcagcaactgttaaactcgtgatcattcggttggttaggtcgtggt
ggtgatttattcaaacatgattttgcgccaaagccgaatatcgatccatatcaagcgatt
aatagtcaaaacggcatcagaagggttttttagatgggtttatttgatcaaaatggcgatttc
aatttaccgaatacaggtgaaatagtgaagcggaacttggtaccggttggtatttttagtc
gttgcaatcggtgtactgatcttaacgggtgagatttcataaaaaaacacgcaacaataa

The protein sequence translated from SEQ ID NO 15 is designated SEQ ID NO: 16
and is shown below:

SEQ ID NO: 16

MTERKSPSSQ NMRHRLVKAG TVLLLVGSGL QMPSTLSHEM TAIAQTDATD
 DLKTLRENAD KKVKALQYLN TDYKNEFLAL IREYDTSSKN IEVVVDEAEA
 ANRLAHDAQS DDEIQPELDA IDEKISALKA KVDEGQREST EARQDVTSTE
 TKSAESEGRE PSTEGESKVK ESSSAQTIVA PHHGQQDVSA LKDHINKNDVD
 TLKQDYATQD KQVTPLQGID SAITRIDHFV SESVDHKSDN YFEEKRQHLQ
 NFEQDIKRT DISGTEKATL LDDAKTVANQ LNAQNNTILT ELQQHDDKRA
 AVESILGEIF NAQEAAERAK QIDVKGKTDQ QLANEIHQQA DGLIKTSSDD
 LLLGMLNNS NTQGLVESIL RTRFDKQEAH KIAGEIMQ GK PSNTAILDRL
 KDHFKANGKA SGDDILNALI NNTDADAEVI ESILGGRLNA ENAKLIADRV
 QQDKKKTHQN LKAIEDELSA QANRLLTLRK QLQQIRHNTQ TDMNDLFAPL
 RRIANILGGG LNRDDIHSSG RTNDKLQQLL NRDHSLLRG GDLFKHDFAP
 KPNIDPYQAI NSQTASEGFL DGLFDQNGDF NLPNTGEIVK RTWLPLGILV
 VAIGVLILTV RFHKKTRKQ

SEQ ID NO: 17

atgttaaaaaaattaattgttacaggtttgattgctacagcggcgacacaagtttatgcg
 catgacacgcaagcggcggaagggtgctacagatgctccgaatgtgatggttaaggat
 gaggcgaaaaaagaagtgcacccgataatccataaaccgacttgcatttaccgcacatcta
 gaaggcgaaagatgatgctgctgtatttaaaacgtatggcaacgaatccaccagaaggcgca
 gtgccgtacgggtgtattgaataaagatggatcgattacagaaccgaatacaaatccacat
 tttgatgttttaaaaaattgaagatccaaatgcgatgaaagatttggttgatacaccggca
 gatgatcaagatacgggtaccgagtgatttataaattgaaccaccagcattaataggacca
 gctactaaacatacggatggtacgggagacgcaaaatctaataatgatgaccacaaagtaaca
 aaatcttcgggagcgtcagcccaagatatgaagaaaaagacgtgacaacacaaactgca
 caacccaaaagcagataaaaagatggcgactgcaaaagtagcaccagcgaaacaacaagat
 aaagcagccaaaatgttaccagcagcaggggaaccacaagtgaatgcaatcagtcaaaaca
 gcacttgcactttcaatgatgcattaggtgtcatcgcttctttacacgacgacgcaaa
 acaaattaa

The protein sequence translated from SEQ ID NO 17 is designated SEQ ID NO: 18 and is shown below:

SEQ ID NO: 18

MLKKLIVTGL IATAATQVYA HDTQAAEKGA TDAPNVMVKD EAKKEVTPII
 HKPTCIYPHL EGEDDAAYLK RMA TNPPGA VPYGVLNKDG SITEPNTNPH
 FDLVKIEDPN AMKDLVDTPA DDQDTVPSDL QIEPPALIGP ATKHTDGTGD
 AKSNDHVKVT KSSGASAQDM KKKDVTQTQA QPKADKKMAT AKVAPAKQQD
 KAAKMLPAAG EPQVNAISQT ALALSMIALG VIAFFTRRRK TN

SEQ ID NO: 19

atggtagaatataaaaaagaacatagcgtaaagcgactattaaaattaggaatcggttca
 acgagtattttatgtgttgatcacctcttttattaacacatgacgttggttcaagcagca
 gatatcaataacaggatgccagctttgaatacattgaagaccacttcttcatatgatcaa
 agggcacacatggatgaattacgaaacgccattacttcagatagtgacactactcaaaca
 ccatcattcaatgagataactgtgtcttcaactaatgaaacggatgcagcgtcaacggaa
 aatgtgaacccgagtgatgaggtcccggaaggaatgaaagtgaatcaacgacaccgagt
 acagaacaagacacatctatagaagaaacgggtactgaagaagtgccatctcatgaagac
 aatcatcacaacaccccaagtcaagaagagcaaccgtctccgcctgatcaaccaggaaca
 aacaaagatgaagagagtggagaaaaaccgaataaagaaaatcatcggaagccgaatcaa
 ccgaacaaagaccaaccttcaaaagatgagaataaaaaacctgacaaaggaaacaaacca
 gcaccaccgtctaaaatgccaaatcgcccgatcaaaaggaagatgggttcaaacaacacc
 ccaccacctgccactgataacggtggaaacagtaatgacggtacaacaacgggtcccaat
 ggtggaggtggcagtgaagcaagtcaccaccgaatgagcaaccgtcaaattggcaatgca
 agcgatacccatcaaaacgggttcagtttcaagcaccaatcattcgaatcagtatggtaca
 tcggcttatgatgaatacgcaggtttattgaataataattataaatataatccattgttt
 aaagaagaggttgcgctttaagtcaatttggaagtcaagatcaacatgatattgcaagt
 ttgagtcgtaaagaacaattttctcaaaatgcatttttagatgacttgcaacaaagtaca
 gattatttttagatatcaatattttaaccgctttccacagagcaatactatcatcggtta
 gataaacaagtatttagcactcgttacgggggaatttggttcgatgccagatttcaagaaa
 agtgggtgataagtcattgggttaataagcatcagcaagataaagtgaagaaaattgaacag
 caaggagaaaatattaatacgcacatcatatgaaaaatacgaaagaagatacaggaaaatca
 ttaagttacaagccgatgatatatattggcattgtcatggtcgggtttgtcgggcctgatc
 agtatgattttatggaaacgactgcatcatttttggaataa

The protein sequence translated from SEQ ID NO 19 is designated SEQ ID NO: 20
 and is shown below:

SEQ ID NO: 20

MVEYKKEHSV KRLLKLGIGS TSILCVVSPL LLTHDVVQAA DINNRMPALN
 TLKTTSSYDQ RAHMDELRNA ITSDSDTTQT PSFNEITVSS TNETDAASTE
 NVNPSDEVPA KDESESTTPS TEQDTSIEET GTEEVPSHED NHHNTPSQEE
 QPSPPDQPGT NKDEESGEKP NKENHRKPNQ PNKDQPSKDE NKKPDKGNKP
 APPSKMPNRP DQKEDGSNNT PPPATDNNGN SNDGTTTGPV GGGGSEASPP
 PNEQPSNGNA SDTHQNGSVS STNHSNQYGT SAYDEYAGLL NNNYKYNPLF
 KEEVARLSQF GSQDQHDIAS LSRKEQFSQN AFLDDLQOST DYFRYQYFNP
 LSTEQYYHRL DKQVLALVTG EFGSMPDFKK SGDKSLVNHK QQDKVKKIEQ
 QGENINTHHM KNTKEDTGKS LSYKPMIYIG IVMVGFVGLI SMILWKRLHH
 FWK

SEQ ID NO: 21

gtgattacaaataaaaaatatatatagttatcgaaagcataaaacttggcgtggcatcattc
ttattggggacattatgtttagggcatgcaaataatgctgaagcttcagaagtgagc
gcaacaacacaagaacataatgtcgagactgagcaacaaaaactgagggcggaactaaca
actgaggtagcacaacaagcagtcagcgaatcagcacctatagctgaaaacatgcagaaa
acaacatcagtggaagtgaaaatgcgaaagaggttacagcttctgatagcacacaagaa
gtcacaaaaactgaagcaaaagatacagcaacaatgaaagattcagaaattgcacaacct
gtatcagaagtgaataaacctgttactcaaacagctgcacccgtagcagaacctcaaca
gcaaacaaaacaaacttcaccacgacaagtacaagaacttactgcaccaatggacacaaaa
gtaattaatgtagaaaacggaacagatgtgacaagtaagtgaaagtgaaaaatcgtca
attacagggcatcagaataaagataaaacatatcatcaatcgaacactgtaaatccacat
aaagctgaacgtgtgacattaaattatgattggctatttgaaaatggaattaaagctggt
gattatgtttagcttccaattaagcgataatgtcgatacaaatggaatatcaacaataaaa
aaagtcacacacattatggatagtcaaaatagcgaacaaattattgcttacggggaaatt
aatgaaaacaaccgtgtccgttaccgatttatggactatgtaaatcaaaaagaaaattta
aaaggtaattgtcattaaacttatttattaaaccagataaagttcaagatgaaggaaaa
atcactgtcacttcacaattgggcaaggaaatgacaagtcaggaatttgacattaaatat
attgatggtgtaaaaagcccttcaggtatcacattaaacggtcgtcttgatgaattatca
aaagcagatcaatcatttacgcattattctatatatttaaaccctaagcataataacttaact
aatgtaactttaagaggcacagtttcaaataacgcacagcaaaatgaaaaaatggtcaa
gttaatgtttacgaatatattgggtcaaggagaattgccacaaagtgttatgccaatgta
aatgatacgaagcagttcaatgacattactaagagtatgaaatcaatcaaaaataacagt
aatggctatgaaattacttttgacatgaacaaagacaatcatccttatatcatagtatat
caaggtcactttaacaataatgcaaaagactttgatttctcaacaaatgcgacaggttat
caaaatttaaatcaatcggaatatagttattattggccttacaattattcattcaattta
acatgggataatgggtgttcttactctaataatgcaagtggggaagggaacgacaaa
cctgtaccgcgacttatggatatagtcgcagagtaaatacaattcaagatactcatgcg
gattatcctgtaatgactttccaacaacctggaaactctagaggagacagaagacagtatg
ccaatcactacacttacgaatctgggtgaggatcgtggtgaaaatacttctccaattatc
gagacaacagaagattcacagcctgttgagtttgaagaagagacaaatcatggcattcaa
gacgtgacacttcagtcagatgctgttgattttgaggaagaaacaaaccatggtgaacaa
gacacggtacaccactctgatgtcgttgaatacgcgaagatacgcacaactggcatgtta
acaggtgccatttctgaccatacaacagaagaaggcacgatggagtacacaactgatggc
ttattgattgagtttgatgatgaaatgaatcctaattgtgagcggtcagtacgatgacatc
acaacggatacगतagaggaatcatctcatattgacacattcactgaacttgaatctgaa
tttgggtcaacatgacggtatagtgacatttgaagaagatactatcgttgagaagccgaaa
acagaaaagggttaaccgagtaccacttgtaattgatttatcaacacccaaaacataaccat
cagttcaatattcaacctaccgatccaaatattgatacctctgctacgtatcgaattggc
aattttgtatggcgcatgaagatcacaatggcgtacaaaatgatgggtgaacatggtctt
gaaggtgttctgtcacacttaaaacagctgatgggtgtcgttttaatacaacgacaagt
gatgccaatggacactaccagttcactaatgttcaaaaaggaaaaatatattgttgaaattc

actacacctgaagggttatgaagcaacaagcaaacatactacagcgaatactgaaaaagac
tctgatgggttaatcgcaaatatcgatgttactcaagatgatatgtcaatcgatgctggt
ttcttcccgtagaaaaactggaatcctcagccagagccgaaaaaccctgatgatagagag
aaaccggcacctgagcaacctgatgtacctcagccagaaccgaaaaaccctgatgataga
gagaaaccggcacctgagcaacctgatgtacctcagccagaaccgaaaaatcctgatgat
agagagaaaccggcacctgagcaacctgatgtacctcaaccagagccgaaaaatcctgat
gataaagagaaaccggcacctgagcaacctgatgtacctcaaccagagccgaaaaatcct
gatgataaagagaaaccggcacctgagcaacctgatgtacctcaaccagagccgaaaaatcct
ccaggtgaaaaggtgaaacccaaaccaactcatcccggtgaagctatgcaaacaacacct
caggacaaatcaacatctcaaacagatgaagcacttcctaaacaggtgaatcatcatca
caatcatctgctttaatcttcggtgggtttactcagtctattaggacttggtttattacgt
cgatcatctaaacaaaaccggttcttcaatgaaataa

The protein sequence translated from SEQ ID NO 21 is designated SEQ ID NO: 22
and is shown below:

SEQ ID NO: 22

VITNKNIYSI RKHKLGVASF LLGTLFVVGH ANNAEASEVS ATTQEHNVET
EQTKTEGELT TEVAQQAVSE SAPIAENMQK TTSVASENAK EVTASDSTQE
VTKTEAKDTA TMKDSEIAQP VSEVNKPVTQ TAAPVAEPST ANKQTSRQV
QELTAPMDTK VINVENGTDV TSKVKVEKSS ITGHQNKDKT YHQSNTVNP
KAERVTNLND WSFENGKAG DYDFQLSDN VDTNGISTIK KVPHIMDSQN
SEQIIAYGEI NENNRVRYRF MDYVNQKENL KGKLSLNLFI KPDKVQDEGK
ITVTSQLGKE MTSQEFDIKY IDGVKSPSGI TLNGRLDELS KADQSFTHYS
IFKPKHNNLT NVTLRGT VSN NAQQNEKNGQ VNVY EYIGQG ELPQSAYANV
NDTKQFNDIT KSMKSIKNS NGYEITFDMN KDNHPYIIVY QGHFNNAKD
FDFSTNATGY QNLNQSEYSY YWPYNYSFNL TWDNGVAFYS NNASGEGNDK
PVPPTYGYSP TVNTIQDTHA DYPVMTFQQP GTLEETEDSM PITTLTESGE
DRGENTSPII ETTEDSQPVE FEEETNHGIQ DVTLHADAVD FEEETNHGEQ
DTVHSDVVE YDEDTTGM L TGAISDHTTE EGTMEYTTDG LLIEFDDEM N
PNVSGQYDDI TTD TIESSH IDTFTELESE FGQHDGIVTF EEDTIVEKPK
TEKGNRVPLV IDLSTPKHNH QFNIQPTDPN IDTSATYRIG NFVWRDEDHN
GVQNDGEHGL EGVLVTLKTA DGVVLNTTTS DANGHYQFTN VQKGKYIVEF
TTPEGYEATS KHTTANTEKD SDGLIANIDV TQDDMSIDAG FFPLENWN PQ
PEPKNPDDRE KPAPEQPDVP QPEPKNPDDR EKPAPEQPDV PQPEPKNPDD
REKPAPEQPD VPQPEPKNP DKEKPAPEQP DVPQPEPKNP DDKEKPAPEQ
PDAPQPKPML PGEKVKPKPT HPGEAMQ TTP QDKSTSQTDE ALPKTGESS
QSSALIFGGL LSLGLGLLR RSSKQNRSSM K

SEQ ID NO: 23

atggcatttgatgggtatgtttacaagaaaaatggtagaagatttacaatttctcgtttct
 gggcgatttcataaaatcaatcaaccggaaaacgatacaatcatcatggttataagacag
 caacgccaaaatcatcaattgttgtgtcgattcaccgaattttgcacggattcacctc
 actacaaaaaatatgataatccatttgaaccgccgatgtttgcgcgcgtctttcgtaaa
 catttagaaggtggacgtatccttgccattcgccaaatcggaatgaccgtcgcacgaa
 atggacgtggaaagtaagatgaaattggtgacacgattcatcgtagtgatttttagaa
 attatgggcaaacatagtaatctcatttctcgttaatgaagaacgtaaaattttagaaggt
 tttaaacaccttacaccaaatacgaatcaatttagaacggatgccaggttttcaatat
 gaagtgcgcgaacacacataaacagaaacccttatgcatatactggtgcgcaagtgtc
 caacatatgtatttcaatgcgggcaaaattgatcgccaaactgcttcaaacgtttgaaggt
 ttttcaccgttaatcacaaaagaaatcacatcaagacgccattttatgaccacacaaact
 ttacctgaagcttttgacgaagtgtgcccgaacgaaagcgacaccccaaccggtattt
 cataaaaaataacgaaacaggtaaagaagacttttattttatgaagttacatcagttttac
 gatgattgcgtcacatatgattcactccatgaactgctcgaccgtttttatgatgcacgc
 ggtgaacgtgaacgcgtcaaacaacgtgcaaacgatttagtcaaactcgccaacaatta
 cttcaaaaatatcaaaaataaattaagtaagctcgtcgatgaacaagcggggactgaagaa
 aaagaaaatcaacaattgtacggcgagttaatcacagcgaatatttatcaactcaaact
 ggagatcgccaggttagaaacagtaaaattattatacaggagaaaacgtgactattccgtta
 aatccacaaaagtcacctgctgaaaatgcgcaatactattacaagcaatacaaccgaatg
 aaaacacgtgagcgcgaattgacctcatcaaatctttaacggaagaaaatatcgcttat
 tttgaaaatatcgagcaacagttgtcacacattcaagttcatgaaattgacgatattcgt
 gaagaactagcagaacaaggctttatcaaaacaaaagaaacagcagaaaaagaaaaagcaa
 caaaaaatccagttacaatcctacgtttcgactgatggcgatacgatttttagtcggtaaa
 aataataagcaaaatgattatttaacgaataaacgtgcgcaaaaatcgatttatggttc
 cataaaaagatatcccaggaagccatgtcgtgatttttaaatgatgcgccaagtgcacaaa
 acgattgaagaagcggcgatgattgcagcgtacttttcaaaggcggggcaatcgggacaa
 attccagtggtattatacaacaattcgcaatgtgcataagccgagtggcagtaaacctgga
 tttgtaacgtacgataaccagaagacgctttacgcaacgccgattatgacatgattcgt
 cgattgaaagctgaagaagcgtaa

The protein sequence translated from SEQ ID NO 23 is designated SEQ ID NO: 24
 and is shown below:

SEQ ID NO: 24

MAFDGMFTRK MVEDLQFLVS GRIHKINQPE NDTIIMVIRQ QRQNHQLLLS
 IHPNFARIHL TTKKYDNPFE PPMFARVFRK HLEGGRILAI RQIGNDRRIE
 MDVESKDEIG DTIHRTVILE IMGKHSNLIL VNEERKILEG FKHLTPNTNQ
 FRTVMPGFQY EVPPTQHKQN PYAYTGAQVL QHIDFNAGKI DRQLLQTFEG
 FSPLITKEIT SRRHFMTTQT LPEAFDEVMA ETKATPQPVF HKNNETGKED
 FYFMKLHQFY DDCVTYDSLH ELLDRFYDAR GERERVKQRA NDLVKLVQQL

LQKYQNKLSK LVDEQAGTEE KENQQLYGEL ITANIYQLKP GDRQLETVNY
YTGENVTIPL NPQKSPAENA QYYYKQYNRM KTRERELTHQ ITLTEENIAY
FENIEQQLSH IQVHEIDDIR EELAEQGFIL QKKQKKKKKQ QKIQLQSYVS
TDGDTILVGK NNKQNDYLTN KRAQKSHLWF HTKDIPGSHV VILNDAPSDK
TIEEAAMIAA YFSKAGQSGQ IPVDYTTIRN VHKPSGSKPG FVTYDNQKTL
YATPDYDMIR RLKAEAA

SEQ ID NO: 25

atgggtcaaaaaatttgggtataaaacacctaacaatcggtgcacttactttggctggaact
gcattttctgcacaccaagccaatgccgctgaacaagttgcacctgaaaaaacactacg
aatgtacttgatgatcaatacgcattaaaacaagctgatgatgcgaaacaaacgcacaaa
ggaacaacacttgtaggttcaaaagaatacaaggatccttcacaaattgatacgactcaa
gtcgatacagcagcacaactgaaacgcccgtagaaggagggaacaagacgcacacaaa
cctactacaactgatgaagcgacatcaacagatcatactgtatcaaaaggtacaaacgaa
agtgcatacctgcaacagcttctatagatgaaggaaacattaaacgcacaagtcaattca
gatgaaacggctactaacggtacacaagacgctcactgaaaatgtgacaaaatatccttat
cattcaagtgaatcgatacacatgaagacgcaactgtgtcaccagatacatatcatgca
ctggacacgcatgcgcaacaaccttcagcaatggatgtaagcgattcaacatcagcaca
actgaagcgacgcaagtaaatagctcaacaaatgtaaatgacaaagaggccgtttcgaca
acagaagatgcacctactacacaacttcaagcagctgtacaatctgaagccaacaaagaa
gcgaaggcaactactgaaacagctcaaaaataaaacacctaagttgaaaagaaagcaaca
gcaactcaaaaatacagcagcttagcaacggggcatcaggatattactgacaaagtctca
aaacgcgttagcagtgacaaatgaaacgaaagcggtgccacaacagcgaaaacacaagca
cctacttcagtgacacatcaagctgatacacaagcaaaaacgataacagacaagaaggca
acaacttacagtgacaaaaccgcaactgaccaagacataaatgcgaatccggacgggtcca
acacctccacgcgttggcggtaaaggggtccccctgcttcactttcactccaatcgact
gggtcaaacagcattccgttcagctgtcgctagtaaacgagtgcatatcaacctaaagtg
aaatcgctctattaatgactatattcgtaagcaaaactacaaagtgccgtgtatatgaagaa
gattattcaagttacttccctaaatacgggttatcgtaatgggtgcggtaaacctgagggc
atcatcgatgatgatacagcaaatgacaactctacaattgatggcgaaatcagttacatg
aaaagaaattatcaaaatgctttcgatcatggctttattaatgggtcaacgtattgttgaa
acgcaacctacagattatttagcatgggggtgcaggtgcgattgcgaatgaacgctttatt
catatcgaaactcggtcatgttcacagtaaaagaagatttcgcacgtcaaatgaacaatatg
gcagattatggggcgacgaacttacaatatattatggcctttctccagatagtgcggaatat
gatggctcggtgggacagtttggacacatgatgctgtttctagatttttaggtggtacagac
cataccgatccgcacggctatttaaaacaacatgggtatttcctttgatgcgttgatgat
ttaatcaatgaaaaatatcaagtgaaaatgggttatgcctcacctgctaactcgtcttca
aaaccatcaacaaatactggcttaacagttaaaaacacaacaggtttcgccgtatttaac
acaacaaatagcgggttatatacgacgggttatgatcaaaaaggtaagcgacgaatcaa
acgaatcaaacgttaaaagttacaaaagaagcgacgttaaatggcaacaaattctattta
atgagtgatgcaaaatctaatacaaacactcggttgggtcaaatcaaacgacgcaacatat

caagctgccaagctgagaaaaagtaacgaaaacgtataactgtcaaaccaggaacaaca
gtatatcaagtgccttgggtgcctcatctcaaacagtaggcaaagctccaggtacgtca
aaccaatcattcaaatacaagaaagaaactggtgcgaaaacgaaatggctttatggg
acagttggcaaagtgcaggtggattaatgcaagtagtgtttagcaaagatgacaaaa
ccatcgacgaataccgcactaaaagtaacaactgacactgggtctcggtcgcattaaagac
aaaaatagtggtttatagcaacgggtatatgataaaactggtaaaagcacttcagccact
aaccaaacattaaaagtaacgaaaaaagcaagtgatcaatggccaatcattctatttagta
tcagattatgctaaaggtacaaatggtgggtgggtgaaacagtcagatgtcgaatatcaa
acaagtaaagcccccttctaaagtgaatcaaaattatacagattaaatcgggtgcgaaattg
tatcaagtgccttgggtgacaagtaacaagttgccgggtacagtgacaggtgctgcgaca
caaacatttaaggcaacacaatctcaaactgtaggtaaagcaacatacttgtatgggaca
gttggcaaattatctggttggttaattcaacagcatttagcagctcaaaaaacaacaacg
aatgttactaaaacaatttctcaaactcgggtcaactgaacacgaaaaatagcgggtgtcaaa
gcttctattttatgacaaaacagcaaaagatgcatccaatgggcaggtcaaacttataaa
attactaaaacagcttctgccataacgaagactatgtattactgcaaaatagtacagga
ggcacgccactcgggttggttcaatgttaaagacgtcacacacgcaacttaggtgctgaa
acagctgttaaagggcggtacactgttaatatgtaaaacatctggactctacgctatgcct
tgggtgacaacgaagcaacgtgtcgatacattaaaaaatgccacaagtcggtttatttaca
gcttcaaaatcagttaaagtcggtaatgatacattcttattcgggtacagtgaaatcaaaaa
ttgggtggtgattaatcaaaaagacttaacagctgtagcagcaaaagttgcaaacatgaaa
actgcatcgaatagcgcagtcaaaaggtgccgcaatcacaaactttgaaaaagtagaagat
tatgtgattacgaataaaaaatggttattattacactaaagttggagattcaaaaacagct
ggtgctttaaaagggtttttatcaacaaatttttaaagtcgaaaaaacatctttactgaac
ggcattacttgggtactatggcgcatctcaaaaacgggacgaaaggatggattaaagcagct
gacatacgttcatcattcattcaacatactgcgggtcagtagcacattgaaagcagcactc
gataaacaatggcgctgacttaccgcctcaagttcaacgtgtagccggtaaatgggtc
aatgcgaatcgtgcagaaactgaaaaagcaatgaataccgcagcaattgaaaaagatccg
actctcattttaccaattttttaaacttgataaataccaaggtcttggcgtagaagaactt
aataaattgttaagaggcaaggcattttagaagggtcaagggtgccgcatttaagaagcc
gcacaaaaacacaatatattaatgaggtttacttaatgtctcacgcatttttagaaacaggt
aacgggacttctcaattagccaatggcggtcacgtagataaaaaataataaagtcgtaaca
aacggtaaacggaagtattacaacatgttcgggtatcggggcaattgatacagacgcttta
cgcaatggctttaaaactgctgaaaaatatggttgaatacgggtcagcaaagcgattatc
ggtggcgcaaaattcatccgtgatcagtagacatcggttcaggacaaaacacattgtatcgt
atgcgttggaatccagaacacctgccacacatcagtagatgcgactgatattaattgggca
aatgtaaacgcacaacgcagatgaaatatttctatgatcaaattgggtgaaacagggtaaatat
ttcgacgtcgatgtatataagaagtag

The protein sequence translated from SEQ ID NO 25 is designated SEQ ID NO: 26
and is shown below:

SEQ ID NO: 26

MVKKFGYKTP TIVALTLAGT AFSAHQANAA EQVAPEKTPT NVLDDQYALK
 QADDAKQTTQ GTTLAGSKEY KDPSQIDTTQ VDTAAQTETP VEGGQQDAQQ
 PTTTDEATST DHTVSKGTNE SASPATASID EGTLNAQVNS DETATNRTQD
 VTENVTKYPY HSSEIDTHED ATVSPDTHA LDTHAQQPSA MDVSDSTSAQ
 TEATQVNTST NVNDKEAVST TEDAPTTQLQ AAVQSEANKE AKATTETAQN
 KTPQVEKKAT ATQNTAQLAT GHQDITDKVS KRVAVTNETK ADATTAKTQA
 PTSVTHQADT QAKTITDKKA TTYSAQATD QDINANPDGP TPPRVGGKGG
 PPASLSLQST GQTAFRSABA SKPSAYQPKV KSSINDYIRK QNYKVPVYEE
 DYSSYFPKYG YRNGVGKPEG IIVHDTANDN STIDGEISYM KRNYQNAFVH
 GFINGQRIVE TQPTDYLAWG AGAIANERFI HIELVHVHVK EDFARQMNMN
 ADYAATNLQY YGLSPDSAAY DGRGTVWTHD AVSRFLGGTD HTDPHGYLKQ
 HGYSFDALYD LINEKYQVKM GYASPANSSS KPSTNTGLTV KNTTGFRIN
 TTNSGLYTTV YDQKGKATNQ TNQTLKVTKE ATLNKNKFYL MSDAKSNQTL
 GWVKSNDATY QAAQAEKKVT KTYTVKPGTT VYQVPWGASS QTVGKAPGTS
 NQSFKSTKEQ TVAKTKWLYG TVGKVTGWIN ASSVVANDQK PSTNTALKVT
 TDTGLGRIKD KNSGLYATVY DKTGKSTSAT NQTLKVTKKA SVNGQSFYLV
 SDYAKGTNVG WVKQSDVEYQ TSKAPSKVNQ NYTIKSGAKL YQVPWGTSKQ
 VAGTVTGAAT QTFKATQSQT VGKATYLYGT VGKLSGWINS TALAAQKTTT
 NVTKTISQIG QLNTKNSGVK ASIYDKTAKD ASKWAGQTYK ITKTASANNE
 DYVLLQNSTG GTPLGWFNVK DVTTRNLGAE TAVKGRYTVN SKTSGLYAMP
 WGTTKQRVDT LKNATSRLFT ASKSVKVGND TFLFGTVNQK LGWINQKDLT
 AVAAKVANMK TASNSAVKGA AITTLKKVED YVITNKNYY YTKVGDSTKA
 GALKGFYQQI FKVEKTSLLN GITWYYGAFQ NGTKGWIKAA DIRSSFIQHT
 AVSSTLKAAL DKQMALTYPP QVQRVAGKWV NANRAETKA MNTAAIEKDP
 TLIYQFLKLD KYQGLGVEEL NKLLRGKGIL EGQGAAFKEA AQKHNINEVY
 LMSHAFLETG NGTSQLANGG HVDKNNKVVT NGKPKYYNMF GIGAITD DAL
 RNFKTAKEY GWNTVSKAII GGAKFIRDQY IGSQNTLYR MRWNPEHPAT
 HQYATDINWA NVNAQRMKYF YDQIGETGKY FDVDVYKK

SEQ ID NO: 27

gtgtcgacagaaaaacaagatgatacacaagcaaaagcgaatgcactttctacagatgat
 tcaacacctacaacagaacaatcaaaaagtgataccgaaccaacgcaaaatcaagaagtg
 aatgaaaaagaagcaacacaagttgagcaaaactccagataatgcatcatcagaatttaa
 gacagtgcagcacaagatgaaacaacatcgaaagacgctgacattgctcaaacaaaagaa
 gcaaaaaatgaagcattgcaaagtgactcatcagcaaacctatcaaatacaagaagcagaa
 aaagaaaacacaactaacagtgaatctcaagtaaatgaacaacctaagcagatacaact
 tctgattcacaagtttcaaatacacctcaacaagatcctacatcgacagtaccttcacca
 gaaacatcagaagacaatcgaccttcaacagaattaaaaaatagtgaaacaactgcttct

caaacaactttaaacgaacaacactactgaatcaacatccaatcaaactgaaacgacaaaa
gcaccaacaaatacaacagtcgcaaacaaaaagcacctgcacaattaaaagacattaaa
ggtacaactcaacttcgcgcagtcagtgcaagtcacactactgctggtgcagctggtggg
acaaacgtaaatgacaaagtaacagcatcaaatatgaaaataactgaatcttatatcgag
ccaaacaactcaggaaacttttattttaaaaagtaactttaacgtaaacgggactgttaaa
gaaggtgactactttactgtaaaaatgcctgacactgtcaatacttttggtgacacgcgc
cattcacctgacttttagagaaaaattacaaatcaaaaaggtgaagttgtggctttaggt
gaatatgatggtgccaccatactatgacatacacggttcactaatgtcgtaataattta
gaaaatgtgtccggttcgtttaacttgactcaatttatggatcgtaaaagtggcaacagat
tctcaaacatatccattaaaaatcgcacattgcaggcgaatctttagatacaciaaattaaa
gtgaattacggtcaatattacagtggaaggtgatttctaacttaaaatcaatgatcacttca
gaagatcctaaaactggggaatatgatcaatacatttatgtcaaccattacaaaaaacg
gcaaacggtacagttgtaagagttcaagggttccaagttgatccaactaagagtaatggg
caagtgaaaccagatacaacgcagatcaagattttaaaagttgctgatggtcaaccactt
aatagtagtttcggtgtgaatgacagtggaatatgaagatgtcacaaaacaatttaatat
gtttatcgtgataataatttggcagatatttactttggaaacttaaatgggcaacgctat
atcgttaagtgacgagcaaagaaaatttggattctaaagaggatttaaacttgcggtgct
attatggccactcaaaaccgatatggtcaatataactatattacttgggataacgatatt
gtgaaaagctcttctggtggtacagccgacggaaatgaagcatcatatcaattaggcgac
aaagtttggaatgatgtgaataaaaaatggtatccaagatcaaggtgaaactggtattgct
gatgtaaaggttactttaaaagatcttgatggcaacattttggatacaacttatacaaac
acgaatggtaaatatatctttgataattttaaaaaatggttaattatcaagtgggttttgaa
acaccggaaggctatgctgcaagtcacccaaccaaggtaatgacgccttgactctgat
ggtcctacaaatgtacaagctgtcattagtgatgggaacaacttaactatcgaccaaggt
ttttaccaaactgaaacaccaacacacacacgctcggcgacaaagtttggaagacttaaat
aaagatggcatccaagacaaaaatgaaccaggtatcgctaacgttaaggtcactttaaaa
gacgcggatggtaacggttggtgatacacgtacgactgatgataaagggaattacttatc
gaaaaagttaaagaaggcgaatatacaattgaatttgaaacgcctgaaggttatacaccg
acacaaacaggccaaggcagagtcagcactgactctaattgggacatcttccttatattta
gtcgaaggtaacgatgacttaacaatcgatagcggtttctacaaagaacctgttacacac
aaagttggcgacaaagtttgggatgacttaataaaagacgggtatccaagatgacaatgaa
ccaggcatctctgacgttaaagtcactttaaaagatgcggatggtaacgtcgtagataca
cgtacaactgatgctaacggtaactatttatgtgaaaacgtgaaagaaggcgactatacg
attgaatttgaaacgcctgaaggttacacaccgactgttacaggtcaaggtacagctgat
aatgactctaacggtacatctacaaaagttacagttaaagatggcgatgacttaacaatt
gacagtgggtttcactcaagttacacctgagccaccgacacataatgttggcgacaaagtt
tgggatgacttaataaaagacgggtatccaagatgacaatgaaccaggcatctctgacgtt
aaagtcactttaaaagatgcggatggtaacgtcgtagatacacgtacaactgatgctaac
ggtaactatttatgtgaaaacgtgaaagaaggcgactatacgattgaatttgaaacgcct
gaaggttacacaccgactgttacaggtcaaggtacagctgataatgactctaacggtaca
tctacaaaagttacagttaaagatggcgatgacttaacaattgacagtgggtttcactcaa

gttacacctgagccaccgactgaacctgaaaaccctagtcagagcaaccttctgaaccg
 ggtcaacctgaaaatcctagtcagagcaaccttctgaaccaggtcaacctgaaaatcct
 agtcagagcaaccttctgaaccaggtcaacctgaaaatcctagtcagagaacaaccttct
 gaaccgggtcaacctgaaaatcctagtcagaaacagccttctgagccaggacaacctaaa
 aatcctagtcagaaacagccaaataatccaagtgtgccaggtgttcaaatcctgaaaaa
 ccaagcttaactccagtcacacaaccgggttcattcaaacggcaataaagcaaaaccatct
 caacaacaaaaagctttacctgaaacaggtgaaactgaatcacatcaaggtacattattc
 ggtgggtatttttagctgcttttaggcgcattactctttgcacgtaaaaaacgccacgataaa
 aaacaatcacactaa

The protein sequence translated from SEQ ID NO 27 is designated SEQ ID NO: 28
 and is shown below:

SEQ ID NO: 28

VSTEQDDTQ AKANALSTDD STPTTEQSKS DTEPTQNQEV NEKEATQVEQ
 TPDNASSEFK DSAAQDETTS KDADIAQTKE AKNEALQSDS SANLSNQEAE
 KENTTNSESQ VNEQPKADTT SDSQVSNTPO QDPTSTVPSP ETSEDNRPS
 ELKNSETTAS QTTLNEQPTE STSNQTETTK APTNTTVANK KAPAQLKDIK
 GTTQLRAVSA SQPTAVAAGG TNVNDKVTAS NMKITESYIE PNNSGNFYLK
 SNFNVNGTVK EGDYFTVKMP DTVNTFGDTR HSPDFREKIT NQKGEVVALG
 EYDVANHTMT YTFNTVVNNL ENVSGSFNLT QFMDRKVATD SQTYPKYDI
 AGESLDTQIK VNYGQYYSEG DSNLKSMTS EDPKTGEYDQ YIYVNPLQKT
 ANGTVVRVQG FQVDPTKSNG QVKPDTTQIK ILKVADGQPL NSSFGVNDSE
 YEDVTQFNI VYRDNNLADI YFGNLNGQRY IVKVTSKENL DSKEDLNLRA
 IMATQNRYGQ YNYITWDNDI VKSSSGGTAD GNEASYQLGD KWNVDVNKNG
 IQDQGETGIA DVKVTCLKDLG GNILDTTYTN TNGKYIFDNL KNGNYQVGFE
 TPEGYAASPS NQGNDAALDS GPTNVQAVIS DGNNLTIDQG FYQTETPTHN
 VGDKVWEDLN KDGIQDQNEP GIANVKVTLK DADGNVVDTR TTDDKGNLYF
 EKVKEGEYTI EFETPEGYTP TQTGQGRVST DNGTSSLIL VEGNDDLITD
 SGFYKEPVTH KVGDKVWDDL NKDGIQDDNE PGISDVKVTI KDADGNVVD
 RTTDANGNYL FENVKEGDYT IEFETPEGYT PVTGQGTAD NDSNGTSTKV
 TVKDGDDLT I DSGFTQVTPE PPTHNVGDKV WDDLNKDGIQ DDNEPGISDV
 KVTCLKDADGN VVDTRTTDAN GNYLFENVKE GDYTI EFETP EGYTPTVTGQ
 GTADNDSNGT STKVTVKDGD DLTIDSGFTQ VTPEPPTEPE NPSPEQPSEP
 GQPENPSPEQ PSEPGQPENP SPEQPSEPGQ PENPSPEQPS EPGQPENPSP
 EQPSEPGQPK NPSPEQPNNP SVPGVQNPEK PSLTPVTQPV HSNGNKAKPS
 QQQKALPETG ETESHQGTLF GGILAAALGAL LFARKKRHDK KQSH

SEQ ID NO: 29

atgaagaaaacaatttcagtacttgggtctagggctattagcaacattttttgtaagtaac

gaatcatatgccgcagaaacgattcaaaacaatacgtcatcaagtgaacgaatcaaaat
 tcagatcagacgccgttagatcattatattcgaaaagcagatggcacactgggtgaaccg
 aacgtgtaccacataaagattatgtagagaatgaaggacctttaccagagtttaattt
 caagttgactctaagaaagattcatctgatccaaatcaagcacggttagatcattatatt
 cgaaaagcggatggcacggttggttgaaccgaatgtatatccacacaaagattatgtcgaa
 aatgaagggcctttaccagagtttaattttatgtatgctgacaaacaaaatcatcatgac
 caacagagtaaaaaacaacaaggataagcagcgtgcaaattacagtgacaaaaagcataat
 gatcagccgggtcatccaaaagcagtcacgccagctgtacaacatgataaagcagtcact
 tcaaacgctactgtaaaagcattgccaacacaggtgaatctgataaaacaacacaatta
 ccaatcgtattatcattgttatctgtggggatttttagttttattaaaattgagaaaataa

The protein sequence translated from SEQ ID NO 29 is designated SEQ ID NO: 30
 and is shown below:

SEQ ID NO: 30

MKKTISVLGL GLLATFFVSN ESYAAETIQN NTSSETNQ N SDQTPLDHYI
 RKADGTLVEP NVYPHKDYVE NEGPLPEFKF QVDSKKDSSD PNQAPLDHYI
 RKADGTLVEP NVYPHKDYVE NEGPLPEFKF MYADKQNHHD QQSKNNKDKQ
 RANYSDDKKHN DQPGHPKAVT PAVQHDKAVT SNATVKALPN TGESDKTTQL
 PIVLSLLSVG ILVLLKLRK

SEQ ID NO: 31

atgaaaagtaaatatgattttttacctaatagacttaataaattttctatacgaaaattt
 actgttggttagtgatcagtgctaataaggagccactttattattcgggttgtagaagga
 gaagcatcagcatcagtaaaagaaggtcaacaaagtataaattctagtgagaaagaaagc
 gccgatcctacagtagttgatttaattagtaagaaagaaacaaatttagatggactagat
 gtatcaagagaagaaacgaccaaagtaaccaataaatgaaaacaaaagaggtgaggaacaa
 agtattttctgataaagctataacagaaaaagctgatacaccagtaagcaatttatcaagt
 aaggaagttgaggagcaaggtgtttctgataaagctataacagaaaaagctgatacacca
 gtaaccaatttatcaagtaaggaagctaaggagcaaggtgcttctgatagagttataaca
 gaaaaagctgatacaccagtaagcaatttatcaagtaaggaagctaaggagcaaggtgct
 tctgatagagttataacagaaaaagctgatacaccagtaagcaatttatcaagtaaggaa
 gttgaggagcaaggtgtttctgataaagctatagagaaaatagctgatgcatcagctact
 gatttgtcaagtaaggaagaagtagaacaagatatctacacaaggtaaaagtaaaatca
 aaggaagcagtacaagtagaaagtagtcagttacaaaatttaaatagtgaaataaatgct
 gaacctaatgaaattaaggcaatagatagaagttcaatattacctttaaatttaaattgat
 gaagaaaataacaaaaaagttaataaagggactcgggttcagaagctacattaagaaat
 gcctctaataaccaactcaatacacgaatgagatcagtgagtttatttagagttgctaga
 ctaacagaaatcaatagaaatgttaatgataaagtaaaaggtttcgatatcgacatcgca
 atagccccaccgcataactaaccctaaaactggaaaaagaagaattttgggcgacatcttct

tcagttttaaagttaaaggcaagctatgaattggataatagcattttctaaaggggatcaa
tttactattcaatttgggtcaaaatattcggtccaggtggattaaatttaccaagaccttat
aatttttttatatgataaggataaaaaattagttgcaactggccgttacaataaagaatca
aatacaatcacatatacattttacggattatgtagataaacatcaaaacattaaaggtagt
tttgagatgaatgcattttctagaaaggaaaatgctactactgacaaaacagcatatcca
atggatgttactattgccaatcaaaaatatagtgaataattattgtagactatggtaat
aaaaagaatgctgctatttttcaagtacagaatatattgatttagatggtagtagaaaa
atgacaacatatattaatcaaaatggtagtaaaaattccatctatcggtgctgatatgcaa
attgatttgaacggttataaatttgatccatccaaaaacaatttttaaatttatgaagtg
gaaaatagcagtgactttgtggatagcttttaccagatgtgagcaagttaagggtggt
acgagtcaatttaatttcaatatatacaataataatacaatggcaaaagtggattttgggt
actaacctttggaggggtaaaaaatatattattcagcaagtggcgaatatagacgacagt
aaattagtgaaaaatgcttcaatcaattatacattgaataaaatggatttttaataataaa
agaacggtagaaacacataacaatacttattctacagtgaagataaatcaacagcacta
ggtgacgtacaggaaagtcaatctattagtgagagccaatcagttagtgaaagcgagtca
ctaagtgagagccaatcaatcagtgaaagcgaatcattaagtgagagccaatcaatcagt
gaaagcgaatcattaagtgaaagtcaatcaatctcagagagcgaatcactaagtgaaagt
cagtcaatttcagaaagcgaatcattaagtgaaagccaatcaatctcagagagtgaatca
ttaagtgaaagtcaagtcaatttcagagagtgaatcactaagtgaaagtcaagtcaatttca
gaaagcgaatcattaagcgagagtcaagtcaatttcagaaagcgaatcattaagcgagagt
cagtcaatttcagaaagcgaatcattaagtgaaagccaatcaatcagtgaaagcgaatca
ctaagcgagagccaatcaatctcagagagtgaatcattaagcgagagtcaatcaatctca
gagagcgaatcattaagtgagagtcaatcaatcagtgaaagcgagtcaactaagtgagagt
caatcaatttcagagagcgaatcattaagtgaaagccaatcaatctcagagagtgaatca
ctaagtgagagccaatcaatctcagagagtgaatcattaagtgagagccaatcaatctca
gagagcgagtcaactaagcgagagccaatcaatttcagagagtgaatcactaagtgaaagt
caatcaatttcagagagcgaatcactaagtgagagccaatcaatctcagagagcgaatca
ctaagtgaaagtcaatcaatttcagagagtgaatcactaagcgagagccaatcaatctca
gagagtgaatcattaagtgaaagtcaagtcaatttcagagagtgaatcactaagtgaaagt
cagtcaatttcagaaagcgaatcattaagtgaaagccaatcaatcagtgaaagcgaatca
ctaagcgagagtcaatcaatctcagagagcgaatcattaagtgaaagtcaatcaatttca
gaaagcgagtcaattaagcgagagtcaagtcaatctcagagagcgaatcactaagcgagagt
caatcaatctcagagagtgaatcattaagtgagagccaatcagttagtgaaagcgaatca
ctaagtgaaagtcaagtcaatttcagaaagcgaatcattaagtgagagtcaatcaatttca
gaaagcgaatcattaagtgaaagccaatcaatcagtgaaagcgaatcactaagcgagagc
caatcaatcagtgaaagcgaatcattaagtgagagtcaatcaatctcagaaagcgaatca
ttaagtgagagtcaatcaatcagtgaaagcgaatcactaagcgagagccaatcaatctca
gagagcgaatcactaagcgagagccaatcaatctcagagagcgagtcaactaagcgagagc
caatcaatcagtgaaagcgaatcattaagtgagagtcaatcaatcagtgaaagcgagtca
ctaagtgagagccaatcaatctcagagagtgaatcattgagtgaagagccaatcaatctca
gagagcgagtcaactaagtgagagtcaatcaatttcagagagcgaatcattaagtgaaagc

caatcaatctcagagagtggaatcattgagtgagagccaatcagttagtgaaagcgagtgca
ctaagtgagagtgcaatcaatcagtgaaagcgagtgactaagtgagagtgcaatcaatttca
gagagcgaatcattaagcgagagtgagtgcaatctcagagagtggaatcactaagtgagagc
caatcaatctcagagagtggaatcattaagtgagagccaatcaatctcagagagtggaatca
ctaagtgagagtgcaatcaatcagtgaaagcgaatcactaagcgagagccaatcaatttca
gagagtggaatcattaagtgagagccaatcagttagtgaaagcgaatcactaagcgagagc
caatcaatctcagagagcgaatcattgagtgagagccaatcaatctcagagagtggaatca
ttgagtgagagtgcaatcaatcagtgaaagcgaatcactaagcgaaagtcaatcaatttca
gagagtggaatcattgagtgagagccaatcaatttcagagagtggaatcactaagtgaaagt
cagtgcaatttcagaaagcgaatcactaagcgagagccaatcaatctcagagagcgaatca
ctaagtgaaagtcagtgcaatttcagaaagcgaatcattaagtgaaagccaatcaatctca
gagagtggaatcattaagtgaaagtcagtgcaatttcagagagtggaatcactaagtgaaagt
cagtgcaatttcagaaagcgaatcattaagcgagagtgagtgcaatttcagaaagcgaatca
ttaagtgaaagccaatcaatcagtgaaagcgaatcactaagcgagagccaatcaatctca
gagagcgaatcactaagcgagagccaatcaatctcagagagcgaatcactaagtgaaagt
caatcaatttcagagagtggaatcattgagtgagagtgcaatcaatttcagagagtggaatca
ctaagtgaaagtcgaatcaatttcagagagtggaatcactaagcgagagccaatcaatctca
gagagtggaatcattaagtgaaagtcagtgcaatttcagagaggggaatcactaagtgaaagt
cagtgcaatttcagaaagcgaatcattaagtgaaagccaatcaatcagtgaaagcgaatca
ctaagtgaaagtcgaatcaatctcagagagtggaatcactaagtgagagccaatcaatctca
gagagtggaatcattgagtgagagccaatcaatctcagagagcgaatcactaagtgaaagt
caatcaatttcagaaagcgagtgattaagcgagagtgagtgcaatctcagagagtggaatca
ctaagtgagagccaatcaatctcagagagtggaatcactaagtgagagtgcaatcaatcagtg
gaaagcgaatcactaagcgagagccaatcaatttcagagagtggaatcattaagtgagagc
caatcagttagtgaaagcgaatcactaagcgagagccaatcaatctcagagagcgagtgca
ctaagcgagagtgcaatcaatctcagagagtggaatcactaagtgaaagtcagtgcaatttca
gaaagcgagtgactaagcgagagtgcaatcaatctcagagagtggaatcattgagtgagagc
caatcaatctcagagagcgaatcattgagtgagagccaatcaatctcagagagtggaatca
ttgagtgagagccaatcaatttcagagagcgaatcactaagcgagagccaatcaatcagtg
gaaagcgaatcattaagtgagagtgagtgcaattagcgaaagcgaatcactaagtgagagtg
caatcaatctcagagagtggaatcactaagtgaaagtcagtgcaatcagcgaaagcgaatct
aaatctttacctaataccggtaactggagaaaagatttctaattatccaggtatttttagga
ggattattaagcatattaggtataagtttgcttaaaagaaaagacagagagaaaaaatta
ggacaaaaatctaataagtag

The protein sequence translated from SEQ ID NO 31 is designated SEQ ID NO: 32
and is shown below:

SEQ ID NO: 32

MKSKYDFLPN RLNKFSIRKF TVGSVSVLIG ATLLFGFVEG EASASVKEGQ
 QSINSSEKES ADPTVVDLIS KKETNLDGLD VSREETTKVP INENKRGEEQ
 SISDKAITEK ADTPVSNLSS KEVEEQGVSD KAITEKADTP VTNLSSKEAK
 EQGASDRVIT EKADTPVSNL SSKEAKEQGA SDRVITEKAD TPVSNLSSKE
 VEEQGVSDKA IEKIADASAT DLSSKEEVEQ DISTQGKVKS KEAVQVESSQ
 LQNLNSEINA EPNEIKAIDR SSILPLNLND EENNKKVNKG TRVPEATLRN
 ASNNQLNTRM RSVSLFRVAR LTEINRNVND KVKVSDIDIA IAPPHTNPKT
 GKEEFWATSS SVLKLKASYE LDNSISKGDQ FTIQFGQNIR PGGLNLPRPY
 NFLYDKDKKL VATGRYNKES NTITYTFTDY VDKHQNIKGS FEMNAFSRKE
 NATTDKTAYP MDVTIANQKY SENIIVDYGN KKNAAIISST EYIDL DGSRK
 MTTYINQNGS KNSIYRADMQ IDLNGYKFDP SKNNFKIYEV ENSSDFVDSF
 SPDVSKLRDV TSQFNIQYTN NNTMAKVDFG TNLWRGKKYI IQQVANIDDS
 KLVKNASINY TLNKMDFNNK RTVETHNNTY STVKDKSTAL GDVQESQSIG
 ESQSVSESES LSESQSISES ESLSESQSIG ESESLSESQS ISESESLSES
 QSISESESLS ESQSISESES LSESQSISES ESLSESQSIG ESESLSESQS
 ISESESLSES QSISESESLS ESQSISESES LSESQSISES ESLSESQSIG
 ESESLSESQS ISESESLSES QSISESESLS ESQSISESES LSESQSISES
 ESLSESQSIG ESESLSESQS ISESESLSES QSISESESLS ESQSISESES
 LSESQSISES ESLSESQSIG ESESLSESQS ISESESLSES QSISESESLS
 ESQSISESES LSESQSISES ESLSESQSIG ESESLSESQS ISESESLSES
 QSISESESLS ESQSVSESES LSESQSISES ESLSESQSIG ESESLSESQS
 ISESESLSES QSISESESLS ESQSISESES LSESQSISES ESLSESQSIG
 ESESLSESQS ISESESLSES QSISESESLS ESQSISESES LSESQSISES
 ESLSESQSIG ESESLSESQS ISESESLSES QSISESESLS ESQSVSESES
 LSESQSISES ESLSESQSIG ESESLSESQS ISESESLSES QSISESESLS
 ESQSISESES LSESQSISES ESLSESQSIG ESESLSESQS VSESESLSES
 QSISESESLS ESQSISESES LSESQSISES ESLSESQSIG ESESLSESQS
 ISESESLSES QSISESESLS ESQSISESES LSESQSISES ESLSESQSIG
 ESESLSESQS ISESESLSES QSISESESLS ESQSISESES LSESQSISES
 ESLSESQSIG ESESLSESQS ISESESLSES QSISESESLS ESQSISESES
 LSESQSISES ESLSESQSIG ESESLSESQS ISERESLSES QSISESESLS
 ESQSISESES LSESQSISES ESLSESQSIG ESESLSESQS ISESESLSES
 QSISESESLS ESQSISESES LSESQSISES ESLSESQSIG ESESLSESQS
 ISESESLSES QSVSESESLS ESQSISESES LSESQSISES ESLSESQSIG
 ESESLSESQS ISESESLSES QSISESESLS ESQSISESES LSESQSISES
 ESLSESQSIG ESESLSESQS ISESESLSES QSISESESLS ESQSISESES
 KSLPNTGTGE KISNYPGILG GLLSILGISL LKRKDREKKL GQKSNK

SEQ ID NO: 33

atgttaagaacaaattataaaactaagaaagcttaaagtaggttttagtatcgacaggtgtg

gcgttgacttttgtgatggcaagtgggaatgcagagggcgtcggagaacgagcagactgaa
gtaaaaggggagggcgcaagttgcttctgtgaatgaaaaagagagtgaagcagaattacct
gtagcgcaacaagaagcatctattcaactagacaaagtacaaccaggcgatgcacagctt
tcaggctatacacagccaaacaaagcgatttctgtaaagatcgacaataaagatattgtg
tctgtagatgatggctatgaagaggtattatcggatgatacaggtaaatttgtatatgat
ttgaaagggcgtaaatgtttacaatcaaaaagttgatgttgaagcgatgacgccattt
aattttgaagattttggatgaatcagcacttgagagcggaagaggcattggaggcggttaggt
caattggaagacgaagaaacagcgacagcttctgtgacgacgcctagatatgaaggtgcg
tatacagttcctgaagaacgcttgacacccattcaaggccaacagcaagtattcatcgaa
cctatttttagaaggggcaagtaaaatcaaaggacatacatctgtacaaggtaaagtcgag
ttagcaatcaatcaagaacatgtgcacctaggtgatacgttagaagaacaagcagcactc
actgatcaagagtggaaggtcggttatgacgggatttggcgccatattgatgatcaaggg
tttttcgagtttgacttgaaccgtctttacaataaatcttaccattgaagtctggcgat
ttagtgaactttatcttttaaatctaataagcgaagtaggcccattattcaatgtgaacgtt
gagcctttcgaacgtgtggcacaagctaaaacaaagtatgagcagaatgacagtcagta
gtcaacaaattggatgatactaaaagtgaacttggaggttcaacctatctatggagacctt
acacaagcagcagtagcatggcgagtcgaaagtgttgataccggggacgtcaaaagttgaa
ggacgtacgaattatgcacatgcatggatagagatggcatctaatttaggggaatatcgt
agtttccctaaattacaagctgatgacgaggtgagtttatatttgatttaaaagcggca
gacatacaattgttaaacgggagaacgtttgacattcagagccgttgaccacatacaaaa
caacagtttagctgaaactacatcagaagtacgcccagtagatatgcaagatgaagagtca
gaggttgtgcagacttcaagcactgagaaatcagcacttgcggatgaaattcttcgttct
atgacaattgacaaatcatttaattcctgaagttaccgagataccgggtcatgtatatcct
aagaaaacagaggataaaggtgctgaaaatacagaacaagcctcagagaattctgagaag
ccatctcagactacagaatctcaaaatgatgccgtacaagatgtagagaaatcctctgtt
aatgaggaggttacgccaccttcaacagaatctgctcaagttgaaaaggggcaaaataca
gaaggggcttctgcttccaaaaaatgtagaacaacatgtagagagtataccatacaaaaa
cgtaaagcgttgataggactgacaaaacatcaaggatcagggcacatgccgccattttct
ttaagctttaataataaagaagatgacgtatccacaaaggttaacgaagcaaacgagcat
gaacgtaagcaggggtacagtttatccagagcaaatagaacaattacctcaaacaggttta
actgaaaaatcgccattctgggcattgttatgtgtatcaggcacaggtttattatta
ttcaaacgttctagacgacaacgccaatcttaa

The protein sequence translated from SEQ ID NO 33 is designated SEQ ID NO: 34
and is shown below:

SEQ ID NO: 34

MLRTNYKLRKLVGLVSTGVALTFVMSGNAEASENEQTEVKGEAQVASV
NEKESEAELPVAQQEASIQLDKVQPGDAQLSGYTQPNKAISVKIDNKDIV
SVDDGYEEVLSDDTGKFVYDLKGRQIVYNQKVDVEAMTPFNFEDFDESAL

ESEEALEALGQLEDEETATASVTTPRYEGAYTVPEERLTPIQGQQQVFIE
PILEGASKIKGHTSVQGKVALAINQEHVHLGDTLEEQAALTDQEWQGRYD
GIWRHIDDQGFFFDLNLRLYNKSYPLKSGDLVTLSEKSNDEVGPLFNVNV
EPFERVAQAKTKYEQNDSPVVNKLDDTKSDLEVQPIYGDLTQAAVHGESK
VLIPGTSKVEGRNTNYAHAWIEMASNLGEYRSFPKLQADATGAFIFDLKAA
DIQLLNGERLTFRAVDPHTKQQLAETTSEVRPVDQMDEESEVVQTSSTEK
SALADEILRSMTIDKSFNPEVTEIPGHVYPKKTEDKGAENTEQASENSEK
PSQTTESQNDQVQDVEKSSVNEEVTTPSTESAQVEKGQNTGALLPKNVE
QHVESIPYQKRKALIGLTKHQGSGHMPPFSLSFNNKEDDVSTKVNEANEH
ERKQGTVYPEQIEQLPQTGLTEKSPFWALLFVVSGTGLLLFKRSRRQRQS

SEQ ID NO: 35

atgaaaactaaatacacagcaaaattattaattggggcagcaacaatatcttttagcaaca
tttatttcacaagggaacgcacatgcgagcgaacaaactacaggactcgacccggcaca
cctgtcaactttgattcaatcaatgtaacgccagaccacaaaaacattctatcaagtctta
catatggaaggcatttcagaagaccaacgtgaacaatatgtgaaacaattgcacgaagac
ccaagtagcgcacaaaaatgttttttcagaatcaattaaagatgccatccacccggaacgt
cgtgttgcgcaacaaaatgcgtttttacagcgtattacacaacgatgacttatccgaagag
caacgtgatgcatacattggtagaattaaagaagatccagatcaaagccaagaagtattt
gttgagtcctttaaatgtggcacctaagcagaatcacatgaagatcgccctcattgaatta
caaaacaaaaatttaattggaagcgaatgaagcacttaaacggttacaacaagaagacagc
attcagaatagacgtgcggtcaacgtgctgtcaacaaattgacgcccgatagcgcgaaac
gcattccaaaaagaattagatcaaataatgccccacgcgacgctaaaattaaagctgac
gctgaagcaaaaaacaagcacctgaagtaagcgcaccacaaattgaagatgcacctact
actgaagttgcaccatctccaaaacaagatatgccaaaagtagataaaaaagaagaagat
aaagtagaaagtgatactgaggtcaaagaagtaacctaaagctgatacagagaaaaaccct
caatctaaagacacttctaaaactgaacaagctaaagaaacacctaaagtagagcaatca
cctaaaacagaaaaggctgaagaagcacctaaagcagaaacacctcaaatggaaataaa
gcacaaactgaagaagctaaaccagaagtaaaagacaatgtgaaaaacactccatctgca
cctgtgttacctgaaacaggaaaagcaacaacttcaacacttgaaagctactggaattct
ttcaaagacagtgtgaataaagggttatacttacattaaacaaagcttagaaagtgggtat
caatattttaaagggtcaatacgactatatcactaaaaatacaatgatgcgaaatactat
acaaaaatgtattcaaatcataagctctacaattgatcagtcgttattagctatattaggt
aaaactggatctagcgcatatatcaagccattaaatatcgaagaaaattcaaacgtattt
tacaagcttatgcaaaaaacaagaaactttgctacagaaagcattaacacaggaaaagta
ttatacacattatatcaaaaccctactgtagttaaatctgctttcactgcaattgaaaca
gcaataacagtaaaaaatgcaataagcaatcttttctctctcttcaaataa

The protein sequence translated from SEQ ID NO 35 is designated SEQ ID NO: 36 and is shown below:

SEQ ID NO: 36

MKTKYTAKLLIGAATISLATFISQGNHASEQTTGLAPAQPVNFDSINVT
PDQKTFYQVLHMEGISEDQREQYLKQLHEDPSSAQNVFSES IKDAIHPER
RVAQQNAFYSVLHNDLSEEQRDAYIGRIKEDPDQSQEVFVESLNVAPKA
ESHEDRLIELQNKNLMEANEALKALQQEDSIQNRRAAQRAVNKLTPDSAN
AFQKELDQINAPRDAKIKADAEAKKQAPEVSAPQIEDAPTTEVAPSPKQD
MPKVDKKEEDKVESDTEVKEVPKADTEKNPQSKDTSKTEQAKETPKVEQS
PKTEKAEEAPKAETPQNGNKAQTEEAKPEVKDNVKNTPSAPVLPETGKAT
TSTLESYWNSFKDSVNKGTYTIKQSLESGYQYLKGQYDYITKKYNDKYY
TKMYSNHKSTIDQSVLAILGKTGSSAYIKPLNIEENSNVFKAYAKTRNF
ATESINTGKVLTYTLYQNPTVVKS AFTAIETANTVKNAISNLFSLFK

Since each of the abovementioned proteins/nucleic acid sequences is derived from *Staphylococcus pseudintermedius*, the inventors have designated these (and the corresponding protein sequences) *Staphylococcus pseudintermedius* surface genes/nucleic acids/proteins (Sps). For simplicity, the bulk of this specification will use the term “Sps” or “Sps genes” or “Sps nucleic acids” which are intended to encompass all of the nucleic acid sequences described above (i.e. SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33 and 35).

Furthermore, in addition to encompassing the entire or complete gene/nucleic sequences listed above, it is to be understood that the designation “Sps” also encompasses fragments, portions, mutants, derivatives and/or homologues/orthologues of any of these genes.

In addition, the term “Sps” or “Sps proteins” encompasses the proteinaceous products of the Sps genes/nucleic acids or fragments, portions, analogues, variants or derivatives thereof (for example short peptide fragments). In particular, the term “Sps proteins” encompasses the sequences given as SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 and 36 above.

Typically, the gene/nucleic acid fragments, portions, mutants, variants, derivatives and/or homologues/orthologues of the invention are functional or active – that is, they retain the function and/or activity of the wild type or native Sps genes/nucleic acids. Advantageously, fragments, portions, mutants, variants,

derivatives and/or homologues/orthologues of any of the Sps genes/nucleic acids provided by this invention, encode proteins (or peptides, peptide fragments) retaining the ability to bind to or associate with extracellular matrix proteins such as, for example, fibrinogen, fibronectin and/or collagen. In other embodiments, the proteins and/or peptides encoded by the nucleic acid sequences described herein are immunogenic or antigenic. Furthermore, fragments, portions, variants or derivatives of any of the proteins encoded by the nucleic acid sequences described herein may also retain the immunogenicity and/or antigenicity of a corresponding wild type Sps protein (for example the proteins listed above). Where the invention relates to immunogenic compositions and/or vaccines, the use of proteins and/or peptides which are immunogenic (or antigenic) is important.

The term “mutants” may encompass naturally occurring mutants or those artificially created by the introduction of one or more nucleic acid additions, deletions, substitutions or inversions.

Homologous or identical genes, nucleic acid or protein sequences may exhibit as little as approximately 20 or 30% sequence homology or identity to certain reference sequences, however, in other cases, homologous or identical genes/nucleic acids and/or proteins may exhibit at least 40, 50, 60, 65, 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99% homology or identity to the various sequences given above as SEQ ID NOS: 1-36. It should be understood that mutant, variant, derivative and/or orthologous sequences may exhibit similar levels of homology/identity to each other and/or to the Sps genes/nucleic acids shown as SEQ ID NOS 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33 and/or 35 above.

One of skill in this field will readily understand that genes/nucleic acids homologous/identical to the Sps genes detailed herein may be found in other bacterial species. As such, homologous genes from other species may be included within the scope of this invention. Using the various nucleic acid and amino acid sequences described herein, one of skill in the art could readily identify related sequences in other microbial (particularly bacterial) species. For example, nucleic acid obtained from a particular bacterial species may be probed using the probes derived from the sequences of this invention, to identify homologous or closely related sequences.

It should be understood that Sps nucleic acid sequences of this invention may be single-stranded or double-stranded and a single-stranded nucleic acid molecule may include a polynucleotide fragment having a nucleotide sequence that is

complementary to a nucleotide sequence that encodes a Sps protein or fragment thereof. As used herein, the term “complementary” refers to the ability of two single stranded polynucleotide fragments to base pair with each other.

A single-stranded nucleic acid molecule of the invention may further include a polynucleotide fragment having a nucleotide sequence that is substantially complementary to a nucleotide sequence that encodes a Sps protein or fragment thereof according to the invention, or to the complement of the nucleotide sequence that encodes said Sps protein or fragment thereof. Substantially complementary polynucleotide fragments can include at least one base pair mismatch, such that at least one nucleotide present on a first polynucleotide fragment will not base pair to at least one nucleotide present on a second polynucleotide fragment, however the two polynucleotide fragments will still have the capacity to hybridize. The present invention therefore encompasses polynucleotide fragments which are substantially complementary. Two polynucleotide fragments are substantially complementary if they hybridize under hybridization conditions exemplified by 2xSSC (SSC: 150 mM NaCl, 15 mM trisodium citrate, pH 7.6) at 55°C. Substantially complementary polynucleotide fragments for purposes of the present invention may preferably share at least about 60, 65, 70, 75, 80 or 85% nucleotide identity, preferably at least about 90%, 95% or 99% nucleotide identity. Locations and levels of nucleotide sequence identity between two nucleotide sequences can be readily determined using, for example, CLUSTALW multiple sequence alignment software.

In addition, it should be understood that the present invention also relates to the products of the genes/nucleic acids encompassed by this invention and in particular to proteins or peptides homologous/identical to those having sequences provided by SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 and 36. Furthermore, fragments, portions, analogues, variants, derivatives of any of these or homologous and/or identical or modified proteins are also within the scope of this invention. Typically, fragments, portions, derivatives, variants and/or homologous or modified proteins or peptides of the invention are functional or active – that is they retain the function of a wild type Sps protein. In certain embodiments fragments, portions, derivatives or variants of, and/or modified sequences or sequences with homology or identity to, the amino acid sequences provided by this invention, retain the ability to bind to or associate with extracellular matrix proteins such as, for example, fibrinogen, fibronectin and/or collagen.

Additionally or alternatively, fragments, portions, mutants, variants, derivatives and/or homologues/orthologues of the Sps genes provided by this invention, may encode proteins (or peptide fragments) that are antigenically similar or identical to the proteins encoded by the genes described herein. Similarly, fragments, portions, derivatives and/or variants of and/or modified sequences or sequences with homology or identity to, the amino acid sequences provided by this invention are also antigenically similar or identical to the proteins encoded by the genes described herein. It should be understood that the term “antigenically similar or identical” may encompass proteins or peptides eliciting an immune response similar or identical to the immune response elicited by any of the Sps proteins described herein. In certain embodiments fragments, portions, derivatives and/or variants of and/or modified sequences or sequences with homology or identity to, the amino acid sequences described herein, elicit immune responses which protect against *Staphylococcus pseudintermedius* infection and/or prevent, reduce or neutralise *Staphylococcus pseudintermedius* cell/tissue adhesion and/or colonisation. One of skill will readily understand that the antigenicity of a polypeptide can be evaluated *in vitro* by, for example, performing a Western blot on the purified polypeptide (for example, an affinity purified polypeptide) using polyclonal antisera from an animal, such as a rabbit that was vaccinated with at least an antigenic portion of an Sps protein of the present invention.

One of skill in this field will readily understand that for the various nucleic acid sequences and polypeptides described herein, natural variations due to, for example, polymorphisms, may exist between Sps genes and proteins isolated from different microbial species and even different strains of the same species. Gene or protein variants may manifest as proteins and/or genes that exhibit one or more amino acid/nucleic acid substitutions, additions, deletions and/or inversions relative to a reference sequence (for example any of the sequences described above). As such, it is to be understood that all such natural variants, especially those that are functional or display the desired activity, are to be included within the scope of this invention.

In another embodiment, the invention relates to derivatives of any of the Sps sequences described herein. The term “derivatives” may encompass Sps genes or peptide sequences which, relative to those described herein, comprise one or more amino acid substitutions, deletions, additions and/or inversions.

Additionally, or alternatively, analogues of the various peptides described herein may be produced by introducing one or more conservative amino acid substitutions into the primary sequence. One of skill in this field will understand that the term “conservative substitution” is intended to embrace the act of replacing one or more amino acids of a protein or peptide with an alternate amino acid with similar properties and which does not substantially alter the physico-chemical properties and/or structure or function of the native (or wild type) protein. Analogues of this type are also encompassed with the scope of this invention. In one embodiment, substitute amino acids may be selected from other members of the class to which the amino acid belongs. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and tyrosine. Polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine and glutamine. The positively charged (basic) amino acids include arginine, lysine and histidine. The negatively charged (acidic) amino acids include aspartic acid and glutamic acid. Examples of preferred conservative substitutions include Lys for Arg and vice versa to maintain a positive charge; Glu for Asp and vice versa to maintain a negative charge; Ser for Thr so that a free $-OH$ is maintained; and Gln for Asn to maintain a free NH_2 .

As is well known in the art, the degeneracy of the genetic code permits substitution of one or more bases in a codon without changing the primary amino acid sequence. Consequently, although the sequences described in this application are known to encode the Sps proteins described herein, the degeneracy of the code may be exploited to yield variant nucleic acid sequences which encode the same primary amino acid sequences.

The present invention may further provide modified Sps proteins. For example, a “modified” Sps protein may be chemically and/or enzymatically derivatised at one or more constituent amino acids, including side chain modifications, backbone modifications, and N- and C-terminal modifications including acetylation, hydroxylation, methylation, amidation, phosphorylation and the attachment of carbohydrate or lipid moieties, cofactors, and the like.

One of skill in this field will appreciate that the amino acid and/or nucleic acid sequences described herein may be used to generate recombinant Sps genes/proteins and as such, the present invention further contemplates methods of generating and/or expressing recombinant Sps genes and/or proteins, and products for use in such

methods. Accordingly, in addition to providing substantially purified or isolated recombinant Sps sequences, a second aspect of this invention provides DNA constructs comprising a replicable expression vector and nucleic acid encoding one or more of the Sps protein(s) described herein.

Expression vectors for the production of the molecules of the invention include plasmids, phagemids, viruses, bacteriophages, integratable DNA fragments, and other vehicles, which enable the integration of DNA fragments into the genome of the host. Expression vectors are typically self-replicating DNA or RNA constructs containing the desired gene or its fragments, and operably linked genetic control elements that are recognised in a suitable host cell to effect expression of the desired genes.

Generally, the genetic control elements can include a prokaryotic promoter system or a eukaryotic promoter expression control system. Such systems typically include a transcriptional promoter, an optional operator to control the onset of transcription, transcription enhancers to elevate the level of RNA expression, a sequence that encodes a suitable ribosome binding site, RNA splice junctions, sequences that terminate transcription and translation and so forth. Expression vectors usually contain an origin of replication that allows the vector to replicate independently of the host cell.

A vector may additionally include appropriate restriction sites, antibiotic resistance or other markers for selection of vector containing cells.

Plasmids are the most commonly used form of vector but other forms of vectors which serve an equivalent function and which are, or become, known in the art are suitable for use herein. See, e.g., Pouwels et al. *Cloning Vectors: a Laboratory Manual* (1985 and supplements), Elsevier, N.Y.; and Rodriquez, et al. (eds.) *Vectors: a Survey of Molecular Cloning Vectors and their Uses*, Butterworth, Boston, Mass (1988).

In general, such vectors may contain specific genes, which are capable of providing phenotypic selection in transformed cells. The use of prokaryotic and eukaryotic viral expression vectors to express the nucleic acid sequences coding for the recombinant proteins of the present invention are also contemplated.

The vector is introduced into a host cell by methods known to those of skill in the art. Introduction of the vector into the host cell can be accomplished by any method that introduces the construct into the cell, including, for example,

electroporation, heat shock, chemical compounds such, for example, calcium phosphate, strontium phosphate, microinjection techniques and/or gene guns. See, e.g., Current Protocols in Molecular Biology, Ausubel, F.M., et al., John Wiley & Sons, N.Y. (1989).

Another aspect relates to a host cell transformed with any one of the nucleic acid constructs of the present invention. Suitable host cells include prokaryote cells, lower eukaryotic and higher eukaryotic cells. Prokaryotes include Gram negative and Gram positive organisms, e.g., *E. coli* and *B. subtilis*. Lower eukaryotes include yeast, *S. cerevisiae* and *Pichia*, and species of the genus *Diclyostelium*.

“Host cell” as used herein refers to cell which can be recombinantly transformed with vectors constructed using recombinant DNA techniques.

A drug resistance or other selectable marker is intended in part to facilitate the selection of the transformants. Additionally, the presence of a selectable marker, such as a drug resistance marker may be of use in keeping contaminating microorganisms from multiplying in the culture medium. Such a pure culture of the transformed host cells would be obtained by culturing the cells under conditions which require the induced phenotype for survival.

PCR techniques may be exploited to selectively obtain Sps gene sequences from samples of Staphylococcal DNA. These amplified sequences may be introduced into any of the vectors described above. In one embodiment, the vector may further comprise a nucleotide sequence of a tag or label to assist in protein purification procedures.

Techniques used to purify recombinant proteins generated in this way are known and, where the recombinant protein is tagged or labelled, these may include the use of, for example, affinity chromatography techniques.

In view of the above, a fourth aspect of this invention provides a process for the production of recombinant Sps protein(s) or peptide(s) of the invention, said process comprising the steps of (a) transforming a host cell with the nucleotide sequence of the invention or transfecting a host cell with a nucleic acid construct of the invention; (b) culturing the cells obtained in (a) under conditions in which expression of the protein takes place; and (c) isolating the expressed recombinant protein or peptide from the cell culture or/and the culture supernatant.

The polypeptide may be partially purified from the host and where the polypeptide is secreted from the host cell, the cells may be separated from the media

by centrifugation, the cells being pelleted. Alternatively, the polypeptide may be partially purified from this supernatant, for example using affinity chromatography.

A fifth aspect of this invention provides monoclonal or polyclonal antibodies, whether derived from rodents, mammals, avians, ungulates, or other organisms, that bind to the Sps proteins described herein. Production and isolation of monoclonal and polyclonal antibodies to a selected polypeptide sequence is routine in the art see for example "Basic methods in Antibody production and characterisation" Howard & Bethell, 2000, Taylor & Francis Ltd. Such antibodies may be used in diagnostic procedures, as well as for passive immunisation.

Staphylococcus pseudintermedius is known to cause cutaneous inflammatory diseases in a variety of animals. One such cutaneous inflammatory disease is canine pyoderma which is a major cause of morbidity in dogs. Pyoderma associated with *Staphylococcus pseudintermedius* infection is common among dogs and is often associated with puritis, alopecia, erythema and swelling. At present, the treatment of this infection is difficult, requiring the use of aggressive, systemically administered antibiotics. The present inventors have discovered that Sps genes (and their protein products) play a role in *Staphylococcus pseudintermedius* colonisation and pathogenesis. As such, the Sps genes and proteins described herein may find application in the treatment and/or prevention of cutaneous disorders such as canine pyoderma.

Accordingly, a sixth aspect of this invention provides an Sps protein or gene as substantially defined above, for use in raising an immune response in an organism. The proteins and genes described herein may find particular application as a vaccine, but could also be used to obtain an immune serum potentially useful in passive vaccination techniques.

Advantageously, the invention may provide a vaccine for use in preventing or controlling disease in canine species caused or contributed to by *Staphylococcus pseudintermedius*. In other embodiments, the vaccines provided by this invention may be used to protect against the development of infections caused or contributed to by *Staphylococcus pseudintermedius*. In other embodiments, the vaccines may be used to protect against instances of canine pyoderma.

In one embodiment, the vaccine may be a polypeptide and/or polynucleotide vaccine.

A polynucleotide vaccine may comprise a polynucleotide fragment, preferably a DNA fragment, having a nucleotide sequence encoding an antigenic polypeptide comprising at least an antigenic portion any one or more of the Sps proteins described herein. Vaccines of this type may otherwise be referred to as “DNA vaccines” – such vaccines may be introduced to host cells (such as mammalian, for example, canine cells) where they express antigens which elicit immune responses.

A polypeptide or protein vaccine may comprises one or more of the Sps proteins (or antigenic fragments or portions) described herein. One of skill will appreciate that the one or more Sps protein(s) may be naturally occurring and isolated from *Staphylococcus pseudintermedius*, or recombinant.

A protein vaccine may be administered by any suitable route. Advantageously, a protein vaccine may be administered orally (by ingestion), topically or by direct injection - preferably intraperitoneal or intramuscular injection. A protein subunit vaccine formulated for oral administration can contain the polypeptide encapsulated in for example, a biodegradable polymer as described hereinafter.

In view of the above, the invention further provides a method of immunising a dog against *Staphylococcus pseudintermedius*, said method comprising administering to the dog a DNA or protein vaccine of the invention.

Conveniently, the protein vaccines described herein may further include or comprise one or more adjuvant(s). Further, one or more booster vaccinations are preferably administered at time periods subsequent to the initial administration to create a higher level of immune response in the animal.

In yet another aspect, the vaccine of the invention may comprise a fusion protein comprising a carrier polypeptide and one or more Sps protein(s) of the invention. The Sps protein(s) for use in this aspect of the invention can itself be antigenic or non-antigenic; in embodiments wherein the protein is non-antigenic, the carrier polypeptide is antigenic, stimulating the immune system to react to the fusion protein thereby generating an immune response in an organism – such as, for example a canine immune response to *Staphylococcus pseudintermedius*. A non-antigenic protein thus functions as a hapten. An example of an antigenic carrier polypeptide is keyhole limpet hemocyanin (KLH). Conventional fusion constructs between carriers such as glutathione sulfotransferase (GST) and said Sps protein(s) of the invention are also included as protein vaccines according to the invention, as are fusions of the Sps protein(s) and an affinity tag such as a polyhistidine sequence. A fusion construct

may be preferred for use as a protein vaccine when the antigenic Sps analog, fragment, or modification thereof is small.

In a seventh aspect, the present invention provides a method for immunising dogs against *Staphylococcus pseudintermedius*, said method comprising administering to the dog a vaccine of the invention.

A polynucleotide vaccine may further comprises a promoter, such as the CMV promoter, operably linked to the coding sequence for the Sps polypeptide or antigenic fragment thereof (e.g., U.S. Pat. No. 5,780,44, Davis). The polynucleotide may be cloned within a vector such as a plasmid. There are numerous plasmids known to those of ordinary skill in the art useful for the production of polynucleotide vaccines.

Other possible additions to the polynucleotide vaccine constructs include nucleotide sequences encoding cytokines, such as granulocyte macrophage colony stimulating factor (GM-CSF), interleukin-12 (IL-12) and co-stimulatory molecules such B7-1, B7-2, CD40. The cytokines can be used in various combinations to fine-tune the response of the animal's immune system, including both antibody and cytotoxic T lymphocyte responses, to bring out the specific level of response needed to affect the animal's reproductive system. A polynucleotide vaccine of the invention can also encode a fusion product containing the antigenic polypeptide and a molecule, such as CTLA-4, that directs the fusion product to antigen-presenting cells inside the host.

Plasmid DNA can also be delivered using attenuated bacteria as delivery system, a method that is suitable for DNA vaccines that are administered orally. Bacteria are transformed with an independently replicating plasmid, which becomes released into the host cell cytoplasm following the death of the attenuated bacterium in the host cell. An alternative approach to delivering the polynucleotide to an animal involves the use of a viral or bacterial vector. Examples of suitable viral vectors include adenovirus, polio virus, pox viruses such as vaccinia, canary pox, and fowl pox, herpes viruses, including catfish herpes virus, adenovirus-associated vector, retroviruses and bacteriophage. Exemplary bacterial vectors include attenuated forms of *Salmonella*, *Shigella*, *Edwardsiella ictaluri*, and *Yersinia ruckeri*. Preferably, the polynucleotide is a vector, such as a plasmid, that is capable of autologous expression of the nucleotide sequence encoding said Sps protein or fragment thereof.

In one embodiment, the vaccine may be a DNA vaccine comprising a DNA fragment having a nucleotide sequence that encodes a polypeptide having an amino

acid sequence homologous or identic to a sequence selected from the group consisting of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, and 34 or an antigenic analog, fragment, or modified version thereof.

Polynucleotide-based immunisation induces an immune response to an antigen expressed *in vivo* from a heterologous polynucleotide fragment introduced into a cell. DNA vaccine may be particularly useful as the heterologous nucleic acid expression may continue for a length of time sufficient to induce a relatively strong and sustained immune response without the need for subsequent “booster” vaccinations, as may be required when using protein based vaccines. A polynucleotide vaccine comprising a polynucleotide fragment having a nucleotide sequence encoding said Sps can be administered to dog (or rather to a particular tissue or cells thereof) using biolistic bombardment, ingestion or direct injection, as described for example, in U.S. Pat. No. 5,780,448 (Davis), preferably intraperitoneal or intramuscular injection. A preferred method of administration is biolistic bombardment, as with a “gene gun”. A polynucleotide vaccine formulated for oral administration preferably contains DNA encapsulated in a biodegradable polymer. Examples of a suitable biodegradable polymer include chitosan and homo- or co-polymers of polylactic acid and polyglycolic acid. Accordingly, the present invention further provides a method for immunising dogs against *Staphylococcus pseudintermedius* by administering to the dog a polynucleotide vaccine of the invention, preferably a DNA vaccine.

Other methods of administering nucleic acid vaccines of the type described herein may include, for example, use of the technology described in WO02/076498.

The amount of protein/polynucleotide vaccine to be administered to an animal depends on the type and size of animal, the condition being treated, and the nature of the protein/polynucleotide, and can be readily determined by one of skill in the art. In some applications, one or more booster administrations of the protein/polynucleotide vaccine at time periods subsequent to the initial administration are useful to create a higher level of immune response in the animal.

In one embodiment of the vaccine of the invention and/or Sps proteins described herein (including antigenic fragments, analogs or modified version thereof) may be linked, for example, at its carboxy-terminus, to a further component. The further component may serve to facilitate uptake of the Sps protein, or enhance its immunogenicity/processing.

The immune-stimulating compositions of the invention may be optionally mixed with excipients or diluents that are pharmaceutically acceptable as carriers and compatible with the active component(s). The term “pharmaceutically acceptable carrier” refers to a carrier(s) that is “acceptable” in the sense of being compatible with the other ingredients of a composition and not deleterious to the recipient thereof. Suitable excipients are well known to the person skilled in the art. Examples include; water, saline (e.g. 0.85% sodium chloride; see Ph.Eur. monograph 2001:0062), buffered saline, fish oil with an emulsifier (e.g. a lecithin, Bolec MT), inactivant (e.g. formaldehyde; see Ph.Eur. monograph 1997:0193), mineral oils, such as light mineral oils, alhydrogel, aluminium hydroxide. Where used herein, the term “oil adjuvant” to embraces both mineral oils and synthetic oils. A preferred adjuvant is Montanide ISA 711 (SeppicQuai D’Orsay, 75321 Paris, France) which is a manide oleate in an oil suspension. In addition, if desired, the immune-stimulating composition (including vaccine) may contain minor amounts of auxiliary substances such as wetting or emulsifying agents, pH buffering agents, and/or adjuvants which enhance the effectiveness of the immune-stimulating composition.

A vaccine composition may be administered as a course of a number of discrete doses over a period of time. For example it may be administered over a period of around 2-21 days.

Vaccination may be repeated at daily, twice-weekly, weekly or monthly intervals. For example a boost vaccination may be administered after the initial dose. For example a boost may be administered at around 4-14 weeks after the vaccination. The initial vaccination and any boost may be carried out using the same or different modes of administration. For example, the initial may be by injection and the boost may be by oral administration. An example regime includes a first vaccination by injection, followed by a course of orally administered boost vaccine, or a booster prior to an expected outbreak. However, it will be appreciated that any suitable route of administration(s) and/or regime(s) may be employed.

Additionally, knowledge of the Sps protein nucleotide and amino acid sequences set forth herein opens up new possibilities for detecting, diagnosing and characterising *Staphylococcus pseudintermedius* in canine populations. For example, an oligonucleotide probe or primer based on a conserved region of one or more of the Sps proteins described herein, may be used to detect the presence of the Sps protein in or on a canine host.

Vaccines may contain one or more of the Sps proteins/nucleic acids/genes described herein (i.e. those shown as SEQ ID NOS: 1-38). In one embodiment, the vaccine may comprise a cocktail of Sps proteins/peptides and or nucleic acids. Typically, a cocktail may comprise 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 or 36 Sps nucleic acid and/or protein/peptide components (for example (2 or more) components having sequences homologous or identical to any of SEQ ID NOS: 1-36).

Furthermore, the vaccines may contain bacterial antigens used to control other diseases, for example diseases caused by other *Staphylococcal* species and/or antigens to treat, prevent or control diseases and/or conditions with other aetiologies or caused or contributed to by other pathogens. As such, the vaccine compositions described herein may find application in multivalent vaccines including antigens against other canine diseases.

In addition to vaccines and/or immunogenic compositions comprising one or more of the Sps proteins described herein, the present invention further provides compounds for treating infections caused or contributed to by *Staphylococcus pseudintermedius* or compounds for the preparation of medicaments for treating the same.

In one embodiment, the compound may be a small organic molecule, antibody, peptide or carbohydrate which antagonises the interaction between the Sps protein and its ligand (an extracellular matrix (ECM) protein). For example, the compound may be a synthetic peptide comprising or based on, the sequence of an ECM protein known to interact with a particular Sps protein, or the sequence of a protein given above which may interfere with binding between the wild type *S. pseudintermedius* protein and its ligand. Additionally or alternatively, binding agents, such as for example, antibodies with specificity or affinity for one or more Sps protein ligands, may also be used to antagonise the Sps/ligand interaction. Therapeutic approaches of this type may prevent *Staphylococcus pseudintermedius* colonising or binding/adhering to cells.

In view of the above, the invention may relate to methods of treating infections caused or contributed to by *Staphylococcus pseudintermedius*, said method comprising administering to an animal a therapeutically effective amount of a compound which antagonises Sps/ligand interactions.

In a further aspect, the present invention provides pharmaceutical compositions comprising a compound which antagonises Sps/ligand interactions together with a pharmaceutical excipient, carrier or diluent.

One of skill will appreciate that the vaccines, methods, uses or medicaments comprising any of the Sps genes/nucleic acids and/or proteins and/or antagonistic compounds (for example Sps protein/nucleic acid fragments and/or antibodies) described herein may be combined with one or more other compounds for treating one or more other conditions – in particular one or more other skin conditions. Said other skin condition may be, for example, atopic dermatitis.

In a further aspect, the present invention provides methods of diagnosing infections, diseases and/or conditions caused or contributed to by *S. pseudintermedius*, said methods comprising the steps of identifying in a sample provided by a subject suspected of suffering from an infection, disease and/or conditions *S. pseudintermedius* caused or contributed to by *S. pseudintermedius*, a level of a protein, peptide or nucleic acid (for example a gene) encoded by a sequence provided by SEQ ID NOS: 1-36 or a fragment, portion, mutant, derivative and/or homologue/orthologue thereof.

It should be understood that all methods of diagnosis or detection described herein, may include an optional step in which the results are compared with the results of a control sample, which does not comprise sequences derived from *S. pseudintermedius*, in particular sequences corresponding to those provided as SEQ ID NOS: 1-36 disclosed herein.

The term “sample” may be taken to mean any sample comprising protein and/or nucleic acid. For example, a “sample” may comprise a bodily fluids such as whole blood, plasma, serum, saliva, sweat and/or semen. In other instances “samples” such as tissue biopsies and/or scrapings may be used. In particular, cutaneous (i.e. skin) tissue biopsies and/or scrapings may be used. Advantageously such biopsies may comprise cells obtained from lesions suspected of resulting from or being associated with a *S. pseudintermedius*. Specifically, a biopsy, tissue sample or scraping may comprise cells derived from lesions exhibiting pathology characteristic of the *S. pseudintermedius* disease, pyoderma (particularly canine pyoderma).

In addition, a sample may comprise a tissue or gland secretion and washing protocols may be used to obtain samples of fluid secreted into or onto various tissues, including, for example, the skin. One of skill in this field will appreciate that the

samples described above may yield or comprise quantities of nucleic acid (i.e. DNA or RNA) encoding all or part of the various proteins described herein as well as quantities of proteins or peptides (or fragments thereof) encoded thereby. In one embodiment, the sample may comprise quantities of nucleic acid/peptide having or comprising the sequences given as SEQ ID NOS: 1-36.

One of skill in the art will be familiar with the techniques that may be used to identify levels of certain nucleic acid sequences and/or proteins, such as, for example, levels of the sequences given as SEQ ID NOS: 1-36 described herein (or a fragment, portion, mutant, derivative and/or homologue/orthologue thereof).

For example, PCR based techniques may be used to detect levels of gene expression or gene quantity in a sample. Useful techniques may include, for example, polymerase chain reaction (PCR) or reverse transcriptase (RT)-PCR based techniques in combination with real-time PCR (otherwise known as quantitative PCR).

Additionally, or alternatively, a level of gene/protein expression may be identified by way of microarray analysis. Such a method would involve the use of a DNA micro-array which comprises nucleic acid derived from one or more of the nucleic acid sequences described herein (for example SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35). To identify a level of gene expression, one of skill in the art may extract nucleic acid, preferably mRNA, from a sample and subject it to an amplification protocol such as, for example RT-PCR to generate cDNA. Preferably, primers specific for a certain mRNA sequence – in this case a *S. pseudintermedius* sequence comprised with any of, for example, SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35.

The amplified cDNA may be subjected to a further amplification step, optionally in the presence of labelled nucleotides (as described above). Thereafter, the optionally labelled amplified cDNA may be contacted with the microarray under conditions which permit binding with the DNA of the microarray. In this way, it may be possible to identify a level of *S. pseudintermedius* gene expression.

Further information regarding the PCR based techniques described herein may be found in, for example, PCR Primer: A Laboratory Manual, Second Edition Edited by Carl W. Dieffenbach & Gabriela S. Dveksler: Cold Spring Harbour Laboratory Press and Molecular Cloning: A Laboratory Manual by Joseph Sambrook & David Russell: Cold Spring Harbour Laboratory Press.

In addition, other techniques such as deep sequencing and/or pyrosequencing may be used to detect cSCC sequences in any of the samples described above. Further information on these techniques may be found in “Applications of next-generation sequencing technologies in functional genomics”, Olena Morozovaa and Marco A. Marra, Genomics Volume 92, Issue 5, November 2008, Pages 255-264 and “Pyrosequencing sheds light on DNA sequencing”, Ronaghi, Genome Research, Vol. 11, 2001, pages 3-11.

In addition to the molecular detection methods described above, one of skill will also appreciate that immunological detection techniques such as, for example, enzyme-linked immunosorbent assays (ELISAs) may be used to identify levels of *S. pseudintermedius* proteins in samples. In other embodiments, ELISPOT, dot blot and/or Western blot techniques may also be used. In this way, samples provided by subjects suffering from *S. pseudintermedius* related diseases and/or infections (for example canine subjects suffering from canine pyoderma), may be probed for levels of one or more *S. pseudintermedius* proteins, particularly those encoded by SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34 or 36, so as to detect the presence of such proteins in a sample which may indicate a *S. pseudintermedius* infection.

Immunological detection techniques, may require the use of a substrate to which an antibody and/or antigen may be bound, conjugated or otherwise immobilised.

Suitable substrates may comprise, for example, glass, nitrocellulose, paper, agarose and/or plastic. A substrate which comprises, for example, a plastic material, may take the form of a microtitre plate.

Further information regarding ELISA procedures and protocols relating to the other immunological techniques described herein may be found in Using Antibodies: A Laboratory Manual by Harlow & Lane (CSHLP: 1999) and Antibodies: A Laboratory Manual by Harlow & Lane (CSHLP: 1988).

The present invention also extends to kits comprising reagents and compositions suitable for diagnosing, detecting or evaluating possible *S. pseudintermedius* infections, diseases and/or conditions. Kits according to this invention may be used to identify and/or detect levels of *S. pseudintermedius* gene(s)/ *S. pseudintermedius* protein(s) in samples. Depending on whether or not the kits are intended to be used to identify levels of *S. pseudintermedius* genes and/or *S.*

pseudintermedius proteins in samples, the kits may comprise substrates having *S. pseudintermedius* proteins or agents capable of binding *S. pseudintermedius* proteins, bound thereto. In addition, the kits may comprise agents capable of binding *S. pseudintermedius* proteins – particularly where the kit is to be used to identify levels of one or more *S. pseudintermedius* proteins in samples. In other embodiments, the kit may comprise polyclonal antibodies or monoclonal antibodies which exhibit specificity and/or selectivity for one or more *S. pseudintermedius* proteins. Antibodies for inclusion in the kits provided by this invention may be conjugated to detectable moieties. Kits for use in detecting the expression of genes encoding *S. pseudintermedius* proteins may comprise one or more oligonucleotides/primers for detecting/amplifying/probing samples for *S. pseudintermedius* protein encoding sequences. The kits may also comprise other reagents to facilitate, for example, sequencing, PCR and/or RFLP analysis. In one embodiment, the kits may comprise one or more oligonucleotides/primers for detecting/amplifying/probing nucleic acid samples (for example nucleic acid derived from canine skin) for levels of sequences corresponding to all or part of those described as SEQ ID NOS: 1-36 herein.

DETAILED DESCRIPTION

The invention will now be described in more detail with reference to the following Figures which show:

Figure 1. Genomic location of the 17 genes encoding putative CWA proteins in *S. pseudintermedius* strain ED99. Eight genes are situated in the *oriC* environ, indicated in orange, and nine are located in the core genome. *sps* = *S. pseudintermedius* surface protein.

Figure 2. Distribution of the genes encoding putative CWA proteins among 20 *S. pseudintermedius* strains, representatives of the closely related *S. delphini* and *S. intermedius*, and other staphylococcal species associated with animal skin disease. The diversity of strains is represented a phylogenetic tree; grey squares indicate that the gene is present, blank squares that the gene is absent based on Southern blot analysis (for *spsA* to *spsO*), or PCR amplification (for *spsP* and *spsQ*).

Figure 3. Western blot analysis of cell wall-associated proteins of *S. pseudintermedius* ED99 and *L. lactis* expressing *SpsD*, *SpsL*, and *SpsO* with sera from dogs diagnosed with pyoderma. (A) SDS PAGE analysis and (B) Western blot analysis of protein fractions from *S. pseudintermedius* ED99 in exponential phase of growth (lane 1); *L.*

lactis expressing SpsL (lane 2); *L. lactis* expressing SpsD (lane 3); *L. lactis* expressing SpsO (lane 4); and *L. lactis* with pOri23 alone (lane 5).

Figure 4. Adherence of *L. lactis* expressing specified MSCRAMMs to human Fn. Plates were coated with 1 µg of human Fn or 1 µg of BSA per well. Absorbance was measured at 590 nm and results are expressed as mean values of triplicate samples. Error bars indicate standard deviation. *L. lactis* expressing FnbpA from *S. aureus* and PBS were included as controls.

Figure 5. Adherence of *L. lactis* expressing specified MSCRAMMs to Fg from different animal sources. Plates were coated with 1 µg of Fg or 1 µg of BSA per well. Absorbance was measured at 590 nm and results are expressed as mean values of triplicate samples. Error bars indicate standard deviation. *L. lactis* expressing FnbpA from *S. aureus* and PBS were included as controls.

Figure 6. Adherence of *L. lactis* expressing specified MSCRAMMs to CK10. Plates were coated with 1 µg of recombinant CK10 or 1 µg of BSA per well. Absorbance was measured at 590 nm and results are expressed as mean values of triplicate samples. Error bars indicate standard deviation. *S. aureus* strain SH1000 in exponential and stationary phases of growth and PBS were included as controls.

Figure 7. Adherence of *L. lactis* expressing different MSCRAMMs to canine corneocytes of five dogs. Bacterial adherence is calculated as percentage area covered with bacterial cells per field of corneocytes (ROI = 500 µm²). Results are based on the arithmetic mean of duplicate experiments. The bottom of each box represents the first quartile (Q1), the top of the box the third quartile (Q3), the bold lines the median, and the black circles the mean values. The whiskers define the range of the data.

Figure 8: Reactivity of canine convalescent serum from pyoderma cases to Sps D, Sps L and Sps O recombinant A domain. 1 µg aliquots of rSps D and rSps L, and 10 µl volumes of purified rSps O were subjected to SDS-PAGE under standard conditions and Coomassie stained (A) or Western blot transferred onto a nitrocellulose membrane. Membranes were probed with a 1:1000 dilution of canine serum, followed by a 1:3000 dilution of HRP- conjugated sheep anti-canine IgG. Reactive bands were visualized on Chemi-luminescent Film (B). 5 µl aliquots of recombinant ClfB and the superantigen SEI from *S. aureus* were included in the terminal lanes of each gel as negative controls.

Figure 9: Inhibition of adherence of *L. lactis* expressing SpsD (A) and SpsL (B) to fibrinogen (2 µg per well) by canine convalescent serum from pyoderma cases.

Bacterial cultures, normalised to an OD₆₀₀ of 1 in PBS were pre-incubated for 2 h with doubling dilutions of serum ranging from 2% to ~0.01% (v/v), prior to inoculation into fibrinogen coated wells. Results (n = 3, ± SD) are expressed as absorbance readings at 590 nm minus background levels of fluorescence. Background fluorescence was measured by inoculating control cultures, incubated for 2 h in the absence of serum, into wells coated with BSA (2 µg per well). Incubations of *S. aureus Newman* (C) were included as a negative control.

Materials and Methods

Genome-wide screen for genes encoding for cell-wall anchored proteins

The *S. pseudintermedius* strain ED99 draft genome was interrogated for homologous sequences using position specific iterative basic local alignment search tool (PSI-BLAST), available from the National Center for Biotechnology Information (NCBI), USA (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>), and for the presence of a LPX[TSA][GANS] motif pattern by pattern hit initiated basic local alignment search tool (PHI-BLAST), available from NCBI (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). Signal sequences were predicted by employing the SignalP server (<http://www.cbs.dtu.dk/services/SignalP/>), provided by the Center for Biological Sequence Analysis (CBS), Technical University of Denmark.

In silico structural analysis of cell-wall anchored proteins

The predicted CWA proteins were searched for functional domains using EMBL-EBI InterPro Scan (<http://www.ebi.ac.uk/interpro>). Structural analysis was carried out with the PHYRE (Protein Homology/analogy Recognition Engine) fold recognition server, available from the Structural Bioinformatics Group, Imperial College London, UK (<http://www.sbg.bio.ic.ac.uk/phyre/>). Repeat sequences were predicted by generating nucleic acid dot plots, using software available from Colorado State University, USA (<http://www.vivo.colostate.edu/molkit/dnadot/>), applying tandem repeats finder software from Boston University, USA (<http://tandem.bu.edu/trf/trf.html>), and variable sequence tandem repeats extraction and architecture modelling (XSTREAM), available from the University of California, USA (<http://jimcooperlab.mcdm.ucsb.edu/xstream/>). Sequence alignments and pair-wise sequence comparisons were generated with ClustalW2 (<http://www.ebi.ac.uk/Tools/clustalw2>). Amino acid composition and molecular weight predictions were generated using ProtParam on the ExPASy Proteomics Server (<http://www.expasy.ch/tools/protparam.html>).

Cloning of selected genes encoding putative MSCRAMMs of *S. pseudintermedius* ED99 into *L. lactis* MG1363

Oligonucleotides were designed for PCR amplification of the full-length *spsD*, *spsL* and *spsO* genes and either *Pst*I, *Sal*I or *Bam*HI specific restriction sites were inserted on both sides of the DNA fragments. 50 µl PCR reactions contained 2 µl (approximately 100 ng) genomic DNA template, 0.25 µM forward primer, 0.25 µM reverse primer, 1x *PfuUltra*TM II reaction buffer (Stratagene, USA), 0.25 mM dNTP's (Promega, USA) and 1 µl *PfuUltra*TM II Fusion HS DNA polymerase (Stratagene, USA). The thermocycler programme included an initial denaturation step at 95°C for 2 min, followed by 30 cycles of denaturation at 95°C for 20 s, annealing at 50°C for 20 s and extension at 72°C for 2 min, followed by a final extension step at 72°C for 3 min. PCR products were visualised on 0.8% (w/v) agarose gels, gel extracted under avoidance of UV light exposure and purified using QIAquick Gel Extraction Kit (Qiagen, UK) according to the manufacturer's instructions. Purified DNA fragments were cloned into the StrataCloneTM Blunt PCR cloning vector pSC-B (Stratagene, USA) using the StrataClone UltraTM Blunt PCR Cloning Kit (Stratagene, USA) according to the manufacturer's instructions. Each cloning reaction consisted of 3 µl Strataclone Buffer Blunt (Stratagene, USA), 2 µl purified PCR product and 1 µl StratacloneTM Blunt Vector Mix (Stratagene, USA). StrataCloneTM SoloPack[®] competent cells (Stratagene, USA) were transformed according to the manufacturer's instructions and colonies selected using blue-white screening on LB-ampicillin (100 µg/ml)-X-gal plates. White colonies were transferred into 5 ml LB-ampicillin (100 µg/ml) broth and grown overnight at 37°C with shaking at 200 rpm. Plasmid was isolated using QIAprep Spin Miniprep Kit (Qiagen, UK) according to the manufacturer's instructions. Purified plasmids were digested using appropriate restriction endonucleases (New England Biolabs, UK), and diagnostic digests were analysed on 0.8% (w/v) agarose gels. For generating DNA constructs, the *E. coli* - *L. lactis* shuttle vector pOri23 (kindly provided by P. Moreillon, University of Lausanne, Switzerland) was used. The pOri23 vector carries the *ermAM* gene for erythromycin resistance, the high-copy-number *oriColE1* replicon for autonomous replication in *E. coli* and the constitutive lactococcal promoter *P23* (Que *et al.*, 2000). The multiple cloning site of pOri23 consists of restriction sites for endonucleases *Pst*I, *Sal*I and *Bam*HI (Que *et al.*, 2000).

StrataClone™ plasmids containing the DNA inserts of interest and the *E. coli* - *L. lactis* shuttle vector pOri23 were each digested in a 100 µl total reaction volume containing 10 µl plasmid (approximately 2.5 µg), 20 units appropriate restriction endonucleases (New England Biolabs, UK), and suitable buffers (New England Biolabs, UK) according to the manufacturer's instructions. Restriction digestions were performed at 37°C for 16 h. The restriction fragments to be cloned were extracted from 0.8% (w/v) agarose gels without UV exposure as described in the general Material and Methods and purified using QIAquick Gel Extraction Kit (Qiagen, UK) according to the manufacturer's instructions. DNA inserts and restriction-digested pOri23 plasmid were quantified using spectrophotometry (NanoDrop ND-1000, Thermo Scientific, USA) and ligation reactions were carried out with a plasmid to insert ratio of 1:3 in a 10 µl total ligation reaction volume, consisting of 1 µl vector (approximately 10 ng), 400 units T4 DNA ligase (New England Biolabs, UK), 1x T4 DNA ligase reaction buffer (New England Biolabs, UK), x µl DNA insert (depending on DNA concentration), and x µl sterile water (depending on the volume of DNA insert). Ligations were incubated at 16°C for 16 h.

One 50 µl aliquot of electrocompetent *L. lactis* cells was thawed on ice and 2 µl (~20 ng) pOri23 plasmid carrying the DNA insert of interest was added. Electroporation cuvettes (Sigma-Aldrich, UK) were pre-chilled and *L. lactis* cells plus plasmid were transferred into the cuvettes. Electroporation was performed at standard settings (25 µF, 2.5 kV, 200 Ohm) and 1 ml GM17 was added immediately. Cells were incubated at 30°C in a static incubator for 2 h prior to spreading 250 µl of cell suspension per plate onto GM17 plates containing 5 µg/ml erythromycin. Plates were incubated overnight at 30°C.

For screening of *L. lactis* transformants, plasmid was isolated using the Qiagen MiniPrep Kit (Qiagen, UK) with addition of 100 U/ml mutanolysin (Sigma-Aldrich, UK) and 100 µg/ml lysozyme (Sigma-Aldrich, UK) to buffer P1. Diagnostic digests of purified plasmids were carried out with appropriate restriction enzymes and analysed on 0.8% (w/v) agarose gels.

Additionally, colony PCR was performed for pOri23 carrying *spsD* and *spsO* using gene-specific oligonucleotides (Table 5.3). *L. lactis* colonies were resuspended in 10 µl 10% (v/v) IGEPAL (Sigma-Aldrich, UK) and incubated for 10 min at 95°C in a thermocycler machine. 40 µl master mix containing 0.3 µM forward primer, 0.3 µM reverse primer, 0.2 mM dNTP's (Promega, USA), 1 x reaction buffer (Promega,

USA), 1.5 mM MgCl₂ (Promega, USA) and 0.025 u/μl taq polymerase (Promega, USA) was added. The thermocycler programme included an initial denaturation step at 95°C for 2 min, followed by 30 cycles of denaturation at 95°C for 1 min, annealing at 50°C for 1 min and extension at 72°C for 1 min, followed by a final extension step at 72°C for 7 min. PCR products were visualised on 0.8% (w/v) agarose gels.

Western blot analysis of *L. lactis* constructs

Samples were dissolved in 1x Laemmli sample buffer (Sigma-Aldrich, UK), boiled for 10 min and resolved by SDS-PAGE in 10% polyacrylamide gels by standard procedures, and Western blot analysis was carried out as described in the general Materials and Methods. Three canine sera samples from pyoderma cases (obtained from patients at the Hospital for Small Animals, The Royal (Dick) School of Veterinary Studies, The University of Edinburgh) were pooled and used as primary antibody in a 1:1000 dilution. HRP-conjugated sheep anti-dog antibody was used as a secondary antibody in a 1:5000 dilution (Bethyl Laboratories Inc., USA).

Canine corneocyte adherence assay

For preliminary experiments to confirm adherence of *S. pseudintermedius* ED99 and non-adherence of *L. lactis*, corneocytes were obtained from a seven-year-old male neutered Border collie cross-breed with no history or physical signs of systemic or cutaneous disease. Corneocytes for the *L. lactis* adherence study were obtained from five dogs of different breeds (one Labrador retriever, two Border collies and two cross-breeds). Three dogs were ovariohysterectomised females and two were entire males. The median age was seven years (range one to twelve years). The dogs showed no abnormalities on general physical examination and had no history or physical signs of skin disease at the time of corneocyte collection. All dogs were privately owned by staff or students of the Royal (Dick) School of Veterinary Studies, The University of Edinburgh. None of the dogs had received topical or systemic drug treatments for at least three weeks prior to the day of corneocyte collection.

Samples were taken from the ventral abdomen and inner thigh. If necessary, sample sites were clipped with Oster clippers (Oster Cryotech, USA) using a number 40 blade. For collection of corneocytes, the method described by Forsythe *et al.* (2002) was used. Briefly, the area was cleaned of surface debris and commensal bacteria by applying four strips of single sided adhesive tape (Cellux, Henkel Consumer Adhesives, UK), using each strip once. To collect corneocytes, double-sided, clear, adhesive wig tape (Tropical Tape Super Grip, USA) was mounted onto a

microscope slide in 1 cm² pieces and applied to the same area of skin surface 10 times with gentle force. Slides were investigated by microscopic examination and only slides with at least 75% corneocyte coverage were used in the study.

The corneocyte slides were positioned in moisture chambers (Nunc™, Thermo Fisher Scientific, Denmark) as described by Forsythe *et al.* (2002). The moisture chambers consisted of 30 cm x 30 cm plastic trays with lids and were prepared by lining the trays with moistened paper towels. *S. pseudintermedius* ED99 stationary or exponential (OD₆₀₀ of 0.5) phase cultures and *L. lactis* exponential phase cultures (OD₆₀₀ 0.6 to 0.8) were centrifuged at 4000 rpm for 5 min, washed with PBS and resuspended in PBS to a final OD₆₀₀ of 0.5. The moisture chambers were placed in a static incubator and 250 µl of bacterial suspension was added to each 1 cm² of tape, forming a meniscus on the tape. Slides incubated with 250 µl of sterile PBS were included as a control. The slides were incubated at 37°C for 90 min and washed in PBS. Each slide was stained with 0.5% (w/v) crystal violet (Sigma-Aldrich, UK) for 90 s before rinsing off with PBS. The slides were air-dried and a drop of immersion oil (Cargille Laboratories Inc., USA) and a cover slip (Scientific Laboratory Supplies, UK) were added before microscopic quantification. All slides were prepared in duplicate on the same day and incubated at the same time. Prior to incubation with bacterial suspensions or PBS, each slide was labelled with a letter code to allow identification after the microscopic analysis. The identification code on each slide was hidden by a third party for subsequent image acquisition so that the investigator was blinded to the origin of the slide. For quantification of adherent bacteria, computerised image analysis was used as described previously by Forsythe *et al.* (2002) with minor modifications. For each slide, bright field images of 1000x oil-immersion fields were acquired with a Sony DXC-390P 3CCD colour video camera (Scion Corporation, USA) connected to a Leica Laborlux S microscope (Leica Microsystems UK Ltd., UK). The RGB video signal from the camera was digitised using Scion Image (Scion Corporation, USA) installed in a G4 Macintosh computer (Apple Computer, USA) fitted with a CG-7 frame grabber (Scion Corporation, USA). For image acquisition, fields equivalent to 14.4 mm² were selected randomly by starting in the bottom left corner of each slide and moving through the slide in a defined way using the scale on the microscope stage. A field was discarded if the corneocyte layer was not confluent, the bacteria were poorly stained against the background or the field could not be focused properly. The software used for quantification of bacterial adherence was set

to calculate the percentage area that was covered by bacteria per confluent layer of corneocytes in a defined region of interest (ROI) of 500 μm^2 within each image field acquired. Previous studies by Forsythe *et al.* (2002) using the same technique and software have demonstrated that 15 replicates of each duplicate slide resulted in acceptable coefficients of variation of approximately 10%. In this study, 25 replicates of each slide were acquired and the overall mean percentage area of adherence was determined by calculating the mean of all replicates.

Results

Identification of genes encoding 17 putative cell-wall-anchored proteins in the *S. pseudintermedius* ED99 genome

The initial search for putative CWA proteins identified 34 sequences that fulfilled at least one of the search criteria (homology to characterised MSCRAMMs in the database, predicted LPXTG motif or variant near the C terminus, predicted signal peptide at the N terminus). After gap closure and combination of incomplete sequences, a total of 17 ORFs encoding putative CWA proteins with a predicted minimum length of approximately 250 amino acids was determined. The 17 predicted CWA proteins were designated 'Sps' for '*Staphylococcus pseudintermedius* surface proteins', followed by a capital letter (SpsA to SpsQ). Their position in the *S. pseudintermedius* ED99 genome is indicated in Figure 1. Of note, eight genes encoding putative CWA proteins are located near the *oriC* environ (Figure 1). Homology searches in the database resulted in sequence identities with known staphylococcal proteins ranging from ~30% to ~80% (Table 1). Signal sequences, necessary for Sec-dependent protein secretion (Foster and Hook, 1998), were predicted for 14 putative Sps proteins, consisting of 29 aa for SpsC and SpsK, 33 aa for SpsN, SpsP, and SpsQ, 36 aa for SpsD, 37 for SpsG, 38 aa for SpsA, SpsB, and SpsL, 39 aa for SpsH, 44 aa for SpsO, and 48 aa for SpsF and SpsM. No signal sequence was predicted for SpsE, SpsI, and SpsJ (Figure 4.3).

The putative CWA proteins SpsD, SpsL, and SpsO have several MSCRAMM features.

Out of the 17 putative CWA proteins of *S. pseudintermedius* ED99, SpsD, SpsL, and SpsO contained each of the MSCRAMM features screened for, including a signal sequence at the N-terminus, followed by a non-repeated A domain with two IgG-like folds, dividing the A domain into N1, N2, and N3 subdomains, a tandemly repeated domain at the C-terminus (and at the N-terminus for SpsO), and a C-terminal

LPXTG-anchor motif. The main characteristics of SpsD, SpsL, and SpsO are summarised in Table 2. Of interest, a TYTFTDYVD motif or variant, important for the 'dock, lock and latch' ligand-binding mechanism (Ponnuraj et al., 2003), was found in SpsD, SpsL, and SpsO, and putative latching sequences were identified (Table 2). Further, putative Fn-binding motifs with weak homology to FnbpA-10 of FnbpA in *S. aureus* were detected in the repeat region of SpsL (24% identity in pair-wise alignments for SpsL1-SpsL6, and 21% for SpsL-7). No homology to Fn-binding motifs of FnbpA was detected in the repeat regions of SpsD and SpsO. Of note, the genes encoding for SpsD, SpsL, and SpsO in the *S. pseudintermedius* ED99 genome are situated in different genomic contexts. While *spsD* is located in a well-conserved region of the core genome, *spsL* is part of the *oriC* environ (Takeuchi *et al.*, 2005) (Figure 1). The *spsO* gene appears to be species-specific as it is not present in the genomes of other staphylococcal species. The region contains two putative transposases, suggesting that the whole region might be subjected to horizontal gene transfer.

Distribution of the 17 genes encoding putative cell-wall-anchored proteins among the *S. intermedius* group

In order to investigate the distribution of the 17 genes encoding putative CWA proteins identified in the *S. pseudintermedius* ED99 genome among other members of the SIG and closely related staphylococcal species, Southern blot analysis and PCR amplification were performed. A total of 20 *S. pseudintermedius* strains representing the breadth of diversity within the species, representatives of the closely related *S. delphini* and *S. intermedius* species, and other staphylococcal species associated with animal hosts (Figure 2) were screened for the presence of the putative CWA encoding genes by Southern blot analysis (*spsA* to *spsO*). For the *S. aureus spa* orthologues *spsP* and *spsQ*, PCR amplification was employed, as the genes share 70% nucleotide identity which precluded design of gene-specific probes for Southern blot analysis. For similar reasons, the primers designed for PCR amplification were located upstream of *spsP* (*spsP*-F), in the non-repeated region of *spsP* (*spsP*-R), in the unique region between *spsP* and *spsQ* (*spsQ*-F), and in a region unique for *spsQ* (*spsQ*-R).

Of the 17 genes examined, 13 were found in all *S. pseudintermedius* strains investigated. The remaining 4 (*spsF*, *spsO*, and the *S. aureus spa* orthologues *spsP* and *spsQ*) were present in 11, 6, 7, and 11 of the 20 *S. pseudintermedius* strains, respectively. Furthermore, 8 of the 17 genes were detected in *S. delphini* and 6 in *S.*

intermedius, and 9 genes were exclusive to *S. pseudintermedius*. None of the genes encoding putative CWA proteins was detected in the non-SIG staphylococcal species examined. Results are summarised in Figure 2. Of note, it cannot be excluded at this point that DNA sequence variation in PCR primer annealing sites for *spsP* and *spsQ*, and weak homology (less than approximate 70%) for *spsA* to *spsO* among different strains have influenced the results.

Expression of CWA proteins on the *S. pseudintermedius* bacterial cell surface.

The *in silico* identification of 17 putative CWA proteins in *S. pseudintermedius* ED99 raises questions about the expression of these proteins and their role in colonisation and disease. Surface proteome analysis of early-, mid-, and late exponential phase *S. pseudintermedius* ED99 was performed in collaboration with the Moredun Research Institute, Penicuik, Scotland, UK, using liquid chromatography-electrospray ionisation-tandem mass spectrometry (LC-ESI-MS-MS). Six out of the 17 putative CWA proteins predicted in the *S. pseudintermedius* ED99 genome were detected on the bacterial surface, including SpsD, SpsK, SpsL, SpsN, SpsO, and SpsQ. The putative CWA proteins SpsL, SpsN, and SpsQ were identified in all three phases of growth; SpsK was lacking in early-, SpsO in mid-, and SpsD in late exponential phase. The 11 undetected CWA proteins might not have been expressed under the conditions tested, or the expression level might have been below the detection threshold of the LC-ESI-MS-MS method used.

Cloning and expression of SpsD, SpsL, and SpsO in *L. lactis*.

In order to examine the role of putative selected MSCRAMMs independently on the bacterial cell surface, the full-length *spsD* (3096 bp), *spsL* (2793 bp) and *spsO* (5538 bp) genes were cloned into *L. lactis* using the shuttle vector pOri23 (Que *et al.*, 2000). Positive clones were identified by restriction digestion of purified pOri23 plasmids from single colonies of transformed *L. lactis* cells (data not shown). The pOri23 construct inserts were verified by DNA sequencing for *spsL* and *spsD*. For *spsO*, DNA sequence was generated for approximately 3000 bp of the total length of 5538 bp. A segment of the repeat region corresponding to ~2500 bp could not be determined due to the existence of identical tandem repeats which did not allow directed sequencing. As a negative control for subsequent MSCRAMM characterisation studies, *L. lactis* was transformed with the empty vector pOri23, confirmed by restriction digestion analysis. The predicted molecular weights were 115 kDa for SpsD, 103 kDa for SpsL, and 198 kDa for SpsO.

L. lactis expressing SpsD and SpsL demonstrated seroreactivity with canine sera from pyoderma cases.

The potential antibody response to SpsD, SpsL, and SpsO *in vivo* was investigated by Western blot analysis employing canine sera from staphylococcal pyoderma cases. The pyoderma was clinically manifested at the time of blood sampling and the dogs were also diagnosed with AD (Neuber *et al.*, 2008). Cell wall-associated protein fractions of the *L. lactis* constructs and of *S. pseudintermedius* ED99 were subjected to SDS-PAGE, transferred to nitrocellulose membrane and incubated with pooled canine sera from three pyoderma cases as described in Materials and Methods. An array of immunoreactive bands was detected for *S. pseudintermedius* ED99, ranging from 24 kDa to 102 kDa in molecular weight (Figure 3). For *L. lactis* expressing SpsD and *L. lactis* expressing SpsL, multiple seroreactive bands in the range of 38 kDa to 225 kDa for SpsD, and 38 kDa and 52 kDa for SpsL were detected, which were absent in the protein fractions of *L. lactis* carrying pOri23 alone (Figure 3). In contrast, *L. lactis* expressing SpsO did not demonstrate seroreactivity with sera from dogs diagnosed with pyoderma (Figure 3).

Adherence of *L. lactis* constructs to extracellular matrix proteins.

L. lactis expressing SpsO, SpsD, SpsL, and *L. lactis* carrying the vector pOri23 alone were tested for their ability to adhere to human Fn, human, canine, feline, and bovine Fg, and to recombinant mouse CK10 in solid phase assays.

The putative MSCRAMMs SpsD and SpsL mediate binding of *L. lactis* to fibronectin.

L. lactis expressing SpsD and SpsL demonstrated adherence to human Fn, whereas *L. lactis* expressing SpsO demonstrated increased binding to Fn, but also to BSA, indicative of a non-specific interaction (Figure 4).

The putative MSCRAMMs SpsD and SpsL mediate binding of *L. lactis* to fibrinogen, and SpsL demonstrates canine host-specificity.

L. lactis expressing SpsD strongly adhered to Fg from several animal sources (Figure 5). In contrast, *L. lactis* expressing SpsL adhered to canine and feline Fg only, and did not bind to human and bovine Fg (Figure 5), indicating a host-specific interaction. *L. lactis* expressing SpsO did not bind to Fg from any source compared to *L. lactis* with the pOri23 vector alone (Figure 5).

The putative MSCRAMM SpsD mediates binding of *L. lactis* to cytokeratin 10.

L. lactis expressing SpsD demonstrated strong adherence to CK10, whereas *L. lactis* expressing SpsO and SpsL did not show increased binding compared to *L. lactis* with the vector pOri23 alone (Figure 6).

The putative MSCRAMMs SpsD and SpsO, but not SpsL, mediate adherence of *L. lactis* to *ex vivo* canine corneocytes.

L. lactis expressing SpsD, SpsL, and SpsO were tested for their ability to adhere to *ex vivo* canine corneocytes in comparison to *L. lactis* with the empty vector pOri23 and *S. pseudintermedius* ED99. *L. lactis* carrying the empty vector pOri23 adhered poorly to canine corneocytes (Figure 7). For *S. pseudintermedius* ED99, the mean percentage adherence to canine corneocytes was 4.24% which was significantly different to *L. lactis* carrying pOri23 alone ($P = 0.001$) (Figure 7). *L. lactis* expressing SpsD and *L. lactis* expressing SpsO adhered to *ex vivo* canine corneocytes (Figure 7). The increase in adherence was approaching significance for SpsD ($P = 0.050$), and was significant for SpsO when expressed in *L. lactis* compared to *L. lactis* carrying pOri23 alone ($P = 0.004$). Binding of *L. lactis* expressing SpsL was not significantly different to *L. lactis* carrying pOri23 alone ($P = 0.108$), indicating that SpsL does not promote adherence to canine corneocytes (Figure 7).

Purified recombinant Sps D, Sps L and Sps O demonstrate reactivity with canine convalescent serum.

Reactivity of recombinant A domain from Sps D, Sps L and Sps O with canine convalescent serum from pyoderma cases was examined by Western affinity blot analysis (Figure 8). rSpsD, rSpsL and rSpsO all crossreacted with IgG present in the canine serum (Figure 8).

Pre-incubation with canine convalescent serum inhibits SpsL-mediated binding to fibrinogen.

The ability of the reactive antibody present in convalescent serum to inhibit SpsD and SpsL ligand binding was investigated using a modified solid phase adherence assay. Prior to inoculation into fibrinogen coated wells, PBS normalised cultures of *L. lactis* expressing SpsD and SpsL were incubated for 2 h with doubling dilutions of convalescent serum at 28 °C (Figure 9). Convalescent serum inhibited binding of *L. lactis* expressing SpsL, but not SpsD to canine fibrinogen, with complete inhibition at a final concentration of 2% v/v (Figure 9).

Discussion

In summary, genome-wide analysis of *S. pseudintermedius* ED99 revealed the presence of 17 genes encoding putative CWA proteins based on typical MSCRAMM features. All MSCRAMM characteristics searched for were identified for SpsD, SpsL, and SpsO, including a signal sequence, a non-repeated A domain with two IgG-like folds, tandemly repeated regions, and a C-terminal LPXTG-anchor motif. Interestingly, SpsD, SpsL, and SpsO belong to different groups based on Southern blot analysis, with SpsD being present in all SIG members, SpsL in *S. pseudintermedius* only, and SpsO in only six of the *S. pseudintermedius* strains investigated, and not in the other SIG species tested. Based on *in silico* analysis and *in vitro* expression data, SpsD, SpsL, and SpsO were selected for functional characterisation.

All CWA proteins and in particular, SpsD, SpsL, and SpsO could be employed in passive and active immunisation studies to test their antigenic properties, either singular or in combination, in a similar fashion as proposed for *S. aureus* ClfA (Josefsson *et al.*, 2001; Hall *et al.*, 2003; Patti, 2004; Nanra *et al.*, 2009). Further, a combinatory vaccine of *S. aureus* surface proteins IsdA, iron-regulated surface determinant protein B (IsdB), SdrD, and SdrE has proven to be highly protective in a mouse infection model (Stranger-Jones *et al.*, 2006), demonstrating the promising potential of vaccine preparations containing multiple staphylococcal CWA proteins.

In addition, MSCRAMMs with known ligands could be targets of anti-staphylococcal drug development, e.g. by generating synthetic peptides based on the interacting ECM proteins, which antagonise the MSCRAMM-host protein interaction, but do not interfere profoundly with physiological processes in the host. An excellent example is provided by Ganesh *et al.* (2008) who demonstrated that synthetic peptides, based on the Fg-binding site for ClfA, hinder the ClfA interaction, but do not block binding of the platelet integrin $\alpha_{IIb}\beta_3$ to Fg. Recently, Stranger-Jones *et al.* screened the genome of the human pathogen *S. aureus* for all genes predicted to encode CWA proteins, and immunized mice with each protein to determine their capacity to protect against lethal or invasive infection (Stranger-Jones *et al.*, 2006). Four of the proteins were combined into a multiple protein vaccine which induced high levels of protection against *S. aureus* invasive disease of mice. These data have stimulated renewed optimism in a vaccine for the prevention of human *S. aureus* infections. A similar approach could be used to design an effective vaccine for the prevention of *S. pseudintermedius* canine pyoderma.

References

- Bannoehr J, Ben Zakour NL, Waller AS, Guardabassi L, Thoday KL, van den Broek AH, Fitzgerald JR. (2007). Population genetic structure of the *Staphylococcus intermedius* group: insights into agr diversification and the emergence of methicillin-resistant strains. *J Bacteriol.* 189:8685-92
- Ben Zakour, N. L., Guinane, C. M. & Fitzgerald, J. R. (2008) Pathogenomics of the staphylococci: insights into niche adaptation and the emergence of new virulent strains. *FEMS Microbiol Lett*, 289, 1-12.
- Clarke, S. R. & Foster, S. J. (2006) Surface adhesins of *Staphylococcus aureus*. *Adv Microb Physiol*, 51, 187-224.
- Corrigan, R. M., Miajlovic, H. & Foster, T. J. (2009) Surface proteins that promote adherence of *Staphylococcus aureus* to human desquamated nasal epithelial cells. *BMC Microbiol*, 9, 22.
- Clarke, S. R., Andre, G., Walsh, E. J., Dufrene, Y. F., Foster, T. J. & Foster, S. J. (2009) Iron-regulated surface determinant protein A mediates adhesion of *Staphylococcus aureus* to human corneocyte envelope proteins. *Infect Immun*, 77, 2408-16.
- Curtis, C.F., et al (2006) Masked, controlled study to investigate the efficacy of a *Staphylococcus intermedius* autogenous bacterin for the control of canine idiopathic recurrent superficial pyoderma. *Vet Dermatol* 17, 163-8 (2006).
- Forsythe, P. J., Hill, P. B., Thoday, K. L. & Brown, J. (2002) Use of computerized image analysis to quantify staphylococcal adhesion to canine corneocytes: does breed and body site have any relevance to the pathogenesis of pyoderma? *Vet Dermatol*, 13, 29-36.
- Foster, T. J. & Hook, M. (1998) Surface protein adhesins of *Staphylococcus aureus*. *Trends Microbiol*, 6, 484-8.
- Ganesh, V. K., Rivera, J. J., Smeds, E., Ko, Y. P., Bowden, M. G., Wann, E. R., Gurusiddappa, S., Fitzgerald, J. R. & Hook, M. (2008) A structural model of the *Staphylococcus aureus* ClfA-fibrinogen interaction opens new avenues for the design of anti-staphylococcal therapeutics. *PLoS Pathog*, 4, e1000226
- Guardabassi, L., Schwarz, S. & Lloyd, D. H. (2004b) Pet animals as reservoirs of antimicrobial-resistant bacteria. *J Antimicrob Chemother*, 54, 321-32.
- Hall, A. E., Domanski, P. J., Patel, P. R., Vernachio, J. H., Syribeys, P. J., Gorovits, E. L., Johnson, M. A., Ross, J. M., Hutchins, J. T. & Patti, J. M. (2003)

Characterization of a protective monoclonal antibody recognizing *Staphylococcus aureus* MSCRAMM protein clumping factor A. *Infect Immun*, 71, 6864-70.

Hill, P.B. et al. (2006) Survey of the prevalence, diagnosis and treatment of dermatological conditions in small animals in general practice *Vet Rec* 158, 533-9 (2006).

Josefsson, E., Hartford, O., O'brien, L., Patti, J. M. & Foster, T. (2001) Protection against experimental *Staphylococcus aureus* arthritis by vaccination with clumping factor A, a novel virulence determinant. *J Infect Dis*, 184, 1572-80.

Lindsay, J. A., Moore, C. E., Day, N. P., Peacock, S. J., Witney, A. A., Stabler, R. A., Husain, S. E., Butcher, P. D. & Hinds, J. (2006) Microarrays reveal that each of the ten dominant lineages of *Staphylococcus aureus* has a unique combination of surface-associated and regulatory genes. *J Bacteriol*, 188, 669-76.

Mazmanian, S.K., et al. (1999) *Staphylococcus aureus* sortase, an enzyme that anchors surface proteins to the cell wall. *Science* 285, 760-3 (1999).

Nanra, J. S., Timofeyeva, Y., Buitrago, S. M., Sellman, B. R., Dilts, D. A., Fink, P., Nunez, L., Hagen, M., Matsuka, Y. V., Mininni, T., Zhu, D., Pavliak, V., Green, B. A., Jansen, K. U. & Anderson, A. S. (2009) Heterogeneous in vivo expression of clumping factor A and capsular polysaccharide by *Staphylococcus aureus*: implications for vaccine design. *Vaccine*, 27, 3276-80.

Otto, M. (2008) Targeted immunotherapy for staphylococcal infections : focus on anti-MSCRAMM antibodies. *BioDrugs* 22, 27-36 (2008)

Patti, J. M. (2004) A humanized monoclonal antibody targeting *Staphylococcus aureus*. *Vaccine*, 22 Suppl 1, S39-43.

Pizza, M. et al. (2008) Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. *Science* 287, 1816-1820

Table 1. Sequence homology of the 17 predicted cell-wall-anchored proteins against known proteins in the public domain.

Putative CWA protein	Best Hit (BLAST)	Identity (%)	Similarity (%)
SpsA	LPXTG cell-wall surface anchor family protein of <i>S. aureus</i> COL	31.2	56.9
SpsB	RodA, a rod shape determining protein of <i>S. epidermidis</i> ATCC 12228	69.7	87.8
SpsC	bifunctional autolysin precursor of <i>S. epidermidis</i> ATCC 12228	50.7	65.9
SpsD	Fnbp protein homolog of <i>S. aureus</i> Mu50	40.7	59.1
SpsE	Fibrinogen binding protein of <i>S. epidermidis</i> ATCC 12228	78.6	90.1
SpsF	hypothetical protein, similar to the putative cell-surface adhesin SdrF of <i>S. haemolyticus</i> JCSC1435	52.8	69.3
SpsG	hypothetical protein, cell-wall surface anchor family of <i>Streptococcus pneumoniae</i> D39	47.7	63.6
SpsH	Sdr-repeat family protein SdrH, <i>S. aureus</i> USA300	36.0	53.1
SpsI	serine-aspartate rich, fibrinogen-binding, bone sialoprotein-binding protein <i>S. epidermidis</i> ATCC 12228	37.3	55.5
SpsJ	precursor of a serine-rich adhesin for platelets of <i>S. haemolyticus</i> JCSC1435	52.2	61.2
SpsK	IgG-binding protein of <i>S. aureus</i> COL	50.4	71.1
SpsL	Fnbp protein homolog of <i>S. aureus</i> Mu50	33.4	51.7
SpsM	hypothetical protein, similar to the putative cell-surface adhesin SdrF, <i>S. haemolyticus</i> JCSC1435	44.4	61.7
SpsN	probable exported protein of <i>S. aureus</i> RF122	38.0	60.0
SpsO	serine-aspartate repeat-containing protein C precursor of <i>Staphylococcus warneri</i> L37603	50.0	68
SpsP	LPXTG-motif cell wall anchor domain of <i>S. aureus</i> JH9	60.6	74.3
SpsQ	IgG-binding protein A precursor of <i>S. aureus</i> MRSA252	57.0	71.7

Table 2. Main characteristics of the predicted CWA proteins SpsD, SpsL, and SpsO of *S. pseudintermedius* ED99.

Amino acids		MW (kDa) ^a	Signal peptide	LPXTG motif	Ig-like fold ^b (position)	TYTFTDYVD-like motif (position) ^b	Putative latching sequence (position) ^b	Repeat region (position)	Copy number repeats
SpsD	1031	115	36 aa	LPDTG	167-320 aa 322-519 aa	RYRFMDYVN (267-275 aa)	NNASGEG (491-497 aa)	867-959 aa	5
SpsL	930	103	38 aa	LPKTG	220-363 aa 364-531 aa	VYTFKDYVN (298-306 aa)	NSASGSG (502-508 aa)	543-818 aa	7
SpsO	1846	198	44 aa	LPNTG	339-492 aa 487-659 aa	TYTFTDYVD (424-432 aa)	DKSTALG (635-641 aa)	661-1800 aa 97-216 aa ^c	96 4 ^c

aa = amino acids; ^aMW = predicted approximate molecular weight in kDa (kilo dalton); ^bwithin the A domain; ^cN-terminal repeats

Claims

1. A *Staphylococcus pseudintermedius* protein or nucleic acid comprising a sequence at least 65% homologous or identical to (i) the nucleic acid of SEQ ID NO: 5 or a fragment thereof or (ii) the amino acid sequence of SEQ ID NO: 6 or an antigenic fragment thereof, for use in raising an immune response in a subject.
2. A vaccine for use in treating, preventing or controlling a disease caused or contributed to by *Staphylococcus pseudintermedius* in a subject, said vaccine comprising nucleic acid sequence or peptide/protein having a sequence at least 65% homologous or identical to the nucleic acid of SEQ ID NO: 5 or amino acid sequence of SEQ ID NOS: 6 or an antigenic fragment thereof.
3. A *Staphylococcus pseudintermedius* protein or nucleic acid comprising a sequence at least 65% homologous or identical to a nucleic acid or amino acid sequence selected from the group consisting of SEQ ID NOS: 1-4 and 7-36; and fragments of any of SEQ ID NOS: 1-4 and 7-36, for use in raising an immune response in a subject.
4. A vaccine for use in treating, preventing or controlling disease caused or contributed to by *Staphylococcus pseudintermedius* in a subject, said vaccine comprising one or more amino acid or nucleic acid sequence(s) at least 65% homologous or identical to a nucleic acid or amino acid sequence selected from the group consisting of SEQ ID NOS: 1-4 and 7-36; and fragments of any of SEQ ID NOS: 1-4 and 7-36.
5. The *Staphylococcus pseudintermedius* protein or nucleic acid for use of claim 1 or vaccine of claim 2, further comprising one or more protein or nucleic acid sequence(s) at least 65% homologous or identical to an amino acid or nucleic acid sequence selected from the group consisting of SEQ ID NOS 1-4 and 7-36; and fragments of any of SEQ ID NOS: 1-4 and 7-36.
6. A method of raising an immune response to *Staphylococcus pseudintermedius* in a subject, said method comprising immunising said subject with, or administering said subject a vaccine according to claim 2, 4 or 5

7. The protein/nucleic acid for use of claims 1, 3 and 5, vaccine of claim 2, 4 or 5 or method of claim 6, wherein the subject is a canine.
8. The vaccine of claims 2, 4, 5 or 6, wherein the disease caused or contributed to by *Staphylococcus pseudintermedius* is canine pyoderma.
9. The vaccine of claim 2, 4, 5 or 8, wherein the vaccine further comprises one or more adjuvant(s) and/ or antigens for use in treating or preventing another diseases and/or condition.
10. An isolated and/or substantially purified or recombinant *Staphylococcus pseudintermedius* nucleic acid or protein sequence encoding a cell-wall anchored (CWA) or microbial surface components recognising adhesive matrix molecule (MSCRAMM) wherein said nucleic acid sequence comprises a nucleic acid at least 65% homologous or identical to the nucleic acid sequence of SEQ ID NO: 5 or fragments, portions, mutants, derivatives and/or homologues/orthologues thereof and said protein sequence comprises an amino acid sequence at least 65% homologous or identical to an amino acid sequence of SEQ ID NO: 6 or fragments, portions, mutants, derivatives and/or homologues/orthologues thereof.
11. An isolated and/or substantially purified or recombinant *Staphylococcus pseudintermedius* nucleic acid or protein sequence encoding a cell-wall anchored (CWA) or microbial surface components recognising adhesive matrix molecule (MSCRAMM) wherein said nucleic acid or protein sequence comprises a nucleic acid or amino acid sequence at least 65% homologous or identical to a nucleic acid or amino acid sequences selected from the group consisting of SEQ ID NOS: 1-36 or fragments, portions, mutants, derivatives and/or homologues/orthologues thereof.
12. A nucleic acid comprising a replicable expression vector and nucleic acid according to claims 10 or 11.
13. A host cell transformed with the nucleic acid of claim 12.

14. A process for the production of recombinant *S. pseudintermedius* CWA or MSCRAMM proteins, said process comprising the steps of:
- (a) transforming a host cell with the nucleic acid of claim 12;
 - (b) culturing the cells obtained in (a) under conditions in which expression of a peptide/protein encoded by the nucleic acid occurs; and
 - (c) isolating the expressed peptide/protein from the cell culture or/and a culture supernatant derived therefrom.
15. An antibody that selectively and/or specifically binds to, or exhibits affinity for the *S. pseudintermedius* CWA or MSCRAMM proteins encoded by the nucleic acid of claim 10 or 11 and/or proteins recombinantly produced by the process of claim 14.
16. A method of treating, preventing or controlling a *Staphylococcus pseudintermedius* infection or disease in a canine, said method comprising administering to the canine one or more of the nucleic acid(s)/protein(s) provided by claims 10 or 11.
17. A compound for use in treating an infection or disease caused or contributed to by *Staphylococcus pseudintermedius*, wherein said compound antagonises (i) the function, expression and/or activity of a protein and/or nucleic acid encoded by one or more of the sequences provided by claim 1 or (ii) the interaction between a protein and/or nucleic acid having the sequence of claim 10 or 11 and a native ligand.
18. The compound for use of claim 17, wherein the compound is a small organic molecule, antibody, peptide or carbohydrate
19. The compound for use of claim 17 or 14, wherein the compound comprises an antibody according to claim 5 or a polypeptide or polynucleotide sequence comprising a fragment, portion, mutant, derivative and/or homologue/orthologue of one or more of the sequences given as SEQ ID NOS: 5/6 and/or 1-4, 7-36.
20. A method of diagnosing infections, diseases and/or conditions caused or contributed to by *S. pseudintermedius*, said methods comprising the steps of identifying in a sample a level of a protein, peptide or nucleic acid (for example a gene) having a sequence homologous or identical to a sequence provided by SEQ ID NOS: 5/6 and/or 1-

4, 7-36 or a fragment, portion, mutant, derivative and/or homologue/orthologue thereof, wherein detection of a level of a protein, peptide or nucleic acid (for example a gene) having a sequence homologous or identical to a sequence provided by SEQ ID NOS: 5/6 or 1-4, 7-36 in sample, indicates that the sample may have been provided by a subject suffering from a *S. pseudintermedius* infection, disease and/or condition.

21. The method of claim 20, wherein the sample comprises canine skin cells.

22. The method of claims 20 or 21, wherein the detection of a level of a protein, peptide or nucleic acid comprises the use of a molecular detection technique and/or an immunological detection technique.

23. A kit for diagnosing, detecting and/or evaluating possible *S. pseudintermedius* infections, diseases and/or conditions, said kit comprising one or more components selected from the group consisting of:

(i) substrates having *S. pseudintermedius* proteins or agents capable of binding *S. pseudintermedius* proteins, bound thereto;

(ii) antibodies which exhibit specificity and/or selectivity for one or more *S. pseudintermedius* proteins; and

(iii) oligonucleotides/primers for detecting/amplifying/probing samples for *S. pseudintermedius* protein encoding sequences.

24. The kit of claim 19, wherein the *S. pseudintermedius* proteins or *S. pseudintermedius* protein encoding sequences comprise sequences corresponding to (or homologous/identical to) to SEQ ID NOS: 5/6 and/or 1-4, 7-36 and/or fragments, portions, mutants, derivatives and/or homologues/orthologues thereof.

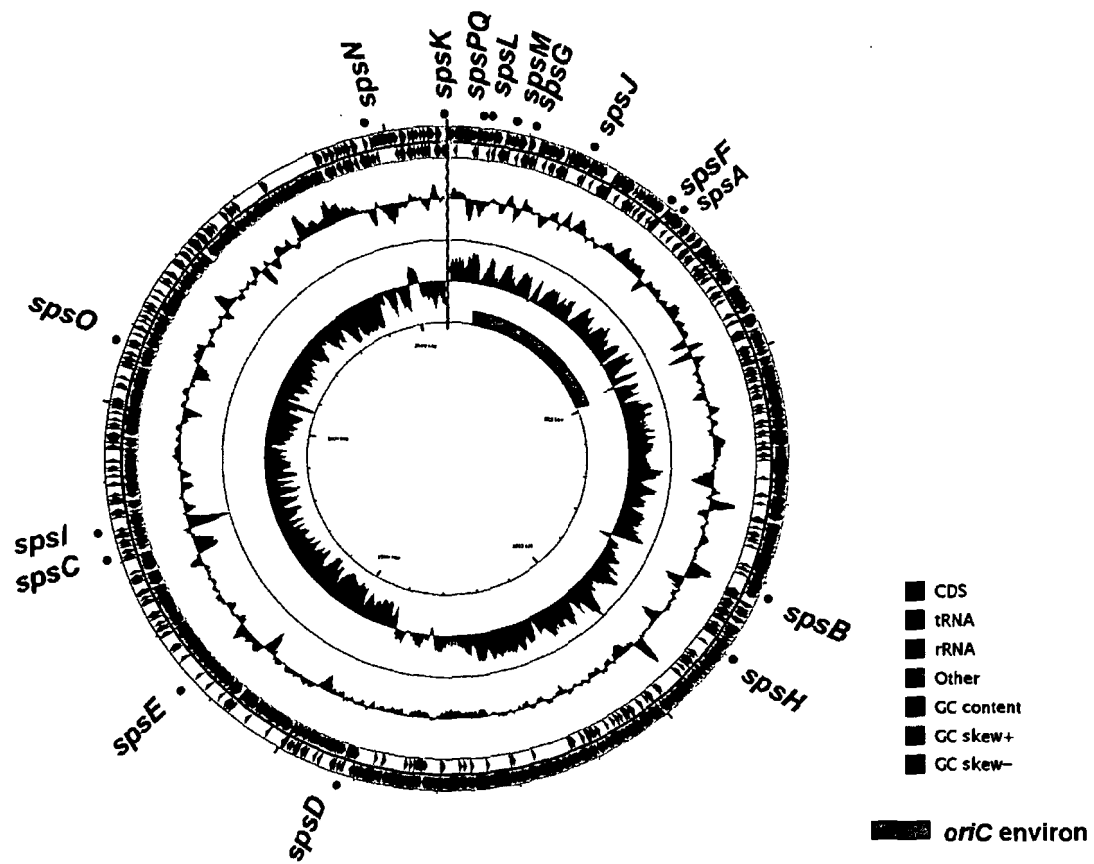


Figure 1

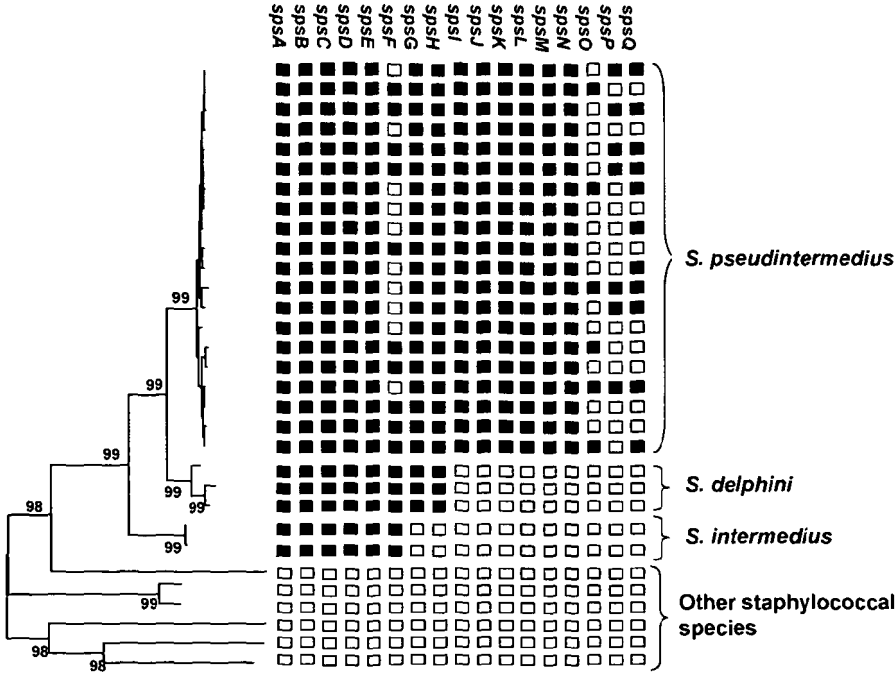


Figure 2

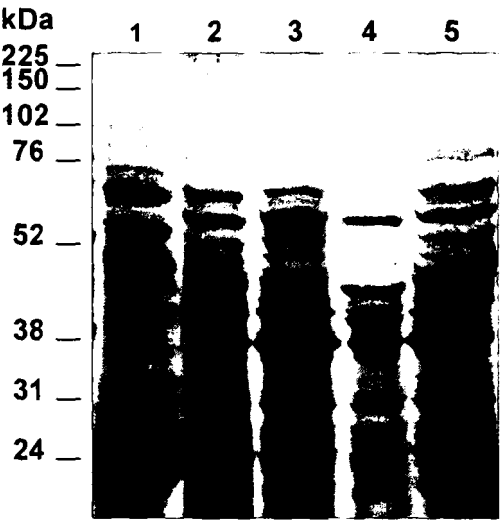


Figure 3 A

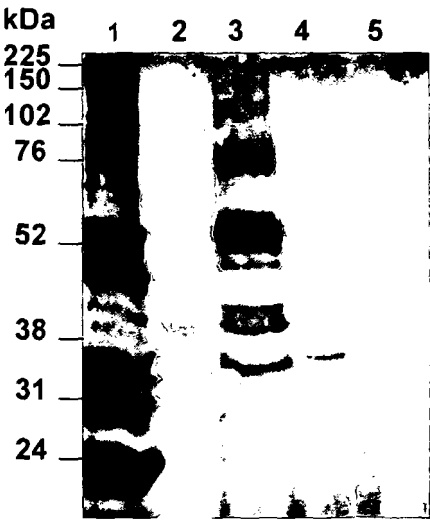


Figure 3B

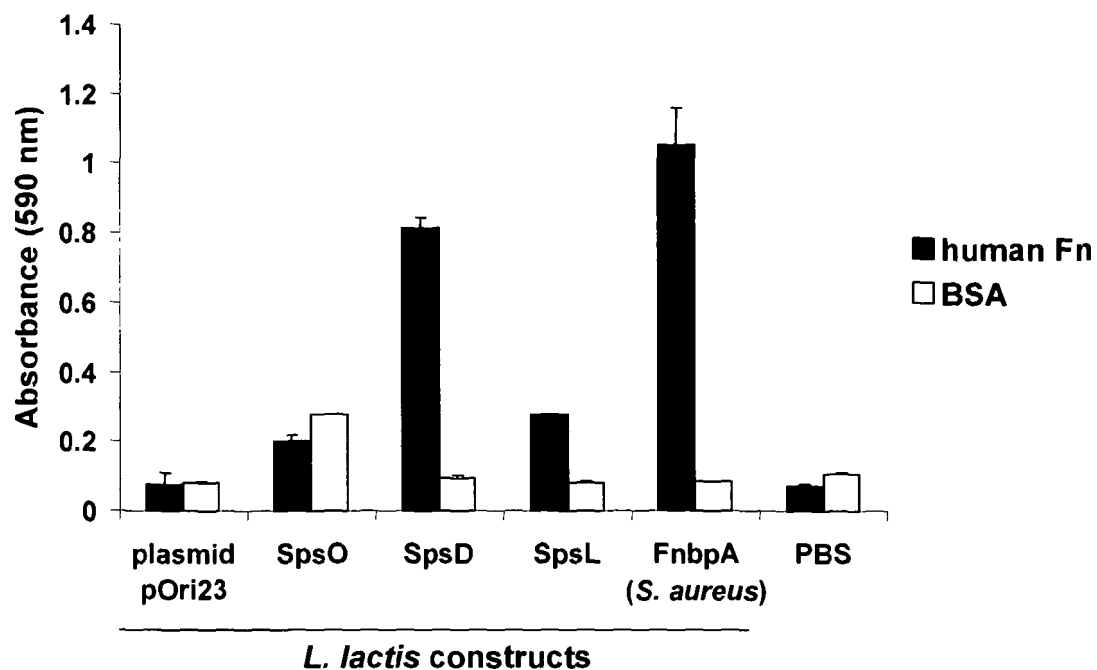


Figure 4

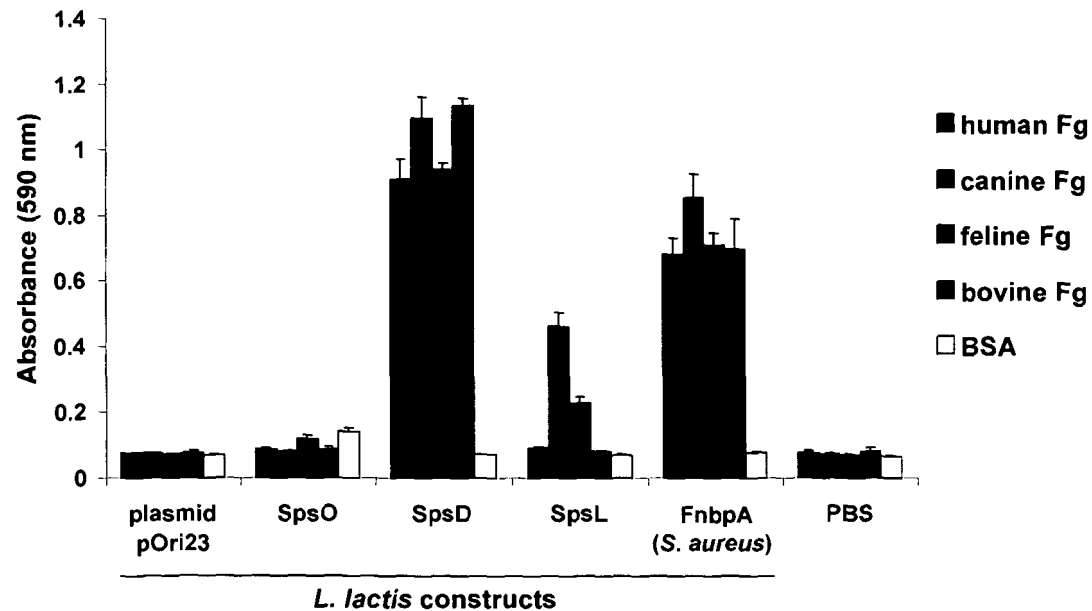


Figure 5

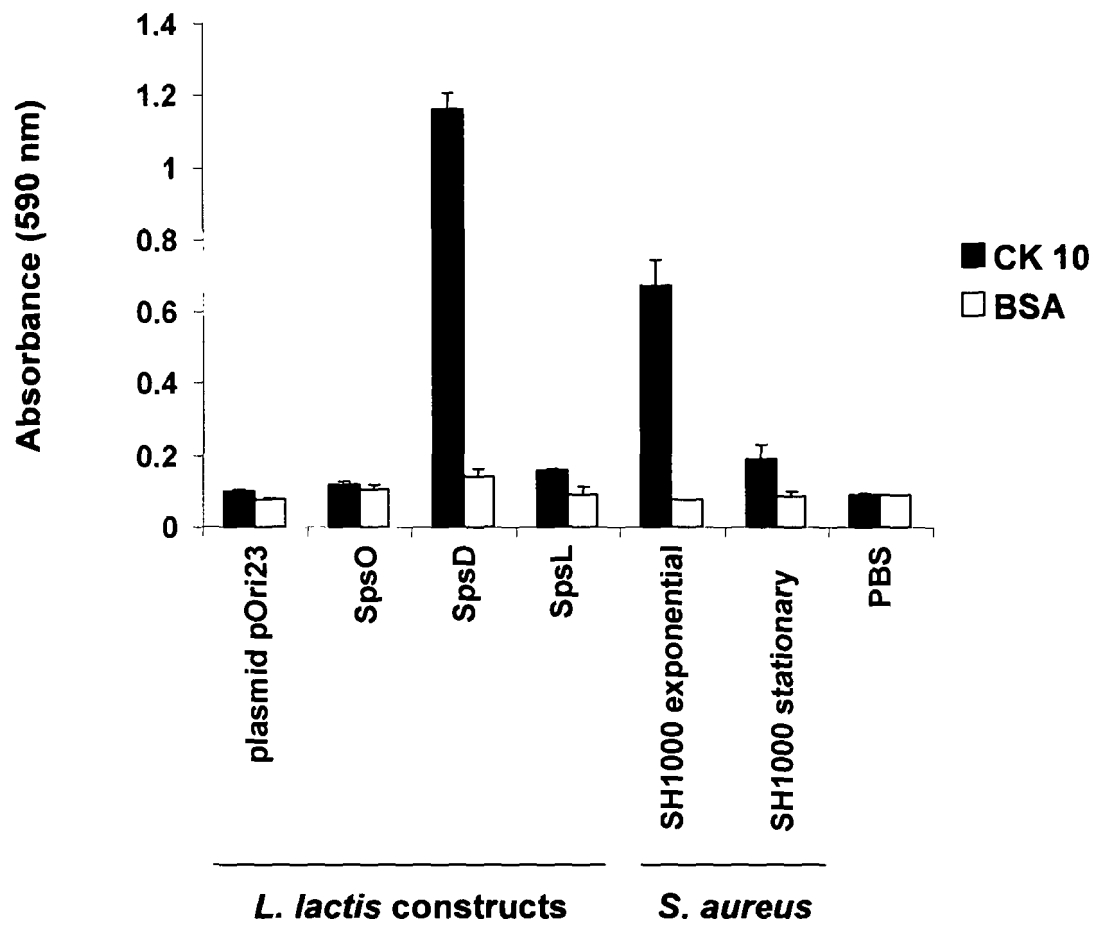


Figure 6

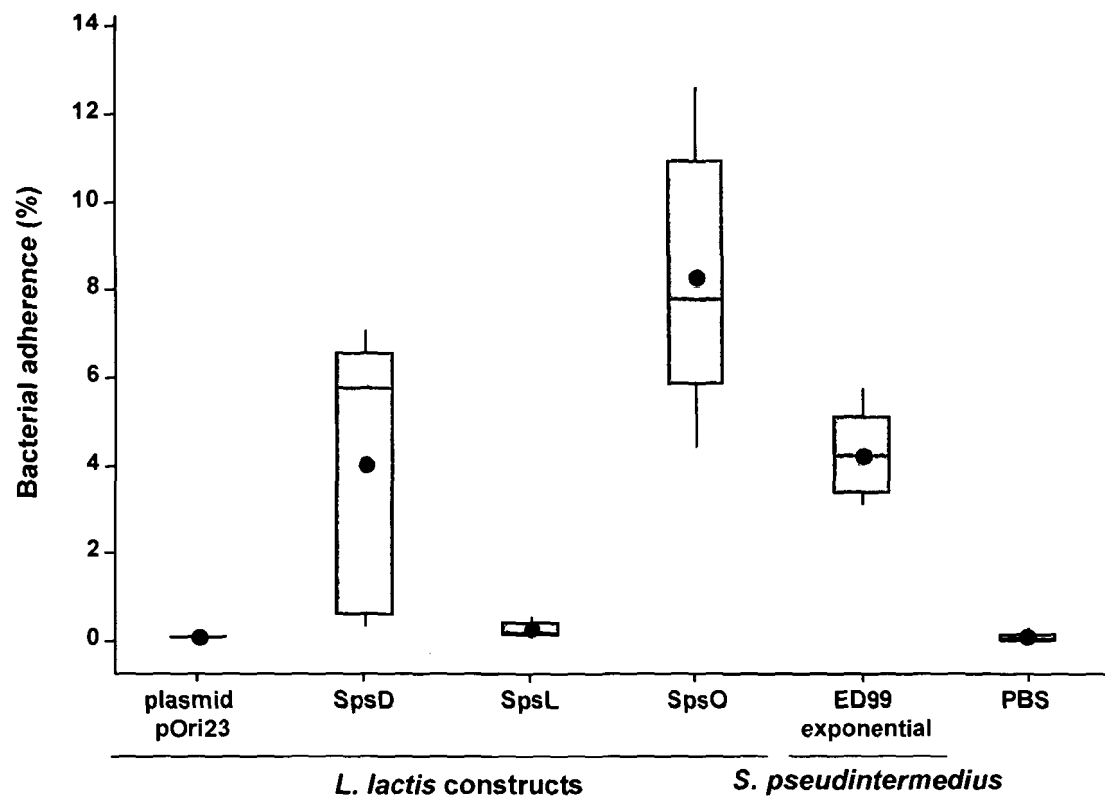


Figure 7

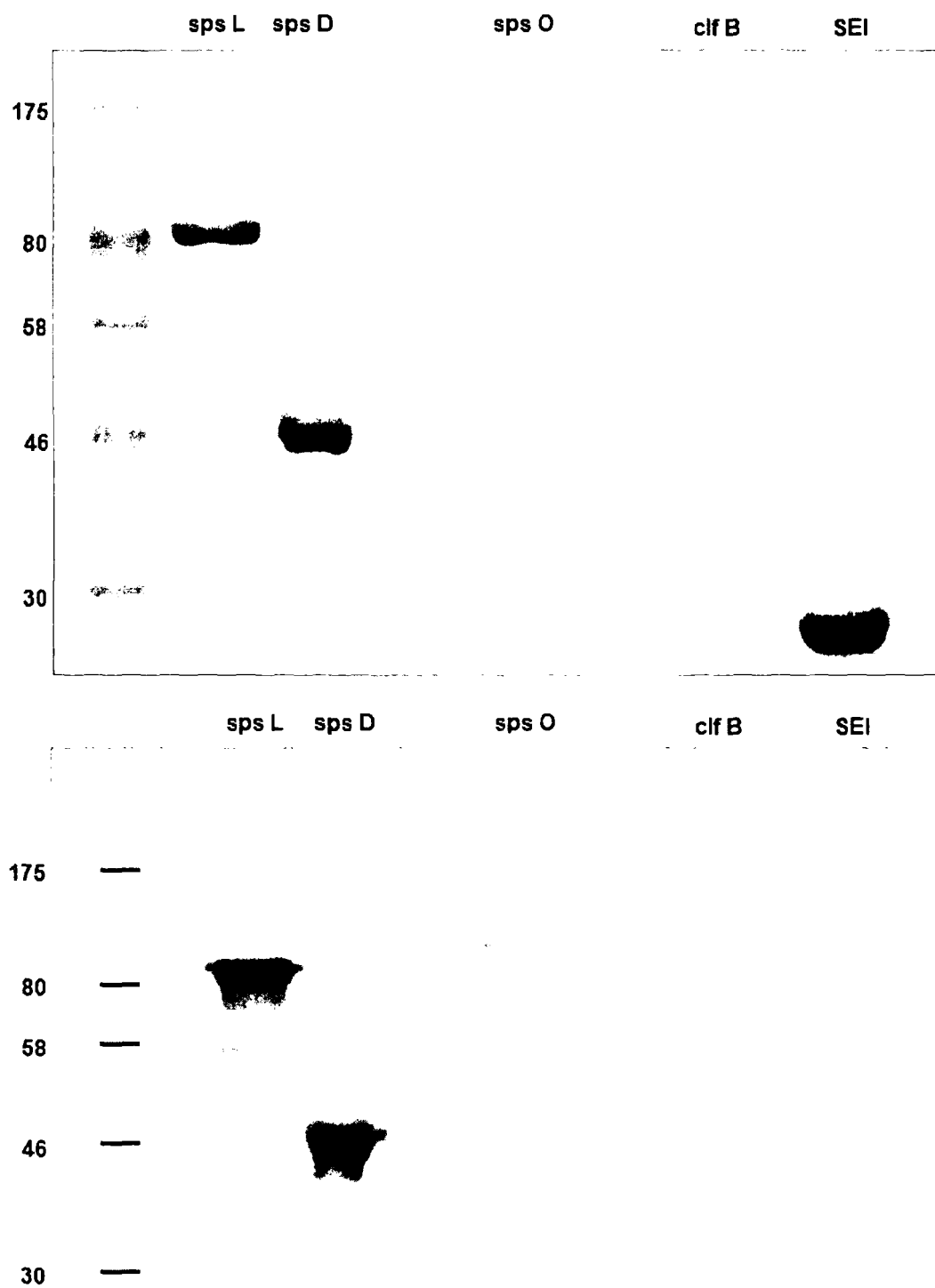
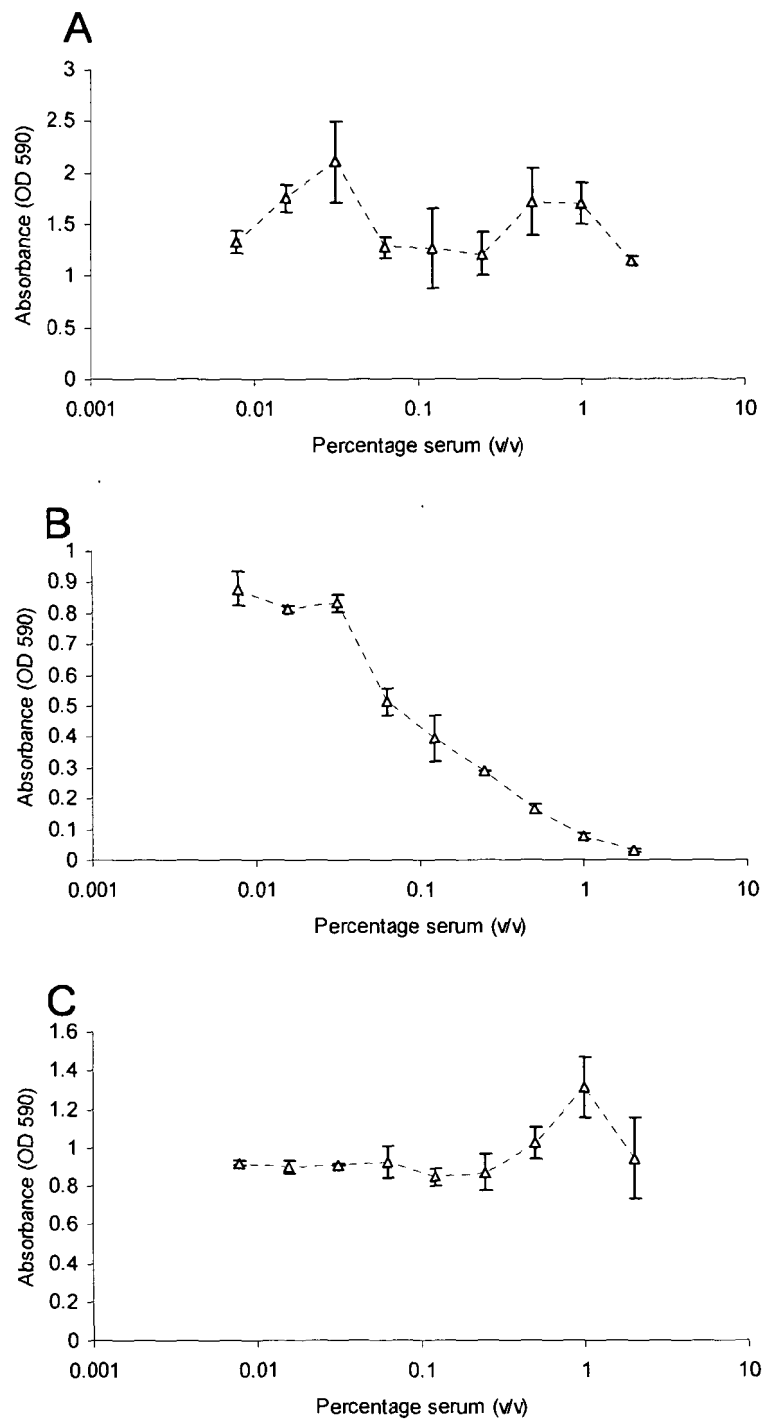


Figure 8 A & B



5 Figure 9