VIRULENCE GENES, PROTEINS, AND THEIR USE

Field of the Invention

This invention relates to virulence genes and proteins, and their use. More particularly, it relates to genes and proteins/peptides obtained from *Streptococcus pyogenes*, and their use in therapy and in screening for drugs.

Background of the Invention

Group A Streptococcus (GAS) is responsible for the majority of Streptococcal illnesses. An organism of particular interest is *S. pyogenes*, which is implicated in a wide range of non-invasive and invasive infections, such as impetigo, pharyngitis, necrotizing fasciitis, bacteraemia, streptococcal toxic shock syndrome (STSS), pneumonia and rheumatic fever.

Some GAS infections can be treated with antibiotics, including penicillin and erythromycin. However, due to the problems associated with resistance to antibiotics, and antibiotic-allergic patients, there is a need for further therapeutics which may be useful in treating of preventing GAS infection.

Summary of the Invention

The present invention is based on the discovery of virulence genes in *S. pyogenes*.

According to a first aspect of the invention, a peptide of the invention is encoded by an operon including any of the nucleotide sequences identified herein as SEQ ID NOS. 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59 and 61 of *S. pyogenes* or a homologue thereof in a Gram-positive bacterium, or a functional fragment thereof, for therapeutic or diagnostic use.

The peptides may have many therapeutic uses for treating Group A Streptococcal infections, including use in vaccines for prophylactic application.

According to a second aspect, a polynucleotide encoding a peptide defined above, may also be useful for therapy or diagnosis.

According to a third aspect, the genes that encode the peptides may be utilised to prepare attenuated microorganisms. The attenuated microorganisms will usually have a mutation that disrupts the expression of one or more of the genes identified herein, to provide a strain that lacks virulence. These microorganisms will also have use in therapy and diagnosis.

According to a fourth aspect, the peptides, genes and attenuated microorganisms according to the invention may be used in the treatment or prevention of a condition associated with infection by *Streptococcal* or Gram-positive bacteria.

Description of the Invention

The present invention is based on the discovery of genes encoding peptides which are implicated in virulence. The peptides and genes of the invention are therefore useful for the preparation of therapeutic agents to treat infection. It should be understood that references to therapy also include preventative treatments, e.g. vaccination. Furthermore, while the products of the invention are intended primarily for treatment of infections in human patients, veterinary applications are also considered to be within the scope of the invention.

The present invention is described with reference to *Streptococcus pyogenes*. However, all the Group A Streptococcal strains, and many other Gram-positive bacterial strains are likely to include related peptides or proteins having amino acid sequence identity or similarity to those identified herein. Organisms likely to contain the peptides include, but are not limited to the genera *Lactococcus*, *Enterococcus*, *Streptococcus* and *Staphylococcus*.

Preferably, the peptides that may be useful in the various aspects of the invention have greater than a 40% similarity with the peptides identified herein. More preferably, the peptides have greater than 60% sequence similarity. Most preferably, the peptides have greater than 80% sequence similarity, e.g. 95% similarity. With regard to the polynucleotide sequences identified herein, related

polynucleotides that may be useful in the various aspects of the invention may have greater than 40% identity with the sequences identified herein. More preferably, the polynucleotide sequences have greater than 60% sequence identity. Most preferably, the polynucleotide sequences have greater than 80% sequence identity, e.g. 95% identity.

The terms "similarity" and "identity" are known in the art. The use of the term "identity" refers to a sequence comparison based on identical matches between correspondingly identical positions in the sequences being compared. The term "similarity" refers to a comparison between amino acid sequences, and takes into account not only identical amino acids in corresponding positions, but also functionally similar amino acids in corresponding positions. Thus similarity between polypeptide sequences indicates functional similarity, in addition to sequence similarity.

Levels of identity between gene sequences and levels of identity or similarity between amino acid sequences can be calculated using known methods. In relation to the present invention, publicly available computer based methods for determining identity and similarity include the BLASTP, BLASTN and FASTA (Atschul *et al.*, J. Molec. Biol., 1990; 215:403-410), the BLASTX program available from NCBI, and the Gap program from Genetics Computer Group, Madison WI. The levels of similarity and identity provided herein, were obtained using the Gap program, with a Gap penalty of 12 and a Gap length penalty of 4 for determining the amino acid sequence comparisons, and a Gap penalty of 50 and a Gap length penalty of 3 for the polynucleotide sequence comparisons.

Having characterised a gene according to the invention, it is possible to use the gene sequence to search for related genes or peptides in other microorganisms. This may be carried out by searching in existing databases, e.g. EMBL or GenBank.

Peptides or proteins according to the invention may be purified and isolated by methods known in the art. In

particular, having identified the gene sequence, it will be possible to use recombinant techniques to express the genes in a suitable host. Active fragments and related molecules can be identified and may be useful in therapy. For example, the peptides or their active fragments may be used as antigenic determinants in a vaccine, to elicit an immune response. They may also be used in the preparation of antibodies, for passive immunisation, or diagnostic applications. Suitable antibodies include monoclonal antibodies, or fragments thereof, including single chain Fv fragments. Methods for the preparation of antibodies will be apparent to those skilled in the art.

Active fragments of the peptides are those that retain the biological function of the peptide. For example, when used to elicit an immune response, the fragment will be of sufficient size, such that antibodies generated from the fragment will discriminate between that peptide and other peptides on the bacterial microorganism. Typically, the fragment will be at least 30 nucleotides (10 amino acids) in size, preferably 60 nucleotides (20 amino acids) and most preferably greater than 90 nucleotides (30 amino acids) in size.

It should also be understood, that in addition to related molecules from other microorganisms, the invention encompasses modifications made to the peptides and polynucleotides identified herein which do not significantly alter the biological function. It will be apparent to the skilled person that the degeneracy of the genetic code can result in polynucleotides with minor base changes from those specified herein, but which nevertheless encode the same peptides. Complementary polynucleotides are also within the invention. Conservative replacements at the amino acid level are also envisaged, i.e. different acidic or basic amino acids may be substituted without substantial loss of function.

The preparation of vaccines based on attenuated microorganisms is known to those skilled in the art. Vaccine compositions can be formulated with suitable

carriers or adjuvants, e.g. alum, as necessary or desired, to provide effective immunisation against infection. The preparation of vaccine formulations will be apparent to the skilled person. The attenuated microorganisms may be prepared with a mutation that disrupts the expression of any of the genes identified herein. The skilled person will be aware of methods for disrupting expression of particular genes. Techniques that may be used include insertional inactivation or gene deletion techniques. Attenuated microorganisms according to the invention may also comprise additional mutations in other genes, for example in a second gene identified herein or in a separate gene required for growth of the microorganism, e.g. an *aro* mutation. Attenuated microorganisms may also be used as carrier systems for the delivery of heterologous antigens, therapeutic proteins or nucleic acids (DNA or RNA). In this embodiment, the attenuated microorganisms are used to deliver a heterologous antigen, protein or nucleic acid to a particular site *in vivo*. Introduction of a heterologous antigen, peptide or nucleic acid into an attenuated microorganism can be carried out by conventional techniques, including the use of recombinant constructs, e.g. vectors, which comprise polynucleotides that express the heterologous antigen or therapeutic protein, and also include suitable promoter sequences. Alternatively, the gene that encodes the heterologous antigen or protein may be incorporated into the genome of the organism and the endogenous promoters used to control expression.

More generally, and as is well known to those skilled in the art, a suitable amount of an active component of the invention can be selected, for therapeutic use, as can suitable carriers or excipients, and routes of administration. These factors would be chosen or determined according to known criteria such as the nature/severity of the condition to be treated, the type and/or health of the subject etc.

In a separate embodiment, the products of the invention may be used in screening assays for the identification of

potential antimicrobial drugs or for the detection for virulence. Routine screening assays are known to those skilled in the art, and can be adapted using the products of the invention in the appropriate way. For example, the products of the invention may be used as the target for a potential drug, with the ability of the drug to inactivate or bind to the target indicating its potential antimicrobial activity.

The various products of the invention may also be used in veterinary applications.

The following is a brief overview of the experimental procedure used to identify the virulence genes. The virulence genes in *S. pyogenes* were identified by using signature-tagged mutagenesis (STM) to screen an *S. pyogenes* mutant bank for attenuated mutants (Hensel *et al.*, 1995. Science 269(5222):400-3).

Mutants were generated via Tn917 transposon insertion, using a plasmid vector. In addition to a fragment of Tn917, the vector comprised a spectromycin-resistance gene, a chloramphenicol-resistance gene (CAT gene) with a synthetic promoter, a Gram-negative origin of replication (rop.ori) and a cloning site for the STM tags. After ligating the tags into the vector, *E. coli* transformation was carried out, and 96 plasmids that hybridised with the original tags were selected.

The *S. pyogenes* strain B514-SM (type M50) was transformed with each of the 96 tagged plasmids, and transformants were selected by resistance to spectinomycin and chloramphenicol. For each of the transformed *S. pyogenes* strains, 20 mutants were generated via Tn917 transposon insertion to create a mutant bank.

The mutant bank was screened in either a skin invasive lesion model of mouse infection (Schrager *et al.*, J. Clinical Investigation 1996; 98:1954-1958) or from a throat colonisation model of mouse infection (Husmann *et al.*, Infection and Immunity, 1997; 65 (4):935-944).

In the skin model, five CD1 mice were each inoculated intradermally with $1x10^8$ cells in a volume of 50 μl

representing the collection of 96 distinct and readily distinguishable mutants. 48 hours after inoculation, samples were taken and bacteria recovered. The skin lesions were macerated in 2xBHI medium and bacteria liberated from the lesion by treatment in a stomacher for 10 minutes. The released bacteria were plated out and a minimum of 10,000 colonies recovered from a minimum of 3 mice. DNA was isolated from these samples and used to amplify the tagged DNA present in the recovered bacteria. The DNA isolated from each recovered pool was used as a hybridisation probe to reveal those mutants in each pool that failed to be recovered from the animals and which were therefore attenuated in this animal model of infection.

In the throat colonisation model, $2x10^8$ cells were inoculated intranasally into six C57BL/6 mice and samples taken after 48 hours. As with the skin model, bacteria containing a transposon Tn917 insertion within a virulence gene failed to be recovered from mice inoculated with a mixed population of mutants, and were therefore likely to be attenuated.

Additional experiments were carried out on mutants identified through the STM screen to determine the competitive index (CI). Individual mutants were tested in mixed infections with the wild-type strain in the skin lesion model of infection (Chiang, S.L. and Mekalanos, J.J. Molecular Microbiology 1998; 27(4):797-805). As for the initial screen, groups of four CD1 mice were inoculated with equal numbers of both wild-type and mutant cells, to a total number of $1x10^8$ cells. Bacteria recovered after 48 hours were plated out onto selective media that allows the wild-type and mutant colonies to be distinguished. The ratio of mutant bacteria to wild-type bacteria seen in the inoculum compared with the ratio in recovered bacteria is the competitive index (ratio of mutants versus wild-type in the inoculum divided by the ratio of mutants versus wild-type bacteria recovered from the animal model).

The following Examples illustrate the invention.

Example 1

A first mutant was identified and the nucleotide sequence immediately following the transposon insertion was cloned.

The nucleotide sequence shows 100% identity at the nucleotide level to a coding sequence within the *S. pyogenes* genome, shown as SEQ ID NO. 1, with the putative protein sequence shown as SEQ ID NO. 2.

The amino acid sequence of the predicted protein product shows 43% identity to the putative NAD(P)H nitroreductase of *H. influenzae* (accession number: SW: Q57431).

Given the similarity of the putative NAD(P)H nitroreductase gene in *S. pyogenes* to the gene in *H. influenzae*, the skilled person will appreciate that similar sequences in other *Streptococci* and Gram-positive bacteria may also be implicated in virulence.

In the test for attenuation of virulence, the mutated microorganism was shown to be attenuated with a competitive index (CI) of 0.644.

Example 2

A further mutant was identified and the nucleotide sequence immediately following the transposon insertion was cloned.

The nucleotide sequence has 98% identity at the nucleotide level to a sequence within the *S. pyogenes* genome, shown as SEQ ID NO. 3. The putative amino acid sequence is shown as SEQ ID NO. 4.

The amino acid sequence shows 81% identity to a probable integrase enzyme of *S. mutans* (accession number: TREMBL: O69155).

This demonstrates that the disrupted gene is at least partially identical to a probable integrase gene of *S. mutans*. However, this gene was previously unknown in *S. pyogenes*, and has not been assigned a role in virulence.

Given the similarity of the *S. pyogenes* gene to the probable integrase gene of *S. mutans*, the skilled person will appreciate that similar sequences in other *Steptococci*

and Gram-positive bacteria may also be virulence determinants.

In the test for attenuation of virulence, the mutated microorganism was shown to be attenuated with a competitive index (CI) of 0.229.

Example 3

A further mutant was identified and the nucleotide sequence immediately following the transposon insertion was cloned.

The nucleotide sequence has 97% identity at the nucleotide level to a sequence within the *S. pyogenes* genome, shown as SEQ ID NO. 5. The predicted amino acid sequence is shown as SEQ ID NO. 6.

The amino acid sequence shows 37% identity at the amino acid level to the GlgP protein of the unicellular cyanobacterium *Synechocystis spp.* (accession number: TREMBL: P73511).

This demonstrates that the disrupted gene is at least partially identical to the *glgP* gene of *Synechocystis spp.*

Given the similarity of the gene of *S. pyogenes* to the *glgP* gene of *Synechocystis spp.*, the skilled person will appreciate that similar sequences in other *Streptococci* and Gram-positive bacteria may also be virulence determinants.

In the test for attenuation of virulence, the mutated microorganism was shown to be attenuated with a competitive index (CI) of 0.011.

Example 4

A further mutant was identified and the nucleotide sequence immediately following the transposon insertion was cloned.

The nucleotide sequence has 97% identity at the nucleotide level to a sequence within the *S. pyogenes* genome, shown as SEQ ID NO. 7. The predicted amino acid sequence is shown as SEQ ID NO. 8.

The amino acid sequence shows 35% identity at the amino acid level to the *BraB* protein of *B. subtilis* (accession number: TREMBL: O34545).

This demonstrates that the disrupted gene is at least partially identical to the *braB* gene of *B. subtilis*.

Given the similarity of the *braB* gene of *S. pyogenes* to the *braB* gene in *B. subtilis*, the skilled person will appreciate that similar sequences in other *Streptococci* and Gram-positive bacteria may also be implicated in virulence.

In the test for attenuation of virulence, the mutated microorganism was shown to be attenuated with a competitive index (CI) of 0.88.

Example 5

A further mutant was identified and the nucleotide sequence immediately following the transposon insertion was cloned.

The nucleotide sequence shows 99% identity at the nucleotide level to a sequence within the *S. pyogenes* genome, shown as SEQ ID NO. 9. The predicted amino acid sequence is shown as SEQ ID NO. 10.

A still further attenuated mutant was also identified with a nucleotide sequence having 98% identity at the nucleotide level to a sequence within the *S. pyogenes* genome, shown as SEQ ID NO. 11. The predicted amino acid sequence is shown as SEQ ID NO. 12.

The predicted protein of the first mutant shows 53% identity at the amino acid level to the AdcR protein of *S. pneumoniae* (accession number: TREMBL:O33703). That of the second mutant shows 79% identity at the amino acid level to the AdcC protein of *S. pneumoniae* (accession number:TREMBL: O87862).

This demonstrates that the disrupted genes are at least partially identical to the *adcR* and *adcC* genes of *S. pneumoniae*. The *adcR and adcC* genes are part of the *adc* operon (including *adcR*, *adcC*, *adcB*, and *adcA*) in *S. pneumoniae*. Therefore the attenuation of *S. pyogenes adcR* and *adcC* mutants could result from a failure to express the *adcR* and *adcC genes*, or the predicted downstream genes (*adcB* and *abcA* homologues). These genes have not previously been assigned a role in virulence in *S. pyogenes*.

Given the similarity of the *S. pyogenes* genes to the *adcR* and *adcC* genes in *S. pneumoniae*, the skilled person will appreciate that similar sequences in other *Steptococci* and Gram-positive bacteria may be implicated in virulence.

In the test for attenuation of virulence, the mutated microorganisms were shown to be attenuated with a competitive index (CI) of 0.548 (*adcR*) and 0.028 (*adcC*).

Example 6

A further mutant was identified and the nucleotide sequence immediately following the transposon insertion was cloned.

The nucleotide sequence has 99% identity at the nucleotide level to a sequence within the *S. pyogenes* genome, shown as SEQ ID NO. 13. The predicted amino acid sequence is shown as SEQ ID NO. 14.

The amino acid sequence shows 39% identity at the amino acid level to the DNA repair protein *radC* homologue (*orfB*) of *B. subtilis* (accession number: SW: Q02170).

This demonstrates that the disrupted gene is at least partially identical to the *orfB* gene of *B. subtilis*.

Given the similarity of the *orfB* gene of *S. pyogenes* to that of *B. subtilis*, the skilled person will appreciate that similar sequences in other *Streptococci* and Gram-positive bacteria may be implicated in virulence.

In the test for attenuation of virulence, the mutated microorganism was shown to be attenuated with a competitive index (CI) of 0.766.

Example 7

A further mutant was identified and the nucleotide sequence immediately following the transposon insertion was cloned.

The nucleotide sequence has 98% identity at the nucleotide level to a sequence within the *S. pyogenes* genome, shown as SEQ ID NO. 15. The predicted amino acid sequence is shown as SEQ ID NO. 16.

The amino acid sequence shows 54% identity at the amino acid level to the biotin carboxyl carrier protein (BCCP) in *S. mutans* (accession number: SW: P29337).

This demonstrates that the disrupted gene is at least partially identical to the gene encoding BCCP of *S. mutans*. This gene was previously unknown in *S. pyogenes* and has not been assigned a role in virulence.

Given the similarity of the translated *S. pyogenes* gene to the BCCP of *S. mutans*, the skilled person will appreciate that similar sequences in other *Streptococci* and Gram-positive bacteria may be implicated in virulence.

In the test for attenuation of virulence, the mutated microorganisms was shown to be attenuated with a competitive index (CI) of 0.044.

Example 8

A series of mutants were identified having sequence similarity to genes involved in citrate fermentation. The nucleotide sequences of eight mutants showed 100% identity to a sequence within the *S. pyogenes* genome, shown to SEQ ID NO. 17. The predicted amino acid sequence is shown as SEQ ID NO. 18.

The predicted protein of the *S. pyogenes* gene shows 58% identity at the amino acid level to the citF protein, citrate lyase alpha chain, of *E. coli* (accession number: SW: P75726).

The nucleotide sequences of three further mutants showed 99% identity to a different sequence within the *S. pyogenes* genome. The nucleotide sequence is shown as SEQ ID NO. 19, and the predicted amino acid sequence is shown as SEQ ID NO. 20.

The amino acid sequences show 55% identity at the amino acid level to the citE protein, citrate lyase beta chain, of *E. coli* (accession number: SW:P77770).

The nucleotide sequences of two further mutants showed 99% identity to the *S. pyogenes* nucleotide sequence identified herein as SEQ ID NO. 21. The predicted amino acid sequence is shown as SEQ ID NO. 22.

The amino acid sequences show 46% identity at the amino acid level to the CitD protein, citrate lyase acyl carrier protein (citrate lyase gamma chain), of *E. coli* (accession number: SW: P77618).

Two further mutants were identified with nucleotide sequences which showed 99% identity to a nucleotide sequence from *S. pyogenes*, identified herein as SEQ ID NO. 23. The predicted amino acid sequence as SEQ ID NO. 24.

The amino acid sequences show 34% identity at the amino acid level to the CitC protein, citrate (pro-3s)-lyase (ligase), of *E. coli* (accession number: SW: P77390).

A further mutant was identified with a nucleotide sequence having 99% sequence identity to the *S. pyogenes* nucleotide sequence identified herein as SEQ ID NO. 25. The predicted amino acid sequence is shown as SEQ ID NO. 26.

The amino acid sequence shows 37% identity at the amino acid level to the CitX protein, of *E. coli* (accession number: SW: P77563).

A further three mutants were identified with nucleotide sequences showing 100% identity to the *S. pyogenes* nucleotide sequence identified herein as SEQ ID NO. 27. The predicted amino acid sequence is shown as SEQ ID NO. 28.

The amino acid sequences show 59% identity at the amino acid level to the OadA protein, oxaloacetate decarboxylase subunit alpha, of *K. pneumoniae* (accession number: SW:P13187).

All the predicted protein sequences of the above mutants show various degrees of identity to gene products of *E. coli* or *K. pneumoniae* required for citrate fermentation.

This demonstrates that the disrupted genes are at least partially identical to the *citF,E,D,C,X* and *oadA* genes of *E. coli* or *K. pneumoniae*. The genes involved in citrate fermentation are located together at a single locus on the chromosome in *E. coli* and *K. pneumoniae*. Therefore, the attenuation of *S. pyogenes* mutants could result from a failure to express its *citFEDCX* and *oadA* genes, or downstream genes. These genes have not previously been assigned a role in virulence.

Given the similarity of the *S. pyogenes* genes to those present in *E. coli* or *K. pneumoniae*, the skilled person will

appreciate that similar sequences in other *Streptococci* and Gram-positive bacteria may be implicated in virulence.

Mutants for each of *citF,E,D,X* and *oadA* were tested for attenuation of virulence. The mutated microorganisms were shown to be attenuated with a competitive index (CI) of 0.122 (*citF*), 0.064 (*citE*), 0.078 (*citD*), 0.025 (*citX*) and 0.251 (*oadA*).

Example 9

A further mutant was identified and the nucleotide sequence immediately following the transposon insertion was cloned.

The nucleotide sequence showed 98% identity at the nucleotide level to a sequence within the *S. pyogenes* genome, identified herein as SEQ ID NO. 29. The predicted amino acid sequence is shown as SEQ ID NO. 30.

The amino acid sequence shows 27% identity at the amino acid level to the *femB* protein of *S. aureas* (accession number: TREMBL: Q9X9D7).

This demonstrates that the disrupted gene is at least partially identical to the *femB* gene of *S. aureas*. However, this gene was previously unknown in *S. pyogenes*, and has not been assigned a role in virulence.

Given the similarity of the *S. pyogenes* gene to the *femB* gene in *S. aureas*, the skilled person will appreciate that similar sequences in other *Steptococci* and Gram-positive bacteria may be implicated in virulence.

Example 10

A further mutant was identified and the nucleotide sequence immediately following the transposon insertion was cloned.

The nucleotide sequence has significant identity to a sequence within the *S. pyogenes* genome, shown as SEQ ID NO. 31. The predicted amino acid sequence is shown as SEQ ID NO. 32.

In the test for attenuation of virulence, the mutated microorganism was shown to be attenuated with a competitive index (CI) of 0.160.

Example 11

A further mutant was identified and the nucleotide sequence immediately following the transposon insertion was cloned.

The nucleotide sequence has 100% identity to a sequence within the *S. pyogenes* genome, shown as SEQ ID NO. 33. The predicted amino acid sequence is shown as SEQ ID NO. 34.

Example 12

A further mutant was identified and the nucleotide sequence immediately following the transposon insertion was cloned.

The nucleotide sequence has 98% identity at the nucleotide level to a sequence within the *S. pyogenes* genome, shown as SEQ ID NO. 35. The predicted amino acid sequence is shown as SEQ ID NO. 36.

The amino acid sequence shows 46% identity at the amino acid level to the subtilin transport ATP-binding protein of *Bacillus subtilis* (accession number: SW: P33116).

This demonstrates that the disrupted gene is at least partially identical to the *braB* gene of *B. subtilis*.

Given the similarity of the gene of *S. pyogenes* to that in *B. subtilis*, the skilled person will appreciate that similar sequences in other *Streptococci* and Gram-positive bacteria may be implicated in virulence.

In the test for attenuation of virulence, the mutated microorganism was shown to be attenuated with a competitive index (CI) of 0.128.

Example 13

A further mutant was identified and the nucleotide sequence immediately following the transposon insertion was cloned.

The nucleotide sequence has 98% identity to a sequence within the *S. pyogenes* genome, shown as SEQ ID NO. 37. The predicted amino acid sequence is shown as SEQ ID NO. 38.

Example 14

A further mutant was identified and the nucleotide sequence immediately following the transposon insertion was cloned.

The nucleotide sequence has 98% identity to a sequence within the *S. pyogenes* genome, shown as SEQ ID NO. 39. The predicted amino acid sequence is shown as SEQ ID NO. 40.

The amino acid sequence shows 29% identity at the amino acid level to a 36.9kDa hypothetical protein from *E. coli* (accession number SW: P33019).

This demonstrates that the disrupted gene is at least partially identical to the gene of *E. coli*. Other similar sequences may be implicated in virulence.

In the test for attenuation of virulence, the mutated microorganism was shown to be attenuated with a competitive index (CI) of 0.747.

Example 15

A further mutant was identified and the nucleotide sequence immediately following the transposon insertion was cloned.

The nucleotide sequence has 99% identity to a sequence within the *S. pyogenes* genome, shown as SEQ ID NO. 41. The predicted amino acid sequence is shown as SEQ ID NO. 42.

The amino acid sequence shows 29% identity at the amino acid level to a 89kDa hypothetical protein from *A. fulgidus* (accession number TREMBL: 028455).

This demonstrates that the disrupted gene is at least partially identical to the gene of *E. coli*. Other, similar sequences may be implicated in virulence.

In the test for attenuation of virulence, the mutated microorganism was shown to be attenuated with a competitive index (CI) of 0.588.

Example 16

A further mutant was identified and the nucleotide sequence immediately following the transposon insertion was cloned.

The nucleotide sequence has 99% identity to a sequence within the *S. pyogenes* genome, shown as SEQ ID NO. 43. The predicted amino acid sequence is shown as SEQ ID NO. 44.

Example 17

A further mutant was identified and the nucleotide sequence immediately following the transposon insertion was cloned.

The nucleotide sequence has 98% identity to a sequence within the *S. pyogenes* genome, shown as SEQ ID NO. 45. The predicted amino acid sequence is shown as SEQ ID NO. 46.

Example 18

A further mutant was identified and the nucleotide sequence immediately following the transposon insertion was cloned.

The nucleotide sequence has 97% identity to a sequence within the *S. pyogenes* genome, shown as SEQ ID NO. 47. The predicted amino acid sequence is shown as SEQ ID NO. 48.

Example 19

A further mutant was identified and the nucleotide sequence immediately following the transposon insertion was cloned.

The nucleotide sequence has 98% identity at the nucleotide level to a sequence within the *S. pyogenes* genome, shown as SEQ ID NO. 49. The predicted amino acid sequence is shown as SEQ ID NO. 50.

The amino acid sequence shows 47% identity at the amino acid level to the ciaH protein of *S. pneumoniae* (accession number:SW: Q54955).

This demonstrates that the disrupted gene is at least partially identical to the *ciaH* gene of *S. pneumoniae*.

Given the similarity of the *ciaH* gene of *S. pyogenes* to the *ciaH* gene of *S. pneumoniae.*, the skilled person will appreciate that similar sequences in other *Streptococci* and Gram-positive bacteria may be implicated in virulence.

In the test for attenuation of virulence, the mutated microorganism was shown to be attenuated with a competitive index (CI) of 0.205.

Example 20

A further mutant was identified and the nucleotide sequence immediately following the transposon insertion was cloned.

The nucleotide sequence has 98% identity at the nucleotide level to a sequence within the *S. pyogenes* genome, shown as SEQ ID NO. 51. The predicted amino acid sequence is shown as SEQ ID NO. 52.

The amino acid sequence shows 39% identity at the amino acid level to the mucA homologue protein *of S. pneumonia* (accession number:TREMBL: Q9ZBB1).

This demonstrates that the disrupted gene is at least partially identical to the *mucA* gene of *S. pneumoniae*.

Given the similarity of the gene of *S. pyogenes* to the *mucA* gene of *S. pneumoniae*, the skilled person will appreciate that similar sequences in other *Streptococci* and Gram-positive bacteria may be implicated in virulence.

In the test for attenuation of virulence, the mutated microorganism was shown to be attenuated with a competitive index (CI) of 0.547.

Example 21

A further mutant was identified and the nucleotide sequence immediately following the transposon insertion was cloned.

The nucleotide sequence has 98% identity at the nucleotide level to a sequence within the *S. pyogenes* genome, shown as SEQ ID NO. 53. The predicted amino acid sequence is shown as SEQ ID NO. 54.

The amino acid sequence shows 69% identity at the amino acid level to the gidA protein of *B. subtilis* (accession number: SW:P39815).

This demonstrates that the disrupted gene is at least partially identical to the gene of *B. subtilis*.

Given the similarity of the *S. pyogenes* to the *gidA* gene of *B. subtilis*, the skilled person will appreciate that similar sequences in other *Streptococci* and Gram-positive bacteria may be implicated in virulence.

In the test for attenuation of virulence, the mutated microorganism was shown to be attenuated with a competitive index (CI) of 0.444.

Example 22

A further mutant was identified and the nucleotide sequence immediately following the transposon insertion was cloned.

The nucleotide sequence has 98% identity at the nucleotide level to a sequence within the *S. pyogenes* genome, shown as SEQ ID NO. 55. The predicted amino acid sequence is shown as SEQ ID NO. 56.

The amino acid sequence shows 74% identity at the amino acid level to the dltA protein of *S. mutans* (accession number:SW:Q53526).

This demonstrates that the disrupted gene is at least partially identical to the *dltA* gene of *S. mutans*.

Given the similarity of the *S. pyogenes* gene to the *dltA* gene of *S. mutans*, the skilled person will appreciate that similar sequences in other Gram-positive bacteria may be implicated in virulence.

In the test for attenuation of virulence, the mutated microorganism was shown to be attenuated with a competitive index (CI) of 0.205.

Example 23

A further mutant was identified and the nucleotide sequence immediately following the transposon insertion was cloned.

The nucleotide sequence has 98% identity at the nucleotide level to a sequence within the *S. pyogenes* genome, shown as SEQ ID NO. 57. The predicted amino acid sequence is shown as SEQ ID NO. 58.

The amino acid sequence shows 41% identity at the amino acid level to the hmgA protein of *A. fulgidus* (accession number: SW: O28538).

This demonstrates that the disrupted gene is at least partially identical to the gene of *A. fulgidus*.

Given the similarity of the gene of *S. pyogenes* to the *hmgA* gene of *A. fulgidus*, the skilled person will appreciate that similar sequences in other *Streptococci* and Gram-positive bacteria may be implicated in virulence.

In the test for attenuation of virulence, the mutated microorganism was shown to be attenuated with a competitive index (CI) of 0.205.

Example 24

A first mutant was identified and the nucleotide sequence immediately following the transposon insertion was cloned.

The nucleotide sequence has 99% identity to a sequence within the *S. pyogenes* genome, shown as SEQ ID NO. 59. The predicted amino acid sequence is shown as SEQ ID NO. 60.

The amino acid sequence shows 34% identity at the amino acid level to the putative JAG protein from *T. maritima* (accession number TREMBL: Q9X1H1).

This demonstrates that the disrupted gene from *S. pyogenes* is at least partially identical to the gene of *T. maritima*. Similar sequences in other Gram-negative bacteria may be implicated in virulence.

Example 25

A further mutant was identified and the nucleotide sequence immediately following the transposon insertion was cloned.

The nucleotide sequence has 99% identity to a sequence within the *S. pyogenes* genome, shown as SEQ ID NO. 61. The predicted amino acid sequence is shown as SEQ ID NO. 62.

The amino acid sequence shows 49% identity at the amino acid level to a 49.4 kDa hypothetical protein from *S. suis* (accession number TREMBL: Q9X4U3).

This demonstrates that the disrupted gene of *S. pyogenes* is at least partially identical to the gene of *S. suis*. Similar sequences in other Gram-negative bacteria may also be implicated in virulence.

CLAIMS

1. A peptide encoded by any of the nucleotide sequences identified herein as SEQ ID NOS. 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59 and 61 of S. pyogenes, or a related molecule having at least 60% sequence similarity or identity at the peptide or nucleotide level in a Gram-positive bacterium, or a functional fragment thereof, for therapeutic or diagnostic use.

2. A peptide according to claim 1, wherein the sequence similarity or identity is at least 80%.

3. A peptide according to claim 1 or claim 2, wherein the sequence similarity or identity is at least 95%.

4. A peptide according to claim 1, comprising the amino acid sequence identified herein as SEQ ID NOS. 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60 or 62.

5. A polynucleotide encoding a peptide according to any preceding claim, for therapeutic or diagnostic use.

6. An attenuated microorganism comprising a mutation that disrupts the expression of any of the nucleotide sequences defined in claim 1.

7. A microorganism according to claim 6, wherein the mutation is insertional inactivation or a gene deletion.

8. A microorganism according to claim 6 or claim 7, wherein the microorganism is a Group A Streptococcus.

9. A microorganism according to any of claims 6 to 8, for therapeutic or diagnostic use.

10. A microorganism according to any of claims 6 to 9, comprising a mutation in a further nucleotide sequence.

11. A microorganism according to any of claims 6 to 10, comprising a heterologous antigen, therapeutic peptide or nucleic acid.

12. A vaccine comprising a peptide according to any of claims 1 to 4.

13. A vaccine comprising a microorganism according to any of claims 6 to 11.

14. Use of a product according to any of claims 1 to 11, for the manufacture of a medicament for use in the treatment or prevention of a condition associated with infection by Streptococcal or Gram-positive bacteria.

15. Use according to claim 14, wherein the condition is pneumonia.

16. Use according to claim 14 or claim 15, for veterinary treatment.

17. An antibody raised against a peptide according to any of claims 1 to 4.

18. Use of a product according to any of claims 1 to 11, in a screening assay for the identification of an antimicrobial drug.