28 DEC 1999

# RECOMBINANT MOLECULES AND USES THEREFOR

## FIELD OF THE INVENTION

The present invention relates generally to nucleic acid molecules encoding a hyaluronidase from an aquatic animal. More particularly, the present invention relates to nucleic acid molecules from an aquatic animal encoding a hyaluronidase or a biologically active fragment thereof, or a variant or derivative of these and to expression vectors comprising such nucleic acid molecules operably linked to regulatory nucleic acids. The invention further relates to recombinant polypeptides encoding a hyaluronidase from the venom gland of an aquatic animal, biologically active fragments of such polypeptides, and variants and derivatives of these. The invention also extends to antibodies against such recombinant polypeptides and to the use of these polypeptides and antibodies for therapeutic and diagnostic purposes. The preferred aquatic animal of the present invention is a stonefish or related species.

## BACKGROUND OF THE INVENTION

Hyaluronidase plays an essential role in the fertilisation of a mammalian egg. The sperm protein PH-20 displays hyaluronidase activities and is found on the plasma membrane of human and mouse sperm (Lin *et al.*, 1994). PH-20 enables sperm to penetrate the egg's cumulus barrier, an essential step in the process of fertilisation.

Hyaluronidases are, therefore, used in *in-vitro* fertilisation (IVF) procedures of mammalian eggs. Currently, in human IVF programs, commercially prepared bovine testicular hyaluronidase is used to remove the surrounding cumulus cells of oocytes prior to micromanipulation procedures, such as partial zona dissection and microinsemination sperm transfer (Ng *et al.*, 1988). Alternative sources of hyaluronidases are required which are safe and efficacious in IVF procedures.

Hyaluronidase has been detected in the venom of *Synanceja horrida*, a marine fish that is found mainly in Singapore, Malaysia and Indonesian waters (Austin *et al.*, 1965). The stonefish hyaluronidase has been purified to homogeneity (Poh *et al.*, 1992) and has been used in murine IVF trials and has shown to be potent and safe in cumulus dispersal of

isolated mature mice oocytes without adversely affecting fertilisation rates nor cytogenetics of the oocytes and embryos (Vasuthevan *et al.*, 1993). This same report further suggested that stonefish hyaluronidase is more specific for its substrate and twice as fast as testicular hyaluronidase in effecting complete cumulus dispersal from mature mice oocytes.

Clinically, pharmaceutical preparations of hyaluronidases (mainly from bovine testes) have been used for a number of purposes. It has been used to increase the effect of local anaesthetics in ophthalmology (Gottlieb *et al.*, 1990). It is also effective in permitting the spread of subcutaneously administered fluids and drugs (Lipschitz *et al.*, 1991), to speed up the resorption of traumatic oedema and haematoma (Raszka *et al.*, 1990) and as an antidote to prevent skin necrosis which can result from inadvertent extravasation of antineoplastic drugs such as *Vinca* alkaloids during intravenous administration (Bertelli *et al.*, 1994). It is also used to enhance the efficacy of chemotherapy against malignant melanoma (Spruss *et al.*, 1995), bladder cancer (Hobarth *et al.*, 1992) and breast cancer (Beckenlehner *et al.*, 1992).

There is a need, therefore, for an abundant source of hyaluronidases for use in IVF and other therapeutic procedures.

## SUMMARY OF THE INVENTION

The present invention arises from the discovery of full-length nucleotide and polypeptide sequences relating to a hyaluronidase of a marine animal, more particularly a hyaluronidase of a stonefish and related species. The discovery of these sequences, thus, allows for the first time the production by recombinant techniques of large quantities of this hyaluronidase and its biologically active fragments as well as variants and derivatives of these. The recombinant polypeptides are useful for therapeutic and diagnostic procedures as described hereinafter.

Accordingly, in one aspect of the present invention, there is provided a recombinant polypeptide encoding a hyaluronidase of an aquatic animal or a biologically active fragment thereof, or a variant or derivative of these, said recombinant polypeptide

comprising the sequence set forth in SEQ ID NO: 6, with the proviso that said biologically active fragment does not consist of the sequence set forth in SEQ ID NO: 9.

The recombinant polypeptide may comprise a leader peptide. Suitably, the leader peptide comprises the sequence set forth in SEQ ID NO: 8, or a biologically active fragment thereof, or variant or derivative of these. In such a case, the recombinant polypeptide preferably comprises the sequence set forth in SEQ ID NO: 4.

In another aspect, the invention provides an isolated polynucleotide encoding a polypeptide, fragment, variant or derivative as broadly described above. Preferably, the polynucleotide comprises the sequence set forth in SEQ ID NO: 5, or a biologically active fragment thereof, or a polynucleotide variant of these.

The polynucleotide preferably comprises a nucleotide sequence encoding a leader peptide. Suitably, said nucleotide sequence comprises the sequence set forth in SEQ ID NO: 7 or a biologically active fragment thereof, or a polynucleotide variant of these. In such a case, the polynucleotide preferably comprises the sequence set forth in SEQ ID NO: 1 or 3.

Preferably, the variant is obtained from a species of stonefish. Suitably, the variant is obtained from a *Synanceja* species. The *Synanceja* species includes, but is not restricted to, *S. verrucosa* and *S. horrida.*

In another aspect, the invention features an expression vector comprising a polynucleotide as broadly described above wherein the polynucleotide is operably linked to one or more regulatory nucleic acids.

In a further aspect, the invention provides a host cell containing a said expression vector.

The invention also contemplates a method of producing a recombinant polypeptide, fragment, variant or derivative as broadly described above, comprising:

(a)culturing a host cell containing an expression vector as broadly described above such that said recombinant polypeptide, fragment, variant or derivative is expressed from said polynucleotide; and

(b)isolating the said recombinant polypeptide, fragment, variant or derivative.

In a further aspect, the invention provides a method of producing a biologically active fragment as broadly described above, comprising:

(a)producing a fragment of said recombinant polypeptide;

(b)contacting a hyaluronidase substrate with the fragment; and

(c)detecting hydrolysis of the hyaluronidase substrate by the fragment which is indicative of said fragment being a said biologically active fragment.

In yet another aspect, the invention provides a method of producing a biologically active fragment as broadly described above, comprising:

(a)producing a fragment of said recombinant polypeptide;

(b)contacting an oocyte-cumulus complex with the fragment; and

(c)detecting dispersal of cumulus cells from the said complex, which is indicative of the fragment being a said biologically active fragment.

Suitably, the method further comprises the step of measuring the rate of dispersal of cumulus cells from the said complex. Preferably, the step of measuring is further characterised in that when the rate of dispersal of cumulus cells from the said complex in the presence of fragment is about half that in the presence of sheep hyaluronidase, then this is indicative of the fragment being a said biologically active fragment. More preferably, the step of measuring is further characterised in that when the rate of dispersal of cumulus cells, preferably murine cumulus cells, from the said complex in the presence of said fragment is about 0.8 minutes per oocyte, then this is indicative of the fragment being a said biologically active fragment.

In yet another aspect, the invention contemplates a method of producing a polypeptide variant of a parent polypeptide comprising the sequence set forth in SEQ ID NO: 2 or 6, or comprising a biologically active fragment of said sequence, comprising the steps of:

(a)replacing at least one amino acid of the parent polypeptide, with a different amino acid to produce a modified polypeptide;

(b)contacting a hyaluronidase substrate with the modified polypeptide; and

(c)detecting hydrolysis of the hyaluronidase substrate by the modified polypeptide, which is indicative of the modified polypeptide being a said polypeptide variant.

In yet another aspect, the invention contemplates a method of producing a polypeptide variant of a parent polypeptide comprising the sequence set forth in SEQ ID NO: 2 or 6, or comprising a biologically active fragment of said sequence, comprising the steps of:

(a)replacing at least one amino acid of the parent polypeptide, with a different amino acid to produce a modified polypeptide;

(b)contacting the modified polypeptide with an antigen-binding molecule that specifically binds to the parent polypeptide; and

(c)detecting the presence of a complex comprising the said modified polypeptide, and the antigen-binding molecule in said contacted sample, which is indicative of the modified polypeptide being a said polypeptide variant.

In yet another aspect, the invention provides a method of producing a polypeptide variant of a parent polypeptide comprising the sequence set forth in SEQ ID NO: 2 or 6, or comprising a biologically active fragment of said sequence, comprising the steps of:

(a)replacing at least one amino acid of the parent polypeptide, with a different amino acid to produce a modified polypeptide;

(b)contacting an oocyte-cumulus complex with the modified polypeptide; and

(c)detecting dispersal of cumulus cells from the said complex, which is indicative of the modified polypeptide being a said polypeptide variant.

Suitably, the method further comprises the step of measuring the rate of dispersal of cumulus cells from the said complex. Preferably, the step of measuring is further characterised in that when the rate of dispersal of cumulus cells from the said complex in the presence of said modified polypeptide is about half that in the presence of sheep hyaluronidase, then this is indicative of the modified polypeptide being a said polypeptide variant. More preferably, the step of measuring is further characterised in that when the rate of dispersal of cumulus cells, preferably murine cumulus cells from the said complex in the presence of said modified polypeptide is about 0.8 minutes per oocyte, then this is indicative of the modified polypeptide being a said polypeptide variant.

In yet another aspect, the invention provides a composition for use in enhancing fertilisation of a mammalian egg, in increasing the effect of anaesthetics, particularly local anaesthetics and preferably in opthalmology, in administrating fluids and drugs preferably subcutaneously, in resorption of traumatic oedema and haematoma, in preventing tissue necrosis, particularly skin necrosis, during intravenous administration, and in enhancing the efficacy of chemotherapy against cancer including melanoma, bladder cancer and breast cancer, wherein said composition comprises a recombinant polypeptide, biologically active fragment, variant or derivative as broadly described above, together with a pharmaceutically acceptable carrier.

According to a further aspect, the invention provides a method for enhancing fertilisation of a mammalian egg, comprising contacting the egg with a fertilisingly effective amount of a composition as broadly described above.

In another aspect, the invention resides in the use of a recombinant polypeptide, fragment, variant or derivative according to the present invention to produce an antigen-binding molecule that binds specifically to the said recombinant polypeptide, fragment, variant or derivative.

In yet another aspect, the invention provides antigen-binding molecules so produced.

The invention also encompasses use of an antigen-binding molecule as broadly described above in treating envenomation by *S. horrida* or related species.

The invention also extends to a method of detecting in a sample a polypeptide, fragment, variant or derivative as broadly described above, comprising:

(a)contacting the sample with an antigen-binding molecule as broadly described above; and

(b)detecting the presence of a complex comprising the said antigen-binding molecule and the said polypeptide, fragment, variant or derivative in said contacted sample.

**Brief Description of the Drawings**

Figure 1 is a representation showing the alignment of the deduced amino acid sequence of hyaluronidases from mouse (GI 2370295; mouse lysosomal hyaluronidase), human (GI 2314820; man plasma hyaluroindase), Macaca (GI 585675; *Macaca fasicularis* sperm surface protein), honeybee (GI 585279; Honeybee venom hyaluroindase) and cele (GI 3879998; *Caenorhabditis elegans* hyaluronidase of unknown tissue origin). The 3 highly conserved regions, suitable for design of PCR primers were marked with ——. This figure is generated with the aid of the multisequence alignment program of DNASTAR software (DNASTAR Inc. Madison Wis.).

Figure 2 is a representation of the nucleic acid sequence (SEQ ID NO: 1) and deduced amino acid sequence (SEQ ID NO: 2) of the stonefish venom hyaluronidase SFHYA1. Bold-faced nucleotide sequences indicate the position of the PCR primers used for amplification of SFHYA1 cDNA segment during RT-PCR. The putative polyadenylation signal, ATTAAA, is underlined. Predicted site of proteolytic cleavage of pre SFHYA1 is indicated by ▼ . Amino-acid sequences derived from N-terminal sequencing of SFHYA1 and its in-gel tryptic fragments are shaded.

Figure 3 is a representation of the alignment of the deduced amino acid sequence of SFHYA1 (SEQ ID NO: 2) with those of highly matching entries at NCBI databases using the multisequence alignment program of DNASTAR software (DNASTAR Inc, Madison Wis.).

## DETAILED DESCRIPTION

### *1. Definitions*

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by those of ordinary skill in the art to which the invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, preferred methods and materials are described. For the purposes of the present invention, the following terms are defined below.

The term *"about"* is used herein to refer to rates of dispersal of cumulus cells from an oocyte-cumulus complex which rates vary by as much as 30%, preferably by as much as 20%, and more preferably by as much as 10% to a reference dispersal rate.

*"Amplification product"* refers to a nucleic acid product generated by nucleic acid amplification techniques.

By *"antigen-binding molecule"* is meant a molecule that has binding affinity for a target antigen. It will be understood that this term extends to immunoglobulins, immunoglobulin fragments and non-immunoglobulin derived protein frameworks that exhibit antigen-binding activity.

As used herein, the term *"binds specifically"* and the like refers to antigen-binding molecules that bind the polypeptide or polypeptide fragments of the invention but do not significantly bind to homologous prior art polypeptides.

By *"biologically active fragment"* is meant a fragment of a full-length parent polypeptide which fragment retains the activity of the parent polypeptide. A biologically active fragment will therefore have hyaluronidase activity and activity for dispersing

cumulus cells from an oocyte-cumulus complex. As used herein, the term *"biologically active fragment"* includes deletion mutants and small peptides, for example of at least 10, preferably at least 20 and more preferably at least 30 contiguous amino acids, which comprise the above activities. Peptides of this type may be obtained through the application of standard recombinant nucleic acid techniques or synthesised using conventional liquid or solid phase synthesis techniques. For example, reference may be made to solution synthesis or solid phase synthesis as described, for example, in Chapter 9 entitled *"Peptide Synthesis"* by Atherton and Shephard which is included in a publication entitled *"Synthetic Vaccines"* edited by Nicholson and published by Blackwell Scientific Publications. Alternatively, peptides can be produced by digestion of a polypeptide of the invention with proteinases such as endoLys-C, endoArg-C, endoGlu-C and staphylococcus V8-protease. The digested fragments can be purified by, for example, high performance liquid chromatographic (HPLC) techniques.

By "*corresponds to*" or "*corresponding to*" is meant a polynucleotide (a) having a nucleotide sequence that is substantially identical or complementary to all or a portion of a reference polynucleotide sequence or (b) encoding an amino acid sequence identical to an amino acid sequence in a peptide or protein. This phrase also includes within its scope a peptide or polypeptide having an amino acid sequence that is substantially identical to a sequence of amino acids in a reference peptide or protein.

By *"derivative"* is meant a polypeptide that has been derived from the basic sequence by modification, for example by conjugation or complexing with other chemical moieties or by post-translational modification techniques as would be understood in the art. The term "*derivative*" also includes within its scope alterations that have been made to a parent sequence including additions, or deletions that provide for functionally equivalent molecules. Accordingly, the term derivative encompasses molecules that will have hyaluronidase activity and disperse cumulus cells from an oocyte-cumulus complex.

*"Homology"* refers to the percentage number of amino acids that are identical or constitute conservative substitutions as defined in Table 1 *infra*. Homology may be determined using sequence comparison programs such as GAP (Deveraux *et al.* 1984, *Nucleic Acids Research* **12**, 387-395). In this way, sequences of a similar or substantially

different length to those cited herein might be compared by insertion of gaps into the alignment, such gaps being determined, for example, by the comparison algorithm used by GAP.

As used herein the term *"hyaluronidase"* refers to the class of enzymes that act on the disaccharide unit of D-glucuronic acid and N-acetyl-D-glucosamine. Such enzymes mediate the hydrolysis of polymers of repeating disaccharides comprising D-glucuronic acid and N-acetyl-D-glucosamine. One example of such polymer is hyaluronic acid. Hyaluronidase catalyses the release of reducing groups of N-acetylglucosamine from a hyaluronic acid.

"*Hybridisation*" is used herein to denote the pairing of complementary nucleotide sequences to produce a DNA-DNA hybrid or a DNA-RNA hybrid. Complementary base sequences are those sequences that are related by the base-pairing rules. In DNA, A pairs with T and C pairs with G. In RNA U pairs with A and C pairs with G. In this regard, the terms "match" and "mismatch" as used herein refer to the hybridisation potential of paired nucleotides in complementary nucleic acid strands. Matched nucleotides hybridise efficiently, such as the classical A-T and G-C base pair mentioned above. Mismatches are other combinations of nucleotides that do not hybridise efficiently.

By *"isolated"* is meant material that is substantially or essentially free from components that normally accompany it in its native state. For example, an "isolated polynucleotide", as used herein, refers to a polynucleotide, which has been purified from the sequences which flank it in a naturally occurring state, *e.g.*, a DNA fragment which has been removed from the sequences which are normally adjacent to the fragment.

By *"obtained from"* is meant that a sample such as, for example, a nucleic acid extract or polypeptide extract is isolated from, or derived from, a particular source of the host. For example, the extract may be obtained from a tissue or a biological fluid isolated directly from the host.

The term "*oligonucleotide*" as used herein refers to a polymer composed of a multiplicity of nucleotide units (deoxyribonucleotides or ribonucleotides, or related structural variants or synthetic analogues thereof) linked via phosphodiester bonds (or

related structural variants or synthetic analogues thereof). Thus, while the term "oligonucleotide" typically refers to a nucleotide polymer in which the nucleotides and linkages between them are naturally occurring, it will be understood that the term also includes within its scope various analogues including, but not restricted to, peptide nucleic acids (PNAs), phosphoramidates, phosphorothioates, methyl phosphonates, 2-O-methyl ribonucleic acids, and the like. The exact size of the molecule may vary depending on the particular application. An oligonucleotide is typically rather short in length, generally from about 10 to 30 nucleotides, but the term can refer to molecules of any length, although the term "polynucleotide" or "nucleic acid" is typically used for large oligonucleotides.

By *"operably linked"* is meant that transcriptional and translational regulatory nucleic acids are positioned relative to a polypeptide-encoding polynucleotide in such a manner that the polynucleotide is transcribed and the polypeptide is translated.

The term "*patient*" refers to patients of human or other mammal and includes any individual it is desired to examine or treat using the methods of the invention. However, it will be understood that "*patient*" does not imply that symptoms are present. Suitable mammals that fall within the scope of the invention include, but are not restricted to, primates, livestock animals (eg. sheep, cows, horses, donkeys, pigs), laboratory test animals (eg. rabbits, mice, rats, guinea pigs, hamsters), companion animals (eg. cats, dogs) and captive wild animals (eg. foxes, deer, dingoes).

By *"pharmaceutically-acceptable carrier"* is meant a solid or liquid filler, diluent or encapsulating substance that may be safely used in topical or systemic administration.

The term "*polynucleotide*" or "*nucleic acid*" as used herein designates mRNA, RNA, cRNA, cDNA or DNA. The term typically refers to oligonucleotides greater than 30 nucleotides in length.

The terms *"polynucleotide variant"* and *"variant"* refer to polynucleotides displaying substantial sequence identity with a reference polynucleotide sequence or polynucleotides that hybridise with a reference sequence under stringent conditions that are defined hereinafter. These terms also encompasses polynucleotides in which one or

more nucleotides have been added or deleted, or replaced with different nucleotides. In this regard, it is well understood in the art that certain alterations inclusive of mutations, additions, deletions and substitutions can be made to a reference polynucleotide whereby the altered polynucleotide retains the biological function or activity of the reference polynucleotide. The terms *"polynucleotide variant"* and *"variant"* also include naturally occurring allelic variants.

*"Polypeptide"*, *"peptide"* and *"protein"* are used interchangeably herein to refer to a polymer of amino acid residues and to variants and synthetic analogues of the same. Thus, these terms apply to amino acid polymers in which one or more amino acid residues is a synthetic non-naturally occurring amino acid, such as a chemical analogue of a corresponding naturally occurring amino acid, as well as to naturally-occurring amino acid polymers.

The term *"polypeptide variant"* refers to polypeptides in which one or more amino acids have been replaced by different amino acids. It is well understood in the art that some amino acids may be changed to others with broadly similar properties without changing the nature of the activity of the polypeptide (conservative substitutions) as described hereinafter. Accordingly, polypeptide variants as used herein encompass polypeptides that will inhibit phospholipase $A_2$ activity and/or alleviate inflammation or conditions associated with envenomation.

By *"primer"* is meant an oligonucleotide which, when paired with a strand of DNA, is capable of initiating the synthesis of a primer extension product in the presence of a suitable polymerising agent. The primer is preferably single-stranded for maximum efficiency in amplification but may alternatively be double-stranded. A primer must be sufficiently long to prime the synthesis of extension products in the presence of the polymerisation agent. The length of the primer depends on many factors, including application, temperature to be employed, template reaction conditions, other reagents, and source of primers. For example, depending on the complexity of the target sequence, the oligonucleotide primer typically contains 15 to 35 or more nucleotides, although it may contain fewer nucleotides. Primers can be large polynucleotides, such as from about 200 nucleotides to several kilobases or more. Primers may be selected to be "substantially

complementary" to the sequence on the template to which it is designed to hybridise and serve as a site for the initiation of synthesis. By "substantially complementary", it is meant that the primer is sufficiently complementary to hybridise with a target nucleotide sequence. Preferably, the primer contains no mismatches with the template to which it is designed to hybridise but this is not essential. For example, non-complementary nucleotides may be attached to the 5' end of the primer, with the remainder of the primer sequence being complementary to the template. Alternatively, non-complementary nucleotides or a stretch of non-complementary nucleotides can be interspersed into a primer, provided that the primer sequence has sufficient complementarity with the sequence of the template to hybridise therewith and thereby form a template for synthesis of the extension product of the primer.

*"Probe"* refers to a molecule that binds to a specific sequence or sub-sequence or other moiety of another molecule. Unless otherwise indicated, the term "probe" typically refers to a polynucleotide probe that binds to another nucleic acid, often called the "target nucleic acid", through complementary base pairing. Probes may bind target nucleic acids lacking complete sequence complementarity with the probe, depending on the stringency of the hybridisation conditions. Probes can be labelled directly or indirectly.

The term *"recombinant polynucleotide"* as used herein refers to a polynucleotide formed *in vitro* by the manipulation of nucleic acid into a form not normally found in nature. For example, the recombinant polynucleotide may be in the form of an expression vector. Generally, such expression vectors include transcriptional and translational regulatory nucleic acid operably linked to the nucleotide sequence.

By *"recombinant polypeptide"* is meant a polypeptide made using recombinant techniques, *i.e.*, through the expression of a recombinant polynucleotide.

By "reporter molecule" as used in the present specification is meant a molecule that, by its chemical nature, provides an analytically identifiable signal that allows the detection of a complex comprising an antigen-binding molecule and its target antigen. The term "reporter molecule" also extends to use of cell agglutination or inhibition of agglutination such as red blood cells on latex beads, and the like.

Terms used to describe sequence relationships between two or more polynucleotides or polypeptides include "reference sequence", "comparison window", "sequence identity", "percentage of sequence identity" and "substantial identity". A "*reference sequence*" is at least 12 but frequently 15 to 18 and often at least 25 monomer units, inclusive of nucleotides and amino acid residues, in length. Because two polynucleotides may each comprise (1) a sequence (*i.e.*, only a portion of the complete polynucleotide sequence) that is similar between the two polynucleotides, and (2) a sequence that is divergent between the two polynucleotides, sequence comparisons between two (or more) polynucleotides are typically performed by comparing sequences of the two polynucleotides over a "comparison window" to identify and compare local regions of sequence similarity. A "*comparison window*" refers to a conceptual segment of typically 12 contiguous residues that is compared to a reference sequence. The comparison window may comprise additions or deletions (*i.e.*, gaps) of about 20% or less as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Optimal alignment of sequences for aligning a comparison window may be conducted by computerised implementations of algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package Release 7.0, Genetics Computer Group, 575 Science Drive Madison, WI, USA) or by inspection and the best alignment (*i.e.*, resulting in the highest percentage homology over the comparison window) generated by any of the various methods selected. Reference also may be made to the BLAST family of programs as for example disclosed by Altschul *et al.*, 1997, *Nucl. Acids Res.* **25**:3389. A detailed discussion of sequence analysis can be found in Unit 19.3 of Ausubel *et al.*, "Current Protocols in Molecular Biology", John Wiley & Sons Inc, 1994-1998, Chapter 15.

The term *"sequence identity"* as used herein refers to the extent that sequences are identical on a nucleotide-by-nucleotide basis or an amino acid-by-amino acid basis over a window of comparison. Thus, a *"percentage of sequence identity"* is calculated by comparing two optimally aligned sequences over the window of comparison, determining the number of positions at which the identical nucleic acid base (*e.g.*, A, T, C, G, I) or the identical amino acid residue (*e.g.*, Ala, Pro, Ser, Thr, Gly, Val, Leu, Ile, Phe, Tyr, Trp, Lys, Arg, His, Asp, Glu, Asn, Gln, Cys and Met) occurs in both sequences to yield the

number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison (*i.e.*, the window size), and multiplying the result by 100 to yield the percentage of sequence identity. For the purposes of the present invention, *"sequence identity"* will be understood to mean the "match percentage" calculated by the DNASIS computer program (Version 2.5 for windows; available from Hitachi Software engineering Co., Ltd., South San Francisco, California, USA) using standard defaults as used in the reference manual accompanying the software.

*"Stringency"* as used herein, refers to the temperature and ionic strength conditions, and presence or absence of certain organic solvents, during hybridisation. The higher the stringency, the higher will be the degree of complementarity between immobilised nucleotide sequences and the labelled polynucleotide sequence.

*"Stringent conditions"* refers to temperature and ionic conditions under which only nucleotide sequences having a high frequency of complementary bases will hybridise. The stringency required is nucleotide sequence dependent and depends upon the various components present during hybridisation. Generally, stringent conditions are selected to be about 10 to 20°C lower than the thermal melting point ($T_m$) for the specific sequence at a defined ionic strength and pH. The $T_m$ is the temperature (under defined ionic strength and pH) at which 50% of a target sequence hybridises to a complementary probe.

By *"fertilisingly effective amount"*, is meant the administration of that amount to a mammalian egg in need of such treatment, either in a single dose or as part of a series, that is effective for enhancing fertilisation of the egg. The effective amount will vary depending upon the age and physical condition of the individual from which the egg is derived, the taxonomic group of individual to be treated, the formulation of the composition, and other relevant factors. It is expected that the amount will fall in a relatively broad range that can be determined through routine trials.

By *"vector"* is meant a nucleic acid molecule, preferably a DNA molecule derived, for example, from a plasmid, bacteriophage, or plant virus, into which a nucleic acid sequence may be inserted or cloned. A vector preferably contains one or more unique restriction sites and may be capable of autonomous replication in a defined host cell including a target cell or tissue or a progenitor cell or tissue thereof, or be integrable with

the genome of the defined host such that the cloned sequence is reproducible. Accordingly, the vector may be an autonomously replicating vector, *i.e.*, a vector that exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, *e.g.*, a linear or closed circular plasmid, an extrachromosomal element, a minichromosome, or an artificial chromosome. The vector may contain any means for assuring self-replication. Alternatively, the vector may be one which, when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. A vector system may comprise a single vector or plasmid, two or more vectors or plasmids, which together contain the total DNA to be introduced into the genome of the host cell, or a transposon. The choice of the vector will typically depend on the compatibility of the vector with the host cell into which the vector is to be introduced. The vector may also include a selection marker such as an antibiotic resistance gene that can be used for selection of suitable transformants. Examples of such resistance genes are well known to those of skill in the art.

Throughout this specification and the appendant claims, unless the context requires otherwise, the words "comprise", "comprises" and "comprising" will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps.

## *2. Recombinant polypeptides, biologically active fragments, polypeptide variants and derivatives*

### 2.1.Recombinant polypeptides of the invention

The invention also features a recombinant polypeptide comprising the sequence set forth in SEQ ID NO: 6, which corresponds to a mature polypeptide obtained from *S. horrida*, as described more fully hereinafter.

In one embodiment, the recombinant polypeptide may include a leader peptide comprising the sequence set forth in SEQ ID NO: 8 or biologically active fragment thereof, or variant or derivative of these. Accordingly, the invention also provides a recombinant precursor polypeptide according to SEQ ID NO: 2, which comprises a leader

peptide according to SEQ ID NO: 8 fused in frame with a polypeptide according to SEQ ID NO: 6.

2.2. Biologically active fragments

Biologically active fragments may be produced according to any suitable procedure known in the art. For example, a suitable method may include first producing a fragment of said recombinant polypeptide and then testing the fragment for the appropriate biological activity. In one embodiment, biological activity of the fragment may be tested by contacting a hyaluronidase substrate (*e.g.*, hyaluronic acid) with the fragment, and detecting hydrolysis of the hyaluronidase substrate by the fragment which is indicative of said fragment being a said biologically active fragment. Hydrolysis may be detected by any suitable procedure. For example, reference may be made to a turbidometric/ spectrophotometric technique disclosed in Poh *et al.* (1992, *supra*).

Alternatively, a suitable fragment may be tested for biological activity by contacting an oocyte-cumulus complex with the fragment and detecting dispersal of cumulus cells from the said complex, which is indicative of the fragment being a said biologically active fragment. In a preferred embodiment, the method further comprises the step of measuring the rate of dispersal of cumulus cells from the said complex. Preferably, the step of measuring is further characterised in that when the rate of dispersal of cumulus cells from the said complex in the presence of fragment is about half that in the presence of sheep hyaluronidase, then this is indicative of the fragment being a said biologically active fragment. More preferably, the step of measuring is further characterised in that when the rate of dispersal of cumulus cells, preferably murine cumulus cells, from the said complex in the presence of said fragment is about 0.8 minutes per oocyte, then this is indicative of the fragment being a said biologically active fragment. Preferably, cumulus dispersal is assessed by methods disclosed by Vasuthevan *et al.* 1993, *supra*).

2.3. Polypeptide variants

The invention also contemplates polypeptide variants of the recombinant polypeptide of the invention wherein said variants have hyaluronidase activity and activity for dispersing cumulus cells from an oocyte-cumulus complex. Suitable methods of

producing polypeptide variants include replacing at least one amino acid of a parent polypeptide comprising the sequence set forth in SEQ ID NO: 2 or 6, or a biologically active fragment thereof, with a different amino acid to produce a modified polypeptide, contacting hyaluronic acid with the modified polypeptide and detecting hydrolysis of hyaluronic acid by the modified polypeptide, which is indicative of the modified polypeptide being a said polypeptide variant. A suitable method of testing for hyaluronidase activity is provided in Section 2.2.

Alternatively, a polypeptide variant according to the invention may be identified by replacing at least one amino acid of a parent polypeptide comprising the sequence set forth in SEQ ID NO: 2 or 6, or a biologically active fragment thereof, with a different amino acid to produce a modified polypeptide, contacting the modified polypeptide with an antigen-binding molecule that specifically binds to the parent polypeptide, and detecting the presence of a complex comprising the said modified polypeptide, and the antigen-binding molecule in said contacted sample, which is indicative of the modified polypeptide being a said polypeptide variant.

In an alternate procedure, a polypeptide variant of a parent polypeptide comprising the sequence set forth in SEQ ID NO: 2 or 6, or a biologically active fragment thereof, may be identified by replacing at least one amino acid of the parent polypeptide, with a different amino acid to produce a modified polypeptide, contacting an oocyte-cumulus complex with the modified polypeptide and detecting dispersal of cumulus cells from the said complex, which is indicative of the modified polypeptide being said polypeptide variant. Suitably, the method further comprises the step of measuring the rate of dispersal of cumulus cells from the said complex. Preferably, the step of measuring is further characterised in that when the rate of dispersal of cumulus cells from the said complex in the presence of said modified polypeptide is about half that in the presence of sheep hyaluronidase, then this is indicative of the modified polypeptide being a said polypeptide variant. More preferably, the step of measuring is further characterised in that when the rate of dispersal of cumulus cells, preferably murine cumulus cells from the said complex in the presence of said modified polypeptide is about 0.8 minutes per oocyte, then this is indicative of the modified polypeptide being a said polypeptide variant. A suitable method of testing for cumulus dispersal activity is provided in Section 2.2.

In general, variants will be at least 80% homologous, more suitably at least 85%, preferably at least 90%, and more preferably at least 95% homologous to a polypeptide as for example shown in SEQ ID NO: 2 or 6.

## 2.4. Methods of producing polypeptide variants

### *2.4.1. Mutagenesis*

Polypeptide variants according to the invention can be identified either rationally, or via established methods of mutagenesis (see, for example, Watson, J. D. *et al.*, "MOLECULAR BIOLOGY OF THE GENE", Fourth Edition, Benjamin/Cummings, Menlo Park, Calif., 1987). Significantly, a random mutagenesis approach requires no *a priori* information about the gene sequence that is to be mutated. This approach has the advantage that it assesses the desirability of a particular mutant based on its function, and thus does not require an understanding of how or why the resultant mutant protein has adopted a particular conformation. Indeed, the random mutation of target gene sequences has been one approach used to obtain mutant proteins having desired characteristics (Leatherbarrow, R. 1986, *J. Prot. Eng.* **1**: 7-16; Knowles, J. R., 1987, *Science* **236**: 1252-1258; Shaw, W. V., 1987, *Biochem. J.* **246**: 1-17; Gerit, J. A. 1987, *Chem. Rev.* **87**: 1079-1105). Alternatively, where a particular sequence alteration is desired, methods of site-directed mutagenesis can be employed. Thus, such methods may be used to selectively alter only those amino acids of the protein that are believed to be important (Craik, C. S., 1985, *Science* **228**: 291-297; Cronin, *et al.*, 1988, *Biochem.* **27**: 4572-4579; Wilks, *et al.*, 1988, *Science* **242**: 1541-1544).

Variant peptides or polypeptides, resulting from rational or established methods of mutagenesis or from combinatorial chemistries as hereinafter described, may comprise conservative amino acid substitutions. Exemplary conservative substitutions in an immuno-interactive polypeptide or polypeptide fragment according to the invention may be made according to the following table:

**TABLE A**

| *Original Residue* | *Exemplary Substitutions* |
|---|---|
| Ala | Ser |
| Arg | Lys |
| Asn | Gln, His |
| Asp | Glu |
| Cys | Ser |
| Gln | Asn |
| Glu | Asp |
| Gly | Pro |
| His | Asn, Gln |
| Ile | Leu, Val |
| Leu | Ile, Val |
| Lys | Arg, Gln, Glu |
| Met | Leu, Ile, |
| Phe | Met, Leu, Tyr |
| Ser | Thr |
| Thr | Ser |
| Trp | Tyr |
| Tyr | Trp, Phe |
| Val | Ile, Leu |

Substantial changes in function are made by selecting substitutions that are less conservative than those shown in TABLE A. Other replacements would be non-conservative substitutions and relatively fewer of these may be tolerated. Generally, the substitutions which are likely to produce the greatest changes in a polypeptide's properties are those in which (a) a hydrophilic residue (*eg*, Ser or Thr) is substituted for, or by, a hydrophobic residue (*eg*, Ala, Leu, Ile, Phe or Val); (b) a cysteine or proline is substituted for, or by, any other residue; (c) a residue having an electropositive side chain (*eg*, Arg, His or Lys) is substituted for, or by, an electronegative residue (*eg*, Glu or Asp) or (d) a residue having a bulky side chain (*eg*, Phe or Trp) is substituted for, or by, one having a smaller side chain (*eg*, Ala, Ser)or no side chain (*eg*, Gly).

What constitutes suitable variants may be determined by conventional techniques. For example, nucleic acids encoding a polypeptide according to SEQ ID NO: 2 or 6 can be mutated using either random mutagenesis for example using transposon mutagenesis, or site-directed mutagenesis as described, for example, in Section 3.2 *infra*.

*2.4.2. Peptide libraries produced by combinatorial chemistry*

A number of facile combinatorial technologies can be utilised to synthesise molecular libraries of immense diversity. In the present case, variants of a polypeptide, or preferably a polypeptide fragment according to the invention, can be synthesised using such technologies. Variants can be screened subsequently using the methods described in Section 2.3.

Preferably, soluble synthetic peptide combinatorial libraries (SPCLs) are produced which offer the advantage of working with free peptides in solution, thus permitting adjustment of peptide concentration to accommodate a particular assay system. SPCLs are suitably prepared as hexamers. In this regard, a majority of binding sites is known to involve four to six residues. Cysteine is preferably excluded from the mixture positions to avoid the formation of disulfides and more difficult-to-define polymers. Exemplary methods of producing SPCLs are disclosed by Houghten *et al.* (1991, *Nature* **354**: 84-86; 1992, *BioTechniques* **13**: 412-421), Appel *et al.* (1992, *Immunomethods* **1**: 17-23), and Pinilla *et al.* (1992, *BioTechniques* **13**: 901-905; 1993, *Gene* **128**: 71-76).

Preparation of combinatorial synthetic peptide libraries may employ either *t*-butyloxycarbonyl (*t*-Boc) or 9-fluorenylmethyloxycarbonyl (Fmoc) chemistries (see Chapter 9.1, of Coligan *et al.*, *supra*; Stewart and Young, 1984, Solid Phase Peptide Synthesis, 2nd ed. Pierce Chemical Co., Rockford, Ill; and Atherton and Sheppard, 1989, Solid Phase Peptide Synthesis: A Practical Approach. IRL Press, Oxford) preferably, but not exclusively, using one of two different approaches. The first of these approaches, suitably termed the "split-process-recombine" or "split synthesis" method, was described first by Furka *et al.* (1988, *14th Int. Congr. Biochem.*, Prague, Czechoslovakia **5**: 47; 1991, *Int. J. Pept. Protein Res.* **37**: 487-493) and Lam *et al.* (1991, *Nature* **354**: 82-84), and reviewed later by Eichler *et al.* (1995, *Medicinal Research Reviews* **15(6)**: 481-496) and Balkenhohl *et al.* (1996, *Angew. Chem. Int. Ed. Engl.* **35**: 2288-2337). Briefly, the split synthesis method involves dividing a plurality of solid supports such as polymer beads into *n* equal fractions representative of the number of available amino acids for each step of the synthesis (e.g., 20 L-amino acids), coupling a single respective amino acid to each polymer bead of a corresponding fraction, and then thoroughly mixing the polymer beads of all the fractions together. This process is repeated for a total of *x* cycles to produce a stochastic collection of up to $N^x$ different compounds. The peptide library so produced may be screened for example with a suitably labelled antigen-binding molecule that binds specifically to a polypeptide according to SEQ ID NO: 2 or 6. Upon detection, some of the positive beads are selected for sequencing to identify the active peptide. Such peptide may be subsequently cleaved from the beads, and assayed using the same antigen-binding molecule to identify the most active peptide sequence.

The second approach, the chemical ratio method, prepares mixed peptide resins using a specific ratio of amino acids empirically defined to give equimolar incorporation of each amino acid at each coupling step. Each resin bead contains a mixture of peptides. Approximate equimolar representation can be confirmed by amino acid analysis (Dooley and Houghten, 1993, *Proc. Natl. Acad. Sci. U.S.A.* **90**: 10811-10815; Eichler and Houghten, 1993, *Biochemistry* **32**: 11035-11041). Preferably, the synthetic peptide library is produced on polyethylene rods, or pins, as a solid support, as for example disclosed by Geysen *et al.* (1986, *Mol. Immunol.* **23**: 709-715). An exemplary peptide library of this type may consist of octapeptides in which the third and fourth position are defined with

each of the 20 amino acids, whereas the remaining six positions are present as mixtures. This peptide library can be represented by the formula Ac-XXO$_1$O$_2$XXXX-S$_n$, where S$_n$ is the solid support. Peptide mixtures remain on the pins when assayed against a soluble receptor molecule. For example, the peptide library of Geysen (1986, *Immun. Today* **6**: 364-369; and Geysen *et al.*, *Ibid*), comprising for example dipeptides, is first screened for the ability to bind to a target molecule. The most active dipeptides are then selected for an additional round of testing comprising linking, to the starting dipeptide, an additional residue (or by internally modifying the components of the original starting dipeptide) and then screening this set of candidates for the desired activity. This process is reiterated until the binding partner having the desired properties is identified.

*2.4.3. Alanine scanning mutagenesis*

In one embodiment, the invention herein utilises a systematic analysis of a polypeptide or polypeptide fragment according to the invention to determine the residues in the polypeptide or fragment that are involved in the interaction with a hyaluronidase substrate (*i.e.*, hyaluronic acid). Such analysis is conveniently performed using recombinant DNA technology. In general, a DNA sequence encoding the polypeptide or fragment is cloned and manipulated so that it may be expressed in a convenient host. DNA encoding the polypeptide or fragment can be obtained from a genomic library, from cDNA derived from mRNA in cells expressing the said polypeptide or fragment, or by synthetically constructing the DNA sequence (Sambrook *et al.*, *supra*; Ausubel *et al.*, *supra*).

The wild-type DNA encoding the polypeptide or fragment is then inserted into an appropriate plasmid or vector as described herein. In particular, prokaryotes are preferred for cloning and expressing DNA sequences to produce variants of the polypeptide or fragment. For example, *E. coli* K12 strain 294 (ATCC No. 31446) may be used, as well as *E. coli* B, *E. coli* X1776 (ATCC No. 31537), and *E. coli* c600 and c600hfl, and *E. coli* W3110 ($F^-$, $^-$, prototrophic, ATCC No. 27325), bacilli such as *Bacillus subtilis*, and other *enterobacteriaceae* such as *Salmonella typhimurium* or *Serratia marcescens*, and various *Pseudomonas* species. A preferred prokaryote is *E. coli* W3110 (ATCC 27325).

Once the polypeptide or fragment is cloned, site-specific mutagenesis as for example described by Carter *et al.* (1986, *Nucl. Acids. Res.*, **13**: 4331) or by Zoller *et al.* (1987, *Nucl. Acids Res.*, **10**: 6487), cassette mutagenesis as for example described by Wells *et al.* (1985, *Gene*, **34**: 315), restriction selection mutagenesis as for example described by Wells *et al.* (1986, *Philos. Trans. R. Soc. London SerA*, **317**: 415), or other known techniques may be performed on the cloned DNA to produce the variant DNA that encodes for the changes in amino acid sequence defined by the residues being substituted. When operably linked to an appropriate expression vector, variants are obtained. In some cases, recovery of the variant may be facilitated by expressing and secreting such molecules from the expression host by use of an appropriate signal sequence operably linked to the DNA sequence encoding the variant. Such methods are well known to those skilled in the art. Of course, other methods may be employed to produce such polypeptides or fragments such as the *in vitro* chemical synthesis of the desired polypeptide variant (Barany *et al.* In *The Peptides*, eds. E. Gross and J. Meienhofer (Academic Press: N.Y. 1979), Vol. 2, pp. 3-254).

Once the different variants are produced, they are contacted with a hyaluronidase substrate and the interaction, if any, between this substrate and each variant is determined. These activities are compared to the activity of the parent polypeptide or fragment with the same hyaluronidase substrate molecule to determine which of the amino acid residues in the active domain are involved in the interaction with the hyaluronidase substrate. The scanning amino acid used in such an analysis may be any different amino acid from that substituted, *i.e.*, any of the 19 other naturally occurring amino acids.

The interaction between the hyaluronidase substrate, and parent and variant polypeptides, respectively, can be measured by any convenient assay as for example described herein. Suitable methods of determining such interaction are given, for example, in Section 2.3. While any number of analytical measurements may be used to compare activities, a convenient one for hydrolysis of a hyaluronidase substrate is the rate of cumulus dispersal of an oocyte-cumulus complex. Generally, a two-fold increase or decrease in this rate per analogous residue substituted by the substitution indicates that the substituted residue(s) is active in the interaction of the parent polypeptide or fragment with the target hyaluronidase substrate.

When a suspected or known active amino acid residue is subjected to scanning amino acid analysis, the amino acid residues immediately adjacent thereto should be scanned. Three residue-substituted polypeptides can be made. One contains a scanning amino acid, preferably alanine, at position N that is the suspected or known active amino acid. The two others contain the scanning amino acid at position N+1 and N-1. If each substituted polypeptide or fragment causes a greater than about two-fold effect on the rate of cumulus dispersal of an oocyte-cumulus complex, the scanning amino acid is substituted at position N+2 and N-2. This is repeated until at least one, and preferably four, residues are identified in each direction which have less than about a two-fold effect on the said rate or either of the ends of the parent polypeptide or fragment are reached. In this manner, one or more amino acids along a continuous amino acid sequence that are involved in the interaction with the particular hyaluronidase substrate molecule can be identified.

The active amino acid residue identified by amino acid scan is typically one that contacts the hyaluronidase substrate directly. However, active amino acids may also indirectly contact the hyaluronidase substrate through salt bridges formed with other residues or small molecules such as $H_2O$ or ionic species such as $Na^+$, $Ca^{+2}$, $Mg^{+2}$, or $Zn^{+2}$.

In some cases, the substitution of a scanning amino acid at one or more residues results in a residue-substituted polypeptide which is not expressed at levels that allow for the isolation of quantities sufficient to carry out analysis of its activity with the hyaluronidase substrate. In such cases, a different scanning amino acid, preferably an isosteric amino acid, can be used.

Among the preferred scanning amino acids are relatively small, neutral amino acids. Such amino acids include alanine, glycine, serine, and cysteine. Alanine is the preferred scanning amino acid among this group because it eliminates the side-chain beyond the beta-carbon and is less likely to alter the main-chain conformation of the variant. Alanine is also preferred because it is the most common amino acid. Further, it is frequently found in both buried and exposed positions (Creighton, *The Proteins*, W. H. Freeman & Co., N.Y.; Chothia, 1976, *J. Mol. Biol.*, **150**: 1). If alanine substitution does not yield adequate amounts of variant, an isosteric amino acid can be used. Alternatively,

the following amino acids in decreasing order of preference may be used: Ser, Asn, and Leu.

Once the active amino acid residues are identified, isosteric amino acids may be substituted. Such isosteric substitutions need not occur in all instances and may be performed before any active amino acid is identified. Such isosteric amino acid substitution is performed to minimise the potential disruptive effects on conformation that some substitutions can cause. Isosteric amino acids are shown in the table below:

**TABLE B**

| *Polypeptide Amino Acid* | *Isosteric Scanning Amino Acid* |
|---|---|
| Ala (A) | Ser, Gly |
| Glu (E) | Gln, Asp |
| Gln (Q) | Asn, Glu |
| Asp (D) | Asn, Glu |
| Asn (N) | Ala, Asp |
| Leu (L) | Met, Ile |
| Gly (G) | Pro, Ala |
| Lys (K) | Met, Arg |
| Ser (S) | Thr, Ala |
| Val (V) | Ile, Thr |
| Arg (R) | Lys, Met, Asn |
| Thr (T) | Ser, Val |
| Pro (P) | Gly |

| | |
|---|---|
| Ile (I) | Met, Leu, Val |
| Met (M) | Ile, Leu |
| Phe (F) | Tyr |
| Tyr (Y) | Phe |
| Cys (C) | Ser, Ala |
| Trp (W) | Phe |
| His (H) | Asn, Gln |

The method herein can be used to detect active amino acid residues within different domains of a polypeptide or fragment according to the invention. Once this identification is made, various modifications to the parent polypeptide or fragment may be made to modify the interaction between the parent polypeptide or fragment and a hyaluronidase substrate.

*2.4.4. Rational drug design*

Variants of a recombinant polypeptide according to the invention or a biologically active fragment thereof may also be obtained using the principles of conventional or of rational drug design as for example described by Andrews, *et al.* (In: "PROCEEDINGS OF THE ALFRED BENZON SYMPOSIUM", volume 28, pp. 145-165, Munksgaard, Copenhagen, 1990), McPherson, A. (1990, *Eur. J. Biochem.* **189**: 1-24), Hol,. *et al.* (In: "MOLECULAR RECOGNITION: CHEMICAL AND BIOCHEMICAL PROBLEMS", Roberts, S. M. (ed.); Royal Society of Chemistry; pp. 84-93, 1989), Hol, W. G. J. (1989, *Arzneim-Forsch.* **39**: 1016-1018), Hol, W. G. J. (1986, *Agnew Chem. Int. Ed. Engl.* **25**: 767-778).

In accordance with the methods of conventional drug design, the desired variant molecules are obtained by randomly testing molecules whose structures have an attribute in common with the structure of a parent polypeptide or biologically active fragment

according to the invention. The quantitative contribution that results from a change in a particular group of a binding molecule can be determined by measuring the capacity of competition or cooperativity between the parent polypeptide or polypeptide fragment and the candidate polypeptide variant.

In one embodiment of rational drug design, the polypeptide variant is designed to share an attribute of the most stable three-dimensional conformation of a polypeptide or polypeptide fragment according to the invention. Thus, the variant may be designed to possess chemical groups that are oriented in a way sufficient to cause ionic, hydrophobic, or van der Waals interactions that are similar to those exhibited by the immuno-interactive polypeptide or polypeptide fragment. In a second method of rational design, the capacity of a particular polypeptide or polypeptide fragment to undergo conformational "breathing" is exploited. Such "breathing" - the transient and reversible assumption of a different molecular conformation - is a well-appreciated phenomenon, and results from temperature, thermodynamic factors, and from the catalytic activity of the molecule. Knowledge of the 3-dimensional structure of the polypeptide or polypeptide fragment facilitates such an evaluation. An evaluation of the natural conformational changes of a polypeptide or polypeptide fragment facilitates the recognition of potential hinge sites, potential sites at which hydrogen bonding, ionic bonds or van der Waals bonds might form or might be eliminated due to the breathing of the molecule, etc. Such recognition permits the identification of the additional conformations that the polypeptide or polypeptide fragment could assume, and enables the rational design and production of mimetic polypeptide variants that share such conformations.

The preferred method for performing rational mimetic design employs a computer system capable of forming a representation of the three-dimensional structure of the polypeptide or polypeptide fragment (such as those obtained using RIBBON (Priestle, J., 1988, *J. Mol. Graphics* **21**: 572), QUANTA (Polygen), InSite (Biosyn), or Nanovision (American Chemical Society)). Such analyses are exemplified by Hol, *et al.* (In: "MOLECULAR RECOGNITION: CHEMICAL AND BIOCHEMICAL PROBLEMS", *supra*, Hol, W. G. J. (1989, *supra*) and Hol, W. G. J., (1986, *supra*).

In lieu of such direct comparative evaluations of candidate polypeptide variants, screening assays may be used to identify such molecules. Such assays will preferably exploit the capacity of the variant to hydrolyse a hyaluronidase substrate or to disperse cumulus cells from an oocyte-cumulus complex as described, for example, in Section 2.2.

2.5. Polypeptide derivatives

With reference to suitable derivatives of the invention, such derivatives include amino acid deletions and/or additions to a polypeptide, fragment or variant of the invention, wherein said derivatives have hyaluronidase activity and activity for dispersing cumulus cells from an oocyte-cumulus complex. *"Additions"* of amino acids may include fusion of the polypeptides, fragments and polypeptide variants of the invention with other polypeptides or proteins. For example, it will be appreciated that said polypeptides, fragments or variants may be incorporated into larger polypeptides, and that such larger polypeptides may also be expected to comprise hyaluronidase activity and activity for dispersing cumulus cells from an oocyte-cumulus complex.

The polypeptides, fragments or variants of the invention may be fused to a further protein, for example, which is not derived from the original host. The further protein may assist in the purification of the fusion protein. For instance, a polyhistidine tag or a maltose binding protein may be used in this respect as described in more detail below. Other possible fusion proteins are those which produce an immunomodulatory response. Particular examples of such proteins include Protein A or glutathione S-transferase (GST).

Other derivatives contemplated by the invention include, but are not limited to, modification to side chains, incorporation of unnatural amino acids and/or their derivatives during peptide, polypeptide or protein synthesis and the use of crosslinkers and other methods which impose conformational constraints on the polypeptides, fragments and variants of the invention.

Examples of side chain modifications contemplated by the present invention include modifications of amino groups such as by acylation with acetic anhydride; acylation of amino groups with succinic anhydride and tetrahydrophthalic anhydride; amidination with methylacetimidate; carbamoylation of amino groups with cyanate;

pyridoxylation of lysine with pyridoxal-5-phosphate followed by reduction with $NaBH_4$; reductive alkylation by reaction with an aldehyde followed by reduction with $NaBH_4$; and trinitrobenzylation of amino groups with 2, 4, 6-trinitrobenzene sulphonic acid (TNBS).

The carboxyl group may be modified by carbodiimide activation via O-acylisourea formation followed by subsequent derivitisation, by way of example, to a corresponding amide.

The guanidine group of arginine residues may be modified by formation of heterocyclic condensation products with reagents such as 2,3-butanedione, phenylglyoxal and glyoxal.

Sulphydryl groups may be modified by methods such as performic acid oxidation to cysteic acid; formation of mercurial derivatives using 4-chloromercuriphenylsulphonic acid, 4-chloromercuribenzoate; 2-chloromercuri-4-nitrophenol, phenylmercury chloride, and other mercurials; formation of a mixed disulphides with other thiol compounds; reaction with maleimide, maleic anhydride or other substituted maleimide; carboxymethylation with iodoacetic acid or iodoacetamide; and carbamoylation with cyanate at alkaline pH.

Tryptophan residues may be modified, for example, by alkylation of the indole ring with 2-hydroxy-5-nitrobenzyl bromide or sulphonyl halides or by oxidation with N-bromosuccinimide.

Tyrosine residues may be modified by nitration with tetranitromethane to form a 3-nitrotyrosine derivative.

The imidazole ring of a histidine residue may be modified by N-carbethoxylation with diethylpyrocarbonate or by alkylation with iodoacetic acid derivatives.

Examples of incorporating unnatural amino acids and derivatives during peptide synthesis include but are not limited to, use of 4-amino butyric acid, 6-aminohexanoic acid, 4-amino-3-hydroxy-5-phenylpentanoic acid, 4-amino-3-hydroxy-6-methylheptanoic acid, t-butylglycine, norleucine, norvaline, phenylglycine, ornithine, sarcosine, 2-thienyl

alanine and/or D-isomers of amino acids. A list of unnatural amino acids contemplated by the present invention is shown in TABLE C.

**TABLE C**

| *Non-conventional amino acid* | *Non-conventional amino acid* |
|---|---|
| α-aminobutyric acid | L-N-methylalanine |
| α-amino--methylbutyrate | L-N-methylarginine |
| aminocyclopropane-carboxylate | L-N-methylasparagine |
| aminoisobutyric acid | L-N-methylaspartic acid |
| aminonorbornyl-carboxylate | L-N-methylcysteine |
| cyclohexylalanine | L-N-methylglutamine |
| cyclopentylalanine | L-N-methylglutamic acid |
| L-N-methylisoleucine | L-N-methylhistidine |
| D-alanine | L-N-methylleucine |
| D-arginine | L-N-methyllysine |
| D-aspartic acid | L-N-methylmethionine |
| D-cysteine | L-N-methylnorleucine |
| D-glutamate | L-N-methylnorvaline |
| D-glutamic acid | L-N-methylornithine |
| D-histidine | L-N-methylphenylalanine |
| D-isoleucine | L-N-methylproline |
| D-leucine | L-N-medlylserine |

| | |
|---|---|
| D-lysine | L-N-methylthreonine |
| D-methionine | L-N-methyltryptophan |
| D-ornithine | L-N-methyltyrosine |
| D-phenylalanine | L-N-methylvaline |
| D-proline | L-N-methylethylglycine |
| D-serine | L-N-methyl-t-butylglycine |
| D-threonine | L-norleucine |
| D-tryptophan | L-norvaline |
| D-tyrosine | α-methyl-aminoisobutyrate |
| D-valine | α-methyl-aminobutyrate |
| D-methylalanine | α-methylcyclohexylalanine |
| D-methylarginine | α-methylcylcopentylalanine |
| D-methylasparagine | α-methyl-napthylalanine |
| D-methylaspartate | α-methylpenicillamine |
| D-methylcysteine | N-(4-aminobutyl)glycine |
| D-methylglutamine | N-(2-aminoethyl)glycine |
| D-methylhistidine | N-(3-aminopropyl)glycine |
| D-methylisoleucine | N-amino-methylbutyrate |
| D-methylleucine | α-napthylalanine |
| D-methyllysine | N-benzylglycine |

| | |
|---|---|
| D-methylmethionine | N-(2-carbamylediyl)glycine |
| D-methylornithiine | N-(carbamylmethyl)glycine |
| D-methylphenylalanine | N-(2-carboxyethyl)glycine |
| D-methylproline | N-(carboxymethyl)glycine |
| D-methylserine | N-cyclobutylglycine |
| D-methylthreonine | N-cycloheptylglycine |
| D-methyltryptophan | N-cyclohexylglycine |
| D-methyltyrosine | N-cyclodecylglycine |
| L-methylleucine | L-methyllysine |
| L-methylmethionine | L-methylnorleucine |
| L-methylnorvatine | L-methylornithine |
| L-methylphenylalanine | L-methylproline |
| L-methylserine | L-methylthreonine |
| L-methyltryptophan | L-methyltyrosine |
| L-methylvaline | L-N-methylhomophenylalanine |
| N-(N-(2,2-diphenylethyl carbamylmethyl)glycine | N-(N-(3,3-diphenylpropyl carbamylmethyl)glycine |
| 1-carboxy-1-(2,2-diphenyl-ethyl amino)cyclopropane | |

Also contemplated is the use of crosslinkers, for example, to stabilise 3D conformations of the polypeptides, fragments or variants of the invention, using homo-bifunctional cross linkers such as bifunctional imido esters having $(CH_2)_n$ spacer groups

with n = 1 to n = 6, glutaraldehyde, N-hydroxysuccinimide esters and hetero-bifunctional reagents which usually contain an amino-reactive moiety such as N-hydroxysuccinimide and another group specific-reactive moiety such as maleimido or dithio moiety or carbodiimide. In addition, peptides can be conformationally constrained, for example, by introduction of double bonds between C and C atoms of amino acids, by incorporation of C and N-methylamino acids, and by formation of cyclic peptides or analogues by introducing covalent bonds such as forming an amide bond between the N and C termini between two side chains or between a side chain and the N or C terminus of the peptides or analogues. For example, reference may be made to: Marlowe (1993, *Biorganic & Medicinal Chemistry Letters* **3**: 437-44) who describes peptide cyclisation on TFA resin using trimethylsilyl (TMSE) ester as an orthogonal protecting group; Pallin and Tam (1995, *J. Chem. Soc. Chem. Comm.* 2021-2022) who describe the cyclisation of unprotected peptides in aqueous solution by oxime formation; Algin *et al.* (1994, *Tetrahedron Letters* **35**: 9633-9636) who disclose solid-phase synthesis of head-to-tail cyclic peptides *via* lysine side-chain anchoring; Kates *et al* (1993, *Tetrahedron Letters* **34**: 1549-1552) who describe the production of head-to-tail cyclic peptides by three-dimensional solid phase strategy; Tumelty *et al.* (1994, *J. Chem. Soc. Chem. Comm.* 1067-1068) who describe the synthesis of cyclic peptides from an immobilised activated intermediate, wherein activation of the immobilised peptide is carried out with *N*-protecting group intact and subsequent removal leading to cyclisation; McMurray *et al.* (1994, *Peptide Research* **7**: 195-206) who disclose head-to-tail cyclisation of peptides attached to insoluble supports by means of the side chains of aspartic and glutamic acid; Hruby *et al.* (1994, *Reactive Polymers* **22**: 231-241) who teach an alternate method for cyclising peptides *via* solid supports; and Schmidt and Langer (1997, *J. Peptide Res.* **49**: 67-73) who disclose a method for synthesising cyclotetrapeptides and cyclopentapeptides. The foregoing methods may be used to produce conformationally constrained polypeptides that have hyaluronidase activity and activity for dispersing cumulus cells from an oocyte-cumulus complex.

The invention also contemplates polypeptides, fragments or variants of the invention that have been modified using ordinary molecular biological techniques so as to

improve their resistance to proteolytic degradation or to optimise solubility properties or to render them more suitable as an immunogenic agent.

*2.6. Methods of preparing the polypeptides of the invention*

Poypeptides of the inventions may be prepared by any suitable procedure known to those of skill in the art. For example, the polypeptides may be prepared by a procedure including the steps of: -

(a)preparing a recombinant polynucleotide comprising a nucleotide sequence encoding a polypeptide comprising the sequence set forth in SEQ ID NO: 2 or 6, or variant or derivative of these, which nucleotide sequence is operably linked to transcriptional and translational regulatory nucleic acid;

(b)introducing the recombinant polynucleotide into a suitable host cell;

(c)culturing the host cell to express recombinant polypeptide from said recombinant polynucleotide; and

(d)isolating the recombinant polypeptide.

Suitably, said nucleotide sequence comprises the sequence set forth in any one of SEQ ID NO: 1, 3 or 5.

The recombinant polynucleotide preferably comprises either an expression vector that may be a self-replicating extra-chromosomal vector such as a plasmid, or a vector that integrates into a host genome.

The transcriptional and translational regulatory nucleic acid will generally be appropriate for the host cell used for expression. Numerous types of appropriate expression vectors and suitable regulatory sequences are known in the art for a variety of host cells.

Typically, the transcriptional and translational regulatory nucleic acid may include, but is not limited to, promoter sequences, leader or signal sequences, ribosomal binding

sites, transcriptional start and stop sequences, translational start and termination sequences, and enhancer or activator sequences.

Constitutive or inducible promoters as known in the art are contemplated by the invention. The promoters may be either naturally occurring promoters, or hybrid promoters that combine elements of more than one promoter.

In a preferred embodiment, the expression vector contains a selectable marker gene to allow the selection of transformed host cells. Selection genes are well known in the art and will vary with the host cell used.

The expression vector may also include a fusion partner (typically provided by the expression vector) so that the recombinant polypeptide of the invention is expressed as a fusion polypeptide with said fusion partner. The main advantage of fusion partners is that they assist identification and/or purification of said fusion polypeptide.

In order to express said fusion polypeptide, it is necessary to ligate a polynucleotide according to the invention into the expression vector so that the translational reading frames of the fusion partner and the polynucleotide coincide.

Well known examples of fusion partners include, but are not limited to, glutathione-S-transferase (GST), Fc potion of human IgG, maltose binding protein (MBP) and hexahistidine ($HIS_6$), which are particularly useful for isolation of the fusion polypeptide by affinity chromatography. For the purposes of fusion polypeptide purification by affinity chromatography, relevant matrices for affinity chromatography are glutathione-, amylose-, and nickel- or cobalt-conjugated resins respectively. Many such matrices are available in "kit" form, such as the QIAexpress™ system (Qiagen) useful with ($HIS_6$) fusion partners and the Pharmacia GST purification system. In a preferred embodiment, the recombinant polynucleotide is expressed in the commercial vector QIAexpress™ pQE-30 as described more fully hereinafter.

Another fusion partner well known in the art is green fluorescent protein (GFP). This fusion partner serves as a fluorescent "tag" which allows the fusion polypeptide of the invention to be identified by fluorescence microscopy or by flow cytometry. The GFP tag is useful when assessing subcellular localisation of the fusion polypeptide of the invention,

or for isolating cells which express the fusion polypeptide of the invention. Flow cytometric methods such as fluorescence activated cell sorting (FACS) are particularly useful in this latter application.

Preferably, the fusion partners also have protease cleavage sites, such as for Factor $X_a$ or Thrombin, which allow the relevant protease to partially digest the fusion polypeptide of the invention and thereby liberate the recombinant polypeptide of the invention therefrom. The liberated polypeptide can then be isolated from the fusion partner by subsequent chromatographic separation.

Fusion partners according to the invention also include within their scope "epitope tags", which are usually short peptide sequences for which a specific antibody is available. Well known examples of epitope tags for which specific monoclonal antibodies are readily available include c-Myc, influenza virus, haemagglutinin and FLAG tags.

The step of introducing into the host cell the recombinant polynucleotide may be effected by any suitable method including transfection, and transformation, the choice of which will be dependent on the host cell employed. Such methods are well known to those of skill in the art.

Recombinant polypeptides of the invention may be produced by culturing a host cell transformed with an expression vector containing nucleic acid encoding a polypeptide, biologically active fragment, variant or derivative according to the invention. The conditions appropriate for protein expression will vary with the choice of expression vector and the host cell. This is easily ascertained by one skilled in the art through routine experimentation.

Suitable host cells for expression may be prokaryotic or eukaryotic. One preferred host cell for expression of a polypeptide according to the invention is a bacterium. The bacterium used may be *Escherichia coli*. Alternatively, the host cell may be an insect cell such as, for example, *SF9* cells that may be utilised with a baculovirus expression system.

The recombinant protein may be conveniently prepared by a person skilled in the art using standard protocols as for example described in Sambrook, *et al.*, MOLECULAR CLONING. A LABORATORY MANUAL (Cold Spring Harbor Press, 1989), in

particular Sections 16 and 17; Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY (John Wiley & Sons, Inc. 1994-1998), in particular Chapters 10 and 16; and Coligan *et al.*, CURRENT PROTOCOLS IN PROTEIN SCIENCE (John Wiley & Sons, Inc. 1995-1997), in particular Chapters 1, 5 and 6.

Alternatively, the polypeptide, fragments, variants or derivatives of the invention may be synthesised using solution synthesis or solid phase synthesis as described, for example, in Chapter 9 of Atherton and Shephard (*supra*) and in Roberge *et al.* (1995, *Science* **269**: 202).

***3. Polynucleotides of the invention***

3.1. Polynucleotides encoding polypeptides of the invention

The invention further provides a polynucleotide that encodes a polypeptide, fragment, variant or derivative as defined above. Suitably, the polynucleotide comprises the entire sequence of nucleotides set forth in SEQ ID NO: 1. SEQ ID NO: 1 corresponds to an 1864 bp cDNA sequence obtained by reverse transcriptase PCR amplification as will be more fully described hereinafter. This sequence defines: (1) a 5' untranslated region from nucleotide 1 through nucleotide 238 of SEQ ID NO: 1; (2) an open reading frame from nucleotide 239 through nucleotide 1670; and (3) a 3' untranslated region from nucleotide 1671 through nucleotide 1864. The aforementioned open reading frame encodes a precursor polypeptide comprising a leader peptide encoded by nucleotides 239 through 322, and a mature polypeptide encoded by nucleotides 323 through 1667. Suitably, the polynucleotide comprises the sequence set forth in SEQ ID NO: 3. SEQ ID NO: 3 defines the aforementioned open reading frame and thus encodes the said precursor polypeptide. Preferably, the polynucleotide comprises the sequence set forth in SEQ ID NO: 5, which corresponds to nucleotide 323 through nucleotide 1670 and thus encodes the said mature polypeptide. SEQ ID NO: 7 corresponds to nucleotide 239 through 322 of SEQ ID NO: 1 and thus encodes the leader polypeptide of the aforementioned precursor polypeptide.

3.2.Polynucleotides variants

In general, polynucleotide variants according to the invention comprise regions that show at least 70%, more suitably at least 80%, preferably at least 90%, and most preferably at least 95% sequence identity over a reference polynucleotide sequence of identical size ("*comparison window*") or when compared to an aligned sequence in which the alignment is performed by a computer homology program known in the art. What constitutes suitable variants may be determined by conventional techniques. For example, a polynucleotide according to any one of SEQ ID NO: 1, 3, 5 or 7 can be mutated using random mutagenesis (*e.g.*, transposon mutagenesis), oligonucleotide-mediated (or site-directed) mutagenesis, PCR mutagenesis and cassette mutagenesis of an earlier prepared variant or non-variant version of an isolated natural promoter according to the invention.

Oligonucleotide-mediated mutagenesis is a preferred method for preparing nucleotide substitution variants of a polynucleotide of the invention. This technique is well known in the art as, for example, described by Adelman *et al.* (1983, *DNA* **2**:183). Briefly, a polynucleotide according to any one of SEQ ID NO: 1, 3, 5 or 7 is altered by hybridising an oligonucleotide encoding the desired mutation to a template DNA, wherein the template is the single-stranded form of a plasmid or bacteriophage containing the unaltered or parent DNA sequence. After hybridisation, a DNA polymerase is used to synthesise an entire second complementary strand of the template that will thus incorporate the oligonucleotide primer, and will code for the selected alteration in said parent DNA sequence.

Generally, oligonucleotides of at least 25 nucleotides in length are used. An optimal oligonucleotide will have 12 to 15 nucleotides that are completely complementary to the template on either side of the nucleotide(s) coding for the mutation. This ensures that the oligonucleotide will hybridise properly to the single-stranded DNA template molecule.

The DNA template can be generated by those vectors that are either derived from bacteriophage M13 vectors, or those vectors that contain a single-stranded phage origin of replication as described by Viera *et al.* (1987, *Methods Enzymol.* **153**:3). Thus, the DNA that is to be mutated may be inserted into one of the vectors to generate single-stranded

template. Production of single-stranded template is described, for example, in Sections 4.21-4.41 of Sambrook *et al.* (1989, *supra*).

Alternatively, the single-stranded template may be generated by denaturing double-stranded plasmid (or other DNA) using standard techniques.

For alteration of the native DNA sequence, the oligonucleotide is hybridised to the single-stranded template under suitable hybridisation conditions. A DNA polymerising enzyme, usually the Klenow fragment of DNA polymerase I, is then added to synthesise the complementary strand of the template using the oligonucleotide as a primer for synthesis. A heteroduplex molecule is thus formed such that one strand of DNA encodes the mutated form of the polypeptide or fragment under test, and the other strand (the original template) encodes the native unaltered sequence of the polypeptide or fragment under test. This heteroduplex molecule is then transformed into a suitable host cell, usually a prokaryote such as *E. coli*. After the cells are grown, they are plated onto agarose plates and screened using the oligonucleotide primer having a detectable label to identify the bacterial colonies having the mutated DNA. The resultant mutated DNA fragments are then cloned into suitable expression hosts such as *E. coli* using conventional technology and clones that retain the desired antigenic activity are detected. Where the clones have been derived using random mutagenesis techniques, positive clones would have to be sequenced in order to detect the mutation.

Alternatively, linker-scanning mutagenesis of DNA may be used to introduce clusters of point mutations throughout a sequence of interest that has been cloned into a plasmid vector. For example, reference may be made to Ausubel *et al.*, *supra*, (in particular, Chapter 8.4) which describes a first protocol that uses complementary oligonucleotides and requires a unique restriction site adjacent to the region that is to be mutagenised. A nested series of deletion mutations is first generated in the region. A pair of complementary oligonucleotides is synthesised to fill in the gap in the sequence of interest between the linker at the deletion endpoint and the nearby restriction site. The linker sequence actually provides the desired clusters of point mutations as it is moved or "scanned" across the region by its position at the varied endpoints of the deletion mutation series. An alternate protocol is also described by Ausubel *et al.*, *supra*, which makes use

of site directed mutagenesis procedures to introduce small clusters of point mutations throughout the target region. Briefly, mutations are introduced into a sequence by annealing a synthetic oligonucleotide containing one or more mismatches to the sequence of interest cloned into a single-stranded M13 vector. This template is grown in an *E. coli* *dut⁻ ung⁻* strain, which allows the incorporation of uracil into the template strand. The oligonucleotide is annealed to the template and extended with T4 DNA polymerase to create a double-stranded heteroduplex. Finally, the heteroduplex is introduced into a wild-type *E. coli* strain, which will prevent replication of the template strand due to the presence of apurinic sites (generated where uracil is incorporated), thereby resulting in plaques containing only mutated DNA.

Region-specific mutagenesis and directed mutagenesis using PCR may also be employed to construct polynucleotide variants according to the invention. In this regard, reference may be made, for example, to Ausubel *et al.*, *supra*, in particular Chapters 8.2A and 8.5.

Alternatively, suitable polynucleotide sequence variants of the invention may be prepared according to the following procedure:

(a)creating primers which are optionally degenerate wherein each comprises a portion of a reference polynucleotide encoding a reference polypeptide or fragment of the invention, preferably encoding the sequence set forth in any one of SEQ ID NO: 2, or 6;

(b)obtaining a nucleic acid extract from a *Synanceja* species, preferably *S. verrucosa* and *S. horrida*; and

(c)using said primers to amplify, via nucleic acid amplification techniques, at least one amplification product from said nucleic acid extract, wherein said amplification product corresponds to a polynucleotide variant.

Suitable nucleic acid amplification techniques are well known to the skilled addressee, and include polymerase chain reaction (PCR) as for example described in Ausubel *et al.* (*supra*); strand displacement amplification (SDA) as for example described in U.S. Patent No 5,422,252; rolling circle replication (RCR) as for example described in Liu *et al.*, (1996, *J. Am. Chem. Soc.* **118**:1587-1594 and International application WO

92/01813) and Lizardi *et al.*, (International Application WO 97/19193); nucleic acid sequence-based amplification (NASBA) as for example described by Sooknanan *et al.*, (1994, *Biotechniques* **17**:1077-1080); and Q- replicase amplification as for example described by Tyagi *et al.*, (1996, *Proc. Natl. Acad. Sci. USA* **93**: 5395-5400).

Typically, polynucleotide variants that are substantially complementary to a reference polynucleotide are identified by blotting techniques that include a step whereby nucleic acids are immobilised on a matrix (preferably a synthetic membrane such as nitrocellulose), followed by a hybridisation step, and a detection step. Southern blotting is used to identify a complementary DNA sequence; northern blotting is used to identify a complementary RNA sequence. Dot blotting and slot blotting can be used to identify complementary DNA/DNA, DNA/RNA or RNA/RNA polynucleotide sequences. Such techniques are well known by those skilled in the art, and have been described in Ausubel *et al.* (1994-1998, *supra*) at pages 2.9.1 through 2.9.20.

According to such methods, Southern blotting involves separating DNA molecules according to size by gel electrophoresis, transferring the size-separated DNA to a synthetic membrane, and hybridising the membrane-bound DNA to a complementary nucleotide sequence labelled radioactively, enzymatically or fluorochromatically. In dot blotting and slot blotting, DNA samples are directly applied to a synthetic membrane prior to hybridisation as above.

An alternative blotting step is used when identifying complementary polynucleotides in a cDNA or genomic DNA library, such as through the process of plaque or colony hybridisation. A typical example of this procedure is described in Sambrook *et al.* ("Molecular Cloning. A Laboratory Manual", Cold Spring Harbour Press, 1989) Chapters 8-12.

Typically, the following general procedure can be used to determine hybridisation conditions. Polynucleotides are blotted/transferred to a synthetic membrane, as described above. A reference polynucleotide such as a polynucleotide of the invention is labelled as described above, and the ability of this labelled polynucleotide to hybridise with an immobilised polynucleotide is analysed.

A skilled addressee will recognise that a number of factors influence hybridisation. The specific activity of radioactively labelled polynucleotide sequence should typically be greater than or equal to about $10^8$ dpm/mg to provide a detectable signal. A radiolabelled nucleotide sequence of specific activity $10^8$ to $10^9$ dpm/mg can detect approximately 0.5 pg of DNA. It is well known in the art that sufficient DNA must be immobilised on the membrane to permit detection. It is desirable to have excess immobilised DNA, usually 10 µg. Adding an inert polymer such as 10% (w/v) dextran sulfate (MW 500,000) or polyethylene glycol 6000 during hybridisation can also increase the sensitivity of hybridisation (see Ausubel *supra* at 2.10.10).

To achieve meaningful results from hybridisation between a polynucleotide immobilised on a membrane and a labelled polynucleotide, a sufficient amount of the labelled polynucleotide must be hybridised to the immobilised polynucleotide following washing. Washing ensures that the labelled polynucleotide is hybridised only to the immobilised polynucleotide with a desired degree of complementarity to the labelled polynucleotide.

It will be understood that polynucleotide variants according to the invention will hybridise to a reference polynucleotide under at least low stringency conditions. Reference herein to low stringency conditions include and encompass from at least about 1% v/v to at least about 15% v/v formamide and from at least about 1 M to at least about 2 M salt for hybridisation at 42°C, and at least about 1 M to at least about 2 M salt for washing at 42°C. Low stringency conditions also may include 1% Bovine Serum Albumin (BSA), 1 mM EDTA, 0.5 M $NaHPO_4$ (pH 7.2), 7% SDS for hybridisation at 65°C, and (i) 2xSSC, 0.1% SDS; or (ii) 0.5% BSA, 1 mM EDTA, 40 mM $NaHPO_4$ (pH 7.2), 5% SDS for washing at room temperature.

Suitably, the polynucleotide variants hybridise to a reference polynucleotide under at least medium stringency conditions. Medium stringency conditions include and encompass from at least about 16% v/v to at least about 30% v/v formamide and from at least about 0.5 M to at least about 0.9 M salt for hybridisation at 42°C, and at least about 0.5 M to at least about 0.9 M salt for washing at 42°C. Medium stringency conditions also may include 1% Bovine Serum Albumin (BSA), 1 mM EDTA, 0.5 M $NaHPO_4$ (pH 7.2),

7% SDS for hybridisation at 65°C, and (i) 2 x SSC, 0.1% SDS; or (ii) 0.5% BSA, 1 mM EDTA, 40 mM $NaHPO_4$ (pH 7.2), 5% SDS for washing at 42°C.

Preferably, the polynucleotide variants hybridise to a reference polynucleotide under high stringency conditions. High stringency conditions include and encompass from at least about 31% v/v to at least about 50% v/v formamide and from at least about 0.01 M to at least about 0.15 M salt for hybridisation at 42°C, and at least about 0.01 M to at least about 0.15 M salt for washing at 42°C. High stringency conditions also may include 1% BSA, 1 mM EDTA, 0.5 M $NaHPO_4$ (pH 7.2), 7% SDS for hybridisation at 65°C, and (i) 0.2 x SSC, 0.1% SDS; or (ii) 0.5% BSA, 1mM EDTA, 40 mM $NaHPO_4$ (pH 7.2), 1% SDS for washing at a temperature in excess of 65°C.

Other stringent conditions are well known in the art. A skilled addressee will recognise that various factors can be manipulated to optimise the specificity of the hybridisation. Optimisation of the stringency of the final washes can serve to ensure a high degree of hybridisation. For detailed examples, see Ausubel *et al.*, *supra* at pages 2.10.1 to 2.10.16 and Sambrook *et al.* (1989, *supra*) at sections 1.101 to 1.104.

While stringent washes are typically carried out at temperatures from about 42°C to 68°C, one skilled in the art will appreciate that other temperatures may be suitable for stringent conditions. Maximum hybridisation typically occurs at about 20°C to 25°C below the $T_m$ for formation of a DNA-DNA hybrid. It is well known in the art that the $T_m$ is the melting temperature, or temperature at which two complementary polynucleotide sequences dissociate. Methods for estimating $T_m$ are well known in the art (see Ausubel *et al.*, *supra* at page 2.10.8).

In general, washing is carried out at T = 69.3 + 0.41 (G + C) % 12°C. However, the $T_m$ of a duplex DNA decreases by 1°C with every increase of 1% in the number of mismatched base pairs.

In a preferred hybridisation procedure, a membrane (*e.g.*, a nitrocellulose membrane or a nylon membrane) containing immobilised DNA is hybridised overnight at 42°C in a hybridisation buffer (50% deionised formamide, 5xSSC, 5x Denhardt's solution (0.1% ficoll, 0.1% polyvinylpyrollidone and 0.1% bovine serum albumin), 0.1% SDS and

200 mg/mL denatured salmon sperm DNA) containing labelled probe. The membrane is then subjected to two sequential medium stringency washes (*i.e.*, 2xSSC/0.1% SDS for 15 min at 45°C, followed by 2xSSC/0.1% SDS for 15 min at 50°C), followed by two sequential high stringency washes (*i.e.*, 0.2xSSC/0.1% SDS for 12 min at 55°C followed by 0.2xSSC and 0.1%SDS solution for 12 min).

Methods for detecting a labelled polynucleotide hybridised to an immobilised polynucleotide are well known to practitioners in the art. Such methods include autoradiography, phosphorimaging, and chemiluminescent, fluorescent and colorimetric detection.

***4. Antigen-binding molecules***

The invention also contemplates antigen-binding molecules against the aforementioned polypeptides, fragments, variants and derivatives. For example, the antigen-binding molecules may comprise whole polyclonal antibodies. Such antibodies may be prepared, for example, by injecting a polypeptide, fragment, variant or derivative of the invention into a production species, which may include mice or rabbits, to obtain polyclonal antisera. Methods of producing polyclonal antibodies are well known to those skilled in the art. Exemplary protocols which may be used are described for example in Coligan *et al.*, CURRENT PROTOCOLS IN IMMUNOLOGY, (John Wiley & Sons, Inc, 1991), and Ausubel *et al.*, (1994-1998, *supra*), in particular Section III of Chapter 11.

In lieu of the polyclonal antisera obtained in the production species, monoclonal antibodies may be produced using the standard method as described, for example, by Köhler and Milstein (1975, *Nature* **256**, 495-497), or by more recent modifications thereof as described, for example, in Coligan *et al.*, (1991, *supra*) by immortalising spleen or other antibody producing cells derived from a production species which has been inoculated with one or more of the polypeptides, fragments, variants or derivatives of the invention.

The invention also contemplates as antigen-binding molecules Fv, Fab, Fab' and $F(ab')_2$ immunoglobulin fragments.

Alternatively, the antigen-binding molecule may comprise a synthetic stabilised Fv fragment. Exemplary fragments of this type include single chain Fv fragments (sFv, frequently termed scFv) in which a peptide linker is used to bridge the N terminus or C terminus of a $V_H$ domain with the C terminus or N-terminus, respectively, of a $V_L$ domain. ScFv lack all constant parts of whole antibodies and are not able to activate complement. Suitable peptide linkers for joining the $V_H$ and $V_L$ domains are those which allow the $V_H$ and $V_L$ domains to fold into a single polypeptide chain having an antigen binding site with a three dimensional structure similar to that of the antigen binding site of a whole antibody from which the Fv fragment is derived. Linkers having the desired properties may be obtained by the method disclosed in U.S. Patent No 4,946,778. However, in some cases a linker is absent. ScFvs may be prepared, for example, in accordance with methods outlined in Kreber *et al.* (Krebber *et al.* 1997, *J. Immunol. Methods*; **201(1)**: 35-55). Alternatively, they may be prepared by methods described in U.S. Patent No 5,091,513, European Patent No 239,400 or the articles by Winter and Milstein (1991, *Nature* **349**:293) and Plückthun *et al.* (1996, In *Antibody engineering: A practical approach*. 203-252).

Alternatively, the synthetic stabilised Fv fragment comprises a disulphide stabilised Fv (dsFv) in which cysteine residues are introduced into the $V_H$ and $V_L$ domains such that in the fully folded Fv molecule the two residues will form a disulphide bond therebetween. Suitable methods of producing dsFv are described for example in (Glockscuther *et al.* *Biochem.* **29**: 1363-1367; Reiter *et al.* 1994, *J. Biol. Chem.* **269**: 18327-18331; Reiter *et al.* 1994, *Biochem. 33*: 5451-5459; Reiter *et al.* 1994. *Cancer Res.* **54**: 2714-2718; Webber *et al.* 1995, *Mol. Immunol.* **32**: 249-258).

Also contemplated as antigen-binding molecules are single variable region domains (termed dAbs) as for example disclosed in (Ward *et al.* 1989, *Nature* **341**: 544-546; Hamers-Casterman *et al.* 1993, *Nature.* **363**: 446-448; Davies & Riechmann, 1994, *FEBS Lett.* **339**: 285-290).

Alternatively, the antigen-binding molecule may comprise a "minibody". In this regard, minibodies are small versions of whole antibodies, which encode in a single chain the essential elements of a whole antibody. Suitably, the minibody is comprised of the $V_H$

and $V_L$ domains of a native antibody fused to the hinge region and CH3 domain of the immunoglobulin molecule as, for example, disclosed in U.S. Patent No 5,837,821.

In an alternate embodiment, the antigen binding molecule may comprise non-immunoglobulin derived, protein frameworks. For example, reference may be made to (Ku & Schultz, 1995, *Proc. Natl. Acad. Sci. USA*, **92**: 652-6556) which discloses a four-helix bundle protein cytochrome b562 having two loops randomised to create complementarity determining regions (CDRs), which have been selected for antigen binding.

The antigen-binding molecule may be multivalent (*i.e.*, having more than one antigen binding site). Such multivalent molecules may be specific for one or more antigens. Multivalent molecules of this type may be prepared by dimerisation of two antibody fragments through a cysteinyl-containing peptide as, for example disclosed by (Adams *et al.*, 1993, *Cancer Res.* **53**: 4026-4034; Cumber *et al.*, 1992, *J. Immunol.* **149**: 120-126). Alternatively, dimerisation may be facilitated by fusion of the antibody fragments to amphiphilic helices that naturally dimerise (Pack P. Plünckthun, 1992, *Biochem.* **31**: 1579-1584), or by use of domains (such as the leucine zippers jun and fos) that preferentially heterodimerise (Kostelny *et al.*, 1992, *J. Immunol.* **148**: 1547-1553). In an alternate embodiment, the multivalent molecule may comprise a multivalent single chain antibody (multi-scFv) comprising at least two scFvs linked together by a peptide linker. In this regard, non-covalently or covalently linked scFv dimers termed "diabodies" may be used. Multi-scFvs may be bispecific or greater depending on the number of scFvs employed having different antigen binding specificities. Multi-scFvs may be prepared for example by methods disclosed in U.S. Patent No. 5,892,020.

The antigen-binding molecules of the invention may be used for affinity chromatography in isolating a natural or recombinant polypeptide or biologically active fragment of the invention. For example reference may be made to immunoaffinity chromatographic procedures described in Chapter 9.5 of Coligan *et al.*, (1995-1997, *supra*).

The antigen-binding molecules can be used to screen expression libraries for variant polypeptides of the invention as described herein. They can also be used to detect polypeptides, fragments, variants and derivatives of the invention as described hereinafter.

***5. Methods of Detection***

The invention also extends to a method of detecting in a sample a polypeptide, fragment, variant or derivative as broadly described above, comprising contacting the sample with an antigen-binding molecule as described in Section 4 and detecting the presence of a complex comprising the said antigen-binding molecule and the said polypeptide, fragment, variant or derivative in said contacted sample.

Any suitable technique for determining formation of the complex may be used. For example, an antigen-binding molecule according to the invention, having a reporter molecule associated therewith may be utilised in immunoassays. Such immunoassays include, but are not limited to, radioimmunoassays (RIAs), enzyme-linked immunosorbent assays (ELISAs) and immunochromatographic techniques (ICTs), Western blotting which are well known those of skill in the art. For example, reference may be made to "CURRENT PROTOCOLS IN IMMUNOLOGY" (1994, *supra*) which discloses a variety of immunoassays that may be used in accordance with the present invention. Immunoassays may include competitive assays as understood in the art or as for example described *infra*. It will be understood that the present invention encompasses qualitative and quantitative immunoassays.

Suitable immunoassay techniques are described for example in US Patent Nos. 4,016,043, 4, 424,279 and 4,018,653. These include both single-site and two-site assays of the non-competitive types, as well as the traditional competitive binding assays. These assays also include direct binding of a labelled antigen-binding molecule to a target antigen.

Two site assays are particularly favoured for use in the present invention. A number of variations of these assays exist, all of which are intended to be encompassed by the present invention. Briefly, in a typical forward assay, an unlabelled antigen-binding molecule such as an unlabelled antibody is immobilised on a solid substrate and the

sample to be tested brought into contact with the bound molecule. After a suitable period of incubation, for a period of time sufficient to allow formation of an antibody-antigen complex, another antigen-binding molecule, suitably a second antibody specific to the antigen, labelled with a reporter molecule capable of producing a detectable signal is then added and incubated, allowing time sufficient for the formation of another complex of antibody-antigen-labelled antibody. Any unreacted material is washed away and the presence of the antigen is determined by observation of a signal produced by the reporter molecule. The results may be either qualitative, by simple observation of the visible signal, or may be quantitated by comparing with a control sample containing known amounts of antigen. Variations on the forward assay include a simultaneous assay, in which both sample and labelled antibody are added simultaneously to the bound antibody. These techniques are well known to those skilled in the art, including minor variations as will be readily apparent. In accordance with the present invention, the sample is one that might contain an antigen including serum, whole blood, and plasma or lymph fluid. The sample is, therefore, generally a circulatory sample comprising circulatory fluid.

In the typical forward assay, a first antibody having specificity for the antigen or antigenic parts thereof is either covalently or passively bound to a solid surface. The solid surface is typically glass or a polymer, the most commonly used polymers being cellulose, polyacrylamide, nylon, polystyrene, polyvinyl chloride or polypropylene. The solid supports may be in the form of tubes, beads, discs of microplates, or any other surface suitable for conducting an immunoassay. The binding processes are well known in the art and generally consist of cross-linking covalently binding or physically adsorbing, the polymer-antibody complex is washed in preparation for the test sample. An aliquot of the sample to be tested is then added to the solid phase complex and incubated for a period of time sufficient and under suitable conditions to allow binding of any antigen present to the antibody. Following the incubation period, the antigen-antibody complex is washed and dried and incubated with a second antibody specific for a portion of the antigen. The second antibody has generally a reporter molecule associated therewith that is used to indicate the binding of the second antibody to the antigen. The amount of labelled antibody that binds, as determined by the associated reporter molecule, is proportional to the amount of antigen bound to the immobilised first antibody.

An alternative method involves immobilising the antigen in the biological sample and then exposing the immobilised antigen to specific antibody that may or may not be labelled with a reporter molecule. Depending on the amount of target and the strength of the reporter molecule signal, a bound antigen may be detectable by direct labelling with the antibody. Alternatively, a second labelled antibody, specific to the first antibody is exposed to the target-first antibody complex to form a target-first antibody-second antibody tertiary complex. The complex is detected by the signal emitted by the reporter molecule.

From the foregoing, it will be appreciated that the reporter molecule associated with the antigen-binding molecule may include the following:

(a)direct attachment of the reporter molecule to the antigen-binding molecule;

(b)indirect attachment of the reporter molecule to the antigen-binding molecule; *i.e.*, attachment of the reporter molecule to another assay reagent which subsequently binds to the antigen-binding molecule; and

(c)attachment to a subsequent reaction product of the antigen-binding molecule.

The reporter molecule may be selected from a group including a chromogen, a catalyst, an enzyme, a fluorochrome, a chemiluminescent molecule, a lanthanide ion such as Europium ($Eu^{34}$), a radioisotope and a direct visual label.

In the case of a direct visual label, use may be made of a colloidal metallic or non-metallic particle, a dye particle, an enzyme or a substrate, an organic polymer, a latex particle, a liposome, or other vesicle containing a signal producing substance and the like.

A large number of enzymes suitable for use as reporter molecules is disclosed in United States Patent Specifications U.S. 4,366,241, U.S. 4,843,000, and U.S. 4,849,338. Suitable enzymes useful in the present invention include alkaline phosphatase, horseradish peroxidase, luciferase, -galactosidase, glucose oxidase, lysozyme, malate dehydrogenase and the like. The enzymes may be used alone or in combination with a second enzyme that is in solution.

Suitable fluorochromes include, but are not limited to, fluorescein isothiocyanate (FITC), tetramethylrhodamine isothiocyanate (TRITC), R-Phycoerythrin (RPE), and Texas Red. Other exemplary fluorochromes include those discussed by Dower *et al.* (International Publication WO 93/06121). Reference also may be made to the fluorochromes described in U.S. Patents 5,573,909 (Singer *et al*), 5,326,692 (Brinkley *et al*). Alternatively, reference may be made to the fluorochromes described in U.S. Patent Nos. 5,227,487, 5,274,113, 5,405,975, 5,433,896, 5,442,045, 5,451,663, 5,453,517, 5,459,276, 5,516,864, 5,648,270 and 5,723,218.

In the case of an enzyme immunoassay, an enzyme is conjugated to the second antibody, generally by means of glutaraldehyde or periodate. As will be readily recognised, however, a wide variety of different conjugation techniques exist which are readily available to the skilled artisan. The substrates to be used with the specific enzymes are generally chosen for the production of, upon hydrolysis by the corresponding enzyme, a detectable colour change. Examples of suitable enzymes include those described *supra*. It is also possible to employ fluorogenic substrates, which yield a fluorescent product rather than the chromogenic substrates noted above. In all cases, the enzyme-labelled antibody is added to the first antibody-antigen complex. It is then allowed to bind, and excess reagent is washed away. A solution containing the appropriate substrate is then added to the complex of antibody-antigen-antibody. The substrate will react with the enzyme linked to the second antibody, giving a qualitative visual signal, which may be further quantitated, usually spectrophotometrically, to give an indication of the amount of antigen which was present in the sample.

Alternately, fluorescent compounds, such as fluorescein, rhodamine and the lanthanide, europium (EU), may be chemically coupled to antibodies without altering their binding capacity. When activated by illumination with light of a particular wavelength, the fluorochrome-labelled antibody adsorbs the light energy, inducing a state to excitability in the molecule, followed by emission of the light at a characteristic colour visually detectable with a light microscope. The fluorescent-labelled antibody is allowed to bind to the first antibody-antigen complex. After washing off the unbound reagent, the remaining tertiary complex is then exposed to light of an appropriate wavelength. The fluorescence observed indicates the presence of the antigen of interest.

Immunofluorometric assays (IFMA) are well established in the art. However, other reporter molecules, such as radioisotope, chemiluminescent or bioluminescent molecules may also be employed.

***6. Compositions***

The invention also provides a composition for use in enhancing fertilisation of a mammalian egg, comprising a recombinant polypeptide, biologically active fragment, variant or derivative as broadly described above (*"fertilising agents"*), together with a pharmaceutically acceptable carrier.

Also within the scope of the present invention use of a composition comprising a recombinant polypeptide, biologically active fragment, variant or derivative as broadly described above (*"therapeutic agents"*), together with a pharmaceutically acceptable carrier for use in a treatment selected from the group consisting of:

(a)increasing the effect of anaesthetics, preferably local anaesthetics and preferably in opthalmology;

(b)administrating fluids and drugs, preferably subcutaneously;

(c)resorption of traumatic oedema and haematoma;

(d)preventing tissue necrosis, particularly skin necrosis, during intravenous administration; and

(e)enhancing the efficacy of chemotherapy against cancer including, but mot restrcited to, melanoma, bladder cancer and breast cancer.

The rationale for use of the agents according to the invention is based in part on the chemical and structural homology between the novel recombinant *S. horrida* hyaluronidase and human PH-20 (GI 585674; Lin *et al*, *supra*). In this regard, Lin *et al.* found that PH-20 enables sperm to penetrate the cumulus cell layer surrounding the egg. A similar hyaluronidase activity in the recombinant polypeptides, biologically active fragments, variants and derivatives of the invention may be useful; for artificially inducing

fertilisation *in vitro* by providing additional enzyme to assist in the breakdown of hyaluronic acid that surrounds mammalian eggs.

Tu *et al.* (1993, *Comp. Biochem. Physiol.* **76**: 377-383) found that hyaluronidase promotes the spread of heamorrhagic area due to lizard venom. Agents that block hyaluronidase activity of the *S. horrida* hyaluronidase (*e.g.*, antigen-binding molecules that specifically bind to *S. horrida* hyaluronidase) may therefore have application in treatment of envenomation by *S. horrida* and related species.

Spruss *et al.* (1995, *J. Cancer Res. Clin. Oncol.* **121**: 193-202) found that hyaluronidase significantly enhances the efficacy of regional vinblastine chemotherapy of malignant melanoma.

Klocker *et al.* (1995, *Am. J. Clin. Oncol.* **18**: 425-428) found that the combined therapy of vindesine, cisplatin, hyaluronidase, and radiation is well tolerated by most patients and highly effective against advanced squamous cell cancer of the head and neck. The therapeutic agents of the invention may be similarly applied in chemotherapeutic regimens for melanomas, head and neck cancers, or other cancers.

Other applications of hyaluronidase discussed *supra* are also encompasses by the present invention.

A variety of pharmaceutically acceptable carriers, well known in the art may be used for the compositions of the invention. These carriers may be selected from sugars, starches, cellulose and its derivatives, malt, gelatine, talc, calcium sulphate, vegetable oils, synthetic oils, polyols, alginic acid, phosphate buffered solutions, emulsifiers, isotonic saline, and pyrogen-free water.

With regard to their use as therapeutic agents, the polypeptides, biologically-active fragments, variants and derivatives may be administered to a patient who requires one of said treatments or procedures mentioned above. Any suitable route of administration may be employed for providing a patient with a composition of the invention. For example, oral, rectal, parenteral, sublingual, buccal, intravenous, intra-articular, intra-muscular, intra-dermal, subcutaneous, inhalational, intraocular, intraperitoneal, intracerebroventricular, transdermal and the like may be employed.

Dosage forms include tablets, dispersions, suspensions, injections, solutions, syrups, troches, capsules, suppositories, aerosols, transdermal patches and the like. These dosage forms may also include injecting or implanting controlled releasing devices designed specifically for this purpose or other forms of implants modified to act additionally in this fashion. Controlled release of a therapeutic agent according to the invention may be effected by coating the same, for example, with hydrophobic polymers including acrylic resins, waxes, higher aliphatic alcohols, polylactic and polyglycolic acids and certain cellulose derivatives such as hydroxypropylmethyl cellulose. In addition, controlled release may be effected by using other polymer matrices, liposomes and/or microspheres.

Compositions suitable for oral or parenteral administration may be presented as discrete units such as capsules, sachets or tablets each containing a pre-determined amount of one or more immunogenic agents of the invention, as a powder or granules or as a solution or a suspension in an aqueous liquid, a non-aqueous liquid, an oil-in-water emulsion or a water-in-oil liquid emulsion. Such compositions may be prepared by any of the methods of pharmacy but all methods include the step of bringing into association one or more immunogenic agents as described above with the carrier which constitutes one or more necessary ingredients. In general, the compositions are prepared by uniformly and intimately admixing the immunogenic agents of the invention with liquid carriers or finely divided solid carriers or both, and then, if necessary, shaping the product into the desired presentation.

The above compositions may be administered in a manner compatible with the dosage formulation, and in such amount as is therapeutically effective. In this regard, the dose of therapeutic agent administered to a patient should be sufficient to effect a beneficial response in the patient over time such as, for example, a reduction in tissue necrosis during intravenous administration or to ameliorate a condition (*e.g.*, cancer) to be treated. The quantity of the therapeutic agent(s) to be administered may depend on the subject to be treated inclusive of the age, sex, weight and general health condition thereof. In this regard, precise amounts of the therapeutic agent(s) for administration will depend on the judgement of the practitioner. In any event, those of skill in the art may readily

determine suitable dosages of the therapeutic agents of the invention. Such dosages may be in the order of nanograms to milligrams of these agents.

In order that the invention may be readily understood and put into practical effect, particular preferred embodiments will now be described by way of the following non-limiting examples.

## EXAMPLES

### *EXAMPLE 1*

Cloning and sequence determination of the cDNA encoding SFHYA1

*MATERIALS AND METHODS*

*Isolation and partial peptide sequencing of SFHYA1.*

Crude venom was first extracted from freshly excised venom sacs of *Synaceja horrida* as decribed (Poh *et al.,* 1992) and lyophilized for storage at -70°C. One hundred and seventy milligrams of lyophilized crude venom was reconstituted in 2.5 mL of 0.05 M sodium phosphate buffer, pH 5.5. The dissolved venom was centrifugerd at 3000 *x* g for 5 minutes to remove any insoluble material. Two hundred microliters of the supernatant was then applied to a Superdex™ 200 HR 10/30 (Pharmacia Biotech) gel permeation column (10 *x* 300 min) previously equilibrated with the same buffer on a FPLC system. Elution was carried out at 60 mL/hr and 1-mL fractions were collected. Absorbance of column eluant was monitored at 280 nm automatically by the FPLC system. Fractions collected were assayed for hyaluronidase activities as described (Poh *et al.,* 1992). Active fractions obtained were pooled and loaded onto a heparin sepharose affinity (HiTrap™ Heparin, Pharmacia Biotech) column (5 mL bed volume) pre-equilibrated with 0.05 M sodium phosphate buffer, pH 5.5. After washing the column with start buffer, the bound proteins were eluted with a linear salt gradient (0- 0.15 M sodium chloride in start buffer, 60 mL) at a flow rate of 60 mL/hr. Two millilitre fractions were collected and hyaluronidases activities were assayed for these fractions. The Affi-Gel™ fractions with hyaluronidase activity were run on SDS PAGE with molecular weight according to the method of

Laemmli (Laemmli, 1970) and stained with Commassie blue to ascertain the purity of the hyaluronidase preparation.

Only hyaluronidase preparation of sufficient purity as shown by the absence of contaminating protein bands on a SDS PAGE gel stained with Commassie blue was used for determination of partial amino acid sequence. N terminal sequence and amino acid sequences of 3 In gel tryptic digestion fragments of SFHYA1 were carried out by microsequencing room of the IMCB as a commercial service.

*Alignment of hyaluronidase amino acid sequences*

Alignment of hyaluronidase amino acid sequences was carried out for the purpose of identifying any conserved stretches of amino acids residues among hyaluronidases of different species and different tissue origins. Amino acid sequences of hyaluronidases from mouse (GI 2370295; lysosomal hyaluronidase), Man (GI 2314820; plasma hyaluronidase), Honeybee (GI 585279; venom hyaluronidase), a crab eating mammal known as *Macaca fasicularis* (GI 585674; sperm surface protein) and *Caenorhabditis elegans* (GI 3879998; unknown tissue origin) were downloaded from databases at NCBI and aligned using the multisequence alignment program of DNASTAR software (DNASTAR Inc, Madison Wis.)

*Isolation of venom gland total RNA.*

Venom sacs excised from freshly killed *Synaceja horrida* were immediately homogenised with TRIAZOL™ (a reagent for RNA isolation from Gibco BRL) and total RNA were isolated according to methods described in the accompanying protocol. RNA agarose gel electrophoresis (Sambrook *et al.,* 1989) was carried out to verify the integrity of the RNA isolated.

*RT PCR*

A 10 μg sample of venom gland total RNA was denatured by heating at 70°C for 10 minutes and immediately quick-chilled on ice for at least 5 minutes prior to reaction. Reverse transcription was then carried out on this sample of denatured RNA in a 20 μL reaction volume consisting of 1 *x* RT buffer (50 mM Tris HCl pH 8.3, 6mM $MgCl_2$ mM KCl), 1 μM oligo $(dT)_{21}$, 1 mM dNTPs, 4 units of RNase inhibitor and 200 units of MMLV

by incubating the reaction mixture at 42°C for 1 hour. Reaction was terminated by heating at 95'C for 5 minutes. The sense primer [5'-G(G/C)IGTIAT(C/T)GA(C/T)(A/T)(G/T)IGA-3'] was designed based on the amino sequence A/GV I D W/F/I-E (as derived from alignment of hyaluronidase shown in Figure 1), while the antisense primer [5' AC(A/G)TC(A/G)TC(A/G)TCCCA(A/G)TG(T/C)TT 3'] was designed based on the amino sequence KHWDDDV (as derived from sequencing a tryptic fragment of SFHYA1), respectively. Inosine residues were included to reduce the number of oligonucleotide species in both the sense and the antisense primer pool to 32. One microliter of the reverse transcription reaction (corresponding to cDNA synthesised from 0.5 µg of total RNA) was used for amplification of a cDNA segment of SFHYA1 in a 50 µL volume polymerase chain reaction consisting of I *x* DynaZyme™ II buffer (10 mM Tris HCl pH 8.8, 1.5 mM $MgCl_2$, 50 mM KCl and 0. 1 % Triton X100), 25 µM of sense primer, 12.5 µM of antisense primer, 250 µM dNTPs and 2 units of DynaZyme™ II DNA polymerase (Finnzymes Oy, Finland). Thermal cycling conditions consisted of 5 stages and were as follows: Stage 1 containing only 1 cycle at 94°C for 2 minutes, Stage 2 containing 4 cycles at 94°C for 1 minute, 52°C for 1.5 minutes and 72°C for 1 minute, Stage 3 containing 4 cycles of 94°C for 1 minute, WC for 1.5 minutes and 72°C for 1 minute, Stage 4 containing 4 cycles of 94°C for 1 minute, 47°C for 1.5 minutes and 72°C for 1 minute and Stage 5 containing 30 cycles of 94°C for 1 minute, 44°C for 1.5 minutes and 72°C for 1 minute.

*Cloning of PCR products*

DNA fragments (of interest) amplified during PCRs were cloned into pCR Script™ Amp SK (+) cloning vector (Stratagene, Stratagene Cloning Systems, 11011 North Torrey Pines Road, La Jolla, CA 92037, USA) according to instructions as specified in the manual accompanying this kit.

*Nucleotide sequence determination*

DNA sequencing reactions were performed using the ABI PRISM BigDye™ Terminator Cycle Sequencing kit (PE Applied Biosystems, Warrington WA 4SR, Britian) according to manufacturer's instructions. Sequencing gel electrophoresis was carried out at National University Medical Institutes (#B1 16, Clinical Research Centre, National

University of Singapore, Block MD11, 25 Medical Drive, Singapore 117601) as a commercial service on Model 377 Automated DNA sequencer (PE Applied Biosystems, 850 Lincoln Centre Drive, Foster City, CA 94404, U. S. A.).

*Search for homologs*

BLASTX programme (Altschul *et al.*, 1990) at NCBI was used for homology comparisons of the 6 possible amino acid sequences encoded by the nucleotide sequences derived from DNA fragments in the process of this invention with protein sequences available.

Amino acid sequences obtained from the N terminal of SFHYA1 and In Gel tryptic digest peptide fragments were used for homology comparisons with peptide sequences available at both NCBI and GenomeNet (Supercomputer Laboratory, Institute for Chemical Research, Kyoto University and the Human Genome Center, Institute of Medical Science, the University of Tokyo) through the use of program BLASTP (Altschul *et al.*, 1990).

*Screening of a venom gland cDNA library*

A cDNA library constructed in gt10 (Ghadessy *et al.*, 1996) was screened for cDNA clones encoding SFHYA1. An approximately 800-bp fragment amplified from pSFHya1 (which has an insert DNA highly homologous to hyaluronidases) with T3 primer (5'-AATTAACCCTCACTAAAGGG-3') and T7 primer (5'-GTAATACGACTCACTATAGGC-3') was radioactively labelled with a $^{32}$P dCTP by using the Megaprime™ kit (Amersham International, U.K) and used as the probe.

DNA hybridisations were carried out overnight as described by (Sambrook *et al.*, 1989), at 37'C using 6 *x* SSC, 50 % formamide, 5 *x* Denhardts, 0.5 % SDS and 100 µg/mL salmon sperm DNA. After hybridisation, the filters were washed for 10 minutes each time with 1 *x* SSC containing 0.5 % SDS at 42°C followed by 0.5 *x* SSC also containing 0.5 % SDS at 42°C. Autoradiography was carried out at 70°C.

*Restriction mapping of cDNAs of sfhya1 gene*

Insert cDNAs amplified from selected cDNA clones by PCRs were electrophoresed on an agarose gel and purified. These cDNAs were then restricted with *Eco*RI, *Pst*I, *Hin*dIII, *Sal*I, *Sph*I and *Sac*I (all enzymes from Promega) according to supplier's instructions. Restricted fragments were then electrophoresed on agarose gels and their sizes analysed for generation of restriction maps for the cDNAs from each clone.

*RESULTS*

The amino acid sequence of the N terminal of SFHYA1 was established as L P W T D P P L H P while that of the 3 In gel tryptic fragments of SFHYA1 were determined to be K H W D D D V FL H L D P R, G G G P L T V T G D L S Q and F G V F P Y V N E K. Of these 4 short sequences, only K H W D D D V F L H L D P R was found to be matching (with 46 % sequence identity) to a region corresponding to amino acid residue 390 402 of human hyaluronidase precursor (GI: 291589) when a homology search was done for all these 4 sequences against all hyaluronidases sequences entries at both NCBI and GenomeNet using the BLASTP programme. Alignment of hyaluronidase amino acid sequences from different species and of different tissue origins (see Figure 1) revealed 3 highly conserved regions suitable for design of oligonucleotides for amplification of hyaluronidase gene fragments. The consensus sequences (as derived from Figure 1) were found to be G G L/V/I P Q A/G V I D W/F/I E and W G F/Y Y G/L/A F/Y P, corresponding respectively to amino acid residue 98 102, 130 135 and 203 209 with respect to mouse hyaluronidase (GI 2370295).

A PCR utilising primers designed based on the amino sequence A/G V I D W/F/I E (as derived from alignment of hyaluronidase shown in Figure 1) as the forward primer and the amino sequence KH W D D D V (as derived from sequencing a tryptic fragment of SFHYA1) as the reverse primer, and as template, cDNAs reversed transcribed from stonefish venom glands total RNA, yielded the expected PCR product of approximately 800-bp among other non specific amplification products. This 800-bp product was excised from agarose gel, purified from the gel and used as the template for PCR amplification under the same conditions. The amplified products were again electrophoresed on an

agarose gel and after purification from the gel were cloned into into pCR Script™ AMP SK (+), a resulting recombinant plasmid was named pSFHYA1.

Nucleotide sequence determination and analyses carried out on pSFHYA1 showed that the insert DNA encodes a peptide highly homologous to other known hyaluronidases. The insert of this plasmid was thus amplified by PCR and radioactively labelled to serve as a probe for picking up cDNA clones encoding SFHYA1 from a cDNA library. Six clones were identified, after 3 rounds of hybridisations, from $3 \times 10^5$ plaques of the cDNA library. Inserts from these clones were amplified by PCR using primers annealing to the edge of the cloning sites of gt10. The sizes of the inserts ranged approximately from 1 kb to 2 kb. Restriction mapping of these insert DNAs coupled with partial nucleotide sequencing on these insert DNAs indicated that one of the clones, sfhya1$_{c\text{-}4}$ (with an insert of approximately 2 kb), contains all the regions found in the other cDNA clones and the entire ORF of *sfhya1* gene.

*DISCUSSION*

A total of 1864 nucleotides have been sequenced from sfhya1$_{c\text{-}4}$ (Figure 2), a cDNA clone of *sfhya1* gene. The deduced amino acid sequences of this nucleotide sequence in all 6 reading frames were compared with amino acid sequences in the database at NCBI by using the program BLASTX. Strong sequence homology with hyaluronidases, particularly hyaluronidases found on the surface of mammalian spermatozoan, was evident in one of the frames. This suggested that the protein encoded by *sfhya1* is a hyaluronidase and is more closely related to the mammalian sperm surface hyaluronidases as compared to other type of hyaluronidases. An alignment of SFHYA1 with some of the highly matching entries at the databases of NCBI is presented in Figure 3.

The ATG codon at position 239 of Figure 2 is most likely to be the translational initiation codon as this is the first ATG codon encountered from the 5' end of this cDNA and there is an A residue at the 3 position (nucleotide 236) upstream of this ATG codon, consistent with Kozak's criteria of translational initiation codon (Kozak, 1999) This is further supported by the fact that an initiation codon at this position predicts a signal peptide with characteristics similar to signal peptides of other secretory proteins (see discussion below).

The translational termination codon of this cDNA was assigned to the TAG codon at position 1670 of Figure 2. This assignment predicts an ORF of 1431 bp with an encoding capacity of 477 amino acid residues. This is comparable to the number of amino acid residues encoded by the cDNAs of known hyaluronidases in other species such as mouse (473 aa, GI 2370295), *Xenopus* (468 aa, GI 4704808) and *Macaca fasicularis* (510 aa, GI 585674). Further downstream of this TAG codon, a putative polyadenylation signal, ATTAAA, is observed at position 1766 1801.

The *sfhya1* gene encodes a precursor of SFHYA1 of 477 amino acid residues, which contains a short prepeptide region, similar to other genes encoding secretory proteins. The position of N terminal sequence of mature SFHYA1 in the predicted amino acid sequence of *sfhya1* (Figure 2) suggested a 28 amino acid residue prepeptide. This peptide have all the characteristics of a 'consensus signal peptide' observed in other secretory proteins (von Heijne, 1985) (such as having a N terminal basic region ($M^1$ to $K^5$), a hydrophobic region ($F^6$–$F^{13}$) and a polar C terminal $S^{14}$-$A^{28}$). Its length of 28 aa residues is also typical of a most eukaryotic signal peptides which have 13 36 amino acid residues. Most importantly, the structural determinants for cleavage of the signal peptide from the mature protein at position 3 and 1 relative to the cleavage site is also observed in the corresponding position of SFHYA1 precursor at $G^{25}$ and $A^{27}$ [see (von Heijne, 1986) for a discussion of the 3 and 1 rule]. This 28 amino acid residue prepeptide is therefore believed to act as a signal peptide that serves to initiate the transport of preSFHYA1 into the endoplasmic reticulum (for glycosylation) and is cleaved off during the process.

Besides the N terminal sequence, the sequences of all other three tryptic fragments were also observed from the predicted amino acid sequence (see Figure 2). Residues essential for substrate binding and catalysis in hyaluronidases had been established by workers in this field prior to this invention on the basis of *in vitro* mutagenesis studies (Arming *et al.*, 1997). These residues with respect to PH 20 hyaluronidase from human sperm are Asp111, Glu 13, Arg176, Glu249 and Arg252. Based on the findings of Arming *et al.*, 1997, we believe Asp139, Glu141, Arg205, Glu280 and Arg283 of SFHYA1 are amino acid residues critical for the functioning this enzyme.

On the basis of the evidence presented so far, we believe we have in hand the cDNA encoding SFHYA1. It is a great advantage, therefore, that the cDNA encoding the venom's gland hyaluronidase from *Synanceja horrida* has been cloned and sequenced, since recombinant expression of the stonefish hyaluronidase should provide an ample source of protein for fertilisation, therapeutic and prophylactic uses herein described.

All references patents and patent applications referred to above are incorporated herein by reference.

Throughout the specification the aim has been to describe the preferred embodiments of the invention without limiting the invention to any one embodiment or specific collection of features. Those of skill in the art will therefore appreciate that, in light of the instant disclosure, various modifications and changes can be made in the particular embodiments exemplified without departing from the scope of the present invention. All such modifications and changes are intended to be included within the scope of the appendant claims.

## REFERENCE

Altschul, SF, Gish W, Miller, W, Myers, E.W. & Lipman, D.J. (1990). Basic local alignment search tool. *Journal of Molecular Biology* 215, 403 4 10.

Arming, S., Strobl, B Wechselberger, C. & Kreil, G. (1997). In vitro mutagenesis of PH 20 hyaluronidase from human sperm. *European Journal of Biochemistry* 247, 810 814.

Au Young, J., Hawkins, P. R.. and Hillman J.L. Human hyaluronidase. (5,854,046). 1998 U, S. A. (GENERIC) Ref Type: Patent

Austin, L., Gillis, R.G. & Youatt, G. (1965). Stonefish venom: some biochemical and chemical observations. *Australian Journal of Experimental Biology and Medical Science* 43, 79 90.

Beekenleliner, K., Bannke, S, Spruss, T, Bernhardt G, Schonenberg, H. & Schiess, W. (1992). Hyaluronidase enhances the activity of adriamycin in breast cancer models in vitro and in vivo. *Journal of Cancer Research & Clinical Oncology 118,* 591 596.

Bertelli, G., Dim, D., Forno, G.B., Gozza, A, Silvestro, S., Venturini, M., Rosso, R. & Pronzato, P. (1994). Hyaluronidase as an antidote to extravasation of Vinca alkaloids: clinical results. *Journal of Cancer Research & Clinical Oncology* 120, 505 506.

Ghadessy, F.J., Chen, D., Kini, R.M., Chung, M.C.M., Jeyaseelan, K, Khoo, H.E. & Yuen, R. (1996). Stonustoxin is a novel lethal factor from stonefish (Synanceja horrida) venom. cDNA cloning and characterization. *Journal of Biological Chemistry* 271, 25575 25581.

Gottlieb, J.L., Antoszyk, AX, Hatchell, D.L. & Saloupis, P. (1990). The safety of intravitreal hyaluronidase. A clinical and histologic study. *Investigative Ophthalmology & Visual Science* 31, 2345 2352.

Hobarth K, Maier, U. & Marberger, M. (1992). Topical chemoprophylaxis of superficial bladder cancer with mitomycin C and adjuvant hyaluronidase. *European Urology* 21, 206 2 10.

Kozak M. (1999). Initiation of translation in prokaryotes and eukaryotes. Gene *8,* 234(2):187 208.

King, T. P. Nucleic acid and recombinant production of vespid venom hyaluronidase. (5,593,877). 1997. (GENERIC) Ref Type: Patent

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680 685.

Lin, Y., Mahan, K., Lathrop, W.F., Myles, D.G. & Primakoff, P. (1994). A hyaluronidase activity of the sperm plasma membrane protein PH-20 enables sperm to penetrate the cumulus cell layer surrounding the egg. *Journal of Cell Biology* **125**, 1157-1163.

Lipschitz, S., Campbell, A.J., Roberts, M.S., Wanwimolruk, S., McQueen, E.G., McQueen, M. & Firth, L.A. (1991). Subcutaneous fluid administration in elderly subjects: validation of an under-used technique [see comments]. *Journal of the American Geriatrics Society* **39**, 6-9.

Ng, S.C., Bongso, A., Ratnam, S.S., Sathananthan, H., Chan, C.L., Wong, P.C., Hagglund, Anandakumar, C., Wong, Y.C. & Goh, V.H. (1988). Pregnancy after transfer of sperm under zona [letter]. *Lancet* **2**, 790

Poh, C.H., Yuen, R., Chung, M.C. & Khoo, H.E. (1992). Purification and partial characterization of hyaluronidase from stonefish (Synanceja horrida) venom. *Comparative Biochemistry & Physiology - B: Comparative Biochemistry* **101**, 159-163.

Raszka, W.V.J., Kueser, T.K., Smith, F.R. & Bass, J.W. (1990). The use of hyaluronidase in the treatment of intravenous extravasation injuries. *Journal of Perinatology* **10**, 146-149.

Sambrooke, J., Fritsch, E.F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual, 2nd Edition.* Cold Spring Harbor Laboratory Cold Spring Harbor, NY,

Spruss, T., Bernhardt, G., Schonenberger, H. & Schiess, W. (1995). Hyaluronidase significantly enhances the efficacy of regional vinblastine chemotherapy of malignant melanoma. *Journal of Cancer Research & Clinical Oncology* **121**, 193-202.

Stern, R., Frost, G. I., Hall, J., Shuster, S., Colbern, G. T., and Fromby, B. BH55 hyaluronidase. (5,827,721).1998. (GENERIC) Ref Type: Patent

Vasuthevan, S., Bongso, S., Khoo, H.E. & Yuen, R. (1993). Effects of stonefish (*Synanceja horrida*) venom hyaluronidase on cumulus dispersal, fertilization rates and cytogenetics of marine oocytes and embryos. *ARTA* **4**, 261-267.

von Heijne, G. (1985). Signal sequences. The limits of variation. *Journal of Molecular Biology* **184**, 99-105.

von Heijne, G. (1986). A new method for predicting signal sequence cleavage sites. *Nucleic Acids Research* **14**, 4683-4690.

# SEQUENCE LISTING

SEQUENCE LISTING

<110> National University of Singapore

<120> Novel polynucleotides and polypeptides and uses therefor

<130> Synanceja horrida hyaluronidase

<140>

<141>

<160> 9

<170> PatentIn Ver. 2.0

<210> 1

<211> 1864

<212> DNA

<213> Synanceia horrida

<220>

<221> CDS

<222> (239)..(1672)

<220>

<221> sig_peptide

<222> (239)..(322)

<220>

<221> mat_peptide

<222> (323)..(1672)

<220>

<221> 5'UTR

<222> (1)..(238)

<220>

<221> 3'UTR

<222> (1673)..(1864)

<400> 1

cacgtttagc agctagtgat ctggcttcag ttataaggcc gctcgctgag agcagcagct 60

cttatcagcc actgggagcc cagtctctgc cacattcaca accctccaat acactttaat 120

tgaatttctt tttagaggtt agccttcgtg acagcttaat aatttcacga ccctccgaga 180

gcacattttg gacagcagga agttggaccc ccagagcaga tcagattaca gtttgaaa 238

atg atc aaa ctg aag ttc ctc tat gtc caa cat ttc ttc tcc att tct 286

Met Ile Lys Leu Lys Phe Leu Tyr Val Gln His Phe Phe Ser Ile Ser

-25 -20 -15

gtc acc atg atg tcc ctg ctg cgc tct ggt ggg gcc tta cca tgg acc 334

Val Thr Met Met Ser Leu Leu Arg Ser Gly Gly Ala Leu Pro Trp Thr

-10 -5 -1 1

gac cct cct ctc cat ccc ggc cac ccc ttt ctc ttc act tgg aat gcc 382

Asp Pro Pro Leu His Pro Gly His Pro Phe Leu Phe Thr Trp Asn Ala

5 10 15 20

cca act gag ctc tgt ggc atc cgc ttc ggc atg cca ctt gac ctc tcc 430

Pro Thr Glu Leu Cys Gly Ile Arg Phe Gly Met Pro Leu Asp Leu Ser

25 30 35

tac ttt gac ttt gta agc agc aca ctg aag tcg gcg acc aac cag agc 478

Tyr Phe Asp Phe Val Ser Ser Thr Leu Lys Ser Ala Thr Asn Gln Ser

40 45 50

atc tcc ata ttc tac acc gat cgc ttc ggc gtg ttc cct tat gtg aac 526

Ile Ser Ile Phe Tyr Thr Asp Arg Phe Gly Val Phe Pro Tyr Val Asn

55 60 65

gaa aaa act ggc aag atg tac aac ggg ggc ctg cca cag ctg ata gac 574

```
Glu Lys Thr Gly Lys Met Tyr Asn Gly Gly Leu Pro Gln Leu Ile Asp
     70                  75                  80
ctg gag cag cac cac gag ctc gct gag gat gac atc gac tat tac atc   622
Leu Glu Gln His His Glu Leu Ala Glu Asp Asp Ile Asp Tyr Tyr Ile
 85                  90                  95                 100
cct ttc aac cag ctg ggc ctt gct gtg ctt gac ttt gag gag tgg agg   670
Pro Phe Asn Gln Leu Gly Leu Ala Val Leu Asp Phe Glu Glu Trp Arg
                105                 110                 115
cca cag tgg atc cga aac tgg ggc agc aaa gac atc tac agg cag tat   718
Pro Gln Trp Ile Arg Asn Trp Gly Ser Lys Asp Ile Tyr Arg Gln Tyr
            120                 125                 130
tcc atc gaa aca gtc cta aag aaa aat tca tca ttg tcc aag gag gag   766
Ser Ile Glu Thr Val Leu Lys Lys Asn Ser Ser Leu Ser Lys Glu Glu
        135                 140                 145
gca gct ggt cag gca aaa atg gct ttt gaa cgt gca gcc aag aag tac   814
Ala Ala Gly Gln Ala Lys Met Ala Phe Glu Arg Ala Ala Lys Lys Tyr
    150                 155                 160
ttt ctc cga tcc atc cgc atc ggg aag cgg ctg agg ccc aac aga ctc   862
Phe Leu Arg Ser Ile Arg Ile Gly Lys Arg Leu Arg Pro Asn Arg Leu
165                 170                 175                 180
tgg ggc tac tac ctg tac cct gag tgc tac aac tat gag tac aaa aag   910
Trp Gly Tyr Tyr Leu Tyr Pro Glu Cys Tyr Asn Tyr Glu Tyr Lys Lys
                185                 190                 195
gac atg gca ggc tac acg ggg gag tgt ccc gcc atc gag aag gac agg   958
Asp Met Ala Gly Tyr Thr Gly Glu Cys Pro Ala Ile Glu Lys Asp Arg
            200                 205                 210
aac aac gaa ctg ctg tgg ctc tgg aga gag tct act gct ctc ttt ccg  1006
```

```
Asn Asn Glu Leu Leu Trp Leu Trp Arg Glu Ser Thr Ala Leu Phe Pro
        215                 220                 225

tcc atc tac ctg gag ctg tta ctc aga gac acc cag cag gcc cgg cag  1054
Ser Ile Tyr Leu Glu Leu Leu Leu Arg Asp Thr Gln Gln Ala Arg Gln
230                 235                 240

tac gtc cga cat cgc atc caa gag tcc ata agg gtg tcg aag ctt ccc  1102
Tyr Val Arg His Arg Ile Gln Glu Ser Ile Arg Val Ser Lys Leu Pro
245                 250                 255                 260

aac agc gcc tac tct atc ccc atc cat gcc tac gtc cga cct gtg tac  1150
Asn Ser Ala Tyr Ser Ile Pro Ile His Ala Tyr Val Arg Pro Val Tyr
                265                 270                 275

aag gac agc act gac aac tac atg tca gag ttt gat ctg gtg aac acc  1198
Lys Asp Ser Thr Asp Asn Tyr Met Ser Glu Phe Asp Leu Val Asn Thr
            280                 285                 290

att gga gaa gct gct gct ctt ggc gcc gcc agc gtt gtt tgt tgg gga  1246
Ile Gly Glu Ala Ala Ala Leu Gly Ala Ala Ser Val Val Cys Trp Gly
        295                 300                 305

gac atg agc gtc gtc gct aca gag gac tcc tgc ttt gat gcc cga cga  1294
Asp Met Ser Val Val Ala Thr Glu Asp Ser Cys Phe Asp Ala Arg Arg
    310                 315                 320

cac ctg gag aag gtc atg aac cca tac atc atg aac gtc tcc aca gca  1342
His Leu Glu Lys Val Met Asn Pro Tyr Ile Met Asn Val Ser Thr Ala
325                 330                 335                 340

aca cag ctc tgc agc aag gcg ctc tgc cag gac caa ggt cgc tgt gtg  1390
Thr Gln Leu Cys Ser Lys Ala Leu Cys Gln Asp Gln Gly Arg Cys Val
                345                 350                 355

agg aag cac tgg gac gat gac gtt ttc ctc cac ctc gac ccg cgg cgt  1438
Arg Lys His Trp Asp Asp Asp Val Phe Leu His Leu Asp Pro Arg Arg
            360                 365                 370
```

tac cgg atc gag cag cag cgc ggc ggc ggc cca ctg act gtg act ggt 1486
Tyr Arg Ile Glu Gln Gln Arg Gly Gly Gly Pro Leu Thr Val Thr Gly
375 380 385

gac ctc tcg cag gat gat gtc aac tgg ttt gac cgc aac ttt gat tgc 1534
Asp Leu Ser Gln Asp Asp Val Asn Trp Phe Asp Arg Asn Phe Asp Cys
390 395 400

atg tgc tac agc gag aag ccg tgt cga tcg gcc ctg acg ttc aac gtc 1582
Met Cys Tyr Ser Glu Lys Pro Cys Arg Ser Ala Leu Thr Phe Asn Val
405 410 415 420

atc aac aag gcc gtc atc aac aag gcc cct cct gct tgt gat ggt act 1630
Ile Asn Lys Ala Val Ile Asn Lys Ala Pro Pro Ala Cys Asp Gly Thr
425 430 435

gct cgg tct gat gta cat tgt gat gtg aac ata aac ttt tag 1672
Ala Arg Ser Asp Val His Cys Asp Val Asn Ile Asn Phe
440 445 450

cacaaacctg atatctatga tgcacaatat gtgaggcagt taaaggacaa ttccactgtt 1732

tttccagcct gtttgtctgt ttgttatttc tgactgtttt tactgtttca tttctgagtt 1792

ttcattaaag ataattttca aatcactttt gatatcactt aaacattaaa acaaacaaat 1852

tttcattaaa aa 1864

<210> 2

<211> 477

<212> PRT

<213> Synanceia horrida

<400> 2

Met Ile Lys Leu Lys Phe Leu Tyr Val Gln His Phe Phe Ser Ile Ser
1 5 10 15

Val Thr Met Met Ser Leu Leu Arg Ser Gly Gly Ala Leu Pro Trp Thr
20 25 30
Asp Pro Pro Leu His Pro Gly His Pro Phe Leu Phe Thr Trp Asn Ala
35 40 45
Pro Thr Glu Leu Cys Gly Ile Arg Phe Gly Met Pro Leu Asp Leu Ser
50 55 60
Tyr Phe Asp Phe Val Ser Ser Thr Leu Lys Ser Ala Thr Asn Gln Ser
65 70 75 80
Ile Ser Ile Phe Tyr Thr Asp Arg Phe Gly Val Phe Pro Tyr Val Asn
85 90 95
Glu Lys Thr Gly Lys Met Tyr Asn Gly Gly Leu Pro Gln Leu Ile Asp
100 105 110
Leu Glu Gln His His Glu Leu Ala Glu Asp Asp Ile Asp Tyr Tyr Ile
115 120 125
Pro Phe Asn Gln Leu Gly Leu Ala Val Leu Asp Phe Glu Glu Trp Arg
130 135 140
Pro Gln Trp Ile Arg Asn Trp Gly Ser Lys Asp Ile Tyr Arg Gln Tyr
145 150 155 160
Ser Ile Glu Thr Val Leu Lys Lys Asn Ser Ser Leu Ser Lys Glu Glu
165 170 175
Ala Ala Gly Gln Ala Lys Met Ala Phe Glu Arg Ala Ala Lys Lys Tyr
180 185 190
Phe Leu Arg Ser Ile Arg Ile Gly Lys Arg Leu Arg Pro Asn Arg Leu
195 200 205
Trp Gly Tyr Tyr Leu Tyr Pro Glu Cys Tyr Asn Tyr Glu Tyr Lys Lys
210 215 220

Asp Met Ala Gly Tyr Thr Gly Glu Cys Pro Ala Ile Glu Lys Asp Arg
225 230 235 240
Asn Asn Glu Leu Leu Trp Leu Trp Arg Glu Ser Thr Ala Leu Phe Pro
245 250 255
Ser Ile Tyr Leu Glu Leu Leu Leu Arg Asp Thr Gln Gln Ala Arg Gln
260 265 270
Tyr Val Arg His Arg Ile Gln Glu Ser Ile Arg Val Ser Lys Leu Pro
275 280 285
Asn Ser Ala Tyr Ser Ile Pro Ile His Ala Tyr Val Arg Pro Val Tyr
290 295 300
Lys Asp Ser Thr Asp Asn Tyr Met Ser Glu Phe Asp Leu Val Asn Thr
305 310 315 320
Ile Gly Glu Ala Ala Ala Leu Gly Ala Ala Ser Val Val Cys Trp Gly
325 330 335
Asp Met Ser Val Val Ala Thr Glu Asp Ser Cys Phe Asp Ala Arg Arg
340 345 350
His Leu Glu Lys Val Met Asn Pro Tyr Ile Met Asn Val Ser Thr Ala
355 360 365
Thr Gln Leu Cys Ser Lys Ala Leu Cys Gln Asp Gln Gly Arg Cys Val
370 375 380
Arg Lys His Trp Asp Asp Asp Val Phe Leu His Leu Asp Pro Arg Arg
385 390 395 400
Tyr Arg Ile Glu Gln Gln Arg Gly Gly Gly Pro Leu Thr Val Thr Gly
405 410 415
Asp Leu Ser Gln Asp Asp Val Asn Trp Phe Asp Arg Asn Phe Asp Cys
420 425 430
Met Cys Tyr Ser Glu Lys Pro Cys Arg Ser Ala Leu Thr Phe Asn Val
435 440 445

Ile Asn Lys Ala Val Ile Asn Lys Ala Pro Pro Ala Cys Asp Gly Thr
450 455 460

Ala Arg Ser Asp Val His Cys Asp Val Asn Ile Asn Phe
465 470 475

<210> 3

<211> 1434

<212> DNA

<213> Synanceia horrida

<220>

<221> CDS

<222> (1)..(1434)

<220>

<221> sig_peptide

<222> (1)..(84)

<220>

<221> mat_peptide

<222> (85)..(1434)

<400> 3

atg atc aaa ctg aag ttc ctc tat gtc caa cat ttc ttc tcc att tct 48
Met Ile Lys Leu Lys Phe Leu Tyr Val Gln His Phe Phe Ser Ile Ser
-25 -20 -15

gtc acc atg atg tcc ctg ctg cgc tct ggt ggg gcc tta cca tgg acc 96
Val Thr Met Met Ser Leu Leu Arg Ser Gly Gly Ala Leu Pro Trp Thr
-10 -5 -1 1

gac cct cct ctc cat ccc ggc cac ccc ttt ctc ttc act tgg aat gcc 144

```
Asp Pro Pro Leu His Pro Gly His Pro Phe Leu Phe Thr Trp Asn Ala
 5                  10                  15                  20
cca act gag ctc tgt ggc atc cgc ttc ggc atg cca ctt gac ctc tcc   192
Pro Thr Glu Leu Cys Gly Ile Arg Phe Gly Met Pro Leu Asp Leu Ser
5                   25                  30                  35
tac ttt gac ttt gta agc agc aca ctg aag tcg gcg acc aac cag agc   240
Tyr Phe Asp Phe Val Ser Ser Thr Leu Lys Ser Ala Thr Asn Gln Ser
            40                  45                  50
atc tcc ata ttc tac acc gat cgc ttc ggc gtg ttc cct tat gtg aac   288
10 Ile Ser Ile Phe Tyr Thr Asp Arg Phe Gly Val Phe Pro Tyr Val Asn
        55                  60                  65
gaa aaa act ggc aag atg tac aac ggg ggc ctg cca cag ctg ata gac   336
Glu Lys Thr Gly Lys Met Tyr Asn Gly Gly Leu Pro Gln Leu Ile Asp
    70                  75                  80
15 ctg gag cag cac cac gag ctc gct gag gat gac atc gac tat tac atc   384
Leu Glu Gln His His Glu Leu Ala Glu Asp Asp Ile Asp Tyr Tyr Ile
 85                  90                  95                 100
cct ttc aac cag ctg ggc ctt gct gtg ctt gac ttt gag gag tgg agg   432
Pro Phe Asn Gln Leu Gly Leu Ala Val Leu Asp Phe Glu Glu Trp Arg
20                 105                 110                 115
cca cag tgg atc cga aac tgg ggc agc aaa gac atc tac agg cag tat   480
Pro Gln Trp Ile Arg Asn Trp Gly Ser Lys Asp Ile Tyr Arg Gln Tyr
            120                 125                 130
tcc atc gaa aca gtc cta aag aaa aat tca tca ttg tcc aag gag gag   528
25 Ser Ile Glu Thr Val Leu Lys Lys Asn Ser Ser Leu Ser Lys Glu Glu
        135                 140                 145
gca gct ggt cag gca aaa atg gct ttt gaa cgt gca gcc aag aag tac   576
```

```
Ala Ala Gly Gln Ala Lys Met Ala Phe Glu Arg Ala Ala Lys Lys Tyr
    150                 155                 160
ttt ctc cga tcc atc cgc atc ggg aag cgg ctg agg ccc aac aga ctc   624
Phe Leu Arg Ser Ile Arg Ile Gly Lys Arg Leu Arg Pro Asn Arg Leu
165                 170                 175                 180
tgg ggc tac tac ctg tac cct gag tgc tac aac tat gag tac aaa aag   672

Trp Gly Tyr Tyr Leu Tyr Pro Glu Cys Tyr Asn Tyr Glu Tyr Lys Lys
                185                 190                 195
gac atg gca ggc tac acg ggg gag tgt ccc gcc atc gag aag gac agg   720
Asp Met Ala Gly Tyr Thr Gly Glu Cys Pro Ala Ile Glu Lys Asp Arg
            200                 205                 210
aac aac gaa ctg ctg tgg ctc tgg aga gag tct act gct ctc ttt ccg   768
Asn Asn Glu Leu Leu Trp Leu Trp Arg Glu Ser Thr Ala Leu Phe Pro
        215                 220                 225
tcc atc tac ctg gag ctg tta ctc aga gac acc cag cag gcc cgg cag   816
Ser Ile Tyr Leu Glu Leu Leu Leu Arg Asp Thr Gln Gln Ala Arg Gln
    230                 235                 240
tac gtc cga cat cgc atc caa gag tcc ata agg gtg tcg aag ctt ccc   864
Tyr Val Arg His Arg Ile Gln Glu Ser Ile Arg Val Ser Lys Leu Pro
245                 250                 255                 260
aac agc gcc tac tct atc ccc atc cat gcc tac gtc cga cct gtg tac   912
Asn Ser Ala Tyr Ser Ile Pro Ile His Ala Tyr Val Arg Pro Val Tyr
                265                 270                 275
aag gac agc act gac aac tac atg tca gag ttt gat ctg gtg aac acc   960
Lys Asp Ser Thr Asp Asn Tyr Met Ser Glu Phe Asp Leu Val Asn Thr
            280                 285                 290
att gga gaa gct gct gct ctt ggc gcc gcc agc gtt gtt tgt tgg gga  1008
```

Ile Gly Glu Ala Ala Ala Leu Gly Ala Ala Ser Val Val Cys Trp Gly

295 300 305

gac atg agc gtc gtc gct aca gag gac tcc tgc ttt gat gcc cga cga 1056

Asp Met Ser Val Val Ala Thr Glu Asp Ser Cys Phe Asp Ala Arg Arg

310 315 320

cac ctg gag aag gtc atg aac cca tac atc atg aac gtc tcc aca gca 1104

His Leu Glu Lys Val Met Asn Pro Tyr Ile Met Asn Val Ser Thr Ala

325 330 335 340

aca cag ctc tgc agc aag gcg ctc tgc cag gac caa ggt cgc tgt gtg 1152

Thr Gln Leu Cys Ser Lys Ala Leu Cys Gln Asp Gln Gly Arg Cys Val

345 350 355

agg aag cac tgg gac gat gac gtt ttc ctc cac ctc gac ccg cgg cgt 1200

Arg Lys His Trp Asp Asp Asp Val Phe Leu His Leu Asp Pro Arg Arg

360 365 370

tac cgg atc gag cag cag cgc ggc ggc ggc cca ctg act gtg act ggt 1248

Tyr Arg Ile Glu Gln Gln Arg Gly Gly Gly Pro Leu Thr Val Thr Gly

375 380 385

gac ctc tcg cag gat gat gtc aac tgg ttt gac cgc aac ttt gat tgc 1296

Asp Leu Ser Gln Asp Asp Val Asn Trp Phe Asp Arg Asn Phe Asp Cys

390 395 400

atg tgc tac agc gag aag ccg tgt cga tcg gcc ctg acg ttc aac gtc 1344

Met Cys Tyr Ser Glu Lys Pro Cys Arg Ser Ala Leu Thr Phe Asn Val

405 410 415 420

atc aac aag gcc gtc atc aac aag gcc cct cct gct tgt gat ggt act 1392

Ile Asn Lys Ala Val Ile Asn Lys Ala Pro Pro Ala Cys Asp Gly Thr

425 430 435

gct cgg tct gat gta cat tgt gat gtg aac ata aac ttt tag 1434

Ala Arg Ser Asp Val His Cys Asp Val Asn Ile Asn Phe
440 445 450

<210> 4
<211> 477
<212> PRT
<213> Synanceia horrida
<400> 4

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

Ser Ile Glu Thr Val Leu Lys Lys Asn Ser Ser Leu Ser Lys Glu Glu
165 170 175

Ala Ala Gly Gln Ala Lys Met Ala Phe Glu Arg Ala Ala Lys Lys Tyr
180 185 190

Phe Leu Arg Ser Ile Arg Ile Gly Lys Arg Leu Arg Pro Asn Arg Leu
195 200 205

Trp Gly Tyr Tyr Leu Tyr Pro Glu Cys Tyr Asn Tyr Glu Tyr Lys Lys
210 215 220

Asp Met Ala Gly Tyr Thr Gly Glu Cys Pro Ala Ile Glu Lys Asp Arg
225 230 235 240

Asn Asn Glu Leu Leu Trp Leu Trp Arg Glu Ser Thr Ala Leu Phe Pro
245 250 255

Ser Ile Tyr Leu Glu Leu Leu Leu Arg Asp Thr Gln Gln Ala Arg Gln
260 265 270

Tyr Val Arg His Arg Ile Gln Glu Ser Ile Arg Val Ser Lys Leu Pro
275 280 285

Asn Ser Ala Tyr Ser Ile Pro Ile His Ala Tyr Val Arg Pro Val Tyr
290 295 300

Lys Asp Ser Thr Asp Asn Tyr Met Ser Glu Phe Asp Leu Val Asn Thr
305 310 315 320

Ile Gly Glu Ala Ala Ala Leu Gly Ala Ala Ser Val Val Cys Trp Gly
325 330 335

Asp Met Ser Val Val Ala Thr Glu Asp Ser Cys Phe Asp Ala Arg Arg
340 345 350

His Leu Glu Lys Val Met Asn Pro Tyr Ile Met Asn Val Ser Thr Ala
355 360 365

Thr Gln Leu Cys Ser Lys Ala Leu Cys Gln Asp Gln Gly Arg Cys Val
370 375 380
Arg Lys His Trp Asp Asp Asp Val Phe Leu His Leu Asp Pro Arg Arg
385 390 395 400
Tyr Arg Ile Glu Gln Gln Arg Gly Gly Gly Pro Leu Thr Val Thr Gly
405 410 415
Asp Leu Ser Gln Asp Asp Val Asn Trp Phe Asp Arg Asn Phe Asp Cys
420 425 430
Met Cys Tyr Ser Glu Lys Pro Cys Arg Ser Ala Leu Thr Phe Asn Val
435 440 445
Ile Asn Lys Ala Val Ile Asn Lys Ala Pro Pro Ala Cys Asp Gly Thr
450 455 460
Ala Arg Ser Asp Val His Cys Asp Val Asn Ile Asn Phe
465 470 475

<210> 5
<211> 1350
<212> DNA
<213> Synanceia horrida
<220>
<221> CDS
<222> (1)..(1350)
<220>
<221> mat_peptide
<222> (1)..(1350)
<400> 5
tta cca tgg acc gac cct cct ctc cat ccc ggc cac ccc ttt ctc ttc 48

Leu Pro Trp Thr Asp Pro Pro Leu His Pro Gly His Pro Phe Leu Phe

1 5 10 15

act tgg aat gcc cca act gag ctc tgt ggc atc cgc ttc ggc atg cca 96

Thr Trp Asn Ala Pro Thr Glu Leu Cys Gly Ile Arg Phe Gly Met Pro

20 25 30

ctt gac ctc tcc tac ttt gac ttt gta agc agc aca ctg aag tcg gcg 144

Leu Asp Leu Ser Tyr Phe Asp Phe Val Ser Ser Thr Leu Lys Ser Ala

35 40 45

acc aac cag agc atc tcc ata ttc tac acc gat cgc ttc ggc gtg ttc 192

Thr Asn Gln Ser Ile Ser Ile Phe Tyr Thr Asp Arg Phe Gly Val Phe

50 55 60

cct tat gtg aac gaa aaa act ggc aag atg tac aac ggg ggc ctg cca 240

Pro Tyr Val Asn Glu Lys Thr Gly Lys Met Tyr Asn Gly Gly Leu Pro

65 70 75 80

cag ctg ata gac ctg gag cag cac cac gag ctc gct gag gat gac atc 288

Gln Leu Ile Asp Leu Glu Gln His His Glu Leu Ala Glu Asp Asp Ile

85 90 95

gac tat tac atc cct ttc aac cag ctg ggc ctt gct gtg ctt gac ttt 336

Asp Tyr Tyr Ile Pro Phe Asn Gln Leu Gly Leu Ala Val Leu Asp Phe

100 105 110

gag gag tgg agg cca cag tgg atc cga aac tgg ggc agc aaa gac atc 384

Glu Glu Trp Arg Pro Gln Trp Ile Arg Asn Trp Gly Ser Lys Asp Ile

115 120 125

tac agg cag tat tcc atc gaa aca gtc cta aag aaa aat tca tca ttg 432

Tyr Arg Gln Tyr Ser Ile Glu Thr Val Leu Lys Lys Asn Ser Ser Leu

130 135 140

tcc aag gag gag gca gct ggt cag gca aaa atg gct ttt gaa cgt gca 480

Ser Lys Glu Glu Ala Ala Gly Gln Ala Lys Met Ala Phe Glu Arg Ala

145 150 155 160

gcc aag aag tac ttt ctc cga tcc atc cgc atc ggg aag cgg ctg agg 528

Ala Lys Lys Tyr Phe Leu Arg Ser Ile Arg Ile Gly Lys Arg Leu Arg

165 170 175

ccc aac aga ctc tgg ggc tac tac ctg tac cct gag tgc tac aac tat 576

Pro Asn Arg Leu Trp Gly Tyr Tyr Leu Tyr Pro Glu Cys Tyr Asn Tyr

180 185 190

gag tac aaa aag gac atg gca ggc tac acg ggg gag tgt ccc gcc atc 624

Glu Tyr Lys Lys Asp Met Ala Gly Tyr Thr Gly Glu Cys Pro Ala Ile

195 200 205

gag aag gac agg aac aac gaa ctg ctg tgg ctc tgg aga gag tct act 672

Glu Lys Asp Arg Asn Asn Glu Leu Leu Trp Leu Trp Arg Glu Ser Thr

210 215 220

gct ctc ttt ccg tcc atc tac ctg gag ctg tta ctc aga gac acc cag 720

Ala Leu Phe Pro Ser Ile Tyr Leu Glu Leu Leu Leu Arg Asp Thr Gln

225 230 235 240

cag gcc cgg cag tac gtc cga cat cgc atc caa gag tcc ata agg gtg 768

Gln Ala Arg Gln Tyr Val Arg His Arg Ile Gln Glu Ser Ile Arg Val

245 250 255

tcg aag ctt ccc aac agc gcc tac tct atc ccc atc cat gcc tac gtc 816

Ser Lys Leu Pro Asn Ser Ala Tyr Ser Ile Pro Ile His Ala Tyr Val

260 265 270

cga cct gtg tac aag gac agc act gac aac tac atg tca gag ttt gat 864

Arg Pro Val Tyr Lys Asp Ser Thr Asp Asn Tyr Met Ser Glu Phe Asp

275 280 285

ctg gtg aac acc att gga gaa gct gct gct ctt ggc gcc gcc agc gtt 912

```
Leu Val Asn Thr Ile Gly Glu Ala Ala Ala Leu Gly Ala Ala Ser Val
    290                 295                 300
gtt tgt tgg gga gac atg agc gtc gtc gct aca gag gac tcc tgc ttt   960
Val Cys Trp Gly Asp Met Ser Val Val Ala Thr Glu Asp Ser Cys Phe
305                 310                 315                 320
gat gcc cga cga cac ctg gag aag gtc atg aac cca tac atc atg aac  1008
Asp Ala Arg Arg His Leu Glu Lys Val Met Asn Pro Tyr Ile Met Asn
                325                 330                 335
gtc tcc aca gca aca cag ctc tgc agc aag gcg ctc tgc cag gac caa  1056
Val Ser Thr Ala Thr Gln Leu Cys Ser Lys Ala Leu Cys Gln Asp Gln
            340                 345                 350
ggt cgc tgt gtg agg aag cac tgg gac gat gac gtt ttc ctc cac ctc  1104
Gly Arg Cys Val Arg Lys His Trp Asp Asp Asp Val Phe Leu His Leu
        355                 360                 365
gac ccg cgg cgt tac cgg atc gag cag cag cgc ggc ggc ggc cca ctg  1152
Asp Pro Arg Arg Tyr Arg Ile Glu Gln Gln Arg Gly Gly Gly Pro Leu
    370                 375                 380
act gtg act ggt gac ctc tcg cag gat gat gtc aac tgg ttt gac cgc  1200
Thr Val Thr Gly Asp Leu Ser Gln Asp Asp Val Asn Trp Phe Asp Arg
385                 390                 395                 400
aac ttt gat tgc atg tgc tac agc gag aag ccg tgt cga tcg gcc ctg  1248
Asn Phe Asp Cys Met Cys Tyr Ser Glu Lys Pro Cys Arg Ser Ala Leu
                405                 410                 415
acg ttc aac gtc atc aac aag gcc gtc atc aac aag gcc cct cct gct  1296
Thr Phe Asn Val Ile Asn Lys Ala Val Ile Asn Lys Ala Pro Pro Ala
            420                 425                 430
tgt gat ggt act gct cgg tct gat gta cat tgt gat gtg aac ata aac  1344
```

Cys Asp Gly Thr Ala Arg Ser Asp Val His Cys Asp Val Asn Ile Asn
435 440 445

ttt tag 1350
Phe
450

<210> 6
<211> 449
<212> PRT
<213> Synanceia horrida
<400> 6

Leu Pro Trp Thr Asp Pro Pro Leu His Pro Gly His Pro Phe Leu Phe
1 5 10 15

Thr Trp Asn Ala Pro Thr Glu Leu Cys Gly Ile Arg Phe Gly Met Pro
20 25 30

Leu Asp Leu Ser Tyr Phe Asp Phe Val Ser Ser Thr Leu Lys Ser Ala
35 40 45

Thr Asn Gln Ser Ile Ser Ile Phe Tyr Thr Asp Arg Phe Gly Val Phe
50 55 60

Pro Tyr Val Asn Glu Lys Thr Gly Lys Met Tyr Asn Gly Gly Leu Pro
65 70 75 80

Gln Leu Ile Asp Leu Glu Gln His His Glu Leu Ala Glu Asp Asp Ile
85 90 95

Asp Tyr Tyr Ile Pro Phe Asn Gln Leu Gly Leu Ala Val Leu Asp Phe
100 105 110

Glu Glu Trp Arg Pro Gln Trp Ile Arg Asn Trp Gly Ser Lys Asp Ile
115 120 125

```
Tyr Arg Gln Tyr Ser Ile Glu Thr Val Leu Lys Lys Asn Ser Ser Leu
    130                 135                 140
Ser Lys Glu Glu Ala Ala Gly Gln Ala Lys Met Ala Phe Glu Arg Ala
145                 150                 155                 160
Ala Lys Lys Tyr Phe Leu Arg Ser Ile Arg Ile Gly Lys Arg Leu Arg
                165                 170                 175
Pro Asn Arg Leu Trp Gly Tyr Tyr Leu Tyr Pro Glu Cys Tyr Asn Tyr
            180                 185                 190
Glu Tyr Lys Lys Asp Met Ala Gly Tyr Thr Gly Glu Cys Pro Ala Ile
        195                 200                 205
Glu Lys Asp Arg Asn Asn Glu Leu Leu Trp Leu Trp Arg Glu Ser Thr
    210                 215                 220
Ala Leu Phe Pro Ser Ile Tyr Leu Glu Leu Leu Leu Arg Asp Thr Gln
225                 230                 235                 240
Gln Ala Arg Gln Tyr Val Arg His Arg Ile Gln Glu Ser Ile Arg Val
                245                 250                 255
Ser Lys Leu Pro Asn Ser Ala Tyr Ser Ile Pro Ile His Ala Tyr Val
            260                 265                 270
Arg Pro Val Tyr Lys Asp Ser Thr Asp Asn Tyr Met Ser Glu Phe Asp
        275                 280                 285
Leu Val Asn Thr Ile Gly Glu Ala Ala Ala Leu Gly Ala Ala Ser Val
    290                 295                 300
Val Cys Trp Gly Asp Met Ser Val Val Ala Thr Glu Asp Ser Cys Phe
305                 310                 315                 320
Asp Ala Arg Arg His Leu Glu Lys Val Met Asn Pro Tyr Ile Met Asn
                325                 330                 335
Val Ser Thr Ala Thr Gln Leu Cys Ser Lys Ala Leu Cys Gln Asp Gln
            340                 345                 350
```

Gly Arg Cys Val Arg Lys His Trp Asp Asp Asp Val Phe Leu His Leu
355 360 365

Asp Pro Arg Arg Tyr Arg Ile Glu Gln Gln Arg Gly Gly Gly Pro Leu
370 375 380

Thr Val Thr Gly Asp Leu Ser Gln Asp Asp Val Asn Trp Phe Asp Arg
385 390 395 400

Asn Phe Asp Cys Met Cys Tyr Ser Glu Lys Pro Cys Arg Ser Ala Leu
405 410 415

Thr Phe Asn Val Ile Asn Lys Ala Val Ile Asn Lys Ala Pro Pro Ala
420 425 430

Cys Asp Gly Thr Ala Arg Ser Asp Val His Cys Asp Val Asn Ile Asn
435 440 445

Phe

<210> 7

<211> 84

<212> DNA

<213> Synanceia horrida

<220>

<221> CDS

<222> (1)..(84)

<220>

<221> sig_peptide

<222> (1)..(84)

<400> 7

atg atc aaa ctg aag ttc ctc tat gtc caa cat ttc ttc tcc att tct 48
Met Ile Lys Leu Lys Phe Leu Tyr Val Gln His Phe Phe Ser Ile Ser
1 5 10 15

gtc acc atg atg tcc ctg ctg cgc tct ggt ggg gcc 84
Val Thr Met Met Ser Leu Leu Arg Ser Gly Gly Ala
20 25

<210> 8
<211> 28
<212> PRT
<213> Synanceia horrida
<400> 8
Met Ile Lys Leu Lys Phe Leu Tyr Val Gln His Phe Phe Ser Ile Ser
1 5 10 15
Val Thr Met Met Ser Leu Leu Arg Ser Gly Gly Ala
20 25

<210> 9
<211> 20
<212> PRT
<213> Synanceia horrida
<400> 9
Ala Pro Ser Xaa Asp Glu Gly Asn Lys Lys Ala Asp Asn Leu Leu Val
1 5 10 15
Lys Lys Ile Asn
20

* C L M S *

## CLAIMS

1. A recombinant polypeptide encoding a hyaluronidase of an aquatic animal or a biologically active fragment thereof, or a variant or derivative of these, said recombinant polypeptide comprising the sequence set forth in SEQ ID NO: 6, with the proviso that said biologically active fragment does not consist of the sequence set forth in SEQ ID NO: 9.

2. The recombinant polypeptide of claim 1, having a leader peptide comprising the sequence set forth in SEQ ID NO: 8 or a biologically active fragment thereof, or variant or derivative of these.

3. The recombinant polypeptide of claim 2, wherein the polypeptide comprises the sequence set forth in SEQ ID NO: 2.

4. An isolated polynucleotide encoding the recombinant polypeptide, biologically active fragment, variant or derivative according to any one of claims 1 to 3.

5. The isolated polynucleotide of claim 4, wherein the polynucleotide comprises the sequence set forth in SEQ ID NO: 5, or a biologically active fragment thereof, or a polynucleotide variant of these.

6. The isolated polynucleotide of claim 5, wherein the polynucleotide comprises a nucleotide sequence encoding a leader peptide, comprising the sequence set forth in SEQ ID NO: 7 or a biologically active fragment thereof, or a polynucleotide variant of these.

7. The isolated polynucleotide of claim 6, wherein the polynucleotide comprises the sequence set forth in SEQ ID NO: 1 or 3.

8. The isolated polynucleotide of any one of claims 4 to 7, wherein the variant is obtained from a species of *Synanceja*.

9. The isolated polynucleotide of any one of claims 8, wherein the species is *S. verrucosa* or *S. horrida*.

10. An expression vector comprising the polynucleotide of any of claims 4 to 9, wherein the polynucleotide is operably linked to one or more regulatory nucleic acids.

11. A host cell containing the expression vector of claim 12.

12. A method of producing a recombinant polypeptide, comprising:

(a)culturing a host cell containing the expression vector of claim 11; and

(b)isolating the said recombinant polypeptide.

13. A method of producing a biologically active fragment of a polypeptide comprising the sequence set forth in SEQ ID NO: 2 or 6, comprising the steps of:

(a)producing a fragment of said polypeptide;

(b)contacting a hyaluronidase substrate with the fragment; and

(c)detecting hydrolysis of the hyaluronidase substrate by the fragment which is indicative of said fragment being a said biologically active fragment.

14. A method of producing a biologically active fragment of a polypeptide comprising the sequence set forth in SEQ ID NO: 2 or 6, comprising the steps of:

(a)producing a fragment of said polypeptide;

(b)contacting an oocyte-cumulus complex with the fragment; and

(c)detecting dispersal of cumulus cells from the said complex, which is indicative of the fragment being a said biologically active fragment.

15. The method of claim 14 further comprising the step of measuring the rate of dispersal of cumulus cells from the said complex.

16. The method of claim 15, wherein the step of measuring is further characterised in that when the rate of dispersal of cumulus cells from the said complex in the presence of fragment is about half that in the presence of sheep hyaluronidase, then this is indicative of the fragment being a said biologically active fragment.

17. The method of claim 15, wherein the step of measuring is further characterised in that when the rate of dispersal of murine cumulus cells from the said complex in the presence of said fragment is about 0.8 minutes per oocyte, then this is indicative of the fragment being a said biologically active fragment.

18. A method of producing a polypeptide variant of a parent polypeptide comprising the sequence set forth in SEQ ID NO: 2 or 6, or comprising a biologically active fragment of said sequence, comprising the steps of:

(a)replacing at least one amino acid of the parent polypeptide, with a different amino acid to produce a modified polypeptide;

(b)contacting a hyaluronidase substrate with the modified polypeptide; and

(c)detecting hydrolysis of the hyaluronidase substrate by the modified polypeptide, which is indicative of the modified polypeptide being a said polypeptide variant.

19. A method of producing a polypeptide variant of a parent polypeptide comprising the sequence set forth in SEQ ID NO: 2 or 6, or comprising a biologically active fragment of said sequence, comprising the steps of:

(a)replacing at least one amino acid of the parent polypeptide, with a different amino acid to produce a modified polypeptide;

(b)contacting the modified polypeptide with an antigen-binding molecule that specifically binds to the parent polypeptide; and

(c)detecting the presence of a complex comprising the said modified polypeptide, and the antigen-binding molecule in said contacted sample, which is indicative of the modified polypeptide being a said polypeptide variant.

20. A method of producing a polypeptide variant of a parent polypeptide comprising the sequence set forth in SEQ ID NO: 2 or 6, or comprising a biologically active fragment of said sequence, comprising the steps of:

(a)replacing at least one amino acid of the parent polypeptide, with a different amino acid to produce a modified polypeptide;

(b)contacting an oocyte-cumulus complex with the modified polypeptide; and

(c)detecting dispersal of cumulus cells from the said complex, which is indicative of the modified polypeptide being a said polypeptide variant.

21. The method of claim 20, further comprising the step of measuring the rate of dispersal of cumulus cells from the said complex.

22. The method of claim 21, wherein the step of measuring is further characterised in that when the rate of dispersal of cumulus cells from the said complex in the presence of said modified polypeptide is about half that in the presence of sheep hyaluronidase, then this is indicative of the modified polypeptide being a said polypeptide variant.

23. The method of claim 21, wherein the step of measuring is further characterised in that when the rate of dispersal of cumulus cells, preferably murine cumulus cells from the said complex in the presence of said modified polypeptide is about 0.8 minutes per oocyte, then this is indicative of the modified polypeptide being a said polypeptide variant.

24. A composition for use in enhancing fertilisation of a mammalian egg, comprising the recombinant polypeptide, biologically active fragment, variant or derivative according to any one of claims 1 to 3, together with a pharmaceutically acceptable carrier.

25. A method for enhancing fertilisation of a mammalian egg, comprising contacting the egg with a fertilisingly effective amount of a composition comprising the recombinant polypeptide, biologically active fragment, variant or derivative according to any one of claims 1 to 3, together with a pharmaceutically acceptable carrier.

26. Use of a composition comprising the recombinant polypeptide, biologically active fragment, variant or derivative according to any one of claims 1 to 3, together with a pharmaceutically acceptable carrier in a treatment selected from the group consisting of:

(a)increasing the effect of anaesthetics;

(b)administrating fluids and drugs;

(c)resorption of traumatic oedema and haematoma;

(d)preventing tissue necrosis; and

(e)enhancing the efficacy of chemotherapy against cancer.

27. An antigen-binding molecule that specifically binds the recombinant polypeptide, biologically active fragment, variant or derivative according to any one of claims 1 to 3.

28. A method of detecting in a sample the recombinant polypeptide, biologically active fragment, variant or derivative according to any one of claims 1 to 3, comprising:

(a)contacting the sample with an antigen-binding molecule that binds specifically to said polypeptide; and

(b)detecting the presence of a complex comprising the said antigen-binding molecule and the said polypeptide, biologically active fragment, variant or derivative in said contacted sample

29. Use of the antigen-binding molecule of claim 27 in treating enenvomation by *S. horrida* or related species.