NOVEL POLYPEPTIDES FOR ANTI-VIRAL TREATMENT

TECHNICAL FIELD

This present invention generally relates to novel polypeptides for use in anti-viral treatment of human diseases.

BACKGROUND

Many naturally occurring cyclic polypeptides exhibit biological activities ranging from uterotonic activity, anti-HIV property, neurotensin inhibition, and antimicrobial activity to insecticidal activity. Particularly well-known are the uterotonic activity of kalata B1 isolated from *Oldenlandia affinis* and the anti-HIV activity of circulins isolated from *Chassalia parvifolia K. Schum*. The traditional use of herbs containing such cyclic polypeptides as indigenous medicine for a range of medical conditions has led to the isolation and identification of the cyclotide polypeptide family (Gran et al., (2000) J Ethnopharmacol. 70(3): 197-203).

Cyclotides are disulphide-rich proteins with the unusual features of a cyclic backbone and a cystine knot topology. Such a protein motif consists of three conserved disulphide bonds, two of which form an embedded ring in the structure that is penetrated by the third disulphide bond. For example, detailed structural studies based on NMR data on kalata B1 revealed that the six cysteine residues in this 29-mer polypeptide form three disulphide bridges that link up the head-to-tail cyclic backbone into the uniquely folded structure designated as a cyclic cystine knot (CCK) (Craik et al. (1999) J Mol Biol. 294(5): 1327-36). Such a unique folding confers to the molecules exceptional structural stability and resistance to thermal and proteolytic degradation.

The structural stability of cyclotides has led to extensive research on their potential use as molecular frameworks for drug delivery in pharmaceutical applications. For example, the cyclotides may provide a molecular framework for drug delivery in anti-viral treatment such as in anti-dengue treatment. No report has however been made on any natural cyclotide having inhibitory activity against the dengue virus. The potential of such an application therefore relies on the use of molecular design to introduce new structural features or sequences into the cyclotide molecular framework, thereby generating new biological activities in the molecules. The basis for any design in turn relies on an in-depth

understanding of mechanisms of viral infection and replication as well as structure-function relationships of the resulting molecules.

Dengue is a mosquito-borne infection, the global prevalence of which has grown dramatically in recent decades. Found in tropical and sub-tropical regions around the world, the disease is now endemic in more than 100 countries in Africa, the Americas, the Eastern Mediterranean, South-east Asia and the Western Pacific. In Singapore, there are about 4000-5000 reported cases of dengue fever or dengue haemorrhagic fever every year. As of October 2005, Singapore reported 12,700 cases in the year with at least 19 deaths (Ministry of Health, Singapore). Other recent dengue outbreaks in South East Asia include the 21,537 cases with 280 dead in the Phillipines (January – October 2005), the 7200 infections in Thailand with at least 12 dead (May 2005), Indonesia's 80,000 infections with 800 deaths (2004) and the 33,203 cases in Malaysia (January 2005).

The clinical features of dengue fever vary according to the age of the patient and may range from non-specific febrile illness with rash in infants and young children to the abrupt onset of high fever, severe headache, pain behind the eyes, muscle and joint pains, and rash in older children and adults. Dengue haemorrhagic fever (DHF), on the other hand, is a potentially deadly complication characterized by high fever, haemorrhagic phenomena--often with enlargement of the liver--and in severe cases, circulatory failure.

At present, there is no commercial vaccine or causative treatment for the prevention or cure of dengue viral diseases, with careful clinical management and supportive therapy frequently being the main treatment options. Furthermore, vaccine development for dengue and DHF is complicated by the possibility of the disease being caused by any of four different viruses, and by the potential increase in risk of more serious disease if protection is only acquired against one or two of the viruses.

Dengue virus contains a single positive-stranded RNA genome, which is first translated into a polypeptide that is then processed by host cell proteases and the virus-encoded NS3 protease to generate three structural and seven non-structural viral proteins. Optimal activity of the NS3 protease is essential for maturation of the virus, and while it has been shown to exhibit activity independently with some model substrates for serine proteases, its activity is markedly enhanced when coupled to NS2B (Yusof *et al*, (2000) J. Biol. Chem. 275:9963-9969). Accordingly, inhibition of NS3 offers the prospect of an effective anti-viral therapy for dengue fever, particularly in severe cases such as those of DHF and dengue shock syndrome.

With the exception of the NS2B/NS3 cleavage site, which contains a glutamine residue at the P2 position, cleavage recognition sites for the NS3 protease comprise a pair of positively charged dibasic amino acids (such as Arg or Lys) at the P1 and P2 positions, followed by a short-chain amino acid (such as Gly, Ala, Asp or Ser) at the P1' site. Table 1 shows cleavage recognition sites of the NS3 protease in the NS2A/NS2B, NS2B/NS3, NS3A/NS4A and NS4B/NS5 regions of the viral polypeptide for dengue, Japanese encephalitis, Kunjin, Murray Valley encephalitis, yellow fever, West Nile, tick-borne encephalitis viruses.

Table 1. Cleavage recognition sites for the NS3 protease

Region / Virus	NS2A-NS2B	NS2B-NS3	NS3A-NS4A	NS4B-NS5
Den 1	GRK↓SWP	KGR↓SGV	GRR↓SVS	GRR↓GTG
Den 2	KKR↓SWP	KQR↓AGV	GRK↓SLT	TRR↓GTG
Den 3	KRR↓SWP	TQR↓SGV	GRK↓SIA	GKR↓GTG
Den 4	SRR↓SWP	TQR↓SGA	GRK↓SIT	PRR↓GTG
JE	KKR↓GWP	TKR↓GGV	GKR↓DAI	LKR↓GRP
KUN	RKR↓GWP	TKR↓GGV	GKR↓SQI	LKR↓GGA
MVE	KKR↓GWP	TKR↓GGV	GKR↓SAI	FKR↓GRA
YF	GRR↓SIP	ARR↓SGD	GRR↓GAA	GRR↓GSA
WN	RKR↓GWP	TKR↓GGV	GKR↓SQI	LKR↓GGA
TBE	GRR↓SFS	SRR↓SDL	GRR↓SFG	GRR↓GGS

Den = dengue; JE = Japanese encephalitis, KUN = Kunjin, MVE = Murray Valley encephalitis, YF = yellow fever, WN = West Nile, TBE = tick-borne encephalitis viruses.

Inhibitors for the NS3 protease may be synthetic polypeptides mimicking uncleavable recognition sequences of the native polypeptide sites, or synthetic polypeptides with cleaved recognition sequences that bind the protease to prevent it from processing other polypeptide substrate molecules. However, the optimal activity of such polypeptides is often dependent on their structural stability as well as pH and temperature conditions. Stabilization of such polypeptides against degradation before they can act on a target therefore presents a serious drug delivery problem.

Clearly, there is a continuous need to provide new modified polypeptides that overcome or at least ameliorate one or more of the disadvantages described above as molecular frameworks for drug delivery systems in pharmaceutical applications such as anti-viral treatment. With such new synthetic molecules, there is a parallel need for

developing convenient and reliable assays for evaluating the efficacy of their biological activities in the intended applications.

SUMMARY

The inventors have designed a series of novel polypeptides comprising viral NS3 protease recognition sequence on a cyclic molecular framework. Advantageously, the backbone of said cyclic molecular framework is "open" with a nick at the viral NS3 protease recognition sequence, enabling the molecule to be recognized and bound by the protease to prevent it from processing other viral polypeptide substrate molecules. With inhibition of the protease, viral replication and maturation are thus inhibited.

The cyclic conformation of the molecules provides structural stability against degradation. The cyclic conformation is maintained and stabilized by disulphide bonds. The disulphide bonds are spontaneously formed during expression of the novel polypeptides as fusion polypeptides, thereby advantageously obviating the need for further cyclization steps.

Efficacy of said novel polypeptides as dengue virus NS3 protease inhibitors in anti-dengue treatment may be measured either using commercially available protease activity assay kits modified for anti-dengue protease activity assay, or with assay methods developed by the inventors. Using these assay methods, efficacy of the novel polypeptides was advantageously shown to be higher than the "open" cyclic kalata B1 molecule.

According to a first aspect, there is provided a polypeptide comprising the following sequence of amino acids or analogues thereof:

X_1-Leu-Pro-Val-Cys-Gly-Glu-Thr-Cys-Val-Gly-Gly-Thr-X_2-Asn-Thr-Pro-Gly-X_3-Thr-X_4-Ser-Trp-Pro-Val-Cys-Thr-X_5-X_6

wherein

X_1 is Ala, Asp, Gly, or Ser

X_2 is Cys or Gly

X_3 is Cys or Gly

X_4 is Cys or Gly

X_5 is Arg, Gln or Lys

X_6 is Arg or Lys.

According to a second aspect, there is provided a polypeptide comprising the following sequence of amino acids or analogues thereof:

X_1-Leu-Pro-Val-Cys-Gly-Glu-Thr-Cys-Val-Gly-Gly-Thr-Gly-Asn-Thr-Pro-Gly-X_2-Thr-X_3-Ser-Trp-Pro-Val-Cys-Thr-X_4-Asn

wherein
X_1 is Ala, Asp, Gly, or Ser
X_2 is Cys or Gly
X_3 is Cys or Gly
X_4 is Arg, Gln or Lys.

According to a third aspect, there is provided a polypeptide comprising an open cyclic backbone, said open cyclic backbone comprising the following sequence of amino acids or analogues thereof:

X_1-Leu-Pro-Val-Cys-Gly-Glu-Thr-Cys-Val-Gly-Gly-Thr-X_2-Asn-Thr-Pro-Gly-X_3-Thr-X_4-Ser-Trp-Pro-Val-Cys-Thr-X_5-X_6

wherein
X_1 is Ala, Asp, Gly, or Ser
X_2 is Cys or Gly
X_3 is Cys or Gly
X_4 is Cys or Gly
X_5 is Arg, Gln or Lys
X_6 is Arg or Lys.

According to a fourth aspect, there is provided a polypeptide comprising an open cyclic backbone, said open cyclic backbone comprising the following sequence of amino acids or analogues thereof:

X_1-Leu-Pro-Val-Cys-Gly-Glu-Thr-Cys-Val-Gly-Gly-Thr-Gly-Asn-Thr-Pro-Gly-X_2-Thr-X_3-Ser-Trp-Pro-Val-Cys-Thr-X_4-Asn

wherein
X_1 is Ala, Asp, Gly, or Ser
X_2 is Cys or Gly
X_3 is Cys or Gly
X_4 is Arg, Gln or Lys.

In one embodiment, the polypeptide of the first and third aspects comprises a sequence set forth in SEQ ID NO. 7, SEQ ID NO. 8, SEQ ID NO. 9, SEQ ID NO. 10 or SEQ ID NO. 11.

In one embodiment, the polypeptide of the second and fourth aspects comprises the sequence set forth in SEQ ID NO. 12.

In one embodiment, the open cyclic backbone of the polypeptide of the third and fourth aspects comprises at least one disulphide bond. The open cyclic backbone may be in the form of a cystine knot.

According to a fifth aspect, there is provided a polypeptide comprising the following sequence of amino acids or analogues thereof:

X_1-Leu-Pro-Val-Cys-Gly-Glu-Thr-Cys-Val-Gly-Gly-Thr-X_2-Asn-Thr-Pro-Gly-X_3-Thr-X_4-Ser-Trp-Pro-Val-Cys-Thr-X_5-X_6

wherein
X_1 is Ala, Asp, Gly, or Ser
X_2 is Cys or Gly
X_3 is Cys or Gly
X_4 is Cys or Gly
X_5 is Arg, Gln or Lys

X_6 is Arg or Lys, wherein said polypeptide is capable of inhibiting viral NS3 protease.

According to a sixth aspect, there is provided a polypeptide comprising an open cyclic backbone, said open cyclic backbone comprising the following sequence of amino acids or analogues thereof:

X_1-Leu-Pro-Val-Cys-Gly-Glu-Thr-Cys-Val-Gly-Gly-Thr-Gly-Asn-Thr-Pro-Gly-X_2-Thr-X_3-Ser-Trp-Pro-Val-Cys-Thr-X_4-Asn

wherein
X_1 is Ala, Asp, Gly, or Ser
X_2 is Cys or Gly
X_3 is Cys or Gly
X_4 is Arg, Gln or Lys, wherein said polypeptide is capable of inhibiting viral NS3 protease.

In one embodiment, there is provided a functional equivalent of the polypeptide of the first, second, third, fourth, fifth or sixth aspects, which retains the inhibitory activity of the reference polypeptide. Preferably, the functional equivalent has at least 60% sequence identity with a polypeptide selected from the group consisting of SEQ ID NOS: 7, 8, 9, 10, 11 and 12. More preferably, the functional equivalent has at least 70%, 80% or 90% sequence identity with a polypeptide selected from the group consisting of SEQ ID NOS: 7, 8, 9, 10, 11 and 12.

According to a seventh aspect, there is provided a polynucleotide comprising a sequence of nucleotides or analogues thereof which encodes a polypeptide of the first, second, third, fourth, fifth or sixth aspects. The polynucleotide may be in the form of RNA or DNA.

In one embodiment, the polynucleotide of the seventh aspect comprises the sequence set forth in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6 or combinations thereof.

According to an eight aspect, there is provided a recombinant vector comprising a polynucleotide of the seventh aspect.

According to a ninth aspect, there is provided a host cell transformed with the recombinant vector of the eighth aspect.

According to a tenth aspect, there is provided a method of producing a polypeptide of the first, second, third, fourth, fifth or sixth aspects, the method comprising the step of culturing a host cell of the ninth aspect under conditions which permit expression of said polypeptide.

According to an eleventh aspect, there is provided a composition for treatment or prophylaxis of a NS3 protease-related condition in a subject, comprising a polypeptide of the first, second, third, fourth, fifth or sixth aspect, and a pharmacologically acceptable carrier, excipient or diluent.

The carriers, diluents and adjuvants must be "acceptable" in terms of being compatible with the other ingredients of the composition, and not deleterious to the recipient thereof. Examples of pharmaceutically acceptable carriers, excipients or diluents are demineralised or distilled water; saline solution; vegetable based oils such as peanut oil, safflower oil, olive oil, cottonseed oil, maize oil, sesame oils, arachis oil or coconut oil; silicone oils, including polysiloxanes such as methyl polysiloxane, phenyl polysiloxane and methylphenyl polysolpoxane; volatile silicones; mineral oils such as liquid paraffin, soft paraffin or squalane; cellulose derivatives such as methyl cellulose, ethyl cellulose, carboxymethylcellulose, sodium carboxymethylcellulose or hydroxypropylmethylcellulose; lower alkanols, for example ethanol or iso-propanol; lower aralkanols; lower polyalkylene glycols or lower alkylene glycols, for example polyethylene glycol, polypropylene glycol, ethylene glycol, propylene glycol, 1,3-butylene glycol or glycerin; fatty acid esters such as isopropyl palmitate, isopropyl myristate or ethyl oleate; polyvinylpyrolidone; agar; gum tragacanth or gum acacia, and petroleum jelly. Typically, the carrier or carriers will form from 10% to 99.9% by weight of the compositions.

In one embodiment, the NS3 protease related condition is selected from the group consisting of dengue, Japanese encephalitis, Kunjin viral disease, Murray Valley encephalitis, yellow fever, West Nile disease and tick-borne encephalitis.

In one embodiment, the subject may be a human.

According to a twelfth aspect, there is provided use of a polypeptide of the first, second, third, fourth, fifth or sixth aspect in the manufacture of a medicament for the

treatment or prophylaxis of a NS3 protease related condition. The NS3 protease related condition may be selected from the group consisting of dengue, Japanese encephalitis, Kunjin viral disease, Murray Valley encephalitis, yellow fever, West Nile disease and tick-borne encephalitis.

According to a thirteenth aspect, there is provided use of a polypeptide of the first, second, third, fourth, fifth or sixth aspect in the study, treatment or prophylaxis of a NS3 protease related condition, particularly dengue.

According to a fourteenth aspect, there is provided a method for treatment or prophylaxis of a NS3 protease related viral disease, comprising administering to a patient in need of such treatment a therapeutically effective dose of a composition of the eleventh aspect.

According to a fifteenth aspect, there is provided a method for assaying the inhibitory activity of a polypeptide of the first, second, third, fourth, fifth or sixth aspect or a composition of the eleventh aspect, the method comprising the steps of:

a) incubating a mixture comprising a NS2B/NS3 fusion polypeptide containing a restriction protease cleavage recognition site and a buffer in the presence or absence of a polypeptide of the first, second, third, fourth, fifth or sixth aspect or a composition of the eleventh aspect;

b) adding to said mixture a restriction protease capable of cleaving said NS2B/NS3 fusion polypeptide at the restriction protease cleavage recognition site to release the NS2B/NS3 polypeptide containing a NS3 cleavage recognition site;

c) measuring the levels of uncleaved NS2B/NS3 polypeptide in the presence and absence of a polypeptide of the first, second, third, fourth, fifth or sixth aspect or a composition of the eleventh aspect; and

d) comparing the levels of uncleaved NS2B/NS3 polypeptide in the presence and absence of a polypeptide of the first, second, third, fourth, fifth or sixth aspect or a composition of the eleventh aspect.

The term "restriction protease" refers to a proteolytic enzyme that cleaves polypeptides or proteins at specific recognition sequence sites. The restriction protease may be enterokinase, Factor Xa, thrombin, subtilisin, rennin, trypsin, chymotrypsin, kallikrein, TEV proteinase, Kex2 proteinase or collagenase. In one embodiment, the restriction protease is enterokinase.

According to a sixteenth aspect, there is provided a kit for assaying the inhibitory activity of a polypeptide of the first, second, third, fourth, fifth or sixth aspect, wherein the kit comprises a polypeptide of the first, second, third, fourth, fifth or sixth aspect.

DEFINITIONS

This section is intended to provide guidance on the interpretation of the words and phrases set forth below (and where appropriate grammatical variants thereof). Further guidance on the interpretation of certain words and phrases as used herein (and where appropriate grammatical variants thereof) may additionally be found in other sections of this specification.

The term "artificially synthesized" when used in reference to a nucleic acid sequence refers to a nucleic acid the sequence of which has been obtained via back translation of a polypeptide, peptide or protein sequence, and synthesized using methods that are well known in the art.

As used herein, the terms "polynucleotide", "oligonucleotide", "nucleic acid' and "nucleic acid molecule" refer to any homo- or hetero-polymer of ribonucleotides or deoxyribonucleotides of any length, that comprise natural, chemically or biochemically modified, non-natural, or derivatized nucleotide bases linked through phosphodiester bonds or modified phosphodiester bonds. The polynucleotide may be a single stranded molecule or a double stranded molecule of genomic or synthetic origin. Modified polynucleotides include methylated nucleotides and nucleotide analogues. The sequence of nucleotides may be interrupted by non-nucleotide components. The terms polynucleotide, oligonucleotide, nucleic acid and nucleic acid molecule further include complements, fragments and variants of the polynucleotide, oligonucleotide, nucleic acid and nucleic acid molecule, or analogues thereof.

As used herein, "fusion polynucleotide" refers to a polynucleotide comprising a plurality of regions. Plurality in this context means at least two.

The terms "polypeptide", "peptide" and "protein" refer to any polymer of amino acid residues (dipeptide or greater) linked through peptide bonds or modified peptide bonds and to variants and synthetic analogues of the same. Thus, these terms apply to naturally-occurring amino acid polymers as well as 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. Polypeptides of the present invention include, but are not limited to, products of chemical synthetic procedures, and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect and mammalian cells. The polypeptides of the invention may comprise non-peptidic components, such as carbohydrate groups. Carbohydrates and other non-peptidic substituents may be added to a polypeptide by the cell in which the polypeptide is produced, and will vary with the type of cell. For polypeptides that are made recombinantly, the nature and extent of the modifications in large part will be determined by the post-translational modification capacity of the particular host cell and the modification signals that are present in the amino acid sequence of the polypeptide in question. For instance, glycosylation patterns vary between different types of host cell. Polypeptides are defined herein, in terms of their amino acid backbone structures; substituents such as carbohydrate groups are generally not specified, but may be present nonetheless. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Proteins may be present as monomeric or as multimeric proteins e.g. as dimers (homo or heterodimers) or trimers.

The term "fusion polypeptide" refers to a polypeptide having a plurality of regions, each corresponding to a distinct peptide. Fusion polypeptides can include linkers connecting the regions thereof. Likewise, the term "fusion protein" as used herein, means a protein having a plurality of regions, each corresponding to a distinct peptide. Fusion proteins can include linkers connecting the regions thereof. Typically, for both fusion polypeptides and fusion proteins, while the plurality of regions are unjoined in their native state, they can be joined by their respective amino and carboxyl termini through a peptide linkage to form a single continuous polypeptide or protein. Plurality in this context means at least two. It will be appreciated that the polypeptide or protein components can be joined directly or joined through a linker. Typically, the linker is not part of the sequence of either the fusion partner as outlined below or the target polypeptide / protein. Typically, the linker is a short sequence of amino acids, for example about one to about 20 amino acids. Typically, the linker includes a cleavage recognition site for an enzymatic or chemical cleavage reagent as outlined below so that the fusion partner may be cleaved and

purified away from the target polypeptide / protein. The linker may also include additional sequences inserted by one skilled in the art, for example, to provide sufficient spacing between the fusion partner and target polypeptide, or to facilitate cloning. Fusion partners may be used, which may include one or more additional amino acid sequences containing secretory or leader sequences, pro-sequences, or sequences which aid in, for instance detection, expression, separation or purification of the protein or to endow the protein with additional properties as desired such as higher protein stability, for example during recombinant production, or for instance to produce an immunomodulatory response. Examples of potential fusion partners include purification tags, such as a polyhistidine tag, epitope tags (short peptide sequences for which a specific antibody is available) and specific binding proteins; enzymes such as ribonuclease S, glutathione S-transferase, beta-galactosidase, luciferase and hemagglutinin; thioredoxin; a secretion signal peptide and a label, which may be, for instance, bioactive, radioactive, enzymatic or fluorescent, or an antibody. The target polypeptide / protein refers to the polypeptide / protein of the present invention.

As used herein, the terms "mutant" and "mutation" include any detectable change in genetic material, e.g. DNA, or any process, mechanism, or result of such a change. This includes gene mutations, in which the structure (e.g. DNA sequence) of a gene is altered, any gene or DNA arising from any mutation process, and any expression product (e.g. protein or enzyme) expressed by a modified gene or DNA sequence.

As used herein, the term "permutant" refers to any arrangement a non-mutant polypeptide may adopt. For example, "open cyclic permutant of kalata B1" refers to a cyclic polypeptide having the same amino acid sequence as the naturally occurring kalata B1 but containing a nick. The term "open cyclic permutant of kalata B1" differs from "open cyclic mutant(s) of kalata B1" in that the latter refers to a nicked cyclic polypeptide having an amino acid sequence that is different (i.e. mutated) from the naturally occurring kalata B1.

As used herein, the term "variant" refers to a polynucleotide or polypeptide that differs from a parent polynucleotide or polypeptide respectively, but retains essential properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, parent polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the parent

polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the parent sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from another, parent polypeptide. Generally, differences are limited so that the sequences of the parent polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and parent polypeptide may differ in amino acid sequence by one or more substitutions, additions and deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code.

As used herein, the term "analogue" refers to a polynucleotide or a polypeptide which is a derivative of a parent polynucleotide or polypeptide respectively, which derivative comprises addition, deletion or substitution of one or more nucleotides or amino acids respectively, such that the polypeptide analogue retains substantially the same function as the parent polypeptide. The term "derivative" refers to a chemically or biologically modified version of the compound that is structurally similar to a parent compound. Derivatives of an amino acid include amino acids which are different from naturally-occurring amino acids and which have functions similar to those of the naturally-occurring amino acids. "Naturally-occurring amino acid" as used herein includes alanine (Ala), arginine (Arg), asparagine (Asn), aspartic acid (Asp), cysteine (Cys), glutamine (Gln), glutamic acid (Glu), glycine (Gly), histidine (His), isoleucine (Ile), leucine (Leu), lysine (Lys), methionine (Met), phenylalanine (Phe), proline (Pro), serine (Ser), threonine (Thr), tryptophan (Trp), tyrosine (Tyr), valine (Val) and other amino acids which may or may not be found in proteins. For example, derivatives of the naturally-occurring amino acid cysteine include D-cysteine, N-acetyl-L-cysteine, DL-homocysteine, L-cysteine methyl ester, L-cysteine ethyl ester, N-carbamoyl cysteine, glutathione and cysteamine.

As used herein, the term "functional equivalent" refers to a polypeptide that retains the biological activity of the parent polypeptide. The functionally-equivalent polypeptide may be homologous to the parent polypeptide or to a natural biological variant or an analogue thereof. Functional equivalents of the polypeptides may also include polypeptides in which relatively short stretches have a high degree of homology (at least 60%, 70%, 80%, 85%, 90%, 92%, 95% or 97%) with the parent polypeptide even though the overall homology between the two polypeptides may be much less. This is because

important recognition or binding sites may be shared even when the general architecture of the polypeptide is different.

As used herein, the term "biological activity" refers to the inhibitory function of a polypeptide of the present invention, or to physiological responses that result upon administration of said polypeptide or a composition containing said polypeptide. Biological activity, thus, encompasses therapeutic effects and pharmaceutical activity of such polypeptides and compositions. Biological activities may be observed in *in vitro* systems designed to test or use such activities. Biological activity of a variant, analogue or functional equivalent of a polypeptide may be retained at the same or different level as the parent polypeptide. In other words the variant, analogue or functional equivalent may be more or less active (or, of course, of similar or identical activity) as the parent polypeptide.

The term "NS3 related condition" refers to a condition which is or can be caused, directly or indirectly, by the activity of the NS3 protease or of the NS3 protease when it is coupled to another viral component such as NS2B (i.e. NS2B/NS3). The term further includes a condition with which the NS3 protease, or the NS3 protease coupled to another viral component such as NS2B (i.e. NS2B/NS3), is in any way associated.

The term "treatment" includes any and all uses which remedy a disease state or symptoms, prevent the establishment of disease, or otherwise prevent, hinder, retard, or reverse the progression of disease or other undesirable symptoms in any way whatsoever. Treatment may be effected prophylactically or therapeutically.

The term "anti-viral" refers to a compound or a method that can inhibit viral infection of cells for example via inhibition of virus-induced cell-cell fusion, or to inhibit viral replication after infection. Viral infection may involve membrane fusion, as occurs in the case of enveloped viruses, or some other fusion event involving a viral structure and a cellular structure (e.g., such as the fusion of a viral pilus and bacterial membrane during bacterial conjugation). Viral replication may be inhibited by for example inhibition of viral enzymes required for replication (e.g., such as inhibition of the dengue NS3 protease that liberates individual viral peptides from the viral polypeptide precursor). The term "anti-viral" further refers to a compound or a method that can reduce the likelihood of a person or animal contracting the viral disease upon exposure to potentially infective viral particles.

The term "therapeutically effective amount" includes a non-toxic but sufficient amount of a compound to provide the desired therapeutic effect. The exact amount

required will vary from subject to subject depending on factors such as the species being treated, the age and general condition of the subject, the severity of the condition being treated, the particular compound being administered and the mode of administration and so forth. Thus, it is not possible to specify an exact "effective amount". However, for any given case, an appropriate "effective amount" may be determined by one of ordinary skill in the art using only routine experimentation.

The term "molecular framework" refers to a polynucleotide or polypeptide molecule that provides rigid or semi-rigid structural and configurational support for a second polynucleotide or polypeptide molecule, such that the biological activity of the second polynucleotide or polypeptide is retained.

The term "cyclic" when used in reference to a peptide, polypeptide or protein of the present invention refers to a peptide, polypeptide or protein that comprises cysteine residues or derivatives thereof at positions suitable for formation of at least one intramolecular disulphide bonds, and that has been folded into a cyclic conformation held together by said intramolecular disulphide bond(s).

The term "close" when used in reference to a cyclic peptide, polypeptide or protein of the present invention refers to a cyclic peptide, polypeptide or protein that contains an intrapeptide bond. This intrapeptide bond is preferably formed between the N-terminal amino group of the peptide and the C-terminal carboxy group of the peptide (referred to as "head-to-tail" cyclization), although peptides that have been cyclized by intrapeptide bond formation involving a side chain group(s) are also included.

The term "open" when used in reference to a cyclic peptide, polypeptide or protein of the present invention refers to a cyclic peptide, polypeptide or protein that contains a nick in its cyclic amino acid backbone, preferably at a NS3 protease recognition site.

The terms "purified", "purification" and grammatical variations thereof means that the material in question has been removed from its natural or recombinant host, and associated impurities reduced or eliminated. Preferably, it means an object species is the predominant species present (ie., on a molar basis it is more abundant than any other individual species in the composition), and more preferably a substantially purified fraction is a composition wherein the object species comprises at least about 30 percent (on a molar basis) of all macromolecular species present. Generally, a substantially pure composition will comprise more than about 80 to 90 percent of all macromolecular species

present in the composition. Most preferably, the object species is purified to essential homogeneity (contaminant species cannot be detected in the composition by conventional detection methods such as Coomassie Brilliant Blue staining following SDS-PAGE analysis) wherein the composition consists essentially of a single macromolecular species.

The term "cleave" and grammatical variations thereof refer to the chemical or enzymatic cutting of a polypeptide at specific amino acid sequence sites and the term "cleavage" includes scission or proteolysis in this invention.

The term "cleavage reagent" as used herein refers to a reagent used to cleave a polypeptide at specific amino acid sequence sites, for example, to release a target polypeptide in a fusion polypeptide from its fusion partner. Suitable cleavage reagents include enzymes, such as proteases, and chemical reagents. The cleavage reagent is selected, taking into consideration the amino acid sequence of the polypeptide to be cleaved. Care must be taken that, if possible, the recognition sequence of the cleavage reagent for cleavage does not occur in a region other than in the intended cleavage region. Suitable specifically cleaving proteases are, for example, enterokinase, Factor Xa, thrombin, subtilisin, renin, collagenase, trypsin, chymotrypsin, endoproteinase Lys-C, kallikrein (Carter, P.: In: Ladisch, M. R.; Willson, R. C.; Painton, C. C.; Builder, S. E., eds., Protein Purification: From Molecular Mechanisms to Large-Scale Processes; ACS Symposium Series No. 427, American Chemical Society, pp. 181-193 (1990)), TEV proteinase (Parks, T. D., et al., Anal. Biochem. (1994) 216: 413-417), IgA protease (Pohlner, J., et al., Nature (1987) 325: 458-462), Kex2p proteinase (EP-A 0 467 839) or *S. aureus* V8 proteinase. Suitable specifically cleaving chemical reagents include halogenated succinimides, cyanogen halides, hydroxylamines or combinations thereof. Exemplary halogenated succinimides include N-chlorosuccinimide (NCS), N-iodosuccinimide (NIS) and N-bromosuccinimide (NBS). Exemplary cyanogen halides include cyanogen chloride (CNCl), cyanogen iodide (CNI) and cyanogen bromide (CNBr). Conditions used for cleavage of a polypeptide will depend on the cleavage reagent used, and the conditions will be readily apparent to one skilled in the art.

It will be understood by the skilled person that "molecular weights" measured as described below using electrophoresis are inherently subject to interpretation since they are measured relative to standard molecular weight markers.

As used in this application, the singular form "a," "an," and "the" include plural references unless the context clearly dictates otherwise. For example, the term "an agent" includes a plurality of agents, including mixtures thereof.

As used herein, the term "comprising" means "including." Variations of the word "comprising", such as "comprise" and "comprises," have correspondingly varied meanings. Thus, for example, a composition "comprising" X may consist exclusively of X or may include one or more additional components.

Throughout this disclosure, various aspects of this invention can be presented in a range format. It should be understood that the description in range format is merely for convenience and brevity and should not be construed as an inflexible limitation on the scope of the invention. Accordingly, the description of a range should be considered to have specifically disclosed all the possible subranges as well as individual numerical values within that range. For example, description of a range such as from 1 to 6 should be considered to have specifically disclosed subranges such as from 1 to 3, from 1 to 4, from 1 to 5, from 2 to 4, from 2 to 6, from 3 to 6 etc., as well as individual numbers within that range, for example, 1, 2, 3, 4, 5, and 6. This applies regardless of the breadth of the range.

As used herein, the term "about" as used in relation to a numerical value means, for example, $\pm$50% or $\pm$30% of the numerical value, preferably $\pm$20%, more preferably $\pm$10%, more preferably still $\pm$5%, and most preferably $\pm$1%. Where necessary, the word "about" may be omitted from the definition of the invention.

Unless otherwise indicated, all technical and scientific terms used herein have the same meaning as commonly understood to one of ordinary skill in the art to which the invention belongs.

BRIEF DESCRIPTION OF DRAWINGS

The accompanying drawings illustrate a disclosed embodiment and serve to explain the principles of the disclosed embodiment. It is to be understood, however, that the drawings are designed for purposes of illustration only, and not as a definition of the limits of the invention.

Figure 1 shows the nucleic acid and amino acid sequences of kalata B1 and its mutants. (A) The synthetic DNA sequence obtained by back translation based on the published amino acid sequence of kalata B1 (GenBank accession No.: 4F393825;

GI:15667740). (B) Amino acid sequence of thioredoxin fused to an open cyclic kalata B1 polypeptide. Residues 1 to 109 specifying thioredoxin are italicized. The polyhistidine tag is highlighted in bold. The enterokinase recognition sequence is underlined and the arrow indicates the site of cleavage to yield the 29-amino acid open cyclic kalata B1 polypeptide with no unwanted residues (C). (D) Amino acid sequences of fusion polypeptides for mutants 1a, 2b, 3d, 4d, 5b and 6c. (E) Amino acid sequences of 29-mer polypeptides of mutants 1a, 2b, 3d, 4d, 5b and 6c.

Figure 2 shows the expression and purification of the open cyclic permutant of kalata B1. (A) SDS-PAGE analysis of the expression level. Lanes 1-3 are the Amersham protein molecular weight marker and the cell extract obtained before and after IPTG induction, respectively. (B) Lanes 1-3 are, respectively, the Amersham protein molecular weight standard, the synthetic His_6-tagged thioredoxin-open cyclic permutant of kalata B1 fusion polypeptide, and the His_6-tagged thioredoxin as cleavage product.

Figure 3 (A) shows the cleavage of synthetic open cyclic permutant of kalata B1 from the fusion polypeptide by enterokinase. 40 µl of reactant solution containing 40 µg of the fusion polypeptide with 0.5 unit of enterokinase was incubated at 25°C. Lanes 3-8 were from samples obtained after 1, 5, 10, 30, 60, and 120 min of incubation, respectively. (B) Same as (A) except that 0.2 M dithiothreitol was included in the reaction mixture; (C) Same as (A) but in the presence of 0.2 M hydrogen peroxide.

Figure 4 shows RP-HPLC chromatograms of the open cyclic permutant of kalata B1. (A) A typical chromatogram of a synthetic open cyclic permutant of kalata B1 in its fully oxidized form. (B) A typical chromatogram of a synthetic open cyclic permutant of kalata B1 in its fully reduced form.

Figure 5 shows the LCMS (A) and MALDI-TOF (B) data for confirmation of the identity of the open cyclic permutant of kalata B1.

Figure 6 shows the LCMS and MALDI-TOF data for confirmation of the identity of (A) mutant 1a; (B) mutant 2b; (C) mutant 3d; (D) mutant 4d; (E) mutant 6c.

Figure 7 shows inhibitory effect of an open cyclic mutant of kalata B1 on the dengue virus NS2B/NS3 protease. Lane 1: Molecular weight standard from Amersham; Lane 2: thioredoxin-NS2B/NS3 without any treatment; Lane 3: thioredoxin-NS2B/NS3

with enterokinase and PBS buffer as control; Lane 4: thioredoxin-NS2B/NS3 with enterokinase treated with an open cyclic permutant of kalata B1; Lane 5: thioredoxin-NS2B/NS3 with enterokinase treated with the open cyclic mutant 1a; Lane 6: thioredoxin-NS2B/NS3 with enterokinase treated with water.

DETAILED DESCRIPTION

The practice of disclosed embodiments will employ, unless otherwise indicated, techniques of molecular biology, microbiology, and recombinant DNA technology which are within the skill of those working in the art. Such techniques are explained fully in the literature. Examples of texts for consultation include the following: Sambrook Molecular Cloning; A Laboratory Manual, Third Edition (2000). For protocols and descriptions of biochemical and biophysical techniques, examples of texts for consultation include those in the series Methods in Enzymology, Academic Press.

Back translation of kalata B1 sequence

The cDNA sequence of kalata B1 was obtained and codon-optimized using back translation tools of Entelechon GmBH (Leto1.0).

Synthetic oligonucleotides and polypeptides

Methods for generating cDNA from mRNA are well known in the art. All oligonucleotides used for plasmid construction were purchased from Proligo (Singapore). *Escherichia coli* strain DH5α (Stratagene, La Jolla, CA) was used as the host for all cloning procedures while protein expression was achieved using *E. coli* strain BL2l (DE3) (Novagen, Madison, WI, USA) as hosts.

Construction of expression vectors

Any expression vector suitable for expression of the target polynucleotide may be used. Examples of vectors include the pET series from Novagen, the pQE series from Qiagen, and pBR322 in the case in which *Escherichia coli* serves as a host and pUB110 in the case in which *Bacillus subtilis* serves as a host. The lac or tac promoter, commonly used for inducing the expression of fused or unfused gene products, may be inserted into

the vector. Alternatively, the T7 promoter may be used. A polynucleotide sequence encoding a fusion partner containing a tag may be cloned downstream of and in operable linkage with the promoter. This sequence may contain the translational start codon ATG. The sequence coding for the cleavage recognition site may be inserted between the fusion partner and target polypeptide encoding sequences. Multiple restriction sites for the insertion of the target gene may be present downstream of the fusion partner. It is preferable to maintain the number of sites in the multiple cloning region (MCR) to a minimum. A nucleotide sequence coding for a translation termination codon, in all 3 reading frames, may be inserted downstream of the target polypeptide encoding sequence. A sequence encoding the target polypeptide of the present invention may be inserted in the multiple cloning site in the appropriate reading frame.

In the present example, an oligonucleotide containing the coding sequence of the open cyclic kalata B1 polypeptide (Fig. 1A) was artificially synthesized with 13 silent codon changes made according to *E. coli* codon preference. To facilitate directional cloning into the bacterial expression vector, pET32a (Novagen), 5' *Bgl* II and 3' *Sal I* sites were incorporated by PCR amplification of the synthetic DNA. The 5' primer also contained a 15-bp sequence encoding an enterokinase recognition site downstream of the *BglI* restriction recognition sequence. Directional insertion of the amplified product places the kalata Bl-coding sequence downstream of the thioredoxin. The resulting amino acid sequence of the expressed fusion polypeptide is shown in Fig. 1B. The incorporated enterokinase recognition site allows separation of the open cyclic kalata Bl polypeptide (or a mutant thereof) from the fusion partner without leaving any unwanted residues on the 29-mer open cyclic kalata Bl polypeptide (or a mutant thereof) (Fig. 1C).

Mutants of the open cyclic kalata B1 polypeptide

Six mutants, labelled 1a, 2b, 3d, 4d, 5b and 6c, with amino acid substitutions were similarly produced. The amino acid sequences of the expressed fusion polypeptides of these mutants are shown in Fig. 1D while their corresponding 29-mer open cyclic polypeptides after cleavage are shown in Fig 1E.

Expression and purification of fusion polypeptide

The expression vector construct may be introduced into an appropriate host cell through conventional methods such as the protoplast method or the competent cell method. Host cells such as Gram-positive bacteria belonging to the genus *Bacillus* and Gram-negative bacteria such as *Escherichia coil* may be used. Selected strains of host cells are inoculated in a medium containing an assimilable carbon source, a nitrogen source and essential nutrients, and are cultured through conventional fermentation methods.

In the present example, an overnight culture of the transformed *E. coli* strain BL21 (DE3) with the expression vector was diluted in a ratio of 1:25 into 100 mL of growth medium (LB medium containing 100 ug/mL of ampicillin). The culture was incubated at 37°C until its optical density at 600nm was 0.6. The polypeptide expression was induced by the addition of 1 mM isopropyl-1-thio-p-D-galactoside (IPTG) for 3 h.

Collection and purification of a target polypeptide from the thus-obtained culture broth can be performed according to conventional methods applicable to the collection and purification of common proteins. For example, cells are separated from the culture broth by centrifugation or filtration, and the target polypeptide can be obtained from the supernatant through conventional purification procedures, examples of which include: fractionation on an ion-exchange column; ethanol precipitation; affinity chromatography; ultracentrifugation; reverse phase HPLC; chromatography on silica or on a cation-exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex G-75; protein A Sepharose columns to remove contaminants such as IgG; and metal chelating columns to bind polyhistidine-tagged forms of polypeptides. Various methods of protein purification may be employed and such methods are known in the art and described for example in Deutscher, *Methods in Enzymology,* 182 (1990) and Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag, New York (1982). The purification step(s) selected will depend, for example, on the nature of the production process used and the particular polypeptide produced.

The transformed cells, after lysis, and fractionation into the soluble and the insoluble fractions, can be analyzed to determine the solubility of the fusion polypeptide.

When the fusion polypeptide is produced in an insoluble form, it is present in the form of inclusion bodies. The insoluble pellet which contains the inclusion bodies is washed in a buffer containing sodium phosphate, DPT, and Triton X-100. The contaminating components are thereby removed, leaving the inclusion bodies containing the target polypeptide in the insoluble pellet.

If the fusion polypeptide is present in the soluble fraction (aggregated or non-aggregated form), then the full-length polypeptide can be purified using affinity column chromatography, depending on the type of purification tag in the fusion partner fused to the target polypeptide.

In the present example, the cell pellet was harvested and resuspended in PBS buffer. After sonication, the resulting cell lysate was cleared of bacterial debris by ultracentrifugation. The fusion polypeptide with a His_6-tagged thioredoxin was absorbed onto a Nickel-charged metal affinity column made from chelating sepharose (Qiagen). The polypeptide was eluted from the column using 300 mM of imidazole. The purity and quantity of the fusion polypeptide was examined by SDS-PAGE and using a BCA (bicinchoninic acid) protein assay kit (Pierce). For the SDS-PAGE analysis, the fusion polypeptide was electrophoresed on a 12% SDS-polyacrylamide gel and quantitatively analyzed using a BioRad GS-800 scanning densitometer.

Release of the target polypeptide upon enterokinase cleavage

The cleavage of purified fusion polypeptides with enzymatic or chemical cleavage reagents may be conducted by dissolving the fusion polypeptide, e.g., in a suitable buffer such as Tris buffer or phosphate buffer (e.g., especially at a pH of 6-9), depending on the solubility properties of the particular fusion polypeptide involved. The cleavage reagent(s) and oxidant are added to the solution and incubation ensues at a temperature and time optimized, e.g., by routine, conventional experiments, generally 25°C-40°C, for 0.5-10 hours.

To remove imidazole, the fusion polypeptide preparation obtained from metal chelating affinity chromatography was subjected to buffer change to PBS buffer using Vivaspin ultrafiltration spin columns with a molecular weight cut-off of 30,000 Dalton (Vivascience). The resulting polypeptide solution was treated with enterokinase to release

the 29-mer synthetic polypeptide. For a standard cleavage reaction in this study, 40 μl of reactant solution containing 40 μg of the fusion polypeptide with 0.5 unit of enterokinase (Novagen) was incubated at 25 °C overnight. The efficiency of the cleavage reaction was monitored by SDS-PAGE analysis by monitoring the densitometric changes of the fusion polypeptide band at 22 kDa. The target polypeptide with a molecular weight of 2.9 kDa in size could not be visualized in conventional SDS-PAGE. However, it could easily be detected and analyzed with reversed-phase HPLC.

Purification of target polypeptide from enterokinase cleavage reactant solution using centrifugal ultrafiltration

The reactant solution after enterokinase cleavage contains several polypeptides including thioredoxin, enterokinase, the 29-mer target polypeptide along with a number of less abundant non-specific cleavage products. To purify the 29-mer target polypeptide, ultrafiltration spin columns with MWCO 10000 Dalton (Vivaspin 500 from Vivascience) was employed to retain proteins with a molecular weight larger than 10 kDa. The ultrafiltration was performed by centrifugation at 3000g for 30 minutes. The effectiveness of the separation of the target polypeptide from His_6-tagged thioredoxin was monitored using SDS-PAGE and HPLC analyses. The concentration of the protein solution obtained was examined using a BCA protein assay kit (Pierce).

Further purification and characterization of target polypeptides by reversed-phase HPLC

The technique of reversed-phase HPLC exploits the differences in hydrophobicity among molecules as a means of separation. Theoretically, peptide molecules with identical amino acid sequences, produced either synthetically or genetically, should generate the same separation profile in the HPLC chromatogram provided that identical conditions such as the elution gradient, column and buffer, are applied for the chromatography.

In the present example, reversed-phase high performance liquid chromatography (RP-HPLC) was carried out on a Shimadzu LC-10Avp series apparatus with a UV detector and a VP-ODS CIS column (250.0 x 4.6 mm, 5 μm, 300A). A gradient was achieved using buffer A (containing 0.05% trifluoroacetic acid in water) and buffer B (containing 80% acetonitrile and 0.05% trifluoroacetic acid in water). The flow rate was 1 mL/min with

buffer B being increased from 0% to 90% in 40 minutes. The chromatograms were recorded at a wavelength of 220 nm and the data analysed using LCsolution Version 1.0 (Shimadzu). The collected fractions were either used directly for mass spectrometric (MS) analysis or dried on a Speedvac concentrator (Braun) and reconstituted in PBS buffer for other activity assays.

Mass spectrometric analysis of target polypeptide

Mass spectrometric analysis was carried out using Shimadzu LCMS-2010 A/EV and Axima-CFR*Plus* (MALDI-TOF) systems. Separation in the LCMS was performed on a VP-ODS column (250.0 x 2.0 mm), in a 40°C oven, with a mobile phase of 70% acetonitrile in water containing 0.1% formic acid having a flow rate of 0.2mL/min in isocratic mode. The target polypeptide was reconstituted for MS analysis in the same PBS buffer used in the enterokinase cleavage reaction. The LCMS interface conditions were as follows: interface and operation mode: ESI, scan, m/z 600~1600; probe temperature: room temperature; CDL temperature: 250°C (desolvation); block temperature: 200°C; nebulizing gas flow: N_2, 1.5 L/min; drying gas:N_2, 10 L/min. The MALDI-TOF conditions were as follows: matrix: CHCA, 10 mg/mL in MeCN/Water (1:1) with 0.05% TFA; analysis conditions: positive reflectron mode.

Anti-dengue protease activity assays

One of the following three protocols were used to assay efficacy of the novel molecules in inhibition of the dengue virus NS3 protease:

(a) Protocol using a commercially available protease assay kit

10 µl of a stock solution of a positive control for the protease was mixed with 90 µl of distilled water. Reaction mixtures were prepared according to Table 2 below.

Table 2. Reaction mixtures for protease assay using a commercially available kit (without inhibitors)

Reagents	Blank	Positive control	Negative control (PBS)	Negative control (Thioredoxin)	Test reaction mixture
FTC-Casein	50µL	50µL	50µL	50µL	50µL

Incubation buffer	50µL	50µL	50µL	50µL	50µL
Deionized water	100µL	-	-	-	-
Enzyme (NS2B/NS3 or NS3)	-	-	-	-	100µL
PBS	-	-	100µL	-	-
Thioredoxin	-	-	-	100µL	-
Control protease	-	100µL	-	-	-
Total	200µL	200µL	200µL	200µL	200µL

The mixtures were incubated at 37°C for 1 hr to overnight, after which the reactions were terminated by adding 500 µl of 5% TCA. The mixtures were vortexed briefly and incubated at 37°C for a further 10 mins. The TCA precipitate was removed by centrifugation at 12,000g for 5 minutes while 400 µl of the supernatant was mixed with 600 µl of assay buffer. The absorbance was read against the reagent blank using a UV-spectrometer at 492 nm or a Fluorometer with excitation wavelengths of 490 nm and emission wavelengths of 525 nm.

(b) BAPNA protocol (Yosof et al., (2000) J. Biol. Chem. 275(14):9963-9969)

The buffer was prepared by mixing 20 mL of 100 mM Tris (pH 8.0) with 100 mL of 600 mM sodium chloride and the volume made up to 200 mL with distilled water. A 10 mM solution of BAPNA and a 10% v/v positive control for the protease was prepared. Initial concentrations of 100µg/mL and 65.5µg/mL for the NS2B/NS3 and NS3 proteases, respectively, were concentrated using a Vivaspin column to 0.1 mg/mL and 0.0655 mg/mL, respectively. Reaction mixtures were prepared according to Tables 3 and 4 below.

Table 3. Reaction mixtures for protease assay using the BAPNA protocol (without inhibitors)

Reagents	Blank	Standard (NS2B/ NS3 or NS3)	Positive Control (Protease)	Negative Control (PBS)

2x Buffer (100mM Tris pH8.0 and 600mM NaCl)	50µL	50µL	50µL	50µL
50% Glycerol	20µL	20µL	20µL	20µL
Distilled Water	20µL	10µL	10µL	10µL
10mM BAPNA	10µL	10µL	10µL	10µL
Enzyme	-	10µL (NS2B/ NS3 or NS3)	10µL Protease	10µL PBS
Total	100µL	100µL	100µL	100µL

Table 4. Reaction mixtures for protease assay using the BAPNA protocol (with inhibitors)

Reagents	Blank	Standard (NS2B/ NS3 or NS3)	Positive Control (protease)	Negative Control (thioredoxin)	Negative Control (PBS)
2x Buffer (100mM Tris pH8.0 and 600mM NaCl)	50µL	50µL	50µL	50µL	50µL
50% Glycerol	20µL	20µL	20µL	20µL	20µL
Distilled Water	20µL	-	10µL	10µL	10µL
Inhibitor (either of mutant 1a,2b,3d,4d,5b,or 6c)	-	10µL	-	-	-
10mM BAPNA	10µL	10µL	10µL	10µL	10µL
Enzyme	-	10µL (NS2B/ NS3 or NS3)	10µL (protease)	10µL (thioredoxin)	10µL PBS
Total	100µL	100µL	100µL	100µL	100µL

The reaction mixtures were incubated at 23°C for two time intervals of 3 hrs and 24 hrs before measuring the absorbance at 405 nm.

(c) Protease assay protocol developed by the inventors

One of the six mutants described above is first incubated with a thioredoxin-NS2B/NS3 fusion polypeptide to enable binding of the mutant molecule to the active site of the protease in the NS2B/NS3 region of the fusion polypeptide and subsequent inhibition of the NS2B/NS3 protease activity. If the inhibition is specific and only targets the NS2B/NS3 protease activity, it should lead to the inhibition of the autolytic cleavage of NS2B/NS3 and should not interfere or inhibit the subsequent enterokinase cleavage of NS2B/NS3 from the full-length thioredoxin-NS2B/NS3 fusion polypeptide. Effectiveness and specificity of the mutant molecule can be evaluated by comparing the profile and level of inhibition of the autocatalytic activity of NS2B/NS3 and the proteolytic activity of enterokinase.

To carry out such an assay, 2 µl of the mutant molecule la at a concentration of 0.1 mg/mL was mixed with 1 µl (1 mg/mL) of thioredoxin-NS2B/NS3 fusion polypeptide and 1 µl of 10 × enterokinase reaction buffer (Novagen). 5 µl of water was added before all the components were fully mixed. The reaction mixture was incubated at 37°C for 30 mins. After incubation, 0.3 unit of enterokinase (Novagen) in 1 µl of water was added to the reaction mixture, which was subsequently incubated at 25°C. At a time interval of 1 and 2 hours, 4 µl of reaction mixture was withdrawn and assayed on SDS-PAGE. The profile of products including the percentage of uncleaved NS2B/NS3 was monitored by measuring the densitometric changes of the uncleaved NS2B/NS3 at 20-25 kDa on SDS-PAGE.

Results and Discussion

Back translation of kalata B1 sequence

Based on back translation tools, synthetic codon-optimised cDNA sequences encoding 29-mer open cyclic mutants of the cysteine-rich plant cyclotide, kalata B1 were designed.

Expression and purification of fusion polypeptide

The cDNA sequence encoding the 29-mer open cyclic permutant of kalata B1 was synthesized and fused in-frame to His_6-tagged thioredoxin in the bacterial expression vector, pET32a. Upon induction with IPTG, the fusion polypeptide was overexpressed in the bacterial host, *E. coli* strain BL21 (DE3). The fusion polypeptide was secreted into the cytoplasm as a soluble protein with an apparent molecular weight of 22 kDa based on SDS-PAGE analysis (Fig. 2). Scanning densitometric analysis revealed that the

fusion polypeptide was overexpressed at a level that accounts for approximately 30% of total cellular proteins. Up to 20 mg of soluble fusion polypeptide could be obtained from 1 litre of bacterial culture. Optimal expression of the fusion polypeptide was achieved when the culture was carried out at 37°C and the induction was performed with IPTG at a concentration of 1 mM.

To purify the fusion polypeptide, a metal chelating Sepharose column was employed to remove cellular proteins while the absorbed His_6-tagged thioredoxin-open cyclic permutant of kalata Bl fusion polypeptide was eluted from the column using imidazole (300mM). The purity and identity of the fusion polypeptide was assayed by polyacrylamide gel electrophoresis and staining using Pierce GelCode blue stain reagent. The yield and concentration of the fusion polypeptide solution were determined using the BCA protein assay. Typically, a fusion polypeptide solution with a purity of over 95% and a final concentration of 2 to 3 mg/mL in PBS buffer can be obtained using Vivaspin ultrafiltration columns for concentration and buffer exchange.

It has therefore been demonstrated that cysteine-rich polypeptides can be efficiently produced with high yields in a bacterial host using an optimized cDNA sequence.

Enterokinase cleavage of the fusion polypeptide

To facilitate the purification of the target polypeptide from the His_6-tagged thioredoxin fusion partner, an enterokinase recognition sequence was introduced immediately upstream of the target polypeptide, allowing the latter to be released without leaving any unwanted residues upon treatment with enterokinase. To achieve maximal release of the target polypeptide, the cleavage reaction is typically performed overnight at 25°C. The cleavage reaction was monitored by SDS-PAGE analysis by observing the disappearance of the fusion polypeptide band at 22 kDa (Fig. 3). The target polypeptide with molecular weight of 2.9 kDa in size is not detectable in conventional SDS-PAGE analysis. However, it can be easily monitored on reversed-phase HPLC as described below.

Purification of target polypeptide from enterokinase cleavage reactant solution by centrifugal ultrafiltration

The reactant solution from enterokinase cleavage contains mainly four peptide entities including the enterokinase, the remnant uncleaved fusion polypeptide, the His_6-tagged thioredoxin fusion partner and the target polypeptide, the open cyclic permutant of kalata B1,

which have molecular weights of 26, 22, 19 and 2.9 kDa, respectively. The molecular weight of the target polypeptide, the open cyclic permutant of kalata B1, is at least 16 kDa smaller than that of the other three peptide entities in the reactant solution, thus facilitating its separation by centrifugal ultrafiltration using a semipermeable membrane. At the molecular weight cut-off of 10 kDa, molecules of the open cyclic permutant of kalata B1 could readily filter through the membrane while other peptide entities were retained on the membrane. This resulted in a filtrate that mainly comprised the target polypeptide. The separation was fast and efficient as evidenced by the subsequent reversed-phase HPLC analysis (Fig. 4) which shows that the filtrate from ultrafiltration mainly comprised the target polypeptide, the open cyclic permutant of kalata B1, eluting at about 30 minutes.

Further purification and characterisation of target polypeptide by reversed-phase HPLC

The technique of reversed-phase HPLC exploits the differences in hydrophobicity among molecules as a means of separation. Theoretically, peptide molecules with identical amino acid sequences, produced either synthetically or genetically, should generate the same separation profile in the HPLC chromatogram provided that identical conditions such as the elution gradient, column and buffer, are applied for the chromatography. In the present experiments with the open cyclic permutant of kalata B1, the molecule produced chemically via solid phase peptide synthesis showed a shorter retention time when eluting through the VP-ODS CIS column compared to its genetically-produced version. The polypeptide solution collected from HPLC fractionation was directly lyophilized and reconstituted in PBS buffer for further LC-MS and MALDI-TOF analyses.

Mass spectrometric analysis of target polypeptide

LC-MS studies revealed that the major component in the polypeptide solution obtained from reversed-phase HPLC separation has a decoverlution MW of 2909.2 as measured in a single-quadrupole mode (Fig. 5). This is in good agreement with the theoretical molecular weight of the open cyclic permutant of kalata B1 in its fully oxidized form, which is 2910. The difference in the measured and theoretical values is probably due to differences in the ionization state of the two carboxylic groups, one at the side chain of glutamic acid and the other at the carboxylic terminus of the polypeptide. The total ion current chromatogram displayed several peaks at m/z 1455.62, 970.87 and 866.49, respectively. The MS-MS analysis of the polypeptide solution demonstrated that the open cyclic

permutant of kalata B1 has the correct polypeptide sequence as shown in Fig. 1(C). LC-MS analysis also revealed that the synthetic version of the open cyclic permutant of kalata B1 has a molecular weight of 2916, which agrees well with its theoretical value. The absence of fragments in PDS fragmentation of MS indicates that the molecule has a strong structure capable of withstanding electronic bombardment, a condition under which most polypeptides would be fragmented.

Furthermore, the production of a fully oxidized form of a small polypeptide from expression of a thioredoxin fusion polypeptide is highly unusual since the fusion partner, the thioredoxin, normally functions as a disulphide reductant under physiologically neutral conditions such as those in a PBS buffer. As an oxidoreductase, thioredoxin has a highly conserved dithiol motif (-Cys-Gly-Pro-Cys-) as active site for regulating redox reactions. It normally functions as a reducing agent for reduction of disulphide bonds to yield thiol groups. The formation of three disulphide bonds in the open cyclic permutant of kalata B1 in the presence of thioredoxin therefore indicates the presence of strong antioxidant activity rendered by the six cysteine residues forming the three disulphide bonds. This indicates that kalata B1 may be expressed in response to oxidative stress and may be actively involved in antioxidant defence in plants.

The LCMS and MALDI-TOF data for the six mutants of the open cyclic permutant of kalata B1 are shown in Fig. 6. Mutant 1a has a molecular weight of 2954.3 (Fig. 6A); mutant 2b a molecular weight of 2913.7 (Fig. 6B); mutant 3d a molecular weight of 2911.5 (Fig. 6C); mutant 4d a molecular weight of 2909.3 (Fig.6D); mutant 6c a molecular weight of 2775.8 (Fig. 6E). It can be seen that the molecular weights of mutants 2b, 3d and 4d correspond closely to the theoretical molecular weight of the open cyclic permutant of kalata B1 in its fully oxidized form (2910) indicating close structural similarities.

Further molecular modelling studies showed that the synthetic open cyclic permutant of kalata B1 adopts an open circular structure folded in a way similar to its natural cyclic counterpart. Force field calculation using the SYBYL program (Tripose) was carried out to calculate the geometric distances between each amino acid in the polypeptide chain. It was found that the distance between the N- and C-terminus of the open cyclic permutant of kalata B1 is 5.24 Å. The close proximity of the two termini reflects the compact nature of the polypeptide.

Anti-dengue protease activity assay

Fig. 7 shows typical results of an activity inhibition assay of the synthetic mutant 1a on dengue NS2B/NS3 protease activity. The intensity of the band for the uncleaved NS2B/NS3 polypeptide at 20-25 kDa was higher in the presence of the mutant 1a (Lane 5, Fig. 7) or the open cyclic permutant of kalata B1 (Lane 4, Fig. 7); hence, autocleavage of the NS2B/NS3 polypeptide was inhibited by the mutant 1a and the open cyclic permutant of kalata B1. In contrast, the intensity of the band for the uncleaved NS2B/NS3 polypeptide at 20-25 kDa in the control reactions (with water or PBS buffer) was lower due to autocatalytic cleavage of the NS2B/NS3 polypeptide in the absence of the inhibitor (i.e. the mutant 1a or the open cyclic permutant of kalata B1).

The corresponding quantitative data obtained via densitometry is shown in Table 5. With water as control, the amount of uncleaved NS2B/NS3 polypeptide remaining after treatment with mutant 1a increased from 16.1% to 18.7% relative to the total thioredoxin-NS2B/NS3 fusion polypeptide added in the reaction. This represented a 16.1% increase in the formation of the NS2B/NS3 polypeptide. The increased yield of the uncleaved NS2B/NS3 polypeptide is the direct result of the inhibition of the autocatalytic protease activity of the NS2B/NS3 by the synthetic mutant 1a.

Table 5. Densitometric analysis of the protein bands in the SDS-PAGE shown in Figure 7.

Lane Number	Band Number	Mol Wt KDa	Lane No	Average OD	Trace OD x mm	Relative abundance	Average percentage
1	1	220		0.971	1.234	16.278	
1	3	97		0.643	1.429	18.859	
1	4	66		0.569	1.373	18.121	
1	5	45		0.337	0.941	12.422	
1	6	30		0.421	1.178	15.539	
1	7	14.3		0.432	1.289	17.015	
1	8	20.1		0.192	0.134	1.767	
2	1	183.055	#2	0.201	0.115	19.043	
2	2	49.791	#2	0.175	0.489	80.957	
3	1	49.076	#3	0.351	1.049	62.887	
3	2	35.472	#3	0.193	0.196	11.754	
3	3	33.96	#3	0.151	0.096	5.747	**17.5**
3	4	17.855	#3	0.133	0.228	13.664	

3	5	19.341	#3	0.065	0.099	5.948	
4	1	49.076	#4	0.359	1.048	66.228	
4	2	35.591	#4	0.177	0.157	9.944	
4	3	34.074	#4	0.13	0.091	5.72	**15.6**
4	4	17.855	#4	0.106	0.216	13.637	
4	5	19.284	#4	0.062	0.071	4.471	
5	1	48.372	#5	0.289	0.807	62.698	
5	2	35.117	#5	0.113	0.15	11.66	
5	3	33.508	#5	0.094	0.09	6.973	**18.7**
5	4	17.386	#5	0.08	0.172	13.401	
5	5	18.833	#5	0.051	0.068	5.268	
6	1	47.678	#6	0.315	0.961	66.145	
6	2	34.883	#6	0.162	0.144	9.936	
6	3	33.284	#6	0.118	0.09	6.193	**16.1**
6	4	17.283	#6	0.105	0.174	11.967	
6	5	18.721	#6	0.055	0.084	5.758	

Applications

Advantageously, the disclosed 29-mer open cyclic permutant of kalata B1 or its mutants can be overexpressed using a synthetic codon-optimised cDNA sequence in a fusion polypeptide and readily purified by affinity chromatography using metal chelating sepharose. Subsequent cleavage of the fusion polypeptide yields the 29-mer polypeptide that can then be readily purified from the enterokinase cleavage reactant solution, followed by centrifugal ultrafiltration through a semipermeable membrane. Further purification of the 29-mer polypeptide can be readily achieved by applying reversed-phase high performance liquid chromatography. Accordingly, open cyclic mutants of naturally occurring cyclic polypeptides, in fully oxidized and functional form, can be readily produced and purified cost-effectively and in high yield, which are important criteria for use of such molecules in pharmaceutical applications such as in anti-viral treatment.

Equally advantageous, the anti-dengue activity of the open cyclic mutants possess higher levels of inhibitory activity on the dengue NS3 protease and therefore provides a higher level of potency when used in anti-viral treatment. Levels of such activity can furthermore be easily and specifically assayed quantitatively using the disclosed protease activity assay.

It will be apparent that various other modifications and adaptations of the invention will be apparent to the person skilled in the art after reading the foregoing disclosure

without departing from the spirit and scope of the invention and it is intended that all such modifications and adaptations come within the scope of the appended claims.

CLAIMS

1. A polypeptide comprising the following sequence of amino acids or analogues thereof:

X_1-Leu-Pro-Val-Cys-Gly-Glu-Thr-Cys-Val-Gly-Gly-Thr-X_2-Asn-Thr-Pro-Gly-X_3-Thr-X_4-Ser-Trp-Pro-Val-Cys-Thr-X_5-X_6

wherein
X_1 is Ala, Asp, Gly, or Ser
X_2 is Cys or Gly
X_3 is Cys or Gly
X_4 is Cys or Gly
X_5 is Arg, Gln or Lys
X_6 is Arg or Lys

2. A polypeptide comprising the following sequence of amino acids or analogues thereof:

X_1-Leu-Pro-Val-Cys-Gly-Glu-Thr-Cys-Val-Gly-Gly-Thr-Gly-Asn-Thr-Pro-Gly-X_2-Thr-X_3-Ser-Trp-Pro-Val-Cys-Thr-X_4-Asn

wherein
X_1 is Ala, Asp, Gly, or Ser
X_2 is Cys or Gly
X_3 is Cys or Gly
X_4 is Arg, Gln or Lys

3. A polypeptide comprising an open cyclic backbone, said open cyclic backbone comprising the following sequence of amino acids or analogues thereof:

X_1-Leu-Pro-Val-Cys-Gly-Glu-Thr-Cys-Val-Gly-Gly-Thr-X_2-Asn-Thr-Pro-Gly-X_3-Thr-X_4-Ser-Trp-Pro-Val-Cys-Thr-X_5-X_6

wherein

X_1 is Ala, Asp, Gly, or Ser

X_2 is Cys or Gly

X_3 is Cys or Gly

X_4 is Cys or Gly

X_5 is Arg, Gln or Lys

X_6 is Arg or Lys

4. A polypeptide comprising an open cyclic backbone, said open cyclic backbone comprising the following sequence of amino acids or analogues thereof:

X_1-Leu-Pro-Val-Cys-Gly-Glu-Thr-Cys-Val-Gly-Gly-Thr-Gly-Asn-Thr-Pro-Gly-X_2-Thr-X_3-Ser-Trp-Pro-Val-Cys-Thr-X_4-Asn

wherein

X_1 is Ala, Asp, Gly, or Ser

X_2 is Cys or Gly

X_3 is Cys or Gly

X_4 is Arg, Gln or Lys

5. A polypeptide according to any one of claims 1 or 3 comprising the sequence set forth in SEQ ID NO. 7.

6. A polypeptide according to any one of claims 1 or 3 comprising the sequence set forth in SEQ ID NO. 8.

7. A polypeptide according to any one of claims 1 or 3 comprising the sequence set forth in SEQ ID NO. 9.

8. A polypeptide according to any one of claims 1 or 3 comprising the sequence set forth in SEQ ID NO. 10.

9. A polypeptide according to any one of claims 1 or 3 comprising the sequence set forth in SEQ ID NO. 11.

10. A polypeptide according to any one of claims 2 or 4 comprising the sequence set forth in SEQ ID NO. 12.

11. A polypeptide according to any one of claims 3 to 10, wherein said open cyclic backbone comprises at least one disulphide bond.

12. A polypeptide according to any one of claims 3 to 10, wherein said open cyclic backbone is in the form of a cystine knot.

13. A polypeptide comprising the following sequence of amino acids or analogues thereof:

X_1-Leu-Pro-Val-Cys-Gly-Glu-Thr-Cys-Val-Gly-Gly-Thr-X_2-Asn-Thr-Pro-Gly-X_3-Thr-X_4-Ser-Trp-Pro-Val-Cys-Thr-X_5-X_6

wherein
X_1 is Ala, Asp, Gly, or Ser
X_2 is Cys or Gly
X_3 is Cys or Gly
X_4 is Cys or Gly
X_5 is Arg, Gln or Lys
X_6 is Arg or Lys, wherein said polypeptide is capable of inhibiting viral NS3 protease.

14. A polypeptide comprising an open cyclic backbone, said open cyclic backbone comprising the following sequence of amino acids or analogues thereof:

X_1-Leu-Pro-Val-Cys-Gly-Glu-Thr-Cys-Val-Gly-Gly-Thr-Gly-Asn-Thr-Pro-Gly-X_2-Thr-X_3-Ser-Trp-Pro-Val-Cys-Thr-X_4-Asn

wherein

X_1 is Ala, Asp, Gly, or Ser

X_2 is Cys or Gly

X_3 is Cys or Gly

X_4 is Arg, Gln or Lys, wherein said polypeptide is capable of inhibiting viral NS3 protease.

15. A polypeptide comprising a functional equivalent of a polypeptide according to any one of claims 1 to 14, wherein said functional equivalent has at least 60% sequence identity with a polypeptide selected from the group consisting of SEQ ID NOS: 7, 8, 9, 10, 11 and 12, and wherein said functional equivalent retains the inhibitory activity of the reference polypeptide.

16. The polypeptide according to claim 15, wherein the sequence identity is at least 70%, 80% or 90%.

17. A polynucleotide comprising a sequence of nucleotides or analogues thereof which encodes a polypeptide according to any one of claims 1 to 16.

18. A polynucleotide according to claim 17, wherein said polynucleotide is in the form of RNA or DNA.

19. A polynucleotide according to claim 17, wherein the sequence of nucleotides or analogues thereof comprises the sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6 and combinations thereof.

20. A recombinant vector comprising the polynucleotide according to any one of claims 17 to 19.

21. A host cell transformed with the recombinant vector according to claim 20.

22. A method of producing the polypeptide according to any one of claims 1 to 16, the method comprising the step of culturing a host cell according to claim 21 under conditions which permit expression of said polypeptide.

23. A composition for treatment or prophylaxis of a NS3 protease related condition in a subject, comprising the polypeptide according to any one of claims 1 to 16 and a pharmacologically acceptable carrier, excipient or diluent.

24. A composition according to claim 23, wherein the NS3 protease related condition is selected from the group consisting of dengue, Japanese encephalitis, Kunjin viral disease, Murray Valley encephalitis, yellow fever, West Nile disease and tick-borne encephalitis.

25. A composition according to any one of claims 23 or 24, wherein said subject is a human.

26. Use of the polypeptide according to any one of claims 1 to 16 in the manufacture of a medicament for the treatment or prophylaxis of a NS3 protease related condition.

27. Use according to claim 26, wherein the NS3 protease related condition is selected from the group consisting of dengue, Japanese encephalitis, Kunjin viral disease, Murray Valley encephalitis, yellow fever, West Nile disease and tick-borne encephalitis.

28. Use of the polypeptide according to any one of claims 1 to 16 in the study, treatment or prophylaxis of a NS3 protease related condition, particularly dengue.

29. A method for treatment or prophylaxis of a NS3 protease related viral disease, comprising administering to a patient in need of such treatment a therapeutically effective dose of a composition according to any one of claims 23 to 25.

30. A method for assaying the inhibitory activity of a polypeptide according to any one of claims 1 to 16 or a composition according to any one of claims 23 to 25, the method comprising the steps of:

a) incubating a mixture comprising a NS2B/NS3 fusion polypeptide containing a restriction protease cleavage recognition site and a buffer in the presence or absence of a polypeptide according to any one of claims 1 to 16 or a composition according to any one of claims 23 to 25;

b) adding to said mixture a restriction protease capable of cleaving said NS2B/NS3 fusion polypeptide at the restriction protease cleavage recognition site to release the NS2B/NS3 polypeptide containing a NS3 cleavage recognition site;

c) measuring the levels of uncleaved NS2B/NS3 polypeptide in the presence and absence of a polypeptide according to any one of claims 1 to 16 or a composition according to any one of claims 23 to 25; and

d) comparing the levels of uncleaved NS2B/NS3 polypeptide in the presence and absence of a polypeptide according to any one of claims 1 to 16 or a composition according to any one of claims 23 to 25.

31. A method according to claim 30, wherein the restriction protease is enterokinase.

32. A kit for assaying the inhibitory activity of the polypeptide according to any one of claims 1 to 16, wherein the kit comprises the polypeptide according to any one of claims 1 to 16.