



US 20050260758A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2005/0260758 A1**
Rasochova et al. (43) **Pub. Date: Nov. 24, 2005**

(54) **HIGH EFFICIENCY PEPTIDE PRODUCTION
IN PLANT CELLS**

Publication Classification

(76) Inventors: **Lada Rasochova**, San Diego, CA (US);
Philip Phuoc Dao, San Diego, CA (US)

(51) **Int. Cl.⁷** **C12Q 1/70**; C07H 21/04;
C12N 5/04
(52) **U.S. Cl.** **435/468**; 435/419; 536/23.2

Correspondence Address:
KING & SPALDING LLP
191 PEACHTREE STREET, N.E.
45TH FLOOR
ATLANTA, GA 30303-1763 (US)

(57) **ABSTRACT**

(21) Appl. No.: **11/069,601**

(22) Filed: **Feb. 28, 2005**

Related U.S. Application Data

(60) Provisional application No. 60/548,744, filed on Feb.
27, 2004.

The present invention provides an improved process for the production of recombinant peptides. In particular, the present invention provides an improved process for the production of recombinant peptides in the form of viral capsid fusion proteins which can be assembled in vivo in plant cell suspension cultures. The invention also includes plasmids, sequences, and plant cells which allow for non-infectious viral capsid fusion peptide production.

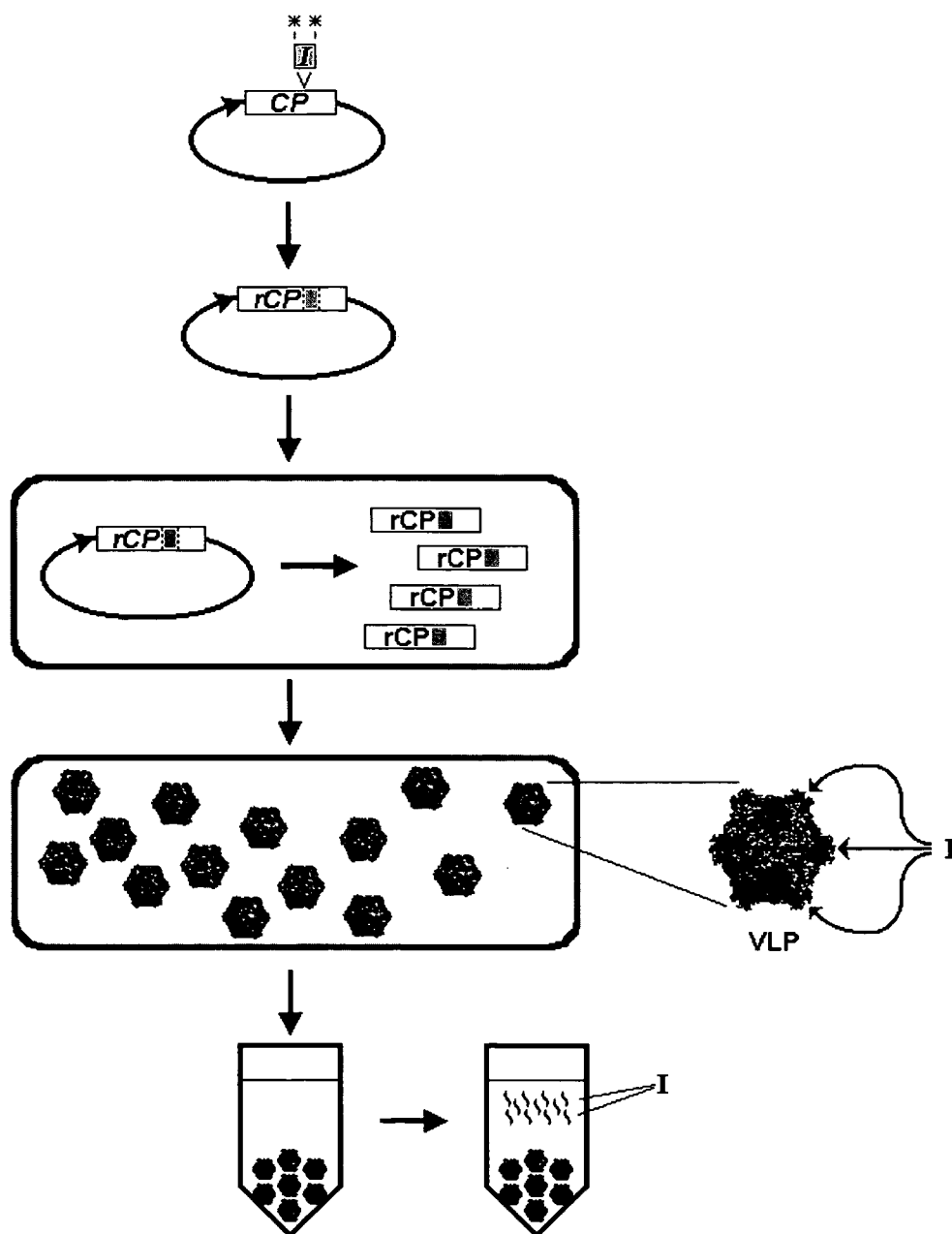


Figure 1

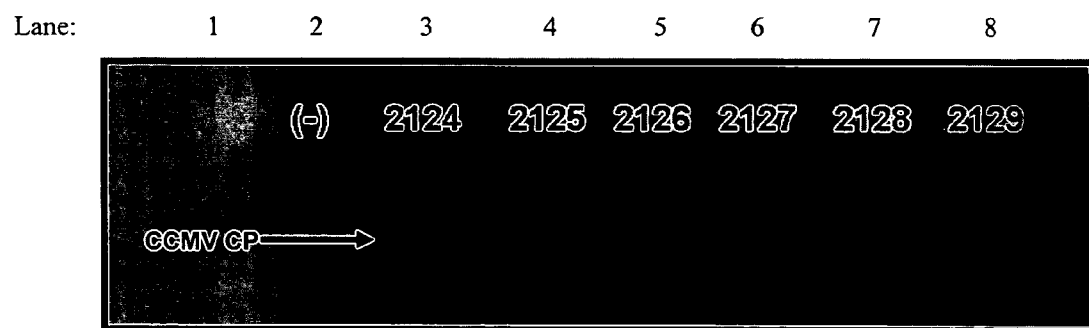


Figure 2

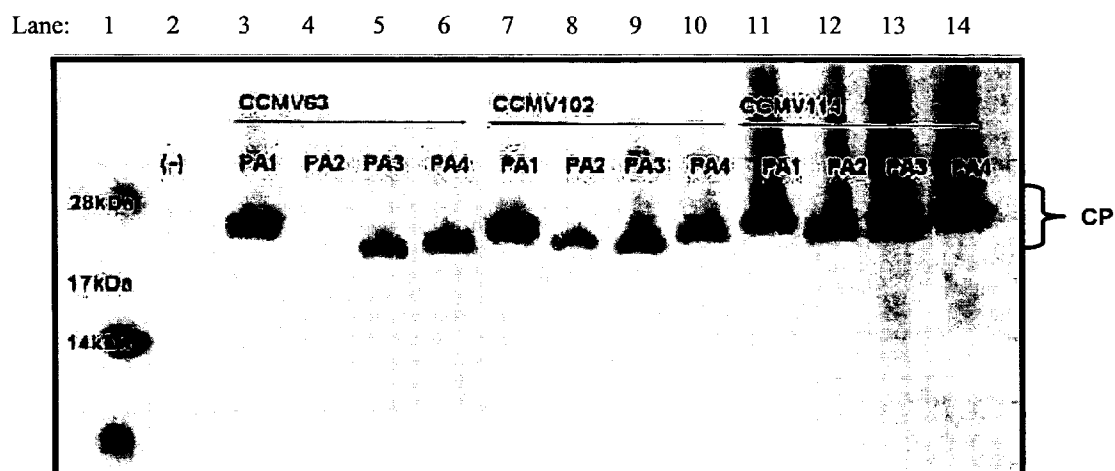


Figure 3

Lane: 1 2 3 4 5 6 7 8 9 10 11

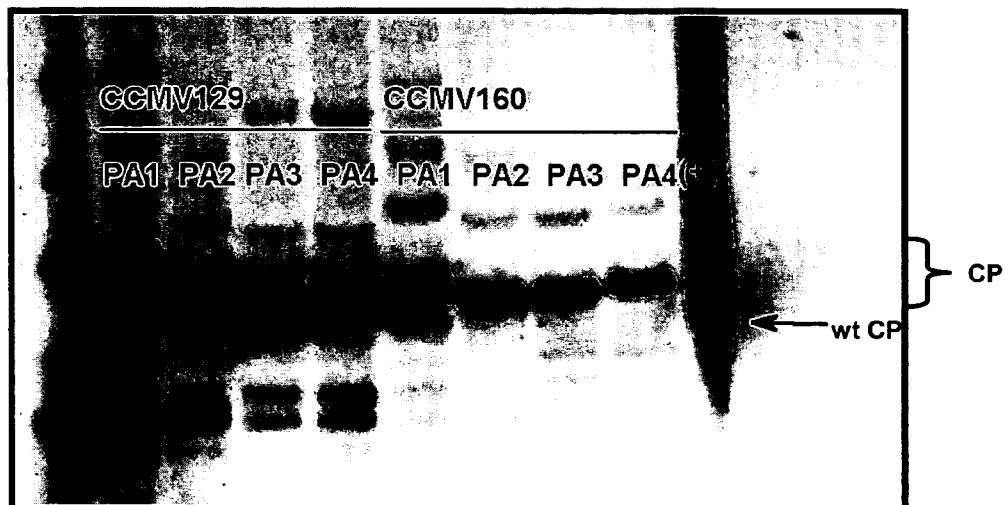


Figure 4

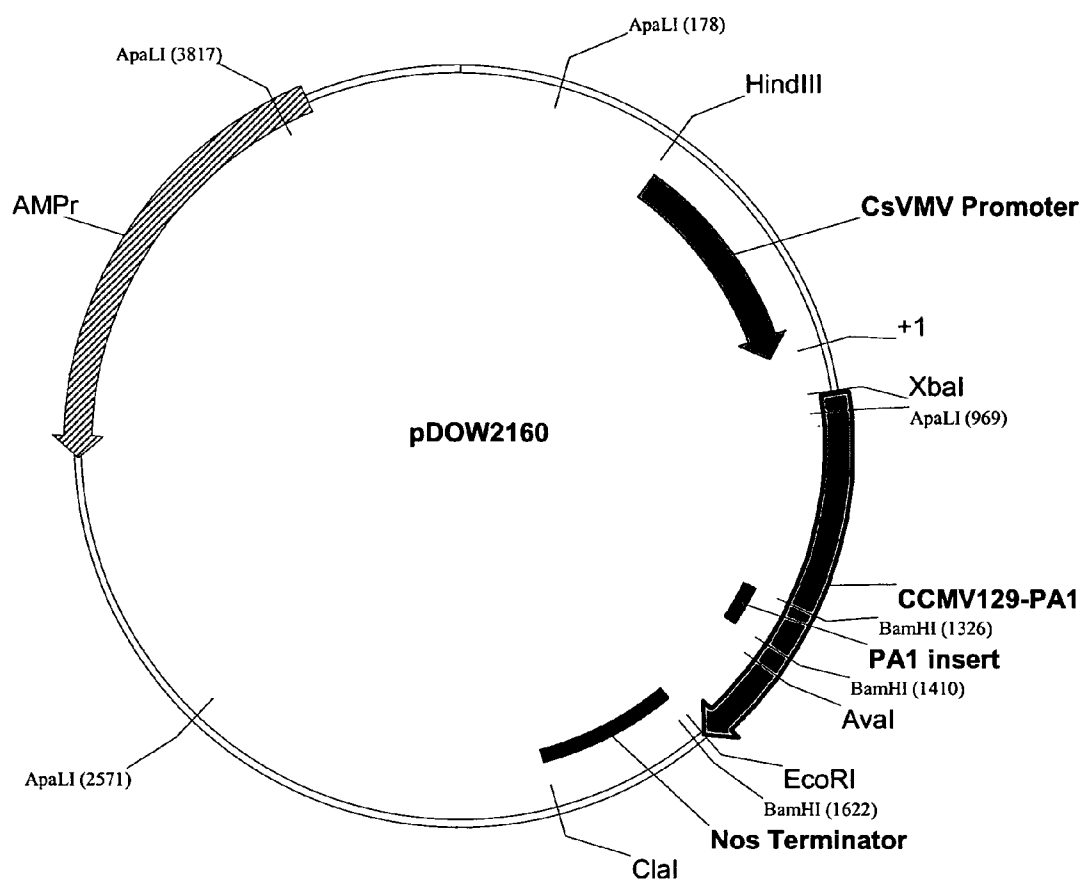


Figure 5

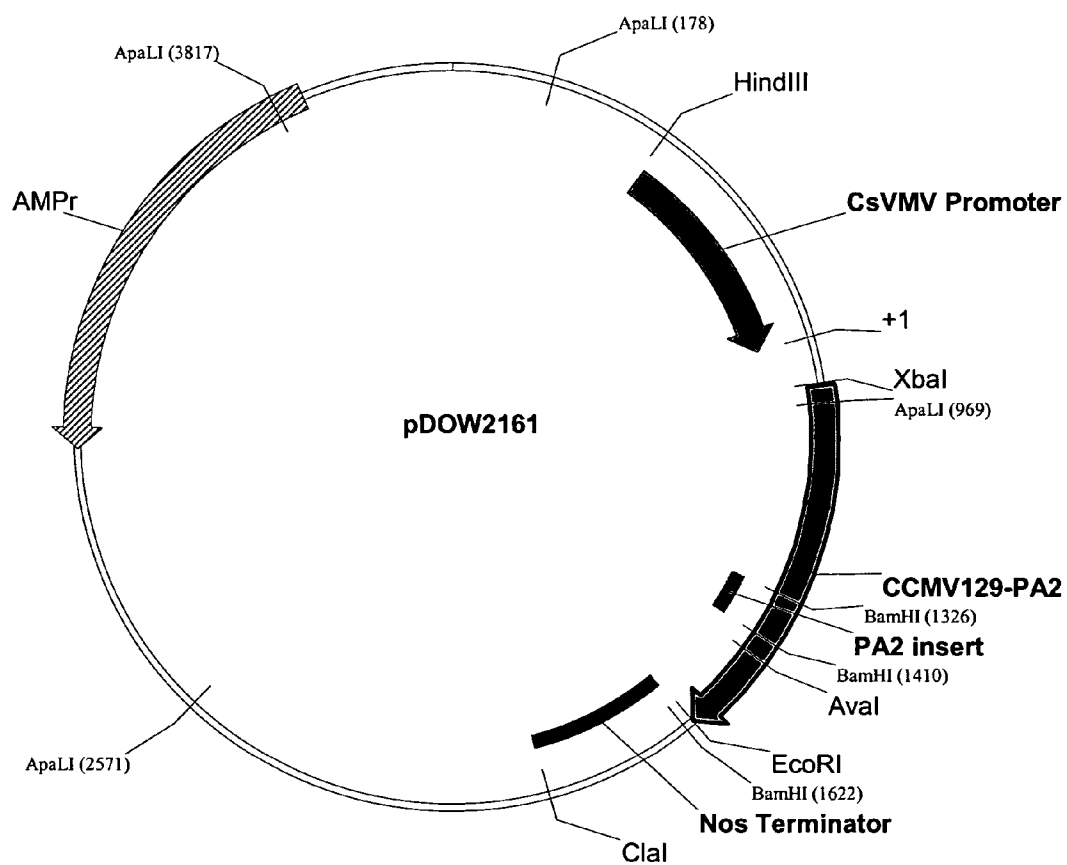


Figure 6.

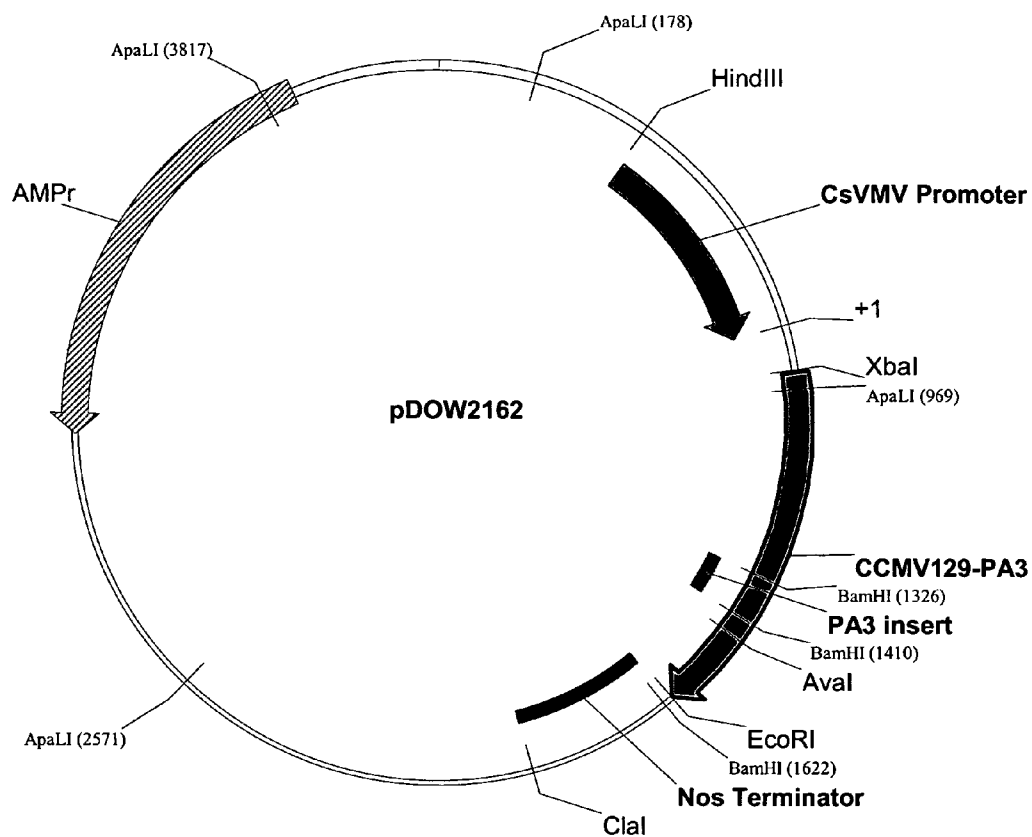


Figure 7

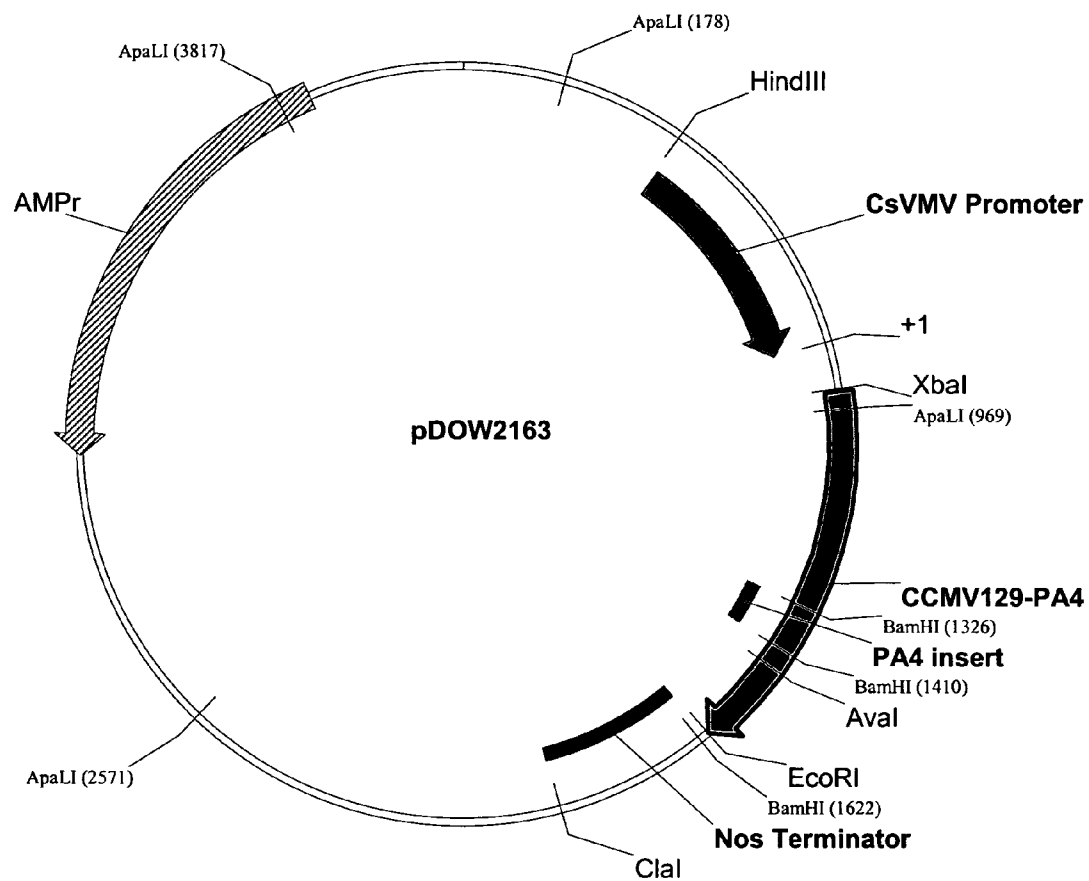


Figure 8

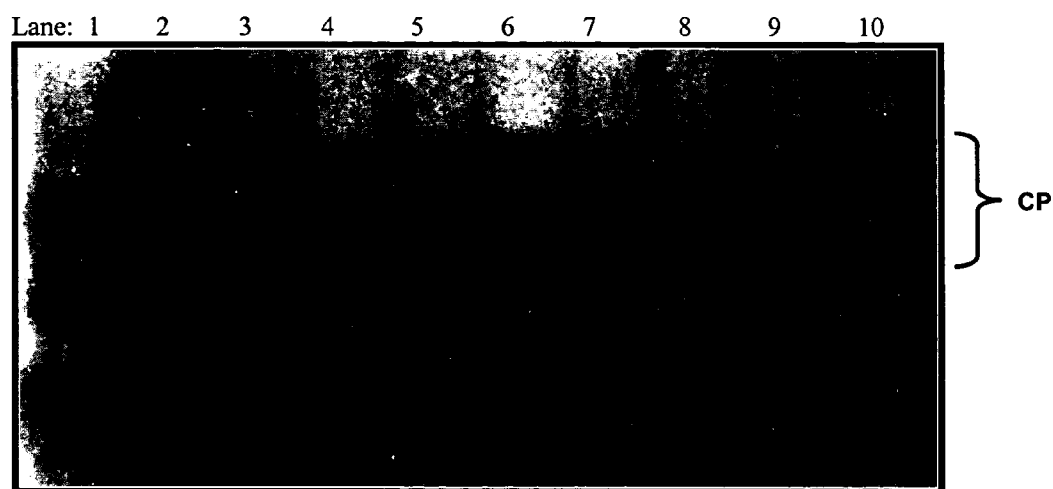


Figure 9

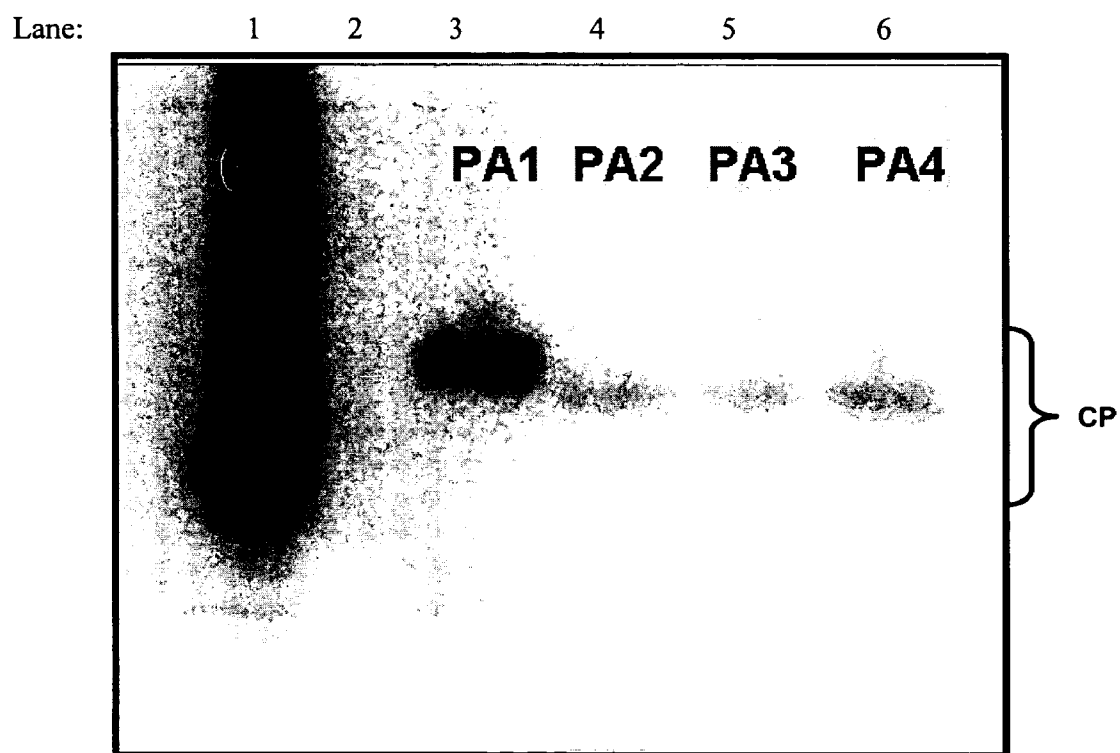


Figure 10

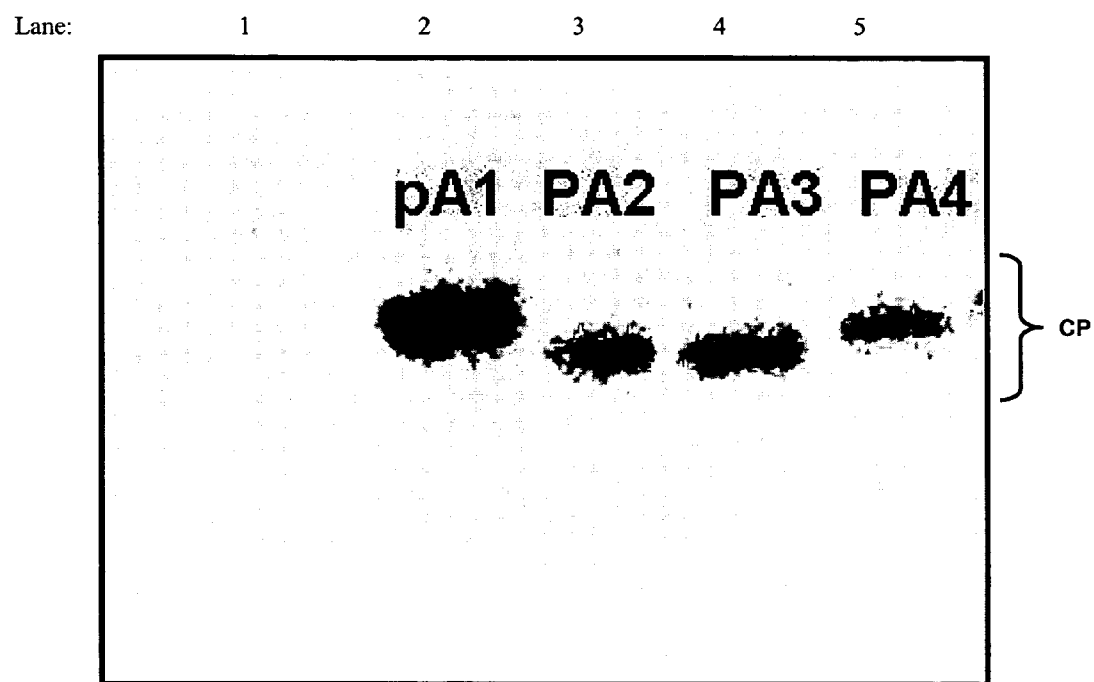


Figure 11

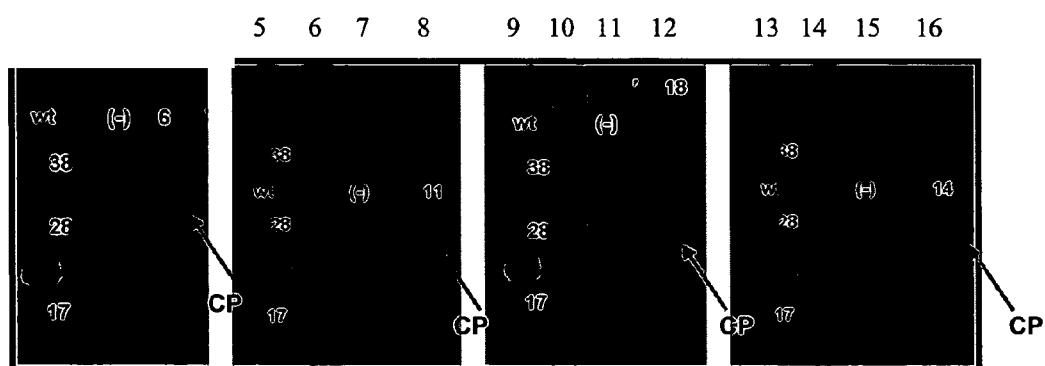


Figure 12

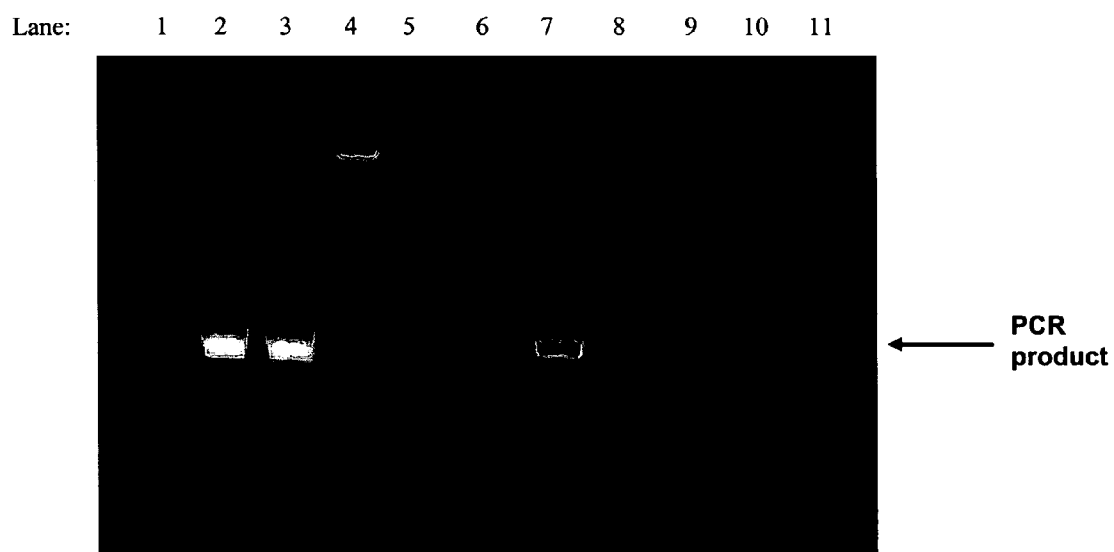


Figure 13

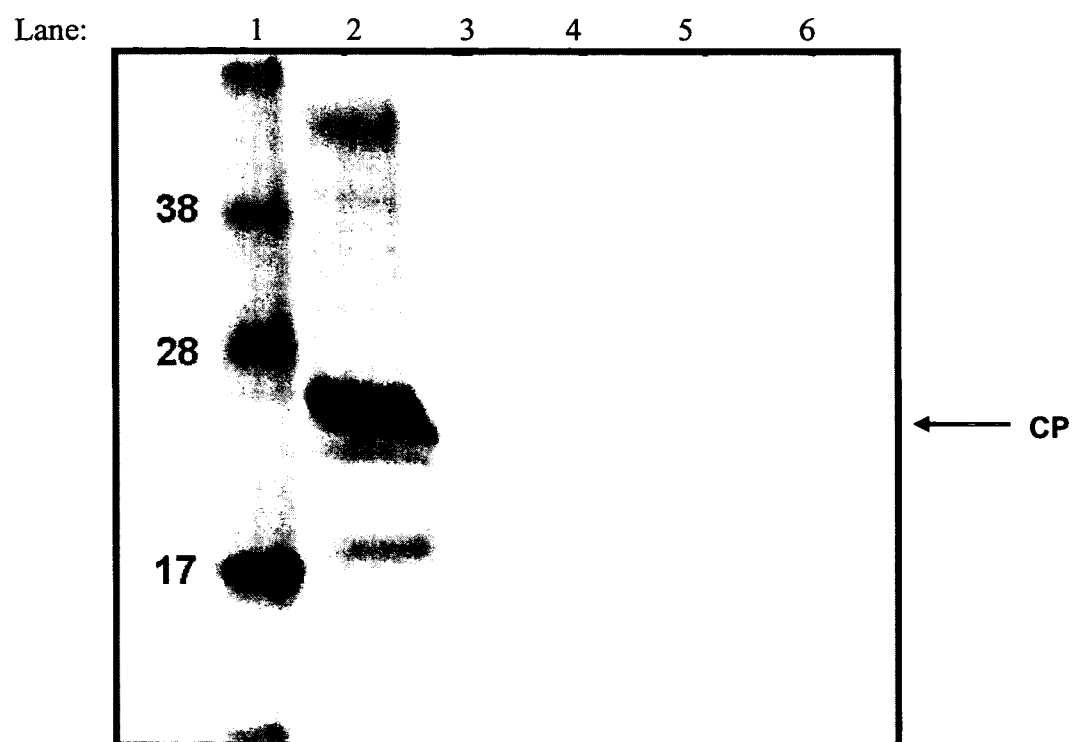
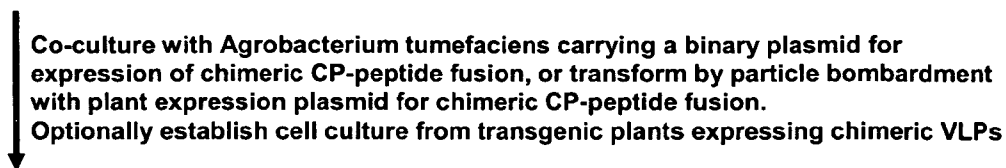


Figure 14

Plant cell culture



Establish transgenic cell culture expressing chimeric VLPs

Select high producing cell lines → **Cryo-preserve master and working cell banks**

Scale-up

Plant cell fermentation

Recover plant cells and homogenize (optionally recover cell supernatant if product secreted)

Purify chimeric VLPs (optionally cleave off the peptide and purify)

Formulate, Fill, QC, QA, Regulatory

Figure 15

HIGH EFFICIENCY PEPTIDE PRODUCTION IN PLANT CELLS

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to U.S. Provisional patent application Ser. No. 60/548,744, filed Feb. 27, 2004, entitled "High Efficiency Peptide Production in Plant Cells."

STATEMENT OF GOVERNMENT INTEREST

[0002] This application is under a United States Government contract with the National Institutes of Health, National Institute of Allergy and Infectious Disease (NIAID), Cooperative Agreement No. 1-U01-AI054641-01.

FIELD OF THE INVENTION

[0003] The present invention provides an improved process for the production of recombinant peptides. In particular, the present invention provides an improved process for the production of recombinant peptides in the form of viral capsid fusion proteins which can be assembled in vivo in plant cell suspension cultures. The invention also includes plasmids, sequences, and plant cells which allow for non-infectious viral capsid fusion peptide production.

BACKGROUND OF THE INVENTION

[0004] Bacterial, yeast, plant, insect, and mammalian cell expression systems are currently used to produce recombinant peptides, with varying degrees of success. One goal in creating expression systems for the production of heterologous peptides is to provide broad based, flexible, efficient, economic, and practical platforms and processes that can be utilized in commercial, therapeutic, and vaccine applications. For example, for the production of certain peptides, it would be ideal to have an expression system capable of producing, in an efficient and inexpensive manner, large quantities of final, desirable products in vivo in order to eliminate or reduce downstream reassembly costs.

[0005] Currently, bacteria are the most widely used expression system for the production of recombinant peptides because of their potential to produce abundant quantities of recombinant peptides. Bacteria generally do not glycosylate, acetylate, acylate, phosphorylate, or gamma-carboxylate eukaryotic proteins, and, therefore, are limited in their capacities to produce, in vivo, certain types of heterologous eukaryotic peptides that require post-translational modifications. Additional steps modifying the bacterially produced eukaryotic peptides are likely to increase the time and reduce the overall yield of peptide production, diluting many of the advantages of the bacterial host expression system. Therefore, alternative non-bacterial expression systems have been utilized to overcome the inherent limitations of bacterial expression systems.

[0006] One particular host expression systems analyzed for their ability to produce heterologous peptides has been whole plants and plant cell suspension cultures. The safe and inexpensive culture of plants provides an improved alternative host for the cost effective production of certain peptides. This alternative is particularly attractive where the protein is complex, requires glycosylation, needs to be free of human or animal infectious viruses, and bacterial toxins.

[0007] Whole plants offer the advantages of being relatively inexpensive to grow in vast quantities, with the potential for a great yield of the desired recombinant protein from the large biomass of the harvested transgenic or viral infected plants. As a result, whole plants have been developed as expression systems for commercial production of biopharmaceutical proteins intended for human or veterinary administration.

[0008] Significant time and resources have been spent on trying to improve the cost of production and yield of heterologous proteins in non-bacterial systems in order to take advantage of the inherent abilities of these systems. While progress has been made in both of these areas, additional processes and platforms for the production of heterologous peptides in non-bacterial expression systems would be beneficial.

[0009] Viruses and Virus Like Particles

[0010] One approach for improving peptide production in host cell expression systems is to make use of the properties of infectious recombinant viruses to produce recombinant peptides of interest. The use of infectious viruses in plant host systems is particularly well known. See, for example, Porta & Lomonosoff, (2002) "Viruses as vectors for the expression of foreign sequences in plants," *Biotechnology and Genetic Engineering Reviews* 19: 245-291.

[0011] Recent strategies have focused on the production of heterologous peptides in virus like particle (VLP) structures. In general, encapsidated viruses include a protein coat or "capsid" that is assembled to contain the viral nucleic acid. Many viruses have capsids that can be "self-assembled" from the individually expressed capsids, both within the cell the capsid is expressed in ("in vivo assembly"), and outside of the cell after isolation and purification ("in vitro assembly"). Ideally, capsids are modified to contain a target recombinant peptide, generating a recombinant viral capsid-peptide fusion. The fusion peptide can then be expressed in a cell, and, ideally, assembled in vivo to form recombinant viral or virus-like particles.

[0012] The production of heterologous proteins via virus capsid fusion proteins assembled into VLPs in plants has been met with varying success. See, for example, C Marusic et al., *J. Virol.* 75(18):8434-39 (September 2001) (use of infectious helical potato virus X in whole *Nicotiana benthamiana* to express virus capsids terminally fused to an antigenic, six amino acid HIV peptide, with in vivo formation of the recombinant virus particles); FR Brennan et al., *Vaccine* 17(15-16):1846-57 (9 Apr. 1999) (use of infectious cowpea mosaic virus or helical potato virus X capsids terminally fused to an antigenic, *Staphylococcus aureus* peptide, with in vivo formation of recombinant virus particles expression in whole cowpea plants (*Vigna unguiculata*)); C Porta et al., *Intervirology* 39(1-2):79-84 (1996) (describing an infectious cowpea mosaic virus expressing a chimeric coat protein including an antigenic HIV sequence in whole plants).

[0013] U.S. Pat. No. 5,874,087 to Lomonosoff & Johnson describes production of infectious plant viruses, in plant cells or whole plants, wherein the viral capsids are engineered to contain a biologically active peptide, such as a hormone, growth factor, or antigenic peptide. A virus selected from the genera Comovirus, Tombusvirus, Sobe-

movirus, and Nepovirus is engineered to contain the exogenous peptide encoding sequence and the entire engineered genome of the virus is expressed to produce the recombinant virus. The specification stresses that multiplication of the modified virus is a central part of the invention.

[0014] U.S. Pat. No. 6,232,099 to Chapman et al. describes the use of infective, rod-shaped viruses to produce foreign proteins connected to viral capsid subunits in whole plants. Rod-shaped viruses, also classified as helical viruses, such as potato virus X (PVX) have recombinant peptides of interest inserted into the genome of the virus to create recombinant viral capsid-peptide fusions. The recombinant, infective virus is then used to infect a plant cell of a whole plant, wherein, the virus actively replicates in the plant cell and further infects other cells, ultimately infecting the entire host plant. Ultimately, the recombinant viral capsid-peptide fusion is purified from the plant.

[0015] Chapman et al. also teaches that a limited insertion size is tolerated by icosahedral viruses. Chapman et al. cite WO 92/18618, which limits the size of the recombinant peptide in an icosahedral virus for expression in a plant host cell to 26 amino acids in length, in supporting his assertion. Chapman et al. theorize that a larger peptide present in the internal insertion site in the capsid of icosahedral viruses may result in disruption of the geometry of the protein and/or its ability to successfully interact with other capsids leading to failure of the chimeric virus to assemble.

[0016] U.S. Pat. No. 6,042,832 to Koprowski et al. describes fusion capsid proteins comprising a plant virus capsid protein fused to an antigenic polypeptide. The resultant particles are produced in whole plants.

[0017] The utilization of infectious recombinant viruses to produce heterologous proteins in capsid fusion proteins is not, however, without its drawbacks. One particularly troubling aspect is the ability of these viruses to mutate in vivo, resulting in capsid proteins that are essentially wild type revertants without the fused heterologous protein of interest, or mutated, non-desirable recombination capsid fusion protein products. See, for example, Porta & Lomonossoff (1996) "Use of viral replicons for the expression of genes in plants," *Molecular Biotechnology* 5:209-221; Dolja et al. (1993) "Spontaneous mutagenesis of a plant potyvirus genome after insertion of a foreign gene," *J. Virol.* 67(10):5968-5975; Dawson et al. (2001) "Assessment of recombinants that arise from the use of a TMV-based transient expression vector," *Virol.* 284(2):182-189 (describing the deletion of the foreign inserted gene in inoculated whole plants). The lack of stability of these viral vectors in whole plants potentially reduces the yield of overall protein product, and may lead to inconsistencies and irregularities in the capsid-fusion product. Such irregularities may be particularly troublesome wherein the integrity of the protein product is essential for a particular desired physio-chemical characteristic in the peptide.

[0018] As a result of the inherent instability of infectious recombinant viruses in plants, there is still a need in the field of commercial recombinant protein production for an efficient peptide production system that offers plant-system-type benefits.

[0019] In addition, the use of whole plants for the production of recombinant peptides also presents potential

problems. For instance, long development times, batch to batch variations in product yield, containment issues, the difficulty of applying good manufacturing practice to the early stages of production, the possibility of contamination with agrochemicals and fertilizers, as well as the impact of pests, disease and variable cultivation conditions due to microclimate and soil differences all result in a potential inconsistent host system for the production of recombinant peptides.

[0020] Therefore, it is an object of the present invention to provide a stable and consistent plant cell expression system for the production of virus like particles containing capsid fusion proteins.

[0021] It is another object of the present invention to provide plant cells for use as host cells in a stable expression system for the production of virus like particles containing capsid fusion proteins.

[0022] It is still another object of the present invention to provide processes for the improved production of virus like particles containing capsid fusion proteins in plant cells, including plant cell suspension cultures.

[0023] It is yet another object of the present invention to provide novel constructs and nucleic acids for use in plant cell expression system for the production of virus like particles containing capsid fusion proteins.

SUMMARY OF THE INVENTION

[0024] The present invention provides a process for the production of recombinant peptides, wherein non-infectious plasmids encoding fusion peptides comprising a viral capsid and a recombinant peptide of interest are stably inserted into the genome of a host plant cell and expressed in a suspension plant cell culture. The viral capsid-heterologous peptide fusion products can be expressed in vivo as virus like particles. The present invention does not require the utilization of an infective viral agent; rather, non-infectious nucleic acid encoding a capsid-heterologous peptide fusion product is stably inserted into the genome of a plant cell, which can be cultured in a fermentation process to produce the peptide of interest. Such a process results in a less variable, and more stable host system for the expression of capsid fusion proteins containing heterologous peptides.

[0025] It has been discovered that infectious viruses containing capsid-heterologous peptide fusion proteins utilized to express heterologous peptides of interest in whole plants exhibit genetic instability in the whole plant that results in mutations in the recombinant capsid protein nucleic acid, and the expressed mutant capsid either cannot assemble into virus particles or contains a mutated target peptide. The present invention provides increased stability of the resultant heterologous peptide, with the additional benefits of precise control over growth conditions, batch to batch consistency, a high level of containment, and the ability to produce recombinant proteins in compliance with good manufacturing practices.

[0026] The capsid-fusion protein products can form virus like particles within the cell. The virus like particle may result in the improved efficiency of achieving a high purity of the recovered peptide. The virus like particles produced in the cell typically are not capable of infecting the plant cell. The viral capsid sequence can be derived from both trophic

and non-trophic viruses, wherein tropism is determined by the specific plant cell utilized as the expression host. In one embodiment, the viral capsid protein is derived from a virus that exhibits a native or natural tropism towards the plant cell utilized to express the fusion product. In one embodiment, the viral capsid protein is derived from a virus that does not exhibit a native or natural tropism towards the plant cell utilized to express the fusion product. In one embodiment, the cell does not include artificially introduced viral proteins other than the desired capsid protein sequences utilized to produce the fusion product. In another embodiment, the cell includes artificially introduced viral proteins or nucleic acids other than the desired capsid protein sequences utilized to produce the fusion product, wherein the additional viral proteins or nucleic acids do not confer infectivity to the nucleic acid sequences. In one embodiment, the viral capsid is derived from a virus with a tropism to a different family of organisms than the plant cell expression host. In another embodiment, the viral capsid is derived from a virus with a tropism to a different genus of organisms than the plant cell expression host. In another embodiment, the viral capsid is derived from a virus with a tropism to a different species of organisms than the cell utilized to express the fusion product. In one embodiment of the present invention, the capsid is derived from a rod shaped plant virus. In a particular embodiment, the capsid is a rod shaped viral capsid derived from the group selected from Tobacco Mosaic Virus and Potato Virus X (PVX). In one embodiment of the present invention, the capsid protein is derived from an icosahedral virus. In a particular embodiment, the capsid is derived from a plant icosahedral virus. In a more particular embodiment, the icosahedral capsid is derived from the group selected from Cowpea Mosaic Virus, Cowpea Chlorotic Mottle Virus, and Alfalfa Mosaic Virus.

[0027] The present invention also provides plant cells that include a non-infectious nucleic acid construct containing an expression cassette encoding for a fusion protein of a virus capsid and a recombinant peptide. The fusion peptide of the virus capsid and recombinant peptide is operably linked to a promoter and terminator that functions in plant cells. In one specific embodiment, the nucleic acid is genomically integrated in the plant cell, wherein the integration results in the stable inheritance and expression of the nucleic acid encoding the capsid fusion protein from generation to generation. In one specific embodiment of the present invention, the capsid protein is derived from an icosahedral virus. In one embodiment the cell produces virus like particles or soluble cage structures. In one embodiment, the plant cell is a Monocot or a Dicot. In a particular embodiment, the plant cell is *Nicotiana tabacum*. In an alternative embodiment, the plant cell is *Oryza sativa*. In one embodiment of the present invention, the recombinant peptide fused to the viral capsid protein is a therapeutic peptide useful for human or animal treatments. In one particular embodiment, the recombinant peptide is an antigenic peptide. In a particular embodiment, the antigenic peptide is a glycosylated antigenic peptide. In one embodiment, the plant cells or extracts containing capsid-recombinant peptide virus like particles containing an antigenic peptide can be administered as a vaccine to a human or animal. In an alternative embodiment, the purified capsid-recombinant peptide virus like particles containing an antigenic peptide can be administered as a vaccine to a human or animal. In one embodiment, the heterologous peptide is an antimicrobial peptide. In another particular

embodiment, the recombinant peptide is a peptide that is toxic to the plant host cell when in free monomeric form. In one embodiment, the recombinant peptide fused to the capsid is at least 7, at least 8, at least 9, at least 10, at least 12, at least 15, at least 20, at least 25, at least 30, at least 35, at least 40, at least 45, at least 50, at least 55, at least 60, at least 65, at least 75, at least 85, at least 95, at least 99, or at least 100 amino acids in length.

[0028] In one embodiment of the present invention, the recombinant peptide fused to the capsid contains at least one monomer of a desired target peptide. In an alternative embodiment, the recombinant peptide contains more than one monomer of a desired target peptide. In certain embodiments, the peptide is composed of at least two, at least 5, at least 10, at least 15 or at least 20 separate monomers that are operably linked as a concatameric peptide to the capsid. In another embodiment, the individual monomers in the concatameric peptide are linked by cleavable linker regions. In still another embodiment, the recombinant peptide is inserted into at least one surface loop of the viral capsid. In one embodiment, the recombinant peptide is inserted into at least one surface loop of an icosahedral viral capsid. In one embodiment, at least one monomer is inserted into more than one surface loops of a viral capsid protein. In still another embodiment, the recombinant peptide is inserted into at least one outer surface loop of the viral capsid. In an alternative embodiment, the recombinant peptide is inserted into at least one inner surface loop of the viral capsid.

[0029] More than one loop of the virus like particle can be modified. In one particular embodiment, the recombinant peptide is expressed on at least two surface loops of the virus-like particle. In another embodiment, at least two different peptides are inserted into at least two surface loops of the viral capsid, cage or virus-like particle. In another embodiment, at least three recombinant peptides are inserted into at least three surface loops of the virus-like particle. The recombinant peptides in the surface loops can have the same amino acid sequence. In separate embodiments, the amino acid sequence of the recombinant peptides in the surface loops differ.

[0030] In still another embodiment, the cell includes at least one additional nucleic acid encoding either a wild-type capsid or different capsid-recombinant peptide fusion peptide, wherein the multiple capsids can be assembled in vivo to produce chimeric virus like particles.

[0031] In one aspect of the present invention, plant cells are provided that include a fusion protein of a viral capsid and a recombinant peptide. In one specific embodiment of the present invention, the plant cell is a monocot. In an alternative embodiment, the plant cell is a dicot. In one embodiment, the capsid-recombinant peptide fusion protein assembles in vivo to form a virus like particle.

[0032] In one embodiment of the present invention, the plant cells containing the nucleic acid construct encoding the capsid protein-recombinant peptide fusion peptide is obtained by nuclear transformation. In one embodiment, the plant cell is obtained by plastid transformation. In still another embodiment, the plant cell is obtained by chloroplast transformation.

[0033] The present invention further provides for non-infectious nucleic acid constructs containing an expression

cassette encoding for a fusion protein of a viral capsid and a recombinant peptide. In one embodiment of the present invention, the expression cassette encoding for a fusion protein of a viral capsid and a recombinant peptide is operably linked to a promoter and terminator. In one embodiment of the present invention, the capsid is derived from a plant virus. In one embodiment of the present invention, the capsid is derived from an icosahedral plant virus. In a particular embodiment, the capsid is an icosahedral viral capsid derived from the group selected from Cowpea Mosaic Virus, Cowpea Chlorotic Mottle Virus, and Alfalfa Mosaic Virus.

[0034] In one embodiment of the present invention, the recombinant peptide contains at least one monomer of a desired target peptide. In an alternative embodiment, the recombinant peptide contains more than one monomer of a desired target peptide. In still another embodiment, the recombinant peptide is inserted into at least one surface loop of the icosahedral virus capsid.

[0035] In another embodiment, the nucleic acid construct includes additional nucleic acid sequences including at least one promoter that functions in plant cell. In another embodiment, the nucleic acid construct includes additional nucleic acid sequences including at least one promoter that functions in plant cells, and at least one terminator that functions in plant cells. In one embodiment, a nucleic acid sequence encoding a selection marker operably linked to a promoter and a terminator sequence is included in the nucleic acid construct. In an alternative embodiment, a selection marker operably linked to a promoter and a termination sequence is provided on a separate nucleic acid construct. In another embodiment, the nucleic acid construct includes additional nucleic acid sequences derived from the 3' untranslated region (3' UTR) of the viral RNA. In still another embodiment, the nucleic acid construct includes at least one encapsidation signal derived from the viral RNA. In still another embodiment, the non-infectious nucleic acid construct includes additional nucleic acid sequences derived from the viral RNA, wherein the additional sequences do not confer infectivity to the viral nucleic acids.

[0036] In one aspect, the present invention provides a process for producing a recombinant peptide including:

- [0037] a) providing a plant cell;
- [0038] b) providing a non-infectious nucleic acid encoding a fusion peptide, wherein the fusion is of at least one heterologous peptide and at least one viral capsid protein;
- [0039] c) transforming the plant cell, wherein the nucleic acid is subsequently integrated into the genome of the plant cell;
- [0040] d) expressing the nucleic acid in the plant cell, wherein the plant cells are grown in a suspension plant cell culture in a fermentation process;
- [0041] e) wherein the expression in the plant cell provides for in vivo assembly of the fusion peptide into virus like particles; and
- [0042] f) isolating the virus like particles.

[0043] In one embodiment, the process further includes: e) cleaving the fusion product to separate the recombinant

peptide from the viral capsid protein. In still another embodiment, the process further provides: f) following step (e), isolating the recombinant peptide. See, for example, FIG. 1. In one embodiment of the present invention, the plant cell is selected from the group consisting of *Oryza sativa* and *Nicotiana tabacum*.

[0044] The nucleic acid encoding a recombinant peptide and a viral capsid protein is operably linked to a promoter sequence and a terminator sequence.

[0045] In one embodiment, the process includes co-expressing another nucleic acid encoding a wild-type capsid or a different capsid-recombinant peptide fusion peptide, wherein the capsids are assembled in vivo to produce chimeric virus like particles.

[0046] In another aspect of the present invention, an expression system for the production of recombinant peptides is provided including:

- [0047] a) a plant cell capable of being propagated in a plant cell medium;
- [0048] b) a non-infectious nucleic acid encoding a fusion peptide; wherein the fusion peptide comprises at least one recombinant peptide, and at least one viral capsid, and wherein the nucleic acid is capable of genomic integration and expression in the plant cell; and
- [0049] c) a growth medium.

[0050] The nucleic acid encoding the fusion peptide can be operably linked to a promoter sequence and a terminator sequence. When expressed the fusion peptide can assemble into virus like particles within the cell. In one embodiment, the promoter is a plant promoter.

BRIEF DESCRIPTION OF THE FIGURES

[0051] FIG. 1 illustrates a scheme for production of peptide monomers in Virus-Like Particles (VLP) in plant cells. A sequence encoding a desired target peptide ("I") is inserted into a sequence encoding a viral coat protein ("CP") constructing a gene encoding recombinant viral coat protein ("rCP"), which, as part of a vector, is transformed into the plant cell and expressed to form recombinant coat proteins ("rCP"). Optional cleavage sites ("**") are also shown. A plant cell is transfected or transformed with a plant expression plasmid for with the rCP gene. rCP can assemble into VLPs in plant cells. The peptide insert is displayed on VLP surfaces. Chimeric VLPs can be purified from the plant cells, and the desired target peptides can optionally be cleaved off the purified VLPs and purified.

[0052] FIG. 2 shows expression of CCMV CP in tobacco NT1 cells transfected with infectious CCMV RNA1, RNA2, and one of the RNA3s transcribed from pDOW2124 (CCMV-RNA3-CP), pDOW2125 (CCMV-RNA3-CP63BamHI), pDOW2126 (CCMV-RNA3-CP102BamHI), pDOW2127 (CCMV-RNA3-CP114BamHI), pDOW2128 (CCMV-RNA3-CP129BamHI), and pDOW2129 (CCMV-RNA3-CP160BamHI). The plant cell extracts were run on SDS-PAGE gel and probed with anti-CCMV polyclonal antibodies. Lane 1 represents Buffer only, Lane 2 represents RNA1 and RNA2 only, Lane 3 represents RNA1 and RNA2+wt RNA3 transcribed from pDOW2124, Lane 4 represents RNA1 and RNA2+RNA3 transcribed from

pDOW2125, Lane 5 represents RNA1 and RNA2+RNA3 transcribed from pDOW2126, Lane 6 represents RNA1 and RNA2+RNA3 transcribed from pDOW2127, Lane 7 represents RNA1 and RNA2+RNA3 transcribed from pDOW2128, Lane 8 represents RNA1 and RNA2+RNA3 transcribed from pDOW2129.

[0053] FIG. 3 shows expression of CCMV CP in tobacco NT1 cells transfected with infectious CCMV RNA1, RNA2, and one of the engineered RNA3s containing the CCMV CP fusion with antigenic peptide PA1, PA2, PA3, or PA4 at the position 63, 102, or 114. The plant cell extracts were run on SDS-PAGE gel and probed with anti-CCMV polyclonal antibodies. Lane 1 represents a size marker, Lane 2 represents Buffer only, Lane 3 represents RNA1+RNA2+chimeric RNA3 with PA1 in CP at position 63, Lane 4 represents RNA1+RNA2+chimeric RNA3 with PA2 in CP at position 63, Lane 5 represents RNA1+RNA2+chimeric RNA3 with PA3 in CP at position 63, Lane 6 represents RNA1+RNA2+chimeric RNA3 with PA4 in CP at position 63, Lane 7 represents RNA1+RNA2+chimeric RNA3 with PA1 in CP at position 102, Lane 8 represents RNA1+RNA2+chimeric RNA3 with PA2 in CP at position 102, Lane 9 represents RNA1+RNA2+chimeric RNA3 with PA3 in CP at position 102, Lane 10 represents RNA1+RNA2+chimeric RNA3 with PA4 in CP at position 102, Lane 11 represents RNA1+RNA2+chimeric RNA3 with PA1 in CP at position 114, Lane 12 represents RNA1+RNA2+chimeric RNA3 with PA2 in CP at position 114, Lane 13 represents RNA1+RNA2+chimeric RNA3 with PA3 in CP at position 114, and Lane 14 represents RNA1+RNA2+chimeric RNA3 with PA4 in CP at position 114.

[0054] FIG. 4 shows expression of CCMV CP in tobacco NT1 cells transfected with infectious CCMV RNA1, RNA2, and one of the engineered RNA3s containing the CCMV CP fusion with antigenic peptide PA1, PA2, PA3, or PA4 at the position 129 or 160. The plant cell extracts were run on SDS-PAGE gel and probed with anti-CCMV polyclonal antibodies. Lane 1 represents a size marker, Lane 2 represents RNA1+RNA2+chimeric RNA3 with PA1 in CP at position 129, Lane 3 represents RNA1+RNA2+chimeric RNA3 with PA2 in CP at position 129, Lane 4 represents RNA1+RNA2+chimeric RNA3 with PA3 in CP at position 129, Lane 5 represents RNA1+RNA2+chimeric RNA3 with PA4 in CP at position 129, Lane 6 represents RNA1+RNA2+chimeric RNA3 with PA1 in CP at position 160, Lane 7 represents RNA1+RNA2+chimeric RNA3 with PA2 in CP at position 160, Lane 8 represents RNA1+RNA2+chimeric RNA3 with PA3 in CP at position 160, Lane 9 represents RNA1+RNA2+chimeric RNA3 with PA4 in CP at position 160, Lane 10 represents RNA1+RNA2+wt RNA3, and Lane 11 represents Buffer only.

[0055] FIG. 5 shows plasmid map of pDOW2160.

[0056] FIG. 6 shows plasmid map of pDOW2161.

[0057] FIG. 7 shows plasmid map of pDOW2162.

[0058] FIG. 8 shows plasmid map of pDOW2163.

[0059] FIG. 9 shows expression of chimeric CCMV CP in tobacco NT1 cells transfected with non-infectious plasmids pDOW2160, pDOW2161, pDOW2162, and pDOW2163. The plant cell extracts were run on SDS-PAGE gel and probed with anti-CCMV polyclonal antibodies. Lane 1 represents Buffer only, Lane 2-6 represent RNA1+RNA2+wt

RNA3 (control), Lane 7 represents pDOW2160, Lane 8 represents pDOW2161, Lane 9 represents pDOW2162, and Lane 10 represents pDOW2163.

[0060] FIG. 10 shows expression of chimeric CCMV CP in tobacco NT1 cells transfected with non-infectious plasmids pDOW2169, pDOW2170, pDOW2171, and pDOW2172. The plant cell extracts were run on SDS-PAGE gel and probed with anti-CCMV polyclonal antibodies. Lane 1 represents RNA1+RNA2+RNA3 (control), Lane 2 represents Buffer, Lane 3 represents pDOW2169, Lane 4 represents pDOW2170, Lane 5 represents pDOW2171, and Lane 6 represents pDOW2172.

[0061] FIG. 11 shows expression of chimeric CCMV CP in tobacco NT1 cells stably transformed with non-infectious plasmids pDOW2160, pDOW2161, pDOW2162, and pDOW2163. After selection for 21 days, calli that had white fluffy cell growth were selected for analysis by western blotting for to test for CP expression. The plant cell extracts were run on SDS-PAGE gel and probed with anti-CCMV polyclonal antibodies. Lane 1 represents Untransformed tobacco cells (negative control), Lane 2 represents Transgenic tobacco cells transformed with pDOW2160, Lane 3 represents Transgenic tobacco cells transformed with pDOW2161, Lane 4 represents Transgenic tobacco cells transformed with pDOW2162, and Lane 5 represents Transgenic tobacco cells transformed with pDOW2163.

[0062] FIG. 12 shows expression of chimeric CCMV CP in rice cells stably transformed with non-infectious plasmids pDOW2160, pDOW2161, pDOW2162, and pDOW2163. After selection for 21 days, calli that had white fluffy cell growth were selected for analysis by western blotting for to test for CP expression. The plant cell extracts were run on SDS-PAGE gel and probed with anti-CCMV polyclonal antibodies. Lane 1 represents wt CCMV CP (positive control), Lane 2 represents Size ladder, Lane 3 represents Non-transgenic rice cells (negative control), Lane 4 represents Transgenic rice cells transformed with pDOW2160, Lane 5 represents wt CCMV CP (positive control), Lane 6 represents Size ladder, Lane 7 represents Non-transgenic rice cells, Lane 8 represents Transgenic rice cells transformed with pDOW2161, Lane 9 represents wt CCMV CP (positive control), Lane 10 represents Size ladder, Lane 11 represents Non-transgenic rice cells (negative control), Lane 12 represents Transgenic rice cells transformed with pDOW2162, Lane 13 represents wt CCMV CP (positive control), Lane 14 represents Size ladder, Lane 15 represents Non-transgenic rice cells (negative control), and Lane 16 represents Transgenic rice cells transformed with pDOW2163.

[0063] FIG. 13 shows detection of PCR products amplified from selected individual rice calli transformed with non-infectious plasmids pDOW2160, pDOW2161, pDOW2162, and pDOW2163. 1.2% agarose gel stained with EtBr showing PCR products amplified from selected individual rice calli. After selection for 21 days, calli that had white fluffy cell growth were selected for analysis by PCR to test for integration of promoter-CP fusion gene-terminator cassette into the plant genome. Lane 1 represents Rice calli 46-5, Lane 2 represents Rice calli 46-11, Lane 3 represents Rice calli 46-17, Lane 4 represents Size Ladder, Lane 5 represents Rice calli 47-6, Lane 6 represents Rice calli 48-12, Lane 7 represents Rice calli 48-18, Lane 8

represents Rice calli 48-20, Lane 9 represents Rice calli 49-11, Lane 10 represents Rice calli 49-18, and Lane 11 represents Negative control (non-transgenic rice cells).

[0064] FIG. 14 shows western blot analysis of chimeric CCMV VLPs purified from rice cells stably transformed with non-infectious plasmid pDOW2160 by PEG precipitation and ultra-filtration and detection of chimeric CP-PA1 fusion proteins in the purified VLP samples. The samples were run on SDS-PAGE gel and probed with anti-CCMV polyclonal antibodies. Lane 1 represents Size ladder, Lane 2 represents Total cell lysate from rice suspension cells transgenic for CCMV CP-PA1 fusion, Lane 3 represents First PEG supernatant, Lane 4 represents First PEG pellet, Lane 5 represents Second PEG pellet, and Lane 6 represents Second PEG supernatant.

[0065] FIG. 15 shows diagram of production process of chimeric VLPs in plant suspension cells by fermentation.

DETAILED DESCRIPTION

[0066] The present invention provides a process for the production of recombinant peptides, wherein non-infectious plasmids encoding fusion peptides comprising a viral capsid and a recombinant peptide of interest are stably inserted into the genome of a host plant cell and expressed in a suspension plant cell culture. The viral capsid-heterologous peptide fusion products can be expressed in vivo as virus like particles. The present invention further provides plant cells and nucleic acid constructs for use in the process. Specifically, the invention provides plant cells, capable of propagation in suspension plant cell cultures, with a nucleic acid construct containing an expression cassette encoding a fusion peptide of a viral capsid and a recombinant peptide. The fusion peptide can be operably linked to a promoter sequence and a termination sequence. In one embodiment, the expression in the plant cell of the fusion peptide produces virus like particles or soluble cage structures. The invention also provides nucleic acid constructs capable of integrating into the genome of the plant cell and encoding the fusion peptide of a viral capsid and a recombinant peptide, which can in one embodiment, be a therapeutic peptide useful for human and animal treatments.

[0067] The invention also provides a process for producing a recombinant peptide in a suspension plant cell culture by providing: a nucleic acid capable of genomic integration in a plant cell containing an expression cassette encoding a fusion peptide of a recombinant peptide and a viral capsid operably linked to a promoter sequence and a termination sequence; expressing the nucleic acid in the plant cell in a suspension plant cell culture, wherein the expression in the plant cell provides for in vivo assembly of the fusion peptide into virus like particles; and isolating the virus like particles.

[0068] The term "infectious" as used herein means the ability of a virus particle to transfer its nucleic acid to a host or introduction of a viral nucleic acid into a host, wherein the viral nucleic acid is replicated, viral proteins are translated, and new viral particles capable of further transfer of nucleic acid to a host are assembled. The term "non-infectious" as used herein means the inability of a virus-derived nucleic acid to replicate in a host after introduction into a host, wherein the viral proteins are translated, and new virus like particles assembled that are not capable of initiating viral infection process in a host. The term "peptide" as used

herein is not limited to any particular molecular weight, and can also include proteins or polypeptides.

I. Recombinant Plant Cells

[0069] The present invention provides plant cells that include a non-infectious nucleic acid construct capable of genomic integration and encoding a fusion peptide of a viral capsid and a recombinant peptide operably linked to a promoter sequence and a termination sequence. The cells can be utilized in a process for producing recombinant peptides.

[0070] The plant cells can be derived by transforming the native plant cell with a nucleic acid construct containing an expression cassette encoding a fusion peptide of a viral capsid and a recombinant peptide operably linked to a promoter sequence and a termination sequence, wherein the nucleic acid construct is stably integrated into the host cell's genome. Stable transformation depends on the integration of the foreign DNA into the genome. The foreign DNA can be integrated into the nuclear or plastid genome of the plant. The nucleic acid construct can be randomly inserted into the genome of the plant, or can be directed to a particular region of the genome through homologous recombination. The plant cell can be obtained by any means known in the art, including nuclear transformation, plastid transformation, and chloroplast transformation. See, for example, U.S. Pat. No. 6,218,145; EP 012345149; WO 0121782; U.S. Pat. No. 6,515,206; WO 99/10513; U.S. Pat. No. 5,693,507; WO 02055651; WO 0170939; U.S. Pat. No. 6,472,586; WO 02057466; U.S. Pat. No. 5,057,422; WO 0120974 Staub, J. M. and Maliga, P, (1992) "Long regions of homologous DNA are incorporated into the tobacco plastid genome by transformation," *Plant Cell* 4: 39-45.

[0071] Viral Capsids

[0072] In one embodiment, the invention provides plant cells for use in a process for producing peptides by expression of the peptide fused to a viral capsid from a non-infectious plasmid stably integrated into the host genome. The expression can result in the formation of at least one virus like particle (VLP) in the cell.

[0073] Viruses can be classified into those with helical symmetry or icosahedral symmetry. Generally recognized capsid morphologies include: icosahedral (including icosahedral proper, isometric, quasi-isometric, and geminate or "twinned"), polyhedral (including spherical, ovoid, and lemon-shaped), bacilliform (including rhabdo- or bullet-shaped, and fusiform or cigar-shaped), and helical (including rod, cylindrical, and filamentous); any of which may be tailed and/or may contain surface projections, such as spikes or knobs.

[0074] Morphology

[0075] In one embodiment of the invention, the amino acid sequence of the capsid is selected from the capsids of viruses classified as having any morphology. In one embodiment, the capsid is derived from a rod shaped virus. In a particular embodiment, the capsid is derived from a rod shaped plant virus. In a more particular embodiment, the capsid is a rod shaped viral capsid derived from the group selected from Tobacco Mosaic Virus (TMV) and Potato Virus X (PVX). TMV consists of a single plus-sense genomic RNA (6.5 kb) encapsidated with a unique coat

protein (17.5 kDa) which results in rod-shaped particles (300 nm). Potato Virus X are filamentous, non enveloped; usually flexuous viruses with a clear modal length of 515 nm and 13 nm wide (Brandes, 1964). The capsid structure forms a basic helix with a pitch of 3.4 nm (Varma et al., 1968).

[0076] In one embodiment, the capsid has an icosahedral morphology. In one embodiment, the capsid amino acid sequence will be selected from the capsids of entities that are icosahedral proper. In another embodiment, the capsid amino acid sequence will be selected from the capsids of icosahedral viruses. In one particular embodiment, the capsid amino acid sequence will be selected from the capsids of icosahedral plant viruses. However, in another embodiment, the viral capsid will be derived from an icosahedral virus not infectious to plants. For example, in one embodiment, the virus is a virus infectious to mammals.

[0077] Generally, viral capsids of icosahedral viruses are composed of numerous protein sub-units arranged in icosahedral (cubic) symmetry. Native icosahedral capsids can be built up, for example, with 3 subunits forming each triangular face of a capsid, resulting in 60 subunits forming a complete capsid. Representative of this small viral structure is e.g. bacteriophage ØX174. Many icosahedral virus capsids contain more than 60 subunits. Many capsids of icosahedral viruses contain an antiparallel, eight-stranded beta-barrel folding motif. The motif has a wedge-shaped block with four beta strands (designated BIDG) on one side and four (designated CHEF) on the other. There are also two conserved alpha-helices (designated A and B), one is between betaC and betaD, the other between betaE and betaF.

[0078] Enveloped viruses can exit an infected cell without its total destruction by extrusion (budding) of the particle through the membrane, during which the particle becomes coated in a lipid envelope derived from the cell membrane (See, e.g.: A J Cann (ed.) (2001) *Principles of Molecular Virology* (Academic Press); A Granoff and R G Webster (eds.) (1999) *Encyclopedia of Virology* (Academic Press); D L D Caspar (1980) *Biophys. J.* 32:103; D L D Caspar and A Klug (1962) *Cold Spring Harbor Symp. Quant. Biol.* 27:1; J Grimes et al. (1988) *Nature* 395:470; J E Johnson (1996) *Proc. Nat'l Acad. Sci. USA* 93:27; and J Johnson and J Speir (1997) *J. Mol. Biol.* 269:665).

[0079] Viruses

[0080] Viral taxonomies recognize the following taxa of encapsidated-particle entities:

[0081] Group I Viruses, i.e. the dsDNA viruses;

[0082] Group II Viruses, i.e. the ssDNA viruses;

[0083] Group III Viruses, i.e. the dsRNA viruses;

[0084] Group IV Viruses, i.e. the ssRNA (+)-stranded viruses with no DNA stage;

[0085] Group V Viruses, i.e. the ssRNA (−)-stranded viruses;

[0086] Group VI Viruses, i.e. the RNA retroid viruses, which are ssRNA reverse transcribing viruses;

[0087] Group VII Viruses, i.e. the DNA retroid viruses, which are dsDNA reverse transcribing viruses;

[0088] Deltaviruses;

[0089] Viroids; and

[0090] Satellite phages and Satellite viruses, excluding Satellite nucleic acids and Prions.

[0091] Members of these taxa are well known to one of ordinary skill in the art and are reviewed in: H. V. Van Regenmortel et al. (eds.), *Virus Taxonomy: Seventh Report of the International Committee on Taxonomy of Viruses* (2000) (Academic Press/Elsevier, Burlington Mass., USA); the Virus Taxonomy web-page of the University of Leicester (UK) Microbiology & Immunology Department at <http://www.micro.msb.le.ac.uk/3035/Virusgroups.html>; and the on-line "Virus" and "Viroid" sections of the Taxonomy Browser of the National Center for Biotechnology Information (NCBI) of the National Library of Medicine of the National Institutes of Health of the US Department of Health & Human Services (Washington, D.C., USA) at <http://www.ncbi.nlm.nih.gov/Taxonomy/tax.html>.

[0092] The amino acid sequence of the capsid may be selected from the capsids of any members of any of these taxa. Amino acid sequences for capsids of the members of these taxa may be obtained from sources, including, but not limited to, e.g.: the on-line "Nucleotide" (Genbank), "Protein," and "Structure" sections of the PubMed search facility offered by the NCBI at <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi>.

[0093] In one embodiment, the capsid amino acid sequence will be selected from taxa members that are specific for at least one of the following hosts: bacteria, fungi including yeasts, plants, protists including algae, invertebrate animals, vertebrate animals, and humans. In one embodiment, the capsid amino acid sequence will be selected from members of any one of the following taxa: Group I, Group II, Group III, Group IV, Group V, Group VII, Viroids, and Satellite Viruses. In one embodiment, the capsid amino acid sequence will be selected from members of any one of these seven taxa that are specific for at least one of the six above-described host types. In a more specific embodiment, the capsid amino acid sequence will be selected from members of any one of Group II, Group III, Group IV, Group VII, and Satellite Viruses; or from any one of Group II, Group IV, Group VII, and Satellite Viruses. In another embodiment, the viral capsid is selected from Group IV or Group VII.

[0094] The viral capsid sequence can be derived from a virus not tropic to the cell. In one embodiment, the cell does not include viral proteins from the particular selected virus other than the desired icosahedral capsids. In one embodiment, the viral capsid is derived from a virus with a tropism to a different family of organisms than the cell. In another embodiment, the viral capsid is derived from a virus with a tropism to a different genus of organisms than the cell. In another embodiment, the viral capsid is derived from a virus with a tropism to a different species of organisms than the cell.

[0095] In a specific embodiment, the viral capsid is selected from a virus of Group IV.

[0096] In one embodiment, the viral capsid is selected from rod-shaped plant viruses. In a more particular embodi-

ment, the viral capsid is selected from the group consisting of Tobacco Mosaic Virus and Potato Virus X.

[0097] In one embodiment, the viral capsid is selected from an icosahedral virus. The icosahedral virus can be selected from a member of any of the Papillomaviridae, Totiviridae, Dicistroviridae, Hepadnaviridae, Togaviridae, Polyomaviridae, Nodaviridae, Tectiviridae, Leviviridae, Microviridae, Sipoviridae, Nodaviridae, Picornoviridae, Parvoviridae, Calciviridae, Tetraviridae, and Satellite viruses.

[0098] In a particular embodiment, the sequence will be selected from members of any one of the taxa that are specific for at least one plant host. In one embodiment the icosahedral plant virus species will be a plant-infectious virus species that is or is a member of any of the Bunyaviridae, Reoviridae, Rhabdoviridae, Luteoviridae, Nanoviridae, Partitiviridae, Sequiviridae, Tymoviridae, Ourmiavirus, Tobacco Necrosis Virus Satellite, Caulimoviridae, Geminiviridae, Comoviridae, Sobemovirus, Tombusviridae, or Bromoviridae taxa. In one embodiment, the icosahedral plant virus species is a plant-infectious virus species that is or is a member of any of the Luteoviridae, Nanoviridae, Partitiviridae, Sequiviridae, Tymoviridae, Ourmiavirus, Tobacco Necrosis Virus Satellite, Caulimoviridae, Geminiviridae, Comoviridae, Sobemovirus, Tombusviridae, or Bromoviridae taxa. In specific embodiments, the icosahedral plant virus species is a plant infectious virus species that is or is a member of any of the Caulimoviridae, Geminiviridae, Comoviridae, Sobemovirus, Tombusviridae, or Bromoviridae. In more particular embodiments, the icosahedral plant virus species will be a plant-infectious virus species that is or is a member of any of the Comoviridae, Sobemovirus, Tombusviridae, or Bromoviridae. In more particular embodiments, the icosahedral plant virus species will be a plant-infectious virus species that is a member of the Comoviridae or Bromoviridae family. In a particular embodiment the viral capsid is derived from a Cowpea Mosaic Virus or a Cowpea Chlorotic Mottle Virus. In another embodiment, the viral capsid is derived from a species of the Bromoviridae taxa. In a specific embodiment, the capsid is derived from an Ilarvirus or an Alfamovirus. In a more specific embodiment, the capsid is derived from a Tobacco streak virus, or an Alfalfa mosaic virus (AMV) (including AMV 1 or AMV 2).

[0099] VLP

[0100] The viral capsid of the invention is non-infective in the host cells described. In one embodiment, a virus like particle (VLP) or cage structure is formed in the host cell during or after expression of the viral capsid. In one embodiment, the VLP or cage structure also includes the peptide of interest, and in a particular embodiment, the peptide of interest is expressed on the surface of the VLP. The expression system typically does not contain additional viral proteins that allow infectivity of the virus. In a typical embodiment, the expression system includes a host cell and a vector which codes for one or more viral capsids and an operably linked peptide of interest. The vector typically does not include additional viral proteins. The invention is derived from the discovery that viral capsids form to a greater extent in certain host cells and allow for more efficient recovery of recombinant peptide.

[0101] In one embodiment, the VLP or cage structure is a multimeric assembly of capsids, including from three to

about 1,000 capsids. In one embodiment, the VLP or cage structure includes at least 30, at least 50, at least 60, at least 90 or at least 120 capsids. In another embodiment, each VLP or cage structure includes at least 150 capsids, at least 160, at least 170, or at least 180 capsids.

[0102] In one embodiment, the VLP is expressed as an icosahedral structure. In another embodiment, the VLP is expressed in the same geometry as the native virus that the capsid sequence is derived of. In a separate embodiment, however, the VLP does not have the identical geometry of the native virus. In certain embodiments, for example, the structure is produced in a particle formed of multiple capsids but not forming a native-type VLP. For example, a cage structure of as few as 3 viral capsids can be formed. In separate embodiments, cage structures of about 6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 36, 39, 42, 45, 48, 51, 54, 57, or 60 capsids can be formed.

[0103] In one embodiment, at least one of the capsids includes at least one peptide of interest. In one embodiment, the peptide is expressed within at least one internal loop, or in at least one external surface loop of the VLP.

[0104] More than one loop of the viral capsid can be modified. In one particular embodiment, the recombinant peptide is expressed on at least two surface loops of the virus-like particle. In another embodiment, at least two different peptides are inserted into at least two surface loops of the viral capsid, cage or virus-like particle. In another embodiment, at least three recombinant peptides are inserted into at least three surface loops of the virus-like particle. The recombinant peptides in the surface loops can have the same amino acid sequence. In separate embodiments, the amino acid sequence of the recombinant peptides in the surface loops differs.

[0105] In certain embodiments, the host cell can be modified to improve assembly of the VLP. The host cell can, for example, be modified to include chaperone proteins that promote the formation of VLPs from expressed viral capsids. In another embodiment, the host cell is modified to include a repressor protein to more efficiently regulate the expression of the capsid to promote regulated formation of the VLPs.

[0106] The nucleic acid sequence encoding the viral capsid or proteins can also be additionally modified to alter the formation of VLPs (see e.g. Brumfield, et al. (2004) J. Gen. Virol. 85: 1049-1053). For example, these modifications are designed to alter the interior, exterior or the interface between adjacent subunits in the assembled protein cage. To accomplish this, mutagenic primers can be used to: (i) alter the interior surface charge of the viral nucleic acid binding region by replacing basic residues (e.g. K, R) in the N terminus with acidic glutamic acids (Douglas et al., 2002b); (ii) delete interior residues from the N terminus (in CCMV, usually residues 4-37); (iii) insert a cDNA encoding an 11 amino acid peptide cell-targeting sequence (Graf et al., 1987) into a surface exposed loop; and (iv) modify interactions between viral subunits by altering the metal binding sites (in CCMV, residues 81/148 mutant).

[0107] Recombinant Peptides Size

[0108] In one embodiment, the peptides operably linked to a viral capsid sequence contain at least two amino acids. In another embodiment, the peptides are at least three, at least

four, at least five, or at least six amino acids in length. In a separate embodiment, the peptides are at least seven amino acids long. The peptides can also be at least eight, at least nine, at least ten, at least 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 45, 50, 60, 65, 75, 85, 95, 96, 99 or more amino acids long. In one embodiment, the peptides encoded are at least 25 kD.

[0109] In one embodiment, the peptide will contain from 2 to about 300 amino acids, or about 5 to about 250 amino acids, or about 5 to about 200 amino acids, or about 5 to about 150 amino acids, or about 5 to about 100 amino acids. In another embodiment, the peptide contains or about 10 to about 140 amino acids, or about 10 to about 120 amino acids, or about 10 to about 100 amino acids.

[0110] In one embodiment, the peptides or proteins operably linked to a viral capsid sequence will contain about 500 amino acids. In one embodiment, the peptide will contain less than 500 amino acids. In another embodiment, the peptide will contain up to about 300 amino acids, or up to about 250, or up to about 200, or up to about 180, or up to about 160, or up to about 150, or up to about 140, or up to about 120, or up to about 110, or up to about 100, or up to about 90, or up to about 80, or up to about 70, or up to about 60, or up to about 50, or up to about 40 or up to about 30 amino acids.

[0111] In one embodiment, the recombinant peptide fused to the capsid is at least 7, at least 8, at least 9, at least 10, at least 12, at least 15, at least 20, at least 25, at least 30, at least 35, at least 40, at least 45, at least 50, at least 55, at least 60, at least 65, at least 75, at least 85, at least 95, at least 99, or at least 100 amino acids.

[0112] In one embodiment of the present invention, the recombinant peptide contains at least one monomer of a desired target peptide. In an alternative embodiment, the recombinant peptide contains more than one monomer of a desired target peptide. In certain embodiments, the peptide is composed of at least two, at least 5, at least 10, at least 15 or at least 20 separate monomers that are operably linked as a concatameric peptide to the capsid. In another embodiment, the individual monomers in the concatameric peptide are linked by cleavable linker regions. In still another embodiment, the recombinant peptide is inserted into at least one surface loop of the icosahedral virus-like particle. In one embodiment, at least one monomer is inserted in a surface loop of the virus-like particle.

[0113] Classification

[0114] The peptides of interest that are fused to the viral capsids can be a heterologous protein that is not derived from the virus and, optionally, that is not derived from the same species as the cell.

[0115] The peptides of interest that are fused to the viral capsids can be functional peptides; structural peptides; antigenic peptides, toxic peptides, antimicrobial peptides, fragments thereof; precursors thereof, combinations of any of the foregoing; and/or concatamers of any of the foregoing. In one embodiment of the present invention, the recombinant peptide is a therapeutic peptide useful for human and animal treatments, including antigenic peptides used in a vaccine strategy. In a particular embodiment, the antigenic peptide is glycosylated in vivo.

[0116] Functional peptides include, but are not limited to, e.g.: bio-active peptides (i.e. peptides that exert, elicit, or otherwise result in the initiation, enhancement, prolongation, attenuation, termination, or prevention of a biological function or activity in or of a biological entity, e.g., an organism, cell, culture, tissue, organ, or organelle); catalytic peptides; microstructure- and nanostructure-active peptides (i.e. peptides that form part of engineered micro- or nanostructures in which, or in conjunction with which, they perform an activity, e.g., motion, energy transduction); and stimulant peptides (e.g., peptide flavorings, colorants, odorants, pheromones, attractants, deterrents, and repellants).

[0117] Bio-active peptides include, but are not limited to, e.g.: immunoactive peptides (e.g., antigenic peptides, allergenic peptides, peptide immunoregulators, peptide immunomodulators); signaling and signal transduction peptides (e.g., peptide hormones, cytokines, and neurotransmitters; receptors; agonist and antagonist peptides; peptide targeting and secretion signal peptides); and bio-inhibitory peptides (e.g., toxic, biocidal, or biostatic peptides, such as peptide toxins and antimicrobial peptides).

[0118] Structural peptides include, but are not limited to, e.g.: peptide aptamers; folding peptides (e.g., peptides promoting or inducing formation or retention of a physical conformation in another molecule); adhesion-promoting peptides (e.g., adhesive peptides, cell-adhesion-promoting peptides); interfacial peptides (e.g., peptide surfactants and emulsifiers); microstructure and nanostructure-architectural peptides (i.e. structural peptides that form part of engineered micro- or nano-structures); and pre-activation peptides (e.g., leader peptides of pre-, pro-, and pre-pro-proteins and -peptides; inteins).

[0119] Catalytic Peptides include, e.g., apo B RNA-editing cytidine deaminase peptides; catalytic peptides of glutamyl-tRNA synthetases; catalytic peptides of aspartate transcarbamoylases; plant Type 1 ribosome-inactivating peptides; viral catalytic peptides such as, e.g., the foot-and-mouth disease virus [FMDV-2A] catalytic peptide; matrix metalloproteinase peptides; and catalytic metallo-oligopeptides.

[0120] The peptide can also be a peptide epitope, hapten, or a related peptide (e.g., antigenic viral peptide; virus related peptide, e.g., HIV-related peptide, hepatitis-related peptide; antibody idiotype domain; cell surface peptide; antigenic human, animal, protist, plant, fungal, bacterial, and/or archaeal peptide; allergenic peptide and allergen desensitizing peptide).

[0121] The peptide can also be a peptide immunoregulators or immunomodulators (e.g., interferons, interleukins, peptide immunodepressants and immunopotentiators); an antibody peptides (e.g., single chain antibodies; single chain antibody fragments and constructs, e.g., single chain Fv molecules; antibody light chain molecules, antibody heavy chain molecules, domain-deleted antibody light or heavy chain molecules; single chain antibody domains and molecules, e.g., a CH1, CH1-3, CH3, CH1-4, CH4, VHCH1, CL, CDR1, or FR1-CDR1-FR2 domain; paratopic peptides; microantibodies); another binding peptide (e.g., peptide aptamers, intracellular and cell surface receptor proteins, receptor fragments; anti-tumor necrosis factor peptides).

[0122] The peptide can also be an enzyme substrate peptide or an enzyme inhibitor peptide (e.g., caspase substrates

and inhibitors, protein kinase substrates and inhibitors, fluorescence-resonance-energy transfer-peptide enzyme substrates).

[0123] The peptide can also be a cell surface receptor peptide ligand, agonist, and antagonist (e.g., caeruleins, dynorphins, orexins, pituitary adenylate cyclase activating peptides, tumor necrosis factor peptides; synthetic peptide ligands, agonists, and antagonists); a peptide hormone (e.g., endocrine, paracrine, and autocrine hormones, including, e.g.: amylin, angiotensins, bradykinins, calcitonins, cardioexcitatory neuropeptides, casomorphins, cholecystokinins, corticotropins and corticotropin-related peptides, differentiation factors, endorphins, endothelins, enkephalins, erythropoietins, exendins, follicle-stimulating hormones, galanins, gastrins, glucagons and glucagon-like peptides, gonadotropins, growth hormones and growth factors, insulins, kallidins, kinins, leptins, lipotropic hormones, luteinizing hormones, melanocyte stimulating hormones, melatonins, natriuretic peptides, neurokinins, neuromedins, nociceptins, osteocalcins, oxytocins (i.e. ocytocins), parathyroid hormones, pleiotrophins, prolactins, relaxins, secretins, serotoninins, sleep-inducing peptides, somatomedins, thymopoietins, thyroid stimulating hormones, thyrotropins, urotensins, vasoactive intestinal peptides, vasopressins); a peptide cytokine, chemokine, virokin, and viroceptor hormone releasing and release-inhibiting peptide (e.g., corticotropin-releasing hormones, cortistatins, follicle-stimulating-hormone-releasing factors, gastric inhibitory peptides, gastrin releasing peptides, gonadotropin-releasing hormones, growth hormone releasing hormones, luteinizing hormone-releasing hormones, melanotropin-releasing hormones, melanotropin-release inhibiting factors; nocistatins, pancreastatins, prolactin releasing peptides, prolactin release-inhibiting factors; somatostatins; thyrotropin releasing hormones); a peptide neurotransmitter or channel blocker (e.g., bombesins, neuropeptide Y, neurotensins, substance P) a peptide toxin, toxin precursor peptide, or toxin peptide portion. In certain embodiments, a peptide toxin contains no D-amino acids. Toxin precursor peptides can be those that contain no D-amino acids and/or that have not been converted by posttranslational modification into a native toxin structure, such as, e.g., by action of a D configuration inducing agent (e.g., a peptide isomerase(s) or epimerase(s) or racemase(s) or transaminase(s)) that is capable of introducing a D-configuration in an amino acid(s), and/or by-action of a cyclizing agent (e.g., a peptide thioesterase, or a peptide ligase such as a trans-splicing protein or intein) that is capable of form a cyclic peptide structure.

[0124] Toxin peptide portions can be the linear or pre-cyclized oligo- and poly-peptide portions of peptide-containing toxins. Examples of peptide toxins include, e.g., agatoxins, amatoxins, charybdotoxins, chlorotoxins, conotoxins, dendrotoxins, insectotoxins, margatoxins, mast cell degranulating peptides, saporins, sarafotoxins; and bacterial exotoxins such as, e.g., anthrax toxins, botulism toxins, diphtheria toxins, and tetanus toxins.

[0125] The peptide can also be a metabolism- and digestion-related peptide (e.g., cholecystokinin-pancreozymin peptides, peptide yy, pancreatic peptides, motilins); a cell adhesion modulating or mediating peptide, extracellular matrix peptide (e.g., adhesins, selectins, laminins); a neuro-protectant or myelination-promoting peptide; an aggregation

inhibitory peptide (e.g., cell or platelet aggregation inhibitor peptides, amyloid formation or deposition inhibitor peptides); a joining peptide (e.g., cardiovascular joining neuropeptides, iga joining peptides); or a miscellaneous peptide (e.g., agouti-related peptides, amyloid peptides, bone-related peptides, cell-permeable peptides, conantokinins, contryphans, contulakins, myelin basic protein, and others).

[0126] The peptide can also be post-translationally modified in vivo. In particular embodiments, the peptide may be, singularly or in combination, glycosylated, acetylated, acylated, phosphorylated, or gamma-carboxylated.

[0127] In certain embodiments, the peptide of interest is exogenous to the selected viral capsid. Peptides may be either native or synthetic in sequence (and their coding sequences may be either native or synthetic nucleotide sequences). Thus, e.g., native, modified native, and entirely artificial sequences of amino acids are encompassed. The sequences of the nucleic acid molecules encoding these amino acid sequences likewise may be native, modified native, or entirely artificial nucleic acid sequences, and may be the result of, e.g., one or more rational or random mutation and/or recombination and/or synthesis and/or selection process employed (i.e. applied by human agency) to obtain the nucleic acid molecules.

[0128] The coding sequence can be a native coding sequence for the target peptide, if available, but will more typically be a coding sequence that has been selected, improved, or optimized for use in the selected expression host cell: for example, by synthesizing the gene to reflect the codon use preference of a host species. In one embodiment of the invention, the host species is *Nicotiana tabacum*, and the codon preference of *Nicotiana tabacum* is taken into account when designing both the signal sequence and the peptide sequence. In an alternative embodiment of the invention, the host species is *Oryza sativa*, and the codon preference of *Oryza sativa* is taken into account when designing both the signal sequence and the peptide sequence.

[0129] Antigenic Peptides (Peptide Epitopes)

[0130] In one embodiment, an antigenic peptide is produced through expression with a viral capsid. The antigenic peptide can be selected from those that are antigenic peptides of human or animal pathogenic agents, including infectious agents, parasites, cancer cells, and other pathogenic agents. Such pathogenic agents also include the virulence factors and pathogenesis factors, e.g., exotoxins, endotoxins, et al., of those agents. The pathogenic agents may exhibit any level of virulence, i.e. they may be, e.g., virulent, avirulent, pseudo-virulent, semi-virulent, and so forth. In one embodiment, the antigenic peptide will contain an epitopic amino acid sequence from the pathogenic agent(s). In one embodiment, the epitopic amino acid sequence will include that of at least a portion of a surface peptide of at least one such agent. In one embodiment, the capsid-recombinant peptide virus like particles can be used as a vaccine in a human or animal application.

[0131] More than one antigenic peptide may be selected, in which case the resulting virus-like particles can present multiple different antigenic peptides. In a particularly embodiment of a multiple antigenic peptide format, the various antigenic peptides will all be selected from a plurality of epitopes from the same pathogenic agent. In a

particular embodiment of a multi-antigenic-peptide format, the various antigenic peptides selected will all be selected from a plurality of closely related pathogenic agents, for example, different strains, subspecies, biovars, pathovars, serovars, or genovars of the same species or different species of the same genus.

[0132] In one embodiment, the pathogenic agent(s) will belong to at least one of the following groups: Bacteria and *Mycoplasma* agents including, but not limited to, pathogenic: *Bacillus* spp., e.g., *Bacillus anthracis*; *Bartonella* spp., e.g., *B. quintana*; *Brucella* spp.; *Burkholderia* spp., e.g., *B. pseudomallei*; *Campylobacter* spp.; *Clostridium* spp., e.g., *C. tetani*, *C. botulinum*; *Coxiella* spp., e.g., *C. burnetii*; *Edwardsiella* spp., e.g., *E. tarda*; *Enterobacter* spp., e.g., *E. cloacae*; *Enterococcus* spp., e.g., *E. faecalis*, *E. faecium*; *Escherichia* spp., e.g., *E. coli*; *Francisella* spp., e.g., *F. tularensis*; *Haemophilus* spp., e.g., *H. influenzae*; *Klebsiella* spp., e.g., *K. pneumoniae*; *Legionella* spp.; *Listeria* spp., e.g., *L. monocytogenes*; Meningococci and Gonococci, e.g., *Neisseria* spp.; *Moraxella* spp.; *Mycobacterium* spp., e.g., *M. leprae*, *M. tuberculosis*; Pneumococci, e.g., *Diplococcus pneumoniae*; *Pseudomonas* spp., e.g., *P. aeruginosa*; *Rickettsia* spp., e.g., *R. prowazekii*, *R. rickettsii*, *R. typhi*; *Salmonella* spp., e.g., *S. typhi*; *Staphylococcus* spp., e.g., *S. aureus*; *Streptococcus* spp., including Group A Streptococci and hemolytic Streptococci, e.g., *S. pneumoniae*, *S. pyogenes*; *Streptomyces* spp.; *Shigella* spp.; *Vibrio* spp., e.g., *V. cholerae*; and *Yersinia* spp., e.g., *Y. pestis*, *Y. enterocolitica*. Fungus and Yeast agents including, but not limited to, pathogenic: *Alternaria* spp.; *Aspergillus* spp.; *Blastomyces* spp., e.g., *B. dermatitidis*; *Candida* spp., e.g., *C. albicans*; *Cladosporium* spp.; *Coccidioides* spp., e.g., *C. immitis*; *Cryptococcus* spp., e.g., *C. neoformans*; *Histoplasma* spp., e.g., *H. capsulatum*; and *Sporothrix* spp., e.g., *S. schenckii*.

[0133] In one embodiment, the pathogenic agent(s) will be from a protist agent including, but not limited to, pathogenic: Amoebs, including *Acanthamoeba* spp., *Amoeba* spp., *Naegleria* spp., *Entamoeba* spp., e.g., *E. histolytica*; *Cryptosporidium* spp., e.g., *C. parvum*; *Cyclospora* spp.; *Encephalitozoon* spp., e.g., *E. intestinalis*; *Enterocytozoon* spp.; *Giardia* spp., e.g., *G. lamblia*; *Isospora* spp.; *Microsporidium* spp.; *Plasmodium* spp., e.g., *P. falciparum*, *P. malariae*, *P. ovate*, *P. vivax*; *Toxoplasma* spp., e.g., *T. gondii*; and *Trypanosoma* spp., e.g., *T. brucei*.

[0134] In one embodiment, the pathogenic agent(s) will be from a parasitic agent (e.g., helminthic parasites) including, but not limited to, pathogenic: *Ascaris* spp., e.g., *A. lumbricoide*; *Dracunculus* spp., e.g., *D. medinensis*; *Onchocerca* spp., e.g., *O. volvulus*; *Schistosoma* spp.; *Trichinella* spp., e.g., *T. spiralis*; and *Trichuris* spp., e.g., *T. trichiura*.

[0135] In another embodiment, the pathogenic agent(s) will be from a viral agent including, but not limited to, pathogenic: Adenoviruses; Arenaviruses, e.g., Lassa Fever viruses; Astroviruses; Bunyaviruses, e.g., Hantaviruses, Rift Valley Fever viruses; Coronaviruses, Deltaviruses; Cytomegaloviruses, Epstein-Barr viruses, Herpes viruses, Varicella viruses; Filoviruses, e.g., Ebola viruses, Marburg viruses; Flaviviruses, e.g., Dengue viruses, West Nile Fever viruses, Yellow Fever viruses; Hepatitis viruses; Influenzaviruses; Lentiviruses, T-Cell Lymphotropic viruses, other leukemia viruses; Norwalk viruses; Papillomaviruses, other tumor

viruses; Paramyxoviruses, e.g., Measles viruses, Mumps viruses, Parainfluenzaviruses, Pneumoviruses, Sendai viruses; Parvoviruses; Picornaviruses, e.g., Cardioviruses, Coxsackie viruses, Echoviruses, Poliomyelitis viruses, Rhinoviruses, Other Enteroviruses; Poxviruses, e.g., Variola viruses, Vaccinia viruses, Parapoxviruses; Reoviruses, e.g., Coltiviruses, Orbiviruses, Rotaviruses; Rhabdoviruses, e.g., Lyssaviruses, Vesicular Stomatitis viruses; and Togaviruses, e.g., Rubella viruses, Sindbis viruses, Western Encephalitis viruses.

[0136] In one particular embodiment, the antigenic peptide is selected from the group consisting of a Canine parvovirus peptide, *Bacillus anthracis* protective antigen (PA) antigenic peptide, and an Eastern Equine Encephalitis virus antigenic peptide. In a particular embodiment, the antigenic peptide is the *Bacillus anthracis*-derived peptide with the amino acid sequence selected from PA1 (SEQ. ID. NO. 1), PA2 (SEQ. ID. NO. 2), PA3 (SEQ. ID. NO. 3), or PA4 (SEQ. ID. NO. 4).

[0137] Host-Cell Toxic Peptide

[0138] In another particular embodiment, the recombinant peptide is a peptide that is toxic to the host cell when in free monomeric form. In certain embodiments, the peptide of interest expressed in conjunction with a viral capsid will be a host cell toxic peptide. A host cell toxic peptide indicates a bio-inhibitory peptide that is biostatic, biocidal, or toxic to the host cell in which it is expressed, or to other cells in the cell culture or organism of which the host cell is a member, or to cells of the organism or species providing the host cells. In one embodiment, the host-cell-toxic peptide will be a bioinhibitory peptide that is biostatic, biocidal, or toxic to the host cell in which it is expressed. Some examples of host-cell-toxic peptides include, but are not limited to: peptide toxins, anti-microbial peptides, and other antibiotic peptides.

[0139] Antimicrobial Peptides

[0140] In another particular embodiment, the recombinant peptide is a peptide that is an antimicrobial peptide. Anti-Microbial Peptides include, e.g., anti-bacterial peptides such as, e.g., magainins, betadefensins, some alpha-defensins; cathelicidins; histatins; anti-fungal peptides; antiprotozoal peptides; synthetic AMPs; peptide antibiotics or the linear or pre-cyclized oligo- or poly-peptide portions thereof; other antibiotic peptides (e.g., anthelmintic peptides, hemolytic peptides, tumoricidal peptides); and anti-viral peptides (e.g., some alpha-defensins; virucidal peptides; peptides that inhibit viral infection). In one embodiment, the antimicrobial peptide is selected from the group consisting of D2A21 antimicrobial peptides.

[0141] Cells for Use in Expressing the VLP

[0142] The cell used as a host for the expression of the viral capsid or viral capsid fusion peptide (also referred to as "host cell") of the invention will be one in which the expression of viral capsid does not allow replication or infection of the cell. In one embodiment, the viral capsid will be derived from a virus that does not infect the specific species of cell that the host cell is derived from. For example, in one embodiment, the viral capsid is derived from an icosahedral plant virus and is expressed in a host cell of a plant species that is not the native trophic host of the virus. In another embodiment, the viral species infects

mammals and the expression system includes a plant host cell. In an alternative embodiment, the virus is a trophic virus of the particular plant species utilized as a host cell for expressing the capsid-recombinant peptide fusion.

[0143] The plant host cell may be any plant cell from the *Tracheophyta*, *Euphyllophyta*, *Spermatophyta*, or *Angiospermophyta*. In a particular embodiment, the plant cell is a Monocot, including, but not limited to, members of, e.g.: *Arecaceae*, e.g., *Cocos* spp., *Elaeis* spp.; *Dioscoreaceae*, e.g., *Dioscorea* spp.; *Gramineae* (*Poaceae*), e.g., *Avena* spp., *Hordeum* spp., *Oryza* spp., *Panicum* spp., *Triticum* spp., *Zea* spp.; and *Musaceae*, e.g., *Musa* spp.

[0144] In an alternative embodiment, the plant cell is a member of the Dicots, including, but not limited to, e.g.: *Amaranthaceae*, e.g., *Amaranthus* spp., *Beta* spp., *Chenopodium* spp.; *Apiaceae* (*Umbelliferae*), e.g., *Daucus* spp., *Pastinaca* spp.; *Brassicaceae* (*Cruciferae*), e.g., *Arabidopsis* spp., *Brassica* spp.; *Compositae*, e.g., *Helianthus* spp.; *Convolvulaceae*, e.g., *Ipomoea* spp.; *Cucurbitaceae*, e.g., *Citrullus* spp., *Cucumis* spp., *Cucurbita* spp.; *Euphorbiaceae*; *Leguminosae* (*Fabaceae*), e.g., *Glycine* spp., *Lens* spp., *Medicago* spp., *Phaseolus* spp., *Pisum* spp., *Trifolium* spp., *Vicia* spp., *Vigna* spp.; *Linaceae*, e.g., *Linum* spp.; and *Solanaceae*, e.g., *Capsicum* spp., *Lycopersicon* spp., *Nicotiana* spp., *Solanum* spp.

[0145] In a particular embodiment, the host cell is selected from the group consisting of *Aradiposis thaliana*, *Taxus cuspidate*, *Catharanthus roseus*, *Nicotiana tabacum*, *Oryza sativa*, *Lycopersicon esculentum*, and *Glycine max*.

II. Nucleic Acid Constructs

[0146] The present invention further provides non-infectious nucleic acid constructs capable of stably integrating into the host cell's genome and encoding a fusion peptide of a capsid and a recombinant peptide. The fusion peptide can be operably linked to a promoter sequence and a termination sequence. In one embodiment, a nucleic acid construct for use in transforming plant host cells including a) a nucleic acid sequence encoding a recombinant peptide, and b) a nucleic acid sequence encoding a viral capsid is provided, wherein the nucleic acid of a) and the nucleic acid of b) are operably linked to form a fusion protein is stably inserted into the genome of the plant host cell, wherein it is expressed in the plant cell.

[0147] In certain embodiments, the vector can include sequence for multiple capsids, or for multiple peptides of interest. In one embodiment, the vector can include at least two different capsid-peptide coding sequences. In one embodiment, the coding sequences are linked to the same promoter. In certain embodiments, the coding sequences are separated by an internal ribosomal binding site. In other embodiments, the coding sequences are linked by a linker sequence that allows the formation of virus like particles in the cell. In another embodiment, the coding sequences are linked to different promoters. These promoters may be constitutive or driven by the same or different induction conditions. In another embodiment, multiple vectors capable of insertion into the genome and encoding different capsid-peptide combinations are provided. The multiple vectors can include promoters that are constitutive or driven by the same induction conditions, or by different induction conditions. In one embodiment, the viral capsid is derived from an icosahedral viral capsid.

[0148] The coding sequence for a peptide of interest can be inserted into the coding sequence for a viral capsid or capsid in a predetermined site. The peptide can also be inserted at a non-predetermined site and cells screened for production of VLPs. In one embodiment, the peptide is inserted into the capsid coding sequence so as to be expressed as a loop during formation of a VLP. In one embodiment, one peptide coding sequence is included in the vector, however in other embodiments, multiple sequences are included. The multiple sequences can be in the form of concatamers, for example concatamers linked by cleavable linker sequences.

[0149] Peptides may be inserted at more than one insertion site in a capsid. Thus, peptides may be inserted in more than one surface loop motif of a capsid; peptides may also be inserted at multiple sites within a given loop motif. The individual functional and/or structural peptide(s) of the insert(s), and/or the entire peptide insert(s), may be separated by cleavage sites, i.e. sites at which an agent that cleaves or hydrolyzes protein can act to separate the peptide(s) from the remainder of the capsid structure or assemblage.

[0150] Peptides may be inserted within external-facing loop(s) and/or within internal-facing loop(s), i.e. within loops of the capsid that face respectively away from or toward the center of the capsid. Any amino acid or peptide bond in a surface loop of a capsid can serve as an insertion for the peptide. Typically, the insertion site will be selected at about the center of the loop, i.e. at about the position located most distal from the center of the tertiary structure of the folded capsid peptide. The peptide coding sequence may be operably inserted within the position of the capsid coding sequence corresponding to this approximate center of the selected loop(s). This includes the retention of the reading frame for that portion of the peptide sequence of the capsid that is synthesized downstream from the peptide insertion site.

[0151] In another embodiment, the peptide can be inserted at the amino terminus of the capsid. The peptide can be linked to the capsid through one or more linker sequences, including the cleavable linkers described above. In yet another embodiment, the peptide can be inserted at the carboxy terminus of the capsid. The peptide can also be linked to the carboxy terminus through one or more linkers, which can be cleavable by chemical or enzymatic hydrolysis. In one embodiment, peptide sequences are linked at both the amino and carboxy termini, or at one terminus and at at least one internal location, such as a location that is expressed on the surface of the capsid in its three dimensional conformation.

[0152] In one embodiment, the peptide can be inserted into the capsid from a Cowpea Chlorotic Mosaic Virus. In one particular embodiment, the peptide can be inserted at amino acid 129 of the CCMV coat protein. In one embodiment, the peptide sequence can be inserted at amino acid 63, 102, 114, 129, or 160 of the CCMV coat protein. In another embodiment, the peptide sequence can be inserted at amino acids 60, 61, 62 or 63 of the CCMV coat protein. In still another embodiment, the peptide can be inserted at both amino acids 129 and amino acids 60-63 of the CCMV protein.

[0153] In a particular embodiment, the present invention provides a nucleic acid construct including a) a nucleic acid

encoding an antigenic peptide, and b) a nucleic acid encoding a viral capsid, wherein the nucleic acid of a) and the nucleic acid of b) are operably linked to form a fusion protein when expressed in a cell. Other capsids and recombinant peptides useful in constructing the nucleic acid construct are disclosed above.

[0154] Promoters

[0155] In one embodiment, the nucleic acid construct includes a promoter sequence operably attached to the nucleic acid sequence encoding the capsid-recombinant peptide fusion peptide. An operable attachment or linkage refers to any configuration in which the transcriptional and any translational regulatory elements are covalently attached to the described sequence so that by action of the host cell, the regulatory elements can direct the expression of the sequence of interest. In alternative embodiment, the nucleic acid construct encoding the capsid fusion product can be inserted in frame with a host cell native promoter utilizing a gene trapping strategy.

[0156] In a fermentation process, once fermentation begins, it is ideal to have a high level of production in order to maximize efficiency of the expression system. The promoter initiates transcription and is generally positioned 1-100 nucleotides upstream of the transcription start site. Ideally, a promoter will be strong enough to allow for a large amount of recombinant peptide accumulation.

[0157] The promoters used in accordance with the present invention may be constitutive promoters or regulated promoters. Common examples of useful regulated promoters used in the art include: ethanol-inducible promoters; iron deficiency-inducible promoters; wound-inducible promoters; and hormone-inducible, e.g., auxin-inducible, promoters. Common examples of useful constitutive promoters include the ubiquitin promoter, the actin promoter, the tubulin promoter, and the Cassava Vein Mosaic Virus promoter. Other types of promoters useful in the present invention include tissue-specific promoters such as corn sucrose synthetase 1 (Yang et al., 1990), corn alcohol dehydrogenase 1 (Vogel et al., 1989), corn light harvesting complex (Simpson, 1986), corn heat shock protein (Odell et al., 1985), pea small subunit RuBP carboxylase (Poulsen et al., 1986; Cashmore et al., 1983), Ti plasmid mannopine synthase (Langridge et al., 1989), Ti plasmid nopaline synthase (Langridge et al., 1989), petunia chalcone isomerase (Van Tunen et al., 1988), bean glycine rich protein 1 (Keller et al., 1989), CaMV ³⁵S transcript (Odell et al., 1985) and Potato patatin (Wenzler et al., 1989). Preferred promoters are the cauliflower mosaic virus (CaMV ³⁵S) promoter, the rice alpha-amylase RAmy3D promoter, and the S-E9 small subunit RuBP carboxylase promoter. In addition, a variety of plant gene promoters that are regulated in response to environmental, hormonal, chemical, and/or developmental signals can be used for expression of an operably linked gene in plant cells, including promoters regulated by (1) heat (Callis et al., Plant Physiol. 88:965, 1988), (2) light (e.g., pea rbcS-3A promoter, Kuhlmeier et al., Plant Cell 1:471, 1989; maize rbcS promoter, Schaffner and Sheen, Plant Cell 3:997, 1991; or chlorophyll a/b-binding protein promoter, Simpson et al., EMBO J. 4:2723, 1985), (3) hormones, such as abscisic acid (Marcotte et al., Plant Cell 1:969, 1989), (4) wounding (e.g., wun1, Siebertz et al., Plant Cell 1:961, 1989); or (5) chemicals such as methyl jasmonate, salicylic

acid, or Safener. It may also be advantageous to employ (6) organ-specific promoters (e.g., Roshal et al., EMBO J. 6:1155, 1987; Scherthaner et al., EMBO J. 7:1249, 1988; Bustos et al., Plant Cell 1:839, 1989). Additional promoters useful in the present invention include those described in WO 97/48819; the phaseolin promoters described in U.S. Pat. No. 5,591,605; rice actin promoters described in U.S. Pat. No. 5,641,876; the per 5 promoter described in WO 98/56921; the gamma zein promoters described in WO 00/12681; and other promoters as described in U.S. Pat. No. 6,825,006, U.S. Pat. No. 6,660,911, and plant promoters described the plant promoter sequence database PlantProm, described in Shahmuradov et al. (2003), "PlantProm: a database of plant promoter sequences," Nucleic Acid Res. 31(1): 114-117, and available at <http://mendel.cs.rhul.ac.uk/mendel.php?topic=plantprom>. Additional promoters may also include cauliflower mosaic virus (CaMV) ³⁵S promoter or its enhanced version, the hybrid (ocs)₃mas promoter, and ubiquitin promoters from maize or *A. thaliana*.

[0158] A promoter having the nucleotide sequence of a promoter native to the selected plant host cell can also be used to control expression of the transgene encoding the target peptide. Tandem promoters may also be used in which more than one promoter is covalently attached to another, whether the same or different in sequence.

[0159] A promoter having the nucleotide sequence derived from a virus can also be used to direct expression of the transgene encoding the target peptide.

[0160] Regulated promoters can utilize promoter regulatory proteins in order to control transcription of the gene of which the promoter is a part. Where a regulated promoter is used herein, a corresponding promoter regulatory protein will also be part of an expression system according to the present invention. Many regulated-promoter/promoter-regulatory-protein pairs are known in the art.

[0161] Other Elements

[0162] Other regulatory elements can be included in a nucleic acid construct. Such elements include, but are not limited to, for example, transcriptional enhancer sequences, translational enhancer sequences, other promoters, activators, translational start and stop signals, transcription terminators, cistronic regulators, polycistronic regulators, tag sequences, such as nucleotide sequence "tags" and "tag" peptide coding sequences, which facilitates identification, separation, purification, or isolation of an expressed peptide, including His-tag, Flag-tag, T7-tag, S-tag, HSV-tag, B-tag, Strep-tag, polyarginine, polycysteine, polyphenylalanine, polyaspartic acid, (Ala-Trp-Trp-Pro)_n, thioredoxin, beta-galactosidase, chloramphenicol acetyltransferase, cyclomaltodextrin gluconotransferase, CTP: CMP-3-deoxy-D-manno-octulosonate cytidyltransferase, trpE or trpLE, avidin, streptavidin, T7 gene 10, T4 gp55, Staphylococcal protein A, streptococcal protein G, GST, DHFR, CBP, MBP, galactose binding domain, Calmodulin binding domain, GFP, KSI, c-myc, ompT, ompA, pelB, NusA, ubiquitin, and hemocytin A.

[0163] In one embodiment, the nucleic acid construct further comprises a tag sequence adjacent to the coding sequence for the recombinant peptide of interest, or linked to a coding sequence for a viral capsid. In one embodiment, this tag sequence allows for purification of the protein. The

tag sequence can be an affinity tag, such as a hexa-histidine affinity tag. In another embodiment, the affinity tag can be a glutathione-S-transferase molecule. The tag can also be a fluorescent molecule, such as YFP or GFP, or analogs of such fluorescent proteins. The tag can also be a portion of an antibody molecule, or a known antigen or ligand for a known binding partner useful for purification.

[0164] The present invention can include, in addition to the capsid-recombinant peptide coding sequence, the following regulatory elements operably linked thereto: a promoter, a transcription terminator, translational start and stop signals.

[0165] Further examples of translation and transcription elements, and other elements useful in the present invention are described in: J D Watson et al. (eds.), *Recombinant DNA*, pp. 273-92 (1992) (Scientific American Books, W. H. Freeman and Co., New York, N.Y., USA); and I Mitsuhara et al., "Efficient promoter cassettes for enhanced expression of foreign genes in dicotyledonous and monocotyledonous plants," *Plant Cell Physiol.* 37(1):49-59 (1996).

[0166] In one particular embodiment, the nucleic acid construct includes additional nucleic acid sequences derived from the 3' untranslated region (3' UTR) of the viral genomic nucleic acid sequence utilized to derive the viral capsid protein. In one particular embodiment, the 3'UTR contains sequences that play a role in regulation of translation or capsid formation in the wild type virus. In one embodiment, the 3'UTR sequences are upstream pseudoknot domains (UPD). See, for example, Leathers et al. (1993) "A phylogenetically conserved sequence within viral 3' untranslated RNA pseudoknots regulates translation," *Mol. Cell Biol.* 13(9): 5331-5347. In a particular embodiment, the 3'UTR are derived from the 3'UTR of the viral capsid encoding region of the viral genome. In a more particular embodiment, the 3'UTR is derived from the cowpea chlorotic mosaic virus RNA3 3' untranslated region. In one embodiment, the 3'UTR sequence includes an encapsidation signal.

[0167] In an alternative embodiment, the nucleic acid construct includes at least one encapsidation signal derived from a viral nucleic acid sequence. The encapsidation signal can be derived from 5'UTR, 3'UTR, or internal sequence of the viral genomic sequence. In one embodiment, the nucleic acid construct is encapsidated in the VLP.

[0168] In one embodiment, the non-infectious nucleic acid construct also includes other sequences derived from the viral or non-viral nucleic acid sequence other than the viral capsid protein sequence.

[0169] Vectors

[0170] Useful expression vectors for use in plant cells in stably integrating into the host cell genome and expressing capsid-recombinant peptide fusion peptides are constructed by inserting a structural DNA sequence encoding a desired target peptide fused with a capsid peptide together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector may also comprise one or more phenotypic selectable or screenable markers to allow for selection of host cells. Alternatively, the selectable or screenable markers can be provided on a separate plasmid.

[0171] A wide variety of vectors and/or carriers are known in the art as useful for transforming a target plant host cell

with recombinant capsid-encoding nucleic acid for expression, and any of these may be used for expressing the genes according to the present invention. See, e.g., U Goderis I J et al., "A set of modular plant transformation vectors allowing flexible insertion of up to six expression units," *Plant Mol. Biol.* 50(1):17-27 (2002); A H Christensen & P H Quail, "Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants," *Transgenic Res.* 5(3):213-18 (1996); JD Jones et al., "Effective vectors for transformation, expression of heterologous genes, and assaying transposon excision in transgenic plants," *Transgenic Res.* 1(6):285-97 (1992); A P Gleave AP, "A versatile binary vector system with a T-DNA organisational structure conducive to efficient integration of cloned DNA into the plant genome," *Plant Mol. Biol.* 20(6):1203-07 (1992).

[0172] Thus, plasmids, transposons, genomic DNA, genomic RNA, plant artificial chromosomes, and other nucleic acid vectors may be used. Examples of some vectors that may be used include, but are not limited to, pRT101, pRT101-MCS, pRT104, pRT104.24LS, pRT104.N-myc, pRT104.C-myc, pRT104.N-3HA, pRT104.N3HAdAsp, pRT104.N-6HA, pRT104.N6HAdAsp, pRT104.C-3HA, pRT104-GST, pRT104 NES(A2), pRT103-3HA, pRTd35S-Luci(-), pRTd35S-Luci-NES, pRTd35S-GFP, pRTd35S-GFP.SacI, pRTdS GFP.BglII, pRTd35S-Ds-Red, pRTd35S-profilin, pBIGal4 DBD, pRT-3HAxHsfA2 DBD, pRT-3HAxGal4 DBD, pUC, pIL-TAB, pET, pME, pBBR, and pROKII. When particle bombardment is used to introduce the vector into the cell, any DNA plasmid or DNA molecule may be utilized in the present invention.

[0173] In a particular embodiment, the vector is pIL-Tab. In a more particular embodiment, the vector is pIL-Tab encoding for the coat protein from Cowpea Chlorotic Mosaic Virus (CCMV). In a particular embodiment, the vector is pIL-Tab, and the heterologous peptide is selected from the group consisting of PA1, PA2, PA3 or PA4. In a still further embodiment, the vector is selected from the group consisting of pDOW2160 (Seq. ID. No. 5), pDOW 2161 (Seq. ID. No. 6), pDOW2162 (Seq. ID. No. 7), pDOW2163 (Seq. ID. No. 8), pDOW 2169 (Seq. ID. No. 9), pDOW2170 (Seq. ID. No. 10), pDOW2171 (Seq. ID. No. 11), and pDOW2172 (Seq. ID. No. 12).

[0174] Alternatively, or in addition, carriers may be used, such as liposomes, dendrimers, cationic polymers, and cationic polymer-lipid complexes, all of which have been employed in widely published methods for delivery of nucleic acids to cells or protoplasts. Alternatively, naked DNA or naked RNA may be delivered to the plant host.

[0175] The vector, carrier, or naked nucleic acid may be directly transformed into a plant cell or protoplast by any method known effective therefore in the art. For example, any of the following methods may be used to transform the plant target host with the recombinant capsid-encoding nucleic acid: *Agrobacterium* spp.; microparticle bombardment as described by V. Vasil et al., *Bio/Technology* 9, 743 (1991); electroporation of protoplasts as described, for example, for lettuce by M C Chupeau et al., *Bio/Technology* 7, 503 (1989); surface abrasion using, e.g., silicon carbide whiskers, glass, or carborundum; liposome fusion with protoplasts as described, for example, by A. Deshayes et al., *EMBO J.* 4, p. 2731-2737 (1985); cell-cell (protoplast-

protoplast) fusion; and endogenous plant virus infection, i.e. transformation of the plant target host with an engineered virus vector, the engineered virus vector being of a type of virus for which the selected plant target host is its native host and wherein the virus has been engineered to contain a recombinant capsid-protein-encoding nucleic acid (wherein the capsid-protein-nucleic acid from which the recombinant is made is selected from a virus for which the plant target host is not its native host); polyethylene glycol mediated transformations as described by, for example, I. Potrykus et al., *Mol. Gen. Genetics*, 197, 183-188; and microinjection as described by, for example, R. Griesbach, *Biotechnology* 3, p. 348-350 and C K Shewmaker *Mol. Gen. Genetics*, 202 p. 179-185 (1986).

III. Expression of Capsid Fusion Products in Plant Cells

[0176] The present invention also provides a process for producing a recombinant peptide. The process includes:

- [0177] a) providing a plant cell;
- [0178] b) providing a non-infectious nucleic acid containing an expression cassette encoding a fusion peptide; wherein the fusion is of a recombinant peptide and an icosahedral capsid;
- [0179] c) expressing the nucleic acid in the plant cell, wherein the expression in the cell provides for in vivo assembly of the fusion peptide into virus like particles; and
- [0180] d) isolating the virus like particles.

[0181] The fusion peptide is operably linked to a promoter sequence and a terminator sequence. In one embodiment, the process further comprises: e) cleaving the recombinant peptide from the viral capsid protein. In still another embodiment, the process further comprises: f) isolating the recombinant peptide from step e. See, for example, **FIG. 1**.

[0182] Peptides may be expressed as single-copy peptide inserts within a capsid peptide (i.e. expressed as individual inserts from recombinant capsid peptide coding sequences that are mono-cistronic for the peptide) or may be expressed as di-, tri-, or multi-copy peptide inserts (i.e. expressed as concatemeric inserts from recombinant capsid peptide coding sequences that are poly-cistronic for the peptide; the concatemeric insert(s) may contain multiple copies of the same exogenous peptide of interest or may contain copies of different exogenous peptides of interest). Concatemers may be homo- or hetero-concatemers.

[0183] In one embodiment, the isolated virus like particle can be administered to a human or animal in a vaccine strategy.

[0184] In another embodiment, the nucleic acid construct can be co-expressed with another nucleic acid encoding a wild type capsid. In a particular embodiment, the co-expressed capsid/capsid-recombinant peptide fusion particles assemble in vivo to form a chimeric virus like particle. The chimeric VLP is a virus like particle including capsids or capsid-peptide fusions encoded by at least two different nucleic acid constructs.

[0185] In still another embodiment, the nucleic acid construct can be co-expressed with another nucleic acid encod-

ing a different capsid-recombinant peptide fusion particle. In a particular embodiment, the co-expressed capsid fusion particles will assemble in vivo to form a chimeric virus like particle.

[0186] In still another embodiment, a second nucleic acid, which is designed to express a different peptide, such as a chaperone protein, can be expressed concomitantly with the nucleic acid encoding the fusion peptide.

[0187] The plant cells, capsids, and recombinant peptides useful for the present invention are discussed above.

[0188] In one embodiment, the expressed viral capsid-heterologous peptide fusion product is expressed in soluble form in the cell.

[0189] In one embodiment, the expressed viral capsid-recombinant heterologous fusion product is assembled into VLPs in the cell.

[0190] In a separate embodiment, a portion of the expressed viral capsid-heterologous peptide fusion product of interest is formed in an insoluble aggregate in the cell. In one embodiment, the peptide of interest can be renatured from the insoluble aggregate.

[0191] Cleavage of Peptide of Interest

[0192] In one embodiment, the process further provides: e) cleaving the fusion product to separate the recombinant peptide from the capsid.

[0193] A cleavable linkage sequence can be included between the viral protein and the recombinant peptide. Examples of agents that can cleave such sequences include, but are not limited to chemical reagents such as acids (HCl, formic acid), CNBr, hydroxylamine (for asparagine-glycine), 2-Nitro-5-thiocyanobenzoate, O-Iodosobenzoate, and enzymatic agents, such as endopeptidases, endoproteases, trypsin, clostripain, and Staphylococcal protease.

[0194] Cleavable linkage sequences are well known in the art. In the present invention, any cleavable linkage sequence recognized by cleavage agents, including dipeptide cleavage sequences such as Asp-Pro, can be utilized.

[0195] Expression

[0196] The process of the invention optimally leads to the increased production of desired capsid-fusion products in a plant host cell in the desired sequence and conformation. The increased production alternatively can be an increased level of desired peptide per gram of protein produced, per gram of host protein, or as a percentage of total recombinant peptide produced. The increased production can also be an increased level of recoverable peptide, such as soluble protein, produced per gram of recombinant or per gram of host cell protein. The increased production can also be any combination of increased total level and increased active or soluble level of protein.

[0197] The improved expression of recombinant protein can be through expression of the protein as a capsid fusion protein and subsequently inserted in VLPs. In certain embodiments, at least 60, at least 70, at least 80, at least 90, at least 100, at least 110, at least 120, at least 130, at least 140, at least 150, at least 160, at least 170, or at least 180 copies of a peptide of interest are expressed in each VLP.

The VLPs can be produced and recovered from the cytoplasm, periplasm or extracellular medium of the host cell.

[0198] In another embodiment, the peptide can be insoluble in the cell. In certain embodiments, the soluble or insoluble peptide is produced in a particle formed of multiple capsids but not forming a native-type VLP. For example, a cage structure of as few as 3 viral capsids can be formed. In certain embodiments, the capsid structure includes more than one copy of a peptide of interest and in certain embodiments, includes at least ten, at least 20, or at least 30 copies. In certain embodiments, the peptide is formed in the vacuole of the plant cell.

[0199] The peptide or viral capsid sequence can also include one or more targeting sequences or sequences to assist purification. These can be an affinity tag. These can also be targeting sequences directing the assembly of capsids into a VLP.

[0200] Cell Growth

[0201] Transformation of the plant host cells with the vector(s) may be performed using any transformation methodology known in the art, and the plant host cells may be transformed as intact cells or as protoplasts. Exemplary transformation methodologies have been described above.

[0202] As used herein, the term "fermentation" includes both embodiments in which literal fermentation is employed and embodiments in which other, non-fermentative culture modes are employed. Fermentation may be performed at any scale. Any method generally known in the art may be utilized in fermentation. See, for example, Hellwig et al. (2004) "Plant cell cultures for the production of recombinant proteins," *Nature Biotech* 22(11): 1415-1422; Sajc et al. (2000) "Bioreactors for plant engineering: an outlook for further research," *Biochem Engin. J.* 4:89-99. In one embodiment, the fermentation medium may be selected from among any viable plant cell culture media. Optionally, additives to the medium can include, for example, PVP, BSA, NaCl, BrefeldinA, reducing manganese, pluronic anti-foam, polyethylene glycol, gelatin, protease inhibitors, redox co-factors, and dimethylsulfoxide.

[0203] The expression system according to the present invention can be cultured in any fermentation format. For example, batch, fed-batch, semi-continuous, and continuous fermentation modes may be employed herein.

[0204] The expression systems according to the present invention are useful for transgene expression at any scale (i.e. volume) of fermentation. Thus, e.g., microliter-scale, centiliter scale, and deciliter scale fermentation volumes may be used; and 1 Liter scale and larger fermentation volumes can be used. In one embodiment, the fermentation volume will be at or above 1 Liter. In another embodiment, the fermentation volume will be at or above 5 Liters, 10 Liters, 15 Liters, 20 Liters, 25 Liters, 50 Liters, 75 Liters, 100 Liters, 200 Liters, 500 Liters, 1,000 Liters, 2,000 Liters, 5,000 Liters, 10,000 Liters or 50,000 Liters.

[0205] In the present invention, growth, culturing, and/or fermentation of the transformed host cells is performed within a temperature range permitting survival of the host cells, preferably a temperature within the range of about 4° C. to about 55° C., inclusive.

[0206] Cell Density (Packed Wet Cells)

[0207] Plant cell expressions systems according to the present invention can provide a cell density or comparable packed wet cell density of about from about 20 g/L (2% packed cell volume) to more than 550 g/L (55% packed cell volume). In one embodiment, the packed cell volume is between 2% to about 60%. In another embodiment, the packed cell volume is 2%, 5%, 10%, 15%, 20%, 25%, 35%, 40%, 50%, 55%, 60%, 70%, or 85%.

[0208] Isolation of VLP or Peptide of Interest

[0209] In certain embodiments, the invention provides a process for improving the recovery of peptides of interest by protection of the peptide during expression through linkage and co-expression with a viral capsid. In certain embodiments, the viral capsid fusion forms a VLP, which can be readily separated from the cell lysate.

[0210] The proteins of this invention may be isolated and purified to substantial purity by standard techniques well known in the art, including, but not limited to, PEG precipitation, ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, nickel chromatography, hydroxylapatite chromatography, reverse phase chromatography, lectin chromatography, preparative electrophoresis, detergent solubilization, selective precipitation with such substances as column chromatography, immunopurification methods, and others. For example, proteins having established molecular adhesion properties can be reversibly fused to a ligand. With the appropriate ligand, the protein can be selectively adsorbed to a purification column and then freed from the column in a relatively pure form. The fused protein can then be removed by enzymatic or other activity. In addition, protein can be purified using immunoaffinity columns or Ni-NTA columns. General techniques are further described in, for example, R. Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag: N.Y. (1982); Deutscher, *Guide to Protein Purification*, Academic Press (1990); U.S. Pat. No. 4,511,503; S. Roe, *Protein Purification Techniques: A Practical Approach* (Practical Approach Series), Oxford Press (2001); D. Bollag, et al., *Protein Methods*, Wiley-Lisa, Inc. (1996); A K Patra et al., *Protein Expr Purif*, 18(2): p/182-92 (2000); and R. Mukhija, et al., *Gene* 165(2): p. 303-6 (1995). See also, for example, Ausubel, et al. (1987 and periodic supplements); Deutscher (1990) "Guide to Protein Purification," *Methods in Enzymology* vol. 182, and other volumes in this series; Coligan, et al. (1996 and periodic Supplements) *Current Protocols in Protein Science* Wiley/Greene, NY; and manufacturer's literature on use of protein purification products, e.g., Pharmacia, Piscataway, N.J., or Bio-Rad, Richmond, Calif. Combination with recombinant techniques allow fusion to appropriate segments, e.g., to a FLAG sequence or an equivalent which can be fused via a protease-removable sequence. See also, for example, Hochuli (1989) *Chemische Industrie* 12:69-70; Hochuli (1990) "Purification of Recombinant Proteins with Metal Chelate Absorbent" in Setlow (ed.) *Genetic Engineering, Principle and Methods* 12:87-98, Plenum Press, NY; and Crowe, et al. (1992) *QIAexpress: The High Level Expression & Protein Purification System* QUIAGEN, Inc., Chatsworth, Calif.

[0211] Similarly, the virus-like particles or cage-like structures can be isolated and purified to substantial purity by

standard techniques well known in the art. Techniques for isolation of VLPs include, in addition to those described above, precipitation techniques such as polyethylene glycol or salt precipitation. Separation techniques include anion or cation exchange chromatography, size exclusion chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, nickel chromatography, hydroxylapatite chromatography, reverse phase chromatography, lectin chromatography, preparative electrophoresis, immunopurification methods, centrifugation, ultracentrifugation, density gradient centrifugation (for example, on a sucrose or on a cesium chloride (CsCl) gradient), ultrafiltration through a size exclusion filter, and any other protein isolation methods known in the art.

[0212] The invention can also improve recovery of active recombinant peptides. Levels of active protein can be measured, for example, by measuring the interaction between an identified and a parent peptide, peptide variant, segment-substituted peptide and/or residue-substituted peptide by any convenient *in vitro* or *in vivo* assay. Thus, *in vitro* assays can be used to determine any detectable interaction between an identified protein and a peptide of interest, e.g. between enzyme and substrate, between hormone and hormone receptor, between antibody and antigen, etc. Such detection can include the measurement of calorimetric changes, changes in radioactivity, changes in solubility, changes in molecular weight as measured by gel electrophoresis and/or gel exclusion processes, etc. *In vivo* assays include, but are not limited to, assays to detect physiological effects, e.g. weight gain, change in electrolyte balance, change in blood clotting time, changes in clot dissolution and the induction of antigenic response. Generally, any *in vivo* assay can be used so long as a variable parameter exists so as to detect a change in the interaction between the identified and the peptide of interest. See, for example, U.S. Pat. No. 5,834, 250.

[0213] Detection of the expressed protein is achieved by methods known in the art and includes, for example, radio-immunoassays, Western blotting techniques or immunoprecipitation.

[0214] An initial salt fractionation can separate many of the unwanted host cell proteins (or proteins derived from the cell culture media) from the recombinant protein of interest. One such example can be ammonium sulfate. Ammonium sulfate precipitates proteins by effectively reducing the amount of water in the protein mixture. Proteins then precipitate on the basis of their solubility. The more hydrophobic a protein is, the more likely it is to precipitate at lower ammonium sulfate concentrations. A typical protocol includes adding saturated ammonium sulfate to a protein solution so that the resultant ammonium sulfate concentration is between 20-30%. This concentration will precipitate the most hydrophobic of proteins. The precipitate is then discarded (unless the protein of interest is hydrophobic) and ammonium sulfate is added to the supernatant to a concentration known to precipitate the protein of interest. The precipitate is then solubilized in buffer and the excess salt removed if necessary, either through dialysis or diafiltration. Other methods that rely on solubility of proteins, such as cold ethanol precipitation, are well known to those of skill in the art and can be used to fractionate complex protein mixtures.

[0215] The molecular weight of a recombinant protein can be used to isolate it from proteins of greater and lesser size using ultrafiltration through membranes of different pore size (for example, Amicon or Millipore membranes). As a first step, the protein mixture can be ultrafiltered through a membrane with a pore size that has a lower molecular weight cut-off than the molecular weight of the protein of interest. The retentate of the ultrafiltration can then be ultrafiltered against a membrane with a molecular cut off greater than the molecular weight of the protein of interest. The recombinant protein will pass through the membrane into the filtrate. The filtrate can then be chromatographed.

[0216] Recombinant proteins can also be separated from other proteins on the basis of its size, net surface charge, hydrophobicity, and affinity for ligands. In addition, antibodies raised against proteins can be conjugated to column matrices and the proteins immunopurified. All of these methods are well known in the art. It will be apparent to one of skill that chromatographic techniques can be performed at any scale and using equipment from many different manufacturers (e.g., Pharmacia Biotech).

[0217] Renaturation and Refolding

[0218] Insoluble protein can be renatured or refolded to generate secondary and tertiary protein structure conformation. Protein refolding steps can be used, as necessary, in completing configuration of the recombinant product. Refolding and renaturation can be accomplished using an agent that is known in the art to promote dissociation/association of proteins. For example, the protein can be incubated with dithiothreitol followed by incubation with oxidized glutathione disodium salt followed by incubation with a buffer containing a refolding agent such as urea.

[0219] Recombinant protein can also be renatured, for example, by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be refolded while immobilized on a column, such as the Ni NTA column by using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, and 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation can be performed over a period of 1.5 hours or more. After renaturation the proteins can be eluted by the addition of 250 mM imidazole. Imidazole can be removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein can be stored at room temperature, 4° C. or frozen at -20° C. to -80° C.

[0220] Other methods include, for example, those that may be described in M H Lee et al., *Protein Expr. Purif.*, 25(1): p. 166-73 (2002), W. K. Cho et al., *J. Biotechnology*, 77(2-3): p. 169-78 (2000), Ausubel, et al. (1987 and periodic supplements), Deutscher (1990) "Guide to Protein Purification," *Methods in Enzymology* vol. 182, and other volumes in this series, Coligan, et al. (1996 and periodic Supplements) *Current Protocols in Protein Science* Wiley/Greene, NY, S. Roe, *Protein Purification Techniques: A Practical Approach* (Practical Approach Series), Oxford Press (2001); D. Bollag, et al., *Protein Methods*, Wiley-Lisa, Inc. (1996)

[0221] Active Peptide Analysis

[0222] Active proteins can have a specific activity of at least 20%, 30%, or 40%, and preferably at least 50%, 60%, or 70%, and most preferably at least 80%, 90%, or 95% that

of the native peptide that the sequence is derived from. Further, the substrate specificity (k_{cat}/K_m) is optionally substantially similar to the native peptide. Typically, k_{cat}/K_m will be at least 30%, 40%, or 50%, that of the native peptide; and more preferably at least 60%, 70%, 80%, or 90%. Methods of assaying and quantifying measures of protein and peptide activity and substrate specificity (k_{cat}/K_m), are well known to those of skill in the art.

[0223] The activity of a recombinant peptide produced in accordance with the present invention by can be measured by any protein specific conventional or standard in vitro or in vivo assay known in the art. The activity of the plant cell host produced recombinant peptide can be compared with the activity of the corresponding native protein to determine whether the recombinant protein exhibits substantially similar or equivalent activity to the activity generally observed in the native peptide under the same or similar physiological conditions.

[0224] The activity of the recombinant protein can be compared with a previously established native peptide standard activity. Alternatively, the activity of the recombinant peptide can be determined in a simultaneous, or substantially simultaneous, comparative assay with the native peptide. For example, an in vitro assays can be used to determine any detectable interaction between a recombinant peptide and a target, e.g. between an expressed enzyme and substrate, between expressed hormone and hormone receptor, between expressed antibody and antigen, etc. Such detection can include the measurement of colorimetric changes, proliferation changes, cell death, cell repelling, changes in radioactivity, changes in solubility, changes in molecular weight as measured by gel electrophoresis and/or gel exclusion methods, phosphorylation abilities, antibody specificity assays such as ELISA assays, etc. In addition, in vivo assays include, but are not limited to, assays to detect physiological effects of the plant host cell produced peptide in comparison to physiological effects of the native peptide, e.g. antigenic response. Generally, any in vitro or in vivo assay can be used to determine the active nature of the recombinant peptide that allows for a comparative analysis to the native peptide so long as such activity is assayable. Alternatively, the peptides produced in the present invention

can be assayed for the ability to stimulate or inhibit interaction between the peptide and a molecule that normally interacts with the peptide, e.g. a substrate or a component of the signal pathway that the native protein normally interacts. Such assays can typically include the steps of combining the protein with a substrate molecule under conditions that allow the peptide to interact with the target molecule, and detect the biochemical consequence of the interaction with the protein and the target molecule.

[0225] Assays that can be utilized to determine peptide activity are described, for example, in Ralph, P. J., et al. (1984) *J. Immunol.* 132:1858 or Saiki et al. (1981) *J. Immunol.* 127:1044, Steward, W. E. II (1980) *The Interferon Systems*. Springer-Verlag, Vienna and New York, Broxmeyer, H. E., et al. (1982) *Blood* 60:595, "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987, AK Patra et al., *Protein Expr Purif.* 18(2): p/182-92 (2000), Kodama et al., *J. Biochem.* 99: 1465-1472 (1986); Stewart et al., *Proc. Nat'l Acad. Sci. USA* 90: 5209-5213 (1993); (Lombillo et al., *J. Cell Biol.* 128:107-115 (1995); (Vale et al., *Cell* 42:39-50 (1985).

EXAMPLES

[0226] Viral Source of Coat Protein-Encoding Nucleic Acid

[0227] In the following examples, the cowpea chlorotic mottle virus (CCMV) has been used as the source of the coat protein for expression of the desired recombinant peptides. CCMV is a member of the bromovirus group of the Bromoviridae. Bromoviruses are 25-28 nm diameter icosahedral viruses with a four-component, positive sense, single-stranded RNA genome. RNA1 and RNA2 code for replicase enzymes. RNA3 codes for a protein involved in viral movement within plant hosts. RNA4 (a subgenomic RNA derived from RNA 3), i.e. sgRNA4, codes for the 20 kDa coat protein (CP) (SEQ ID NO:13) (Table 1). Each CCMV particle contains 180 copies of the CCMV CP. An exemplary DNA sequence encoding the CCMV CP is shown in SEQ ID NO:14 (Table 2).

TABLE 1

Wild type CCMV coat protein encoded by sgRNA4

(Seq. ID No. 13)

```
Met Ser Thr Val Gly Thr Gly Lys Leu Thr Arg Ala Gln Arg Arg Ala Ala Ala Arg Lys
Asn Lys Arg Asn Thr Arg Val Val Gln Pro Val Ile Val Glu Pro Ile Ala Ser Gly Gln
Gly Lys Ala Ile Lys Ala Trp Thr Gly Tyr Ser Val Ser Lys Trp Thr Ala Ser Cys Ala
Ala Ala Glu Ala Lys Val Thr Ser Ala Ile Thr Ile Ser Leu Pro Asn Glu Leu Ser Ser
Glu Arg Asn Lys Gln Leu Lys Val Gly Arg Val Leu Leu Trp Leu Gly Leu Leu Pro
Ser Val Ser Gly Thr Val Lys Ser Cys Val Thr Glu Thr Gln Thr Thr Ala Ala Ala Ser
Phe Gln Val Ala Leu Ala Val Ala Asp Asn Ser Lys Asp Val Val Ala Ala Met Tyr Pro
Glu Ala Phe Lys Gly Ile Thr Leu Glu Gln Leu Thr Ala Asp Leu Thr Ile Tyr Leu Tyr
Ser Ser Ala Ala Leu Thr Glu Gly Asp Val Ile Val His Leu Glu Val Glu His Val Arg
Pro Thr Phe Asp Asp Ser Phe Thr Pro Val Tyr
```

[0228]

TABLE 2

Exemplary DNA Sequence Encoding the CCMV CP																												
(Seq. ID. No. 14)																												
atg	tct	aca	gtc	gga	aca	ggg	aag	tta	act	cgt	gca	caa	cga	agg	gct	gcg	gcc	cgt	aag	aac	aag							
cgg	aac	act	cgt	gtg	gtc	caa	cct	gtt	att	gta	gaa	ccc	atc	gct	tca	ggc	caa	ggc	aag	gct	att	aaa						
gca	tgg	acc	ggt	tac	agc	gta	tcg	aag	tgg	acc	gcc	tct	tgc	gcg	gcc	gcc	gaa	gct	aaa	gta	acc							
tcg	gct	ata	act	atc	tct	ctc	cct	aat	gag	cta	tcg	tcc	gaa	agg	aac	aag	cag	ctc	aag	gta	ggt	aga						
gtt	tta	tta	tgg	ctt	ggg	ttg	ctt	ccc	agt	gtt	agt	ggc	aca	gtg	aaa	tcc	tgt	gtt	aca	gag	acg	cag						
act	act	gct	gct	gcc	tcc	ttt	cag	gtg	gca	tta	gct	gtg	gcc	gac	aac	tcg	aaa	gat	gtt	gtc	gct	gct						
atg	tac	ccc	gag	gcg	ttt	aag	ggt	ata	acc	ctt	gaa	caa	ctc	acc	gcg	gat	tta	acg	atc	tac	ttg	tac						
agc	agt	gcg	gct	ctc	act	gag	ggc	gac	gtc	atc	gtg	cat	ttg	gag	gtt	gag	cat	gtc	aga	cct	acg	ttt						
gac	gac	tct	ttc	act	ccg	gtg	tat	tag																				

[0229] The crystal structure of CCMV has been solved. This structure provides a clearer picture of the coat protein

(San Francisco, Calif., USA) according to the following protocol:

TABLE 3

PCR protocol			
Reaction Mix (100 μ L total volume)		Thermocycling Steps	
10 μ L 10 $\times$ PT HIFI buffer *	Step 1	1 Cycle	2 min. 94° C.
4 μ L 50 mM MgSO ₄ *	Step 2	35 Cycles	30 sec. 94° C.
2 μ L 10 mM dNTPs *			30 sec. 55° C.
0.25 ng Each Primer			1 min. 68° C.
1-5 ng Template DNA	Step 3	1 Cycle	10 min. 70° C.
1 μ L PT HIFI Taq DNA Polymerase *	Step 4	1 Cycle	Maintain 4° C.
Remainder Distilled De-ionized H ₂ O (ddH ₂ O)			

* (from Invitrogen Corp, Carlsbad, CA, USA, hereinafter "Invitrogen")

interactions that appear to be critical to particle stability and dynamics and has been helpful in guiding rational design of insertion sites. Previous studies have demonstrated that CCMV coat proteins can be genetically modified to carry heterologous peptides without interfering with their ability to form particles. A number of suitable insertion sites have been identified. A total of up to 180 copies of a heterologous peptide unit (whether as individual peptide units or in concatameric units) can be inserted into the CCMV particle at a single insertion site in the CCMV CP. Insertion sites identified within CCMV CP to date can accommodate peptides of various lengths.

[0230] Materials and Methods

[0231] Unless otherwise noted, standard techniques, vectors, control sequence elements, and other expression system elements known in the field of molecular biology are used for nucleic acid manipulation, transformation, and expression. Such standard techniques, vectors, and elements can be found, for example, in: Ausubel et al. (eds.), *Current Protocols in Molecular Biology* (1995) (John Wiley & Sons); Sambrook, Fritsch, & Maniatis (eds.), *Molecular Cloning* (1989) (Cold Spring Harbor Laboratory Press, NY); Berger & Kimmel, *Methods in Enzymology* 152: Guide to Molecular Cloning Techniques (1987) (Academic Press); and Bukhari et al. (eds.), *DNA Insertion Elements, Plasmids and Episomes* (1977) (Cold Spring Harbor Laboratory Press, NY).

[0232] Unless noted otherwise, PCR reactions were performed using a PTC225 thermocycler (MJ Research, South

Example 1

Production of Antigenic Peptides in CCMV Virus Particles in Whole Cowpea Plants Inoculated with CCMV RNA1, RNA2, and Chimeric RNA3

[0233] Expression of *Bacillus anthracis* antigenic peptides was performed in whole plants, using cowpea chlorotic mottle virus (CCMV) with capsid proteins (CP) engineered to contain one of four different antigenic peptides.

[0234] DNA having the nucleotide sequence of CCMV RNA 1 and DNA having the nucleotide sequence of CCMV RNA 2 were each separately subcloned in the cloning vector pUC19 downstream from, and under the control of, a T7 promoter and upstream from a unique Xba I site. This produced plasmids pDOW2122 (CCMV RNA1) and pDOW2123 (CCMV RNA2).

[0235] DNA having the nucleotide sequence of CCMV RNA 3, engineered to contain five BamH I restriction enzyme cleavage sites, was further engineered for production of the recombinant capsid protein-encoding nucleic acid. Four DNA molecules, each encoding a different one of four exogenous peptides (four different antigenic peptides from *Bacillus anthracis* Protective Antigen PA) were each synthesized by SOE (splicing-by-overlap-extension) of synthetic oligonucleotides. The resulting nucleic acids contained BamHI recognition site termini. Each PA DNA fragment was restricted with BamHI restriction enzyme and independently inserted into the coat protein at one of the five different engineered restriction enzyme cleavage sites in the

CCMV coat protein coding sequence: BamHI at codon 63, BamHI at codon 102, BamHI at codon 114, BamHI at codon 129, and BamHI at codon 160.

[0236] The four different peptides were PA1 (SEQ ID NO: 1 encoded by SEQ ID NO: 15), PA2 (SEQ ID NO: 2, encoded by SEQ ID NO: 16), PA3 (SEQ ID NO: 3, encoded by SEQ ID NO: 17), and PA4 (SEQ ID NO: 4, encoded by SEQ ID NO: 18).

CHINE T7 Kit (RNA transcription kit, from Ambion, Inc., Austin, Tex., USA). This produced 24 different RNA varieties: one for RNA1, one for RNA2, 20 for chimeric RNA3, and two for RNA3 controls.

[0240] Cocktail mixes of RNA1, RNA2, RNA3 or chimeric RNA3s containing a PA insert were used to infect cowpea plants. Cowpea plants were sprouted from Cowpea California Blackeye #5 seeds (Ferry-Morse Seed Co. KY).

TABLE 4

<i>Bacillus anthracis</i> antigenic nucleic acid and amino acid sequences.	
PA1	
Nucleic Acid Sequence (SEQ ID NO: 15)	5'-agt aat tct cgt aag aaa cgt tct acc tct gct ggc cct acc gtg cct gat cgt gat aat gat ggc aft cct gat-3'
Amino Acid Sequence (SEQ ID NO: 1)	Ser Asn Ser Arg Lys Lys Arg Ser Thr Ser Ala Gly Pro Thr Val Pro Asp Arg Asp Asn Asp Gly Ile Pro Asp
PA2	
Nucleic Acid Sequence (SEQ ID NO: 16)	5'-agt cct gaa gct cgt cat cct ctc gtg gct gcg tat cct att gtg cat gtt gat atg gaa aat aft atc ctc tct-3'
Amino Acid Sequence (SEQ ID NO: 2)	Ser Pro Glu Ala Arg His Pro Leu Val Ala Ala Tyr Pro Ile Val His Val Asp Met Glu Asn Ile Ile Leu Ser
PA3	
Nucleic Acid Sequence (SEQ ID NO: 17)	5'-cgt att aft ftc aat ggc aaa gat ctc aat ctc gtg gaa cgt cgt att gct gct gtg aat cct tct gat cct ctc -3'
Amino Acid Sequence (SEQ ID NO: 3)	Arg Ile Ile Phe Asn Gly Lys Asp Leu Asn Leu Val Glu Arg Arg Ile Ala Ala Val Asn Pro Ser Asp Pro Leu
PA4	
Nucleic Acid Sequence (SEQ ID NO: 18)	5'-cgt caa gat ggc aaa acc ftc aft gat ftc aaa aag tat aat gat aaa ctc cct ctc tat aft tct aat cct aat-3'
Amino Acid Sequence (SEQ ID NO: 4)	Arg Gln Asp Gly Lys Thr Phe Ile Asp Phe Lys Lys Tyr Asn Asp Lys Leu Pro Leu Tyr Ile Ser Asn Pro Asn

[0237] Each of these was inserted into each of the five plasmids: pDOW2125 (pUC-CCMV RNA3-CP63BamHI), pDOW2126 (pUC-CCMV-RNA3-102BamHI), pDOW2127 (pUC-CCMV-RNA3-CP114BamHI), pDOW2128 (pUC-CCMV-RNA3-CP129BamHI), and pDOW2129 (pUC-CCMV-RNA3-CP160BamHI) which had been digested with BamHI restriction enzyme and then dephosphorylated.

[0238] This produced a total of 20 different versions of DNA for "chimeric RNA 3." Plasmids pDOW2135 (pUC-CCMV-RNA3-CP63BamHI-PA1), pDOW2139 (pUC-CCMV-RNA3-CP102BamHI-PA1), pDOW2143 (pUC-CCMV-RNA3-CP114BamHI-PA1), pDOW2147 (pUC-CCMV-RNA3-CP129BamHI-PA1), and pDOW2151 (pUC-CCMV-RNA3-CP160BamHI-PA1) are examples of chimeric RNA3 having a PA1 insert at position 63, 102, 114, 129, and 160. Wild-type CCMV coat protein-encoding RNA 3 and engineered CCMV coat protein-encoding RNA 3 containing BamHI restriction site but no inserts were used as controls. Each modified RNA3 construct was separately sub-cloned in the cloning vector pUC19 downstream from and under the control of a T7 promoter, and upstream from a unique Xba I site.

[0239] Each of the three classes of plasmids was cloned in *E. coli*. Plasmids were isolated, linearized by Xba I restriction enzyme digestion, and then one microgram of each was transcribed into RNA in vitro using a mMESSAGE mMMA-

Sprouts were transplanted singly into 6 inch pots with Miracle-Gro potting mix (Miracle-Gro Lawn Products OH). Cowpea plants were infected at 2-leaf stage (approximately 7 days post germination). A dusting of Carborundum powder 400grit (Fisher Scientific cat.409-21-2) was applied onto one leaf of each plant. RNA cocktail mixes were applied onto the carborundum layer. Leaves were abraded by gentle rubbing with a gloved finger. Infections were established 7-14 days post inoculation.

[0241] Unexpectedly, in this whole plant system, even though the native host plant was used, the chimeric coat proteins reverted to the wild type or in some cases, the desired exogenous antigen peptide was partially deleted. This was verified by sequencing the virus progeny. The RNA was extracted and reverse-transcribed into DNA using ThermoScript RT-PCR system (Invitrogen cat.11146-024) and CCMVRNA4.R gene specific primer (SEQ ID NO:19) (5'-CTC GAG CTA ATA CAC CGG AGT GAA AG-3'). cDNAs were further amplified by PCR using primers CCMVRNA4.F (SEQ ID NO:20) (5'-CTG CAG ATG TCT ACA GTC GGA ACA GG-3') and CCMVRNA4.R. The reactions were purified and sequenced with CCMV-CP-F1 primer, (SEQ ID NO: 21) (5'-AAC CCA TCG CTT CAG GCC AA-3'). Sequences were analyzed with Sequencher software version 4.0.5 (Gene Codes Corporation). For example, upon sequencing of RNA3 progeny derived from pDOW2128 containing the PA1 insert, the coat protein

contained a small, unrecognized nucleotide insert (5'-CG-TATTTCTGATCCTCTC-3') that was translated as RISDPL instead of the PA1 insert amino acid sequence. The inserted nucleotide sequence kept the rest of the coat protein in frame. It has been determined that the engineered CCMV RNA3 is unstable upon expression in the whole plant, easily undergoes recombinations, and that the whole plant system is not useful for production of epitopes on chimeric CCMV virus particles.

Example 2

Production of Antigenic Peptides in CCMV Virus Particles in Tobacco Suspension Culture Inoculated with CCMV RNA1, RNA2, and Chimeric RNA3

[0242] Expression of the same CCMV-peptide-encoding constructs was performed in plant cell suspension culture. *Nicotiana tabacum* NT1 cells were transfected by electroporation with RNA transcripts of CCMV RNA 3 coding for the chimeric CCMV coat proteins and CCMV RNA 1 and 2 coding for the replicase genes. Wild-type CCMV coat protein-encoding RNA 3 and engineered CCMV coat protein-encoding RNA 3 containing the appropriate BamHI restriction site but no inserts were used as controls.

[0243] 24 different RNA varieties were obtained by in vitro RNA transcription as described in the Example 1: one for RNA1, one for RNA2, 20 for chimeric RNA3, and two for RNA3 controls. In 22 different groups, two micrograms of each of three resulting RNAs (RNA1, RNA2, and one of the RNA3s) were transformed into tobacco cells by electroporation.

[0244] The Following Protocol was Used for Plant Cell Transfection:

[0245] 1) Media Preparation:

[0246] NTI Media (1 Liter):

[0247] 4.33 g Murashige & Skoog basal salt (Phyto Technology Laboratories KS cat.M524), 100 mg Myo-Inositol (Sigma cat.I-3011), 1 ml of 1 mg/ml solution of Thiamine HCl (Sigma cat. T-3902), 180 mg Potassium Phosphate Monobasic KH_2PO_4 (Sigma cat. P-8416), 30 g Sucrose (Sigma cat.S-5390), and 200 μl 2,4-D solution of 10 mg/ml (Sigma cat. D-7299) were mixed in a small amount of water. Purified water was added to the solution to bring volume up to 1 liter. The pH was adjusted to 5.8, and the solution was autoclaved.

[0248] Mannitol Wash Solution 0.4M:

[0249] 36.43 Mannitol (Sigma cat. M-1902) was added to small amount of purified water. Purified water was added to bring volume up to 500 ml and the pH was adjusted to 5.5. The solution was autoclaved in order to sterilize.

[0250] Enzyme Solution:

[0251] 0.4M mannitol, and 0.02M MES were mixed in a small amount of water. Purified water was added to bring volume up to 500 ml and the pH was adjusted to 5.5. The solution was autoclaved in order to sterilize, and the solution was stored at 4° C. Prior to use, 1% cellylysin (Calbiochem cat.219466) and 0.3% Macerace Pectinase (Calbiochem cat.441201) was added to solution, and the solution was shaken until cellylysin and Macerace Pectinase were dissolved.

[0252] Electroporation Buffer:

[0253] 0.8% NaCl, 0.02% KCl, 0.02% KH_2PO_4 , and 0.11% Na_2HPO_4 were added to a small amount of purified water. Purified water was added to bring the solution up to 100 mls, and the pH was adjusted to 6.5. The buffer was autoclaved and stored at 4° C.

[0254] 2) Tobacco Cell Partial Digestion Protocol:

[0255] Tobacco cell line NT 1 was partially digested prior to electroporation as following. A robust cell line was maintained by sub-culturing weekly, cells were grown in NTI media at 24° C. or 28° C. with gentle shaking, and three days before digestion, 5 ml of cell suspension was sub-cultured into 50 ml of NTI media and incubated at 28° C. Cells were spun down in Soml tubes for 5 minutes at 800 rpm. The enzyme solution was prepared, and cells were washed in the mannitol wash solution (approx. 40 ml). Cells were spun for 5 minutes at 800 rpm, and 3 volumes of enzyme solution were added to the cells. Cells were resuspended by inversion; transferred into 10 cm petri dishes by pouring, and the dish was shaken very slowly at room temperature while wrapped in aluminum foil for 60-120 minutes. The cells were then transferred back into 50 ml plastic tubes; and spun for 5 minutes at 800 rpm. Cells were washed with 40 ml mannitol wash solution and spun again, washed with electroporation buffer, and spun. Three volumes of electroporation buffer (total volume is usually 20 ml) were added to the cells; and the cells were stored at 4° C. wrapped in aluminum foil.

[0256] 3) Electroporation of Partially Digested Cells Protocol:

[0257] 1 ml of digested cells were aliquoted into electroporation cuvettes (4 mm gap). 2 μg each RNA transcript—CCMV RNA1, RNA2, and Chimeric RNA3—were added to the cuvettes and placed on ice for 5 minutes. 10 ml of NTI plating media were added (NTI+0.4M mannitol) to each Petri dish. The cells were electroporated at 500 μF , 250V, and the cuvettes were placed back on ice. The cells were then transferred into Petri dishes, and incubate at room temperature in darkness with no shaking. The cells were then collected for analysis at 48 hours post transfection.

[0258] Results

[0259] Expression of chimeric coat protein was analyzed by Western blot using anti-CCMV coat protein polyclonal antibody. Also, RNA was extracted from each of the cultures and the chimeric CCMV RNA4 was reverse transcribed into cDNA, which was then amplified by PCR. The PCR products were sequenced as described in Example 1. All samples were positive on the western blot except for the negative control that was transfected with only RNA1 and 2.

[0260] Compared to control CCMV capsid proteins, the chimeric coat proteins had greater size and slower mobility on the gel indicating that the chimeric coat proteins contained PA inserts. This was verified by sequencing of the viral RNA progeny. Results demonstrated that 19 out of 20 chimeric constructs expressed the chimeric CCMV coat protein properly without mutation in the desired antigen peptide. The results indicate that chimeric CCMV RNA3s expressed in suspension cells are stable.

[0261] FIG. 2 shows the expression of CCMV CP in cells transfected with CCMV RNA1, RNA2, and RNA3 tran-

scribed from pDOW2125 (CCMV63BamHI), pDOW2126 (CCMV102BamHI), pDOW2127 (CCMV114BamHI), pDOW2128 (CCMV129BamHI) and pDOW2129 (CCMV160BamHI). The results demonstrate that CCMV CP is expressed from all RNA3s with engineered BamHI site in different positions in the CP.

[0262] All 4 PA peptides (PA1, PA2, PA3, and PA4) fused to CCMV coat protein at the 63, 102, 114, 129, and 160 BamHI site were successfully produced in tobacco cells. **FIG. 3** and **FIG. 4** show the expression of chimeric CCMV CP in cells transfected with CCMV RNA1, RNA2, and chimeric RNA3 containing CP with the four PA peptide inserts in the position 63, 102, 114, 129, and 160. **FIG. 3** demonstrates that the fusion protein of CCMV CP with PA1, PA2, PA3, or PA4 peptide that has been inserted at the position 63, 102, or 114, is expressed in the tobacco plant cells with the exception of PA2 peptide that has been inserted into the CCMV CP at the position 63. **FIG. 4** demonstrates that the fusion protein of CCMV CP with PA1, PA2, PA3, or PA4 peptide that has been inserted at the position 129 or 160, is expressed in the tobacco plant cells.

Example 3

Production of 4 Antigenic Peptides in CCMV Virus-like Particles in Tobacco Suspension Culture Transfected with Plant Expression Plasmids Encoding the Chimeric CCMV CPs

[0263] 1) Vector Construction:

[0264] Plasmid pIL-Tab358 was used as the plant expression vector. Restriction sites chosen for CCMV CP insertion were XbaI and EcoRI. This plasmid contains the Cassava Vein Mosaic Virus promoter upstream of the XbaI site and a Nos terminator downstream of the EcoRI site. The vector was prepared by digestion with XbaI and EcoRI and dephosphorylated before ligation with the inserts.

[0265] 2) Insert Construction:

[0266] CCMV CP-PA fusions were amplified by PCR out of pDOW2147 (PUC-CCMV-RNA3-CP129BamHI-PA1), pDOW2148 (pUC-CCMV-RNA3-CP129BamHI-PA2), pDOW2149 (pUC-CCMV-RNA3-CP 129BamHI-PA3), pDOW2150 (pUC-CCMV-RNA3-CP 129BamHI-PA4) using primers CCMV-CP-XbaI (SEQ ID NO:22) and CCMV-CP-EcoRI (SEQ ID NO: 23) to create pDOW2160 (pIL-Tab-CCMV129BamHI-PA1) (SEQ ID NO: 5), pDOW2161 (pIL-Tab-CCMV129BamHI-PA2) (SEQ ID NO:6), pDOW2162 (pIL-Tab-CCMV129BamHI-PA3) (SEQ ID NO:7), and pDOW2163 (pIL-Tab-CCMV129BamHI-PA4) (SEQ ID NO:8). Plasmid maps for pDOW2160, pDOW 2161, pDOW2162, and pDOW 2163 are shown in **FIGS. 5, 6, 7, and 8.**

[0267] 3) Plant Cell Transfection:

[0268] Plant cells were transfected with 10%g of plasmid pDOW2160, pDOW2161, pDOW2162, and pDOW2163. Plant cell transfection was performed as in Example 2.

[0269] Results

[0270] All four PA peptides-CCMV CP fusions at 129BamHI site were successfully expressed in tobacco cells. **FIG. 9** shows the expression of chimeric CCMV CP in cells transfected with pDOW2160, pDOW2161, pDOW2162,

and pDOW2163. Compared to the control (CCMV coat protein with no insert), all four chimeric coat proteins showed slower mobility indicative of having PA inserts.

Example 4

Production of 4 Antigenic Peptides as CCMV CP fusions in Virus-like Particles in Tobacco Suspension Culture Transfected with Plant Expression Plasmids Encoding the Chimeric CCMV CPs and 3'UTR (CCMV RNA3 3' untranslated region)

[0271] 1) Vector Construction.

[0272] Plasmid pIL-Tab358 was used as the plant expression vector. Restriction sites chosen for CCMV CP insertion were XbaI and EcoRI. This plasmid contains the Cassava Vein Mosaic Virus promoter upstream of the XbaI site and a Nos terminator downstream of the EcoRI site. The vector was prepared by digestion with XbaI and EcoRI and dephosphorylated before ligation with the inserts.

[0273] 2) Insert Construction:

[0274] CCMV CP-PA fusions containing the 3'UTR of RNA3 were amplified by PCR out of pDOW2147, pDOW2148, pDOW2149, and pDOW2150 using primers CCMV-CP-XbaI and CCMV-CP-EcoRI-3'UTR (SEQ ID NO: 24) to create pDOW2169 (pIL-Tab-CP129BamHI-PA1-3'UTR) (SEQ ID NO: 9), pDOW2170 (pIL-Tab-CP129BamHI-PA2-3'UTR) (SEQ ID NO:

[0275] 10), pDOW2171 (pIL-Tab-CP129BamHI-PA3-3'UTR) (SEQ ID NO: 11), and pDOW2172 (pIL-Tab-CP 129BamHI-PA4-3'UTR) (SEQ ID NO: 12).

[0276] 3) Plant Cell Transfection:

[0277] Plant cells were transfected with 10%g of plasmid pDOW2169, pDOW2170, pDOW2171, and pDOW2172. Plant cell transfection was performed as in Example 2.

[0278] Results

[0279] All four PA peptides-CCMV CP fusions at 129BamHI site containing the 3'UTR were successfully expressed in tobacco cells. **FIG. 10** shows the expression of chimeric CCMV CP in cells transfected with pDOW2169, pDOW2170, pDOW2171, and pDOW2172. Compared to the control (CCMV coat protein with no insert), all four chimeric coat proteins showed slower mobility indicative of having PA inserts.

Example 5

Production of 4 Antigenic Peptides as CCMV CP fusions in Virus-like Particles in Transgenic Tobacco Suspension Cell Culture Transformed with Plant Expression Plasmids Encoding the Chimeric CCMV CPs

[0280] Tobacco cell line NT1 was maintained by subculturing weekly. Cells were grown in NT1 media at 24° C. or 28° C. with gentle shaking. Three days before transformation, 5 ml of cell suspension was subcultured into 50 ml of NT1 media and incubated at 28° C.

[0281] 1) Plant Cell Transformation by Particle Bombardment:

[0282] Plasmids pDOW2160, pDOW2161, pDOW2162, pDOW2163 and plant expression plasmid pBBV containing plant selectable marker Pat driven by the Cassava Vein Mosaic Virus promoter (pBBV) were used for NT1 tobacco cells transformation by micro-particle bombardment. The plasmid ratio used was 1:6 pBBV:pDOW2160, pBBV:pDOW2161, pBBV:pDOW2162, or pBBV:pDOW2163. Total of 5 ug of plasmid DNAs were sufficient for 6 bombardments. Biorad bombardment protocol as described in Chen, L et al. Plant Cell Reports (1998) 18: 25-31 was used for transformation.

[0283] 2) Plating

[0284] Bombarded cells were transferred on non-selective NT1 media agar plates for 4 hours before transferring to NT1 media agar plate with 25 ug/ml of Glufosinate-ammonium (Sigma cat#45520).

[0285] 3) Selecting Transgenic Calli.

[0286] After 21 days, calli that had white fluffy cell growth were selected for analysis by western blotting to test for the expression of CP fusions and PCR to test for integration of promoter-CP fusion gene-terminator cassette into the plant genome. Transgenic calli expressing chimeric CCMV CPs were transferred to liquid NT1 media. Cells were grown in NT1 media at 24° C. or 28° C. with gentle shaking and subcultured weekly.

[0287] Results

[0288] FIG. 11 shows the expression of chimeric CCMV CP in cells stably transformed with pDOW2160, pDOW2161, pDOW2162, and pDOW2163. The CP fusion was detected by polyclonal antibodies for CCMV. Expression of all four chimeric coat protein transgenes was detected.

Example 6

Production of 4 Antigenic Peptides as CCMV CP fusions in CCMV Virus-like Particles in Transgenic Rice Suspension Culture Transformed with Plant Expression Plasmids Encoding the Chimeric CCMV CPs

[0289] Rice cell lines were maintained by sub-culturing weekly. Cells were grown in NB media (Li, L et al. Plant Cell Reports. (1993) 12: 250-255) at 28° C. with gentle shaking. Three days before transformation, 5 ml of cell suspension was sub-cultured into 50 ml of NB media and incubated at 28° C.

[0290] 1) Plant Cell Transformation by Particle Bombardment.

[0291] Plasmids pDOW2160, pDOW2161, pDOW2162, and pDOW2163 and plant expression plasmid pBBV containing plant selectable marker Pat driven by the Cassava Vein Mosaic Virus promoter were used for rice cells transformation by microparticle bombardment. The plasmid ratio used was 1:6 pBBV:pDOW2160, pBBV:pDOW2161, pBBV:pDOW2162, or pBBV:pDOW2163. Total of 5 ug of plasmid DNAs were sufficient for 6 bombardments. Biorad

bombardment protocol as described in Chen, L et al. Plant Cell Reports (1998) 18: 25-31 was used for transformation.

[0292] 2) Plating.

[0293] Bombarded cells were transferred to NB media agar plates with 25 ug/ml of Glufosinate-ammonium (Sigma cat#45520) for selection.

[0294] 3) Selecting Transgenic Calli.

[0295] After 21 days, calli that had white fluffy cell growth were selected for analysis by western blotting to test for the expression of CP fusions and PCR to test for integration of promoter-CP fusion gene-terminator cassette into the plant genome.

[0296] Results

[0297] FIG. 12 shows the expression of chimeric CCMV CP in cells stably transformed with pDOW2160, pDOW2161, pDOW2162, and pDOW2163 as detected by polyclonal antibodies to CCMV. Expression of all four chimeric coat protein transgenes was detected.

[0298] The following protocol has been used to extract genomic DNA and test for integration of promoter-CP fusion gene-terminator cassette into the plant genome by PCR. Qiagen DNeasy Plant Mini was used to extract genomic DNA following the manufacturer directions.

[0299] Approximately 50-100 mg of fresh rice callus was placed into a 1.5 ml tube. The tissue was frozen in liquid N₂ for immediate DNA extraction or placed in -80° C. freezer for storage. Prior to DNA extraction the tissue was manually disrupted by grinding with a micropestle. The sample was placed on ice and the genomic DNA was extracted as described in Qiagen DNeasy Plant Mini Kit manual. The Access Quick Master Mix 2x (Promega cat#A1720) was used to set up the PCR reaction as follows:

Master mix	25 ul
CCMV-F (10 mM) primer, SEQ ID NO: 25	1 ul
Nos-term-R (10 mM) primer, SEQ ID NO: 26	1 ul
Genomic DNA	1 ul
water	22 ul
Total	50 ul

[0300] The following PCR cycle has been used to amplify the promoter-CP fusion gene-terminator cassette from the genomic DNA:

1 cycle	95° C.	2 min
35 cycles of	95° C.	30 sec
	55° C.	30 sec
	70° C.	2 min
1 cycle	70° C.	5 min
	4° C.	hold

[0301] 10 ul of PCR reaction was run on 1.2% agarose gel stained with EtBr. FIG. 13 shows a PCR products amplified from selected individual rice calli. The samples that contained a PCR product of predicted size were scored as

positive for stable transformation and integration of the chimeric CP transgene into the plant genome.

[0302] Transgenic calli expressing chimeric CCMV CPs were transferred to liquid NB media to create a cell suspension culture suitable for scale up fermentation. Cells were grown in NB media at 28° C. with gentle shaking and subcultured weekly.

[0303] VLP Extraction:

[0304] Chimeric VLPs were precipitated by lysis of shake-flask culture samples, followed by PEG (polyethylene glycol)-treatment of the resulting cell lysates and ultrafiltration, according to the following protocol:

[0305] 50 mL aliquots of each shake-flask culture were centrifuged to pellet the cells. Pelleted cells were resuspended in virus buffer (0.2M Sodium Acetate pH 5.2; 10 mM EDTA.0) at a 2 volume buffer to 1 volume pellet ratio. Cells were then disrupted by blending for 60 sec multiple times, with 2 minutes resting on ice in between. The resulting homogenate was squeezed through 3 layers of cheese cloth and was then centrifuged for 15 min at 15,000×G at 4° C. The resulting supernatants were removed and their volumes measured. To each supernatant, PEG8000 was added to a final concentration of 10% and the solution was incubated on ice for 1 hr or overnight at 4° C. Then, the

solution was centrifuged at 15,000×G for 10 min at 4° C. Precipitated pellets were then resuspended in 1/10 initial supernatant volume of virus buffer and stored at 4° C. and analyzed by western blotting with polyclonal anti-CCMV antibodies.

[0306] FIG. 14 shows that the chimeric CCMV VLPs were recovered from cells stably transformed with pDOW2160. Chimeric CP-PA1 fusion proteins were detected in the resuspended PEC pellets but not the supernatant indicating the chimeric CP assembled into VLPs.

[0307] Alternatively, the resuspended samples were centrifuged for 10 min at 15,000×G at 4° C., the supernatant was recovered and subjected to the second round of PEG precipitation. PEG8000 was added to final concentration of 15% and stirred at 4° C. for 2 hours. The solution was then centrifuged at 15,000×G for 10 mins and the pellet was resuspended in small volume of virus buffer. The resuspended VLP solution was loaded on to Centricon Plus-20 with 300K molecular weight cut-off and spun at 4,000×G for 5 mins. The concentrated VLP sample was then analyzed by western blotting with polyclonal anti-CCMV antibodies (see FIG. 14). Chimeric CP-PA1 fusion proteins were detected in the resuspended and size-filtered second PEG pellets, but not in the supernatant, indicating the chimeric CP assembled into VLPs.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 26

<210> SEQ ID NO 1

<211> LENGTH: 25

<212> TYPE: PRT

<213> ORGANISM: *Bacillus anthracis*

<400> SEQUENCE: 1

Ser Asn Ser Arg Lys Lys Arg Ser Thr Ser Ala Gly Pro Thr Val Pro
1 5 10 15

Asp Arg Asp Asn Asp Gly Ile Pro Asp
20 25

<210> SEQ ID NO 2

<211> LENGTH: 25

<212> TYPE: PRT

<213> ORGANISM: *Bacillus anthracis*

<400> SEQUENCE: 2

Ser Pro Glu Ala Arg His Pro Leu Val Ala Ala Tyr Pro Ile Val His
1 5 10 15

Val Asp Met Glu Asn Ile Ile Leu Ser
20 25

<210> SEQ ID NO 3

<211> LENGTH: 25

<212> TYPE: PRT

<213> ORGANISM: *Bacillus anthracis*

<400> SEQUENCE: 3

Arg Ile Ile Phe Asn Gly Lys Asp Leu Asn Leu Val Glu Arg Arg Ile
1 5 10 15

-continued

Ala Ala Val Asn Pro Ser Asp Pro Leu
 20 25

<210> SEQ ID NO 4
 <211> LENGTH: 25
 <212> TYPE: PRT
 <213> ORGANISM: Bacillus anthracis

<400> SEQUENCE: 4

Arg Gln Asp Gly Lys Thr Phe Ile Asp Phe Lys Lys Tyr Asn Asp Lys
 1 5 10 15

Leu Pro Leu Tyr Ile Ser Asn Pro Asn
 20 25

<210> SEQ ID NO 5
 <211> LENGTH: 4136
 <212> TYPE: DNA
 <213> ORGANISM: artificial
 <220> FEATURE:
 <223> OTHER INFORMATION: pDOW2160

<400> SEQUENCE: 5

tcgcgcgttt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca	60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg	120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc	180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc	240
attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggccc tcttcgctat	300
tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt	360
tttcccagtc acgacgttgt aaaacgacgg ccagtgcgaa gcttcagaa ggtaattatc	420
caagatgtag catcaagaat ccaatgttta cgggaaaaac tatggaagta ttatgtgagc	480
tcagcaagaa gcagatcaat atgcggcaca tatgcaacct atgttcaaaa atgaagaatg	540
tacagatata agatcctata ctgccagaat acgaagaaga atacgtagaa attgaaaaag	600
aagaaccagg cgaagaaaag aatcttgaag acgtaagcac tgacgacaac aatgaaaaga	660
agaagataag gtcggtgatt gtgaaagaga catagaggac acatgtaagg tggaaaatgt	720
aagggcggaa agtaacctta tcacaaagga atcttatccc ccactactta tccttttata	780
ttttccgctg tcattttttgc ccttgagttt tcctatataa ggaaccaagt tcggcatttg	840
tgaaaacaag aaaaaatttg gtgtaagcta ttttctttga agtactgagg atacaacttc	900
agagaaatgt gtaagtttgt agatctcgat tctagaatgt ctacagtcgg aacagggaag	960
ttaactcgct cacaacgaag ggctgcggcc cgtaagaaca agcgggaacac tcgtgtggtc	1020
caacctgtta ttgtagaacc catcgcttca ggccaaggca aggctattaa agcatggacc	1080
ggttacagcg tatcgaagtg gaccgcctct tgtgcggctg ccgaagctaa agtaacctcg	1140
gctataacta tctctctccc taatgagcta tcgtccgaaa ggaacaagca gctcaaggta	1200
ggtagagttt tattatggct tgggttgctt ccagtggtta gtggcacagt gaaatcctgt	1260
gttacagaga cgcagactac tgctgctgcc tcctttcagg tggcattagc tgtggccgac	1320
aacgggatcc ttagtaattc tcgtaagaaa cgttctacct ctgctggccc tacogtgcct	1380
gatcgtgata atgatggcat tcctgatggg atcctgtcga aagatgttgt cgctgctatg	1440

-continued

taccccgagg	cgtttaaggg	tataaccctt	gaacaactca	ccgcggattt	aacgatctac	1500
ttgtacagca	gtgcggctct	cactgagggc	gacgtcatcg	tgcatTTgga	ggttgagcat	1560
gtcagaccta	cgtttgacga	ctctttcact	ccgtattagt	aagaattcga	gctcggTacc	1620
ggatccaatt	cccgatcggt	caaacatttg	gcaataaagt	ttcttaagat	tgaatcctgt	1680
tgccggtctt	gcgatgatta	tcatataatt	tctgttgaat	tacgttaagc	atgtaataat	1740
taacatgtaa	tgcatgacgt	tatttatgag	atgggttttt	atgattagag	tcccgcaatt	1800
atacatttaa	tacgcgatag	aaaacaaaat	atagcgcgca	aactaggata	aattatcgcg	1860
cgcggtgtca	tctatgttac	tagatcgggg	atcgatcccc	aattcgtaat	catggtcata	1920
gctgtttcct	gttgTaaatt	gttatccgct	cacaattcca	cacaacatac	gagccggaag	1980
cataaagtgt	aaagcctggg	gtgcctaata	cgTgagctaa	ctcacattaa	ttgcgttgcg	2040
ctcactgccc	gctttccagt	cgggaaacct	gtcgtgccag	ctgcattaat	gaatcggccca	2100
acgcgcgggg	agaggcggtt	tgcgtattgg	gcgctcttcc	gcttcctcgc	tactgacttc	2160
gctgcgctcg	gtcgttcggc	tgcggcgagc	ggtatcagct	cactcaaagg	cggtaatacg	2220
gttatccaca	gaatcagggg	ataacgcagg	aaagaacatg	tgagcaaaaag	gccagcaaaa	2280
ggccagggaac	cgtaaaaagg	ccgcgttgct	ggcgTTTTtc	cataggctcc	gccccctga	2340
cgagcatcac	aaaaatcgac	gctcaagtca	gaggtggcga	aaccgcagag	gactataaag	2400
ataccaggcg	tttccccctg	gaagctccct	cgtgcgctct	cctgttccga	ccctgccgct	2460
taccggatac	ctgtccgcct	ttctcccttc	gggaagcgtg	gcgctttctc	aatgctcacg	2520
ctgtagggtat	ctcagttcgg	tgtaggctcg	tcgctccaag	ctgggggtgtg	tgacggaacc	2580
ccccgttcag	cccgaccgct	gcgccttata	cggtaactat	cgtcttgagt	ccaaccgggt	2640
aagacacgac	ttatcgccac	tggcagcagc	cactggtaac	aggattagca	gagcgaggta	2700
tgtaggcgggt	gtacagaggt	tcttgaagtg	gtggcctaac	tacggctaca	ctagaaggac	2760
agtatttggt	atctgcgctc	tgctgaagcc	agttaccttc	ggaaaaagag	ttggtagctc	2820
ttgatccggc	aaacaaacca	ccgctggtag	cggtggtttt	tttgtttgca	agcagcagat	2880
tacgcgcaga	aaaaaaggat	ctcaagaaga	tcctttgatc	ttttctacgg	ggtctgacgc	2940
tcagtggaaac	gaaaactcac	gttaagggat	tttggtcatg	agattatcaa	aaaggatcct	3000
cacctagatc	cttttaaatt	aaaaatgaag	ttttaaatca	atctaaagta	tatatgagta	3060
aacttggtct	gacagttacc	aatgcttaat	cagtgaggca	cctatctcag	cgatctgtct	3120
atttcgttca	tccatagttg	cctgactccc	cgtcgtgtag	ataactacga	tacgggaggg	3180
cttaccatct	ggccccagtg	ctgcaatgat	accgcgagac	ccacgctcac	cggctccaga	3240
tttatcagca	ataaaccagc	cagccggaag	ggccgagcgc	agaagtggtc	ctgcaacttt	3300
atccgcctcc	atccagtcta	ttaattgttg	ccgggaagct	agagtaagta	gttcgccagt	3360
taatagtttg	cgcaacgttg	ttgccattgc	tacaggcatc	gtggtgtcac	gctcgtcggt	3420
tggtatggct	tcattcagct	ccggttccca	acgatcaagg	cgagttacat	gatcccccat	3480
gttgTgcaaa	aaagcggtta	gtccttcggg	tcctccgata	gttgTcagaa	gtaagttggc	3540
cgcagtgtta	tactcatggt	ttatggcagc	actgcataat	tctcttactg	tcatgccatc	3600
cgtaaagtgc	ttttctgtga	ctggtgagta	ctcaaccaag	tcattctgag	aatagtgtat	3660
gcggcgaccg	agttgctctt	gcccggcgct	aatacgggat	aataccgcgc	cacatagcag	3720

-continued

aactttaaaa	gtgctcatca	ttggaaaacg	ttcttcgggg	cgaaaactct	caaggatctt	3780
aggcgtgttg	agatccagtt	cgatgtaacc	cactcgtgca	cccaactgat	cttcagcatc	3840
ttttactttc	accagcgttt	ctgggtgagc	aaaaacagga	aggcaaatg	ccgcaaaaaa	3900
gggaataaag	gcgacacgga	aatgttgaat	actcatactc	ttcctttttc	aatattattg	3960
aagcatttat	caggggttatt	gtctcatgag	cggatacata	tttgaatgta	tttagaaaaa	4020
taaaaaata	ggggttccgc	gcacatttcc	ccgaaaagtg	ccacctgacg	tctaagaaac	4080
cattattatc	atgacattaa	cctataaaaa	taggcgtatc	acgaggccct	ttcgtc	4136

<210> SEQ ID NO 6
 <211> LENGTH: 4136
 <212> TYPE: DNA
 <213> ORGANISM: artificial
 <220> FEATURE:
 <223> OTHER INFORMATION: pDOW2161

<400> SEQUENCE: 6

tcgcgcgttt	cggtgatgac	ggtgaaaacc	tctgacacat	gcagctcccg	gagacgggtca	60
cagcttgtct	gtaagcggat	gccgggagca	gacaagcccg	tcagggcgcg	tcagcgggtg	120
ttggcgggtg	tcggggcttg	cttaactatg	cggcatcaga	gcagattgta	ctgagagtgc	180
accatatgcg	gtgtgaaata	ccgcacagat	gcgtaaggag	aaaataccgc	atcaggcgcc	240
attcgccatt	caggctgcgc	aactgttggg	aagggcgatc	ggtgcgggcc	tcttcgctat	300
tacgccagct	ggcgaaagg	ggatgtgctg	caaggcgatt	aagttgggta	acgccagggt	360
tttcccagtc	acgacgttgt	aaaacgacgg	ccagtgccaa	gcttcagaa	ggtaattatc	420
caagatgtag	catcaagaat	ccaatgttta	cgggaaaaac	tatggaagta	ttatgtgagc	480
tcagcaagaa	gcagatcaat	atgcggcaca	tatgcaacct	atgttcaaaa	atgaagaatg	540
tacagatata	agatcctata	ctgccagaat	acgaagaaga	atacgtagaa	attgaaaaag	600
aagaaccagg	cgaagaaaag	aatcttgaag	acgtaagcac	tgacgacaac	aatgaaaaga	660
agaagataag	gtcgggtgatt	gtgaaagaga	catagaggac	acatgtaagg	tggaaaatgt	720
aaggcgcgaa	agtaacctta	tcacaaagga	atcttatccc	ccactactta	tccttttata	780
tttttccgtg	tcattttttg	ccttgagttt	tcctatataa	ggaaccaagt	tcggcatttg	840
tgaaaaaag	aaaaaatttg	gtgtaagcta	ttttctttga	agtactgagg	atacaacttc	900
agagaaattt	gtaagtttgt	agatctcgat	tctagaatgt	ctacagtcgg	aacagggaag	960
ttaactcgtg	cacaacgaag	ggctgcggcc	cgtaagaaca	agcggaacac	tcgtgtggtc	1020
caacctgtta	ttgtagaacc	catcgcttca	ggccaaggca	aggctattaa	agcatggacc	1080
ggttacagcg	tatcgaagtg	gaccgcctct	tgtgcggctg	ccgaagctaa	ggtaacctcg	1140
gctataacta	tctctctccc	taatgagcta	tcgtccgaaa	ggaacaagca	gctcaaggta	1200
ggtagagttt	tattatggct	tgggttgctt	cccagtgtta	gtggcacagt	gaaatcctgt	1260
gttacagaga	cgcagactac	tgctgctgcc	tccttttcagg	tggcattagc	tgtggccgac	1320
aacgggatcc	ttagtcctga	agctcgtcat	cctctcgtgg	ctgcgtatcc	tattgtgcat	1380
gttgatatgg	aaaatattat	cctctctggg	atcctgtcga	aagatgttgt	cgtgctatg	1440
taccccgagg	cgtttaagg	tataaccctt	gaacaactca	ccgcggattt	aacgatctac	1500
ttgtacagca	gtgcggctct	cactgagggc	gacgtcatcg	tgcatttgga	ggttgagcat	1560

-continued

gtcagaccta cgtttgacga ctctttcact ccgtattagt aagaattcga gtcggtacc	1620
ggatccaatt cccgatcggt caaacatttg gcaataaagt ttcttaagat tgaatcctgt	1680
tgccggtcct gcgatgatta tcatataatt tctgttgaat tacgttaagc atgtaataat	1740
taacatgtaa tgcgatgacgt tatttatgag atgggttttt atgattagag tcccgcatt	1800
atacatttaa tacgcgatag aaaacaaaat atagcgcgca aactaggata aattatcgcg	1860
cgcggtgtca tctatgttac tagatcggg atcgatcccc aattcgtaat catggtcata	1920
gctgtttcct gtgtgaaatt gttatccgct cacaattcca cacaacatac gagccggaag	1980
cataaagtgt aaagcctggg gtgcctaatt cgtgagctaa ctcacattaa ttgcgttgcg	2040
ctcactgccc gctttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcggcca	2100
acgcgcgggg agaggcggtt tgcgtattgg gcgctcttcc gcttcctcgc tcaactgactc	2160
gctgcgctcg gtcgttcggc tgcgcgagc ggtatcagct cactcaaagg cggtaatacg	2220
gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa	2280
ggccaggaac cgtaaaaagg ccgcttgctt ggcgtttttc cataggctcc gccccctga	2340
cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aaccgcagag gactataaag	2400
ataccaggcg tttccccctg gaagtcctct cgtgcgctct cctgttccga cctgccgct	2460
taccggatac ctgtccgcct ttctcccttc gggaaagcgtg gcgctttctc aatgtcacg	2520
ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctggggtgtg tgcacgaacc	2580
ccccgttcag cccgaccgct gcgccttacc cggtaactat cgtcttgagt ccaaccgggt	2640
aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta	2700
tgtaggcggg gctacagagt tcttgaagtgt gtggcctaac tacggctaca ctagaaggac	2760
agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaanaagag ttggtagctc	2820
ttgatccggc aaacaaacca ccgctggtag cgggtggttt tttgtttgca agcagcagat	2880
tacgcgcaga aaaaaggat ctcaagaaga tcctttgatc tttctacgg ggtctgacgc	2940
tcagtggaac gaaaactcac gtttaaggat tttggtcatg agattatcaa aaaggatcct	3000
cacctagatc cttttaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta	3060
aacttggtct gacagttacc aatgcctaat cagtgaggca cctatctcag cgatctgtct	3120
atttcgttca tccatagttg cctgactccc cgtcgtgtag ataactacga tacgggaggg	3180
cttaccatct ggccccagt ctgcaatgat accgcgagac ccacgctcac cggtccaga	3240
tttatcagca ataaaccagc cagccggaag ggccgagcgc agaagtggtc ctgcaacttt	3300
atccgcctcc atccagtcta ttaattgttg ccgggaagct agagtaagta gttcgccagt	3360
taatagtttg cgcaacgttg ttgccattgc tacaggcatc gtggtgtcac gctcgtcgtt	3420
tggtatggct tcattcagct ccggttccca acgatcaagg cgagttacat gatccccc	3480
gttgtgcaaa aaagcgggta gctccttcgg tcctccgacg gttgtcagaa gtaagttggc	3540
cgcagtgtta tcaactcatg ttatggcagc actgcataat tctcttactg tcatgccatc	3600
cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat	3660
gcggcgaccg agttgctctt gcccggcgct aatacgggat aataccgcgc cacatagcag	3720
aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatcct	3780
aggcgtgttg agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc	3840

-continued

ttttactttt accagcggtt ctgggtgagc aaaaacagga aggcaaatg ccgcaaaaaa	3900
gggaataagg gcgacacgga aatgttgaat actcactctc ttcctttttc aatattattg	3960
aagcatttat cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa	4020
taaacaaata ggggttcgcg gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac	4080
cattattatc atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc	4136

<210> SEQ ID NO 7
 <211> LENGTH: 4133
 <212> TYPE: DNA
 <213> ORGANISM: artificial
 <220> FEATURE:
 <223> OTHER INFORMATION: pDOW2162

<400> SEQUENCE: 7

tcgcgcgttt cgggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca	60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg	120
ttggcgggtg tcggggcttg cttaactatg cggcatcaga gcagattgta ctgagagtgc	180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc	240
attcgccatt cagggtcgcg aactgttggg aagggcgatc ggtgcgggccc tcttcgctat	300
tacgccagct ggcgaagggt ggatgtgctg caaggcgatt aagtgggta acgccagggt	360
tttcccagtc acgacgttgt aaaacgacgg ccagtgccaa gcttccagaa ggtaattatc	420
caagatgtag catcaagaat ccaatgttta cgggaaaaac tatggaagta ttatgtgagc	480
tcagcaagaa gcagatcaat atgcccacac tatgcaacct atgttcaaaa atgaagaatg	540
tacagatata agatcctata ctgccagaat acgaagaaga atacgtagaa attgaaaaag	600
aagaaccagg cgaagaaaag aatcttgaag acgtaagcac tgacgacaac aatgaaaaga	660
agaagataag gtcggtgatt gtgaaagaga catagaggac acatgtaagg tggaaaatgt	720
aagggcgcaa agtaacctta tcacaaagga atcttatccc ccactactta tccttttata	780
tttttccgtg tcattttttgc ccttgagttt tcctatatataa ggaaccaagt tcggcatttg	840
tgaaaacaag aaaaaatttg gtgtaagcta ttttctttga agtactgagg atacaacttc	900
agagaaatgt gtaagtgtgt agatctcgat tctagaatgt ctacagtcgg aacagggaag	960
ttaactcgtg cacaacgaag ggctgcggcc cgtaagaaca agcggaacac tcgtgtggtc	1020
caacctgtta ttgtagaacc catcgcttca ggccaaggca aggtatttaa agcatggacc	1080
ggttacagcg tatcgaagtg gaccgcctct tgtgcggctg ccgaagctaa agtaacctcg	1140
gctataacta tctctctccc taatgagcta tcgtccgaaa ggaacaagca gctcaaggta	1200
ggtagagttt tattatggct tgggttgctt ccagtggtta gtggcacagt gaaatcctgt	1260
gttacagaga cgcagactac tgctgctgcc tcctttcagg tggcattagc tgtggccgac	1320
aacgggatcc gtattatttt caatggcaaa gatctcaatc tcgtggaacg tcgtattgct	1380
gctgtgaatc cttctgatcc tctcgggatc ctgtcgaaaag atgttgtcgc tgctatgtac	1440
cccgaggcgt ttaagggtat aacccttgaa caactcaccg cggatttaac gatctacttg	1500
tacagcagtg cggctctcac tgaggggcgc gtcacgtgc atttggagggt tgagcatgtc	1560
agacctacgt ttgacgactc tttcactccg tattagtaag aattcgagct cgttacggga	1620
tccaattccc gatcgttcaa acatttgca ataaagtttc ttaagattga atcctgttgc	1680

-continued

cggtcttgcg atgattatca tataatttct gttgaattac gttaagcatg taataattaa	1740
catgtaatgc atgacgttat ttatgagatg ggtttttatg attagagtcc cgcaattata	1800
catttaatac gcgatagaaa acaaaatata gcgcgcaaac taggataaat tatcgcgcg	1860
ggtgtcatct atgttactag atcggggatc gatccccaat tcgtaatcat ggtcatagct	1920
gtttcctgtg tgaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat	1980
aaagtgtaaa gcctggggtg cctaattgctg gagctaactc acattaattg cgttgcgctc	2040
actgcccgtt ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa tcggccaacg	2100
cgcggggaga ggcggtttgc gtattggcg ctcttccgct tcctcgctca ctgactcgct	2160
gcgctcggtc gttcgctgc ggcgagcgg atcagctcac tcaaggcgg taatacgggt	2220
atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc	2280
caggaaccgt aaaaaggccg cgttgcgtgc gtttttccat aggtccgccc ccctgacga	2340
gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata	2400
ccaggcgttt cccctggaa gctccctcgt gcgctctcct gttccgacct tgcgcttac	2460
cggatacctg tccgccttcc tcccttcggg aagcgtggcg ctttctcaat gtcacgctg	2520
taggtatctc agttcgggtt aggtcgttcg ctccaagctg ggggtgtgtc acgaaccccc	2580
cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccgtaag	2640
acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggatatg	2700
aggcgggtct acagagtctt tgaagtggg gcctaactac ggctacacta gaaggacagt	2760
atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg	2820
atccggcaaa caaaccaccg ctggtagcgg tgggtttttt gtttgcaagc agcagattac	2880
gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca	2940
gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac	3000
ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac	3060
ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt	3120
tcgttcatcc atagttgcct gactccccgt cgtgtagata actacgatac gggagggtt	3180
accatctggc cccagtgcgt caatgatacc gcgagaccca cgctcaccgg ctccagattt	3240
atcagcaata aaccagccag ccggaagggc cgagcgaga agtggtcctg caactttatc	3300
cgctccatc cagtctatta attgttgcg ggaagctaga gtaagtagtt cgcagttaa	3360
tagtttgcg aacgttggtt ccattgctac aggcacgtg gtgtcacgct cgtcgtttg	3420
tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat ccccatggt	3480
gtgcaaaaaa gcggttagct ccttcggtcc tccgatcgtt gtcagaagta agttggccgc	3540
agtgttatca ctcatggtta tggcagcact gcataattct cttactgtca tgccatccgt	3600
aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgtatgcg	3660
gcgaccgagt tgctcttgcc cggcgtaaat acgggataat accgcgccac atagcagaac	3720
tttaaaagt ctcacattg gaaaacgttc ttcggggcga aaactctcaa ggatcttagg	3780
cgtgttgaga tccagttcga tgtaaccac tcgtgcaccc aactgatctt cagcatcttt	3840
tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaatgccg caaaaaagg	3900
aataaggcgc acacggaaat gttgaatact catactcttc ctttttcaat attattgaag	3960

-continued

catttatcag ggttatgtgc tcatgagcgg atacatat	gaatgtat	agaaaaataa	4020
acaaataggg gttccgcgca catttccccg aaaagtgcc	cctgacgtct	aagaaaccat	4080
tattatcatg acattaacct ataaaaatag gcgtatcacg	aggccctttc	gtc	4133

<210> SEQ ID NO 8

<211> LENGTH: 4133

<212> TYPE: DNA

<213> ORGANISM: artificial

<220> FEATURE:

<223> OTHER INFORMATION: pDOW2163

<400> SEQUENCE: 8

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca	60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg	120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc	180
accatatgcg gtgtgaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc	240
attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggccc tcttcgctat	300
tacgccagct gccgaaagg ggatgtgctg caaggcgatt aagttgggta acgccagggt	360
tttcccagtc acgacgttgt aaaacgacgg ccagtgccaa gcttccagaa ggtaattatc	420
caagatgtag catcaagaat ccaatgttta cgggaaaaac tatggaagta ttatgtgagc	480
tcagcaagaa gcagatcaat atgcggcaca tatgcaacct atgttcaaaa atgaagaatg	540
tacagatata agatcctata ctgccagaat acgaagaaga atacgtagaa attgaaaaag	600
aagaaccagg cgaagaaaag aatcttgaag acgtaagcac tgacgacaac aatgaaaaga	660
agaagataag gtcgggtgatt gtgaaagaga catagaggac acatgtaagg tggaaaatgt	720
aaggcgcgaa agtaacctta tcacaaagga atcttatccc ccactactta tccttttata	780
tttttccgtg tcatttttgc ccttgagttt tcctatataa ggaaccaagt tcggcatttg	840
tgaaaacaag aaaaaatttg gtgtaagcta ttttctttga agtactgagg atacaacttc	900
agagaaat	960
ttactcgtg cacaacgaag ggctgcggcc cgtaagaaca agcgggaacac tcgtgtggtc	1020
caacctgtta ttgtagaacc catcgcttca ggccaaggca aggtattaa agcatggacc	1080
ggttacagcg tatcgaagt gaccgcctct tgtgcggctg ccgaagctaa agtaacctcg	1140
gctataacta tctctctccc taatgagcta tcgtccgaaa ggaacaagca gctcaaggta	1200
ggtagagttt tattatggct tgggttgctt ccagtggtta gtggcacagt gaaatcctgt	1260
gttacagaga cgcagactac tgctgctgcc tcctttcagg tggcattagc tgtggccgac	1320
aacgggatcc gtattatttt caatggcaaa gatctcaatc tcgtggaacg tcgtattgct	1380
gctgtgaatc cttctgatcc tctcgggac ctgtcgaaag atgttgtcgc tgctatgtac	1440
cccaggcgt ttaagggat aacccttgaa caactcaccg cggatttaac gatctacttg	1500
tacagcagtg cggctctcac tgaggggcag gtcacgtgac atttgagggt tgagcatgtc	1560
agacctacgt ttgacgactc tttcactccg tattagtaag aatcagagct cggtaaccgga	1620
tccaattccc gatcgttcaa acatttgga ataaagtttc ttaagattga atcctgttgc	1680
cggctctgag atgattatca tataatttct gttgaattac gttgaagcatg taataattaa	1740
catgtaatgc atgacgttat ttatgagatg ggtttttatg attagagtcc cgcaattata	1800

-continued

catttaatac gcgatagaaa acaaaatata gcgcgcaaac taggataaat tatcgcgcg	1860
ggtgtcatct atgttactag atcggggatc gatccccaat tcgtaatcat ggtcatagct	1920
gtttcctgtg tgaattgtt atccgctcac aattccacac aacatacgag ccggaagcat	1980
aaagtgtaaa gcctgggggtg cctaattgcgt gagctaactc acattaattg cgttgcgctc	2040
actgcccgtt ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa tcggccaacg	2100
cgcggggaga ggcggtttgc gtattggcg ctcttcgctt tcctcgctca ctgactcgct	2160
gcgctcggtc gttcggttcg ggcgagcgg atcagctcac tcaaaggcgg taatacggtt	2220
atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc	2280
caggaaccgt aaaaaggcgg cgttgctggc gtttttccat aggcctccgc cccctgacga	2340
gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata	2400
ccaggcggtt cccctggaa gctccctcgt gcgctctcct gttccgacct tgccgcttac	2460
cggataacct tccgccttcc tcccttcggg aagcgtggcg ctttctcaat gtcacgctg	2520
taggtatctc agttcgggtg aggtcggtcg ctccaagctg ggggtgtgtc acgaaccccc	2580
cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggttaag	2640
acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggatgt	2700
aggcgggtgt acagagtctc tgaagtgggt gcctaactac ggctacacta gaaggacagt	2760
atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg	2820
atccggcaaa caaacaccg ctggtagcgg tgggtttttt gtttgcaagc agcagattac	2880
gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggg ctgacgctca	2940
gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggtacttcac	3000
ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac	3060
ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt	3120
tcgttcatcc atagttgcct gactccccgt cgtgtagata actacgatac gggagggtt	3180
accatctggc cccagtgtg caatgatacc gcgagacca cgctcaccgg ctccagattt	3240
atcagcaata aaccagccag ccggaaggcg cgagcgaga agtggtcctg caactttatc	3300
cgcctccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt cgccagttaa	3360
tagtttgccg aacgttgttg ccattgttac aggcacgtg gtgtcacgct cgtcgtttgg	3420
tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat ccccatgtt	3480
gtgcaaaaaa gcggttagct ccttcgggtc tccgatcggt gtcagaagta agttggccgc	3540
agtgttatca ctcatgttga tggcagcact gcataattct cttactgtca tgccatccgt	3600
aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgtatgcg	3660
gcgaccgagt tgctcttgcc cggcgtaaat acgggataat accgcgccac atagcagaac	3720
tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa ggtacttagg	3780
cgtgttgaga tccagttcga tgtaaccac tcgtgcaccc aactgatctt cagcatcttt	3840
tactttcacc agcgtttctg ggtgagcaaa aacagggaagg caaaatgccg caaaaaagg	3900
aataaggcgg acacggaaat gttgaatact catactcttc ctttttcaat attattgaag	3960
catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa	4020
acaaataggg gttccgcgca catttccccg aaaagtgcc cctgacgtct aagaaaccat	4080

-continued

tattatcatg acattaacct ataaaaatag gcgtatcacg aggccctttc gtc 4133

<210> SEQ ID NO 9

<211> LENGTH: 4290

<212> TYPE: DNA

<213> ORGANISM: artificial

<220> FEATURE:

<223> OTHER INFORMATION: pDOW2169

<400> SEQUENCE: 9

tcgcgcgttt	cggtgatgac	ggtgaaaacc	tctgacacat	gcagctcccg	gagacgggtca	60
cagcttgtct	gtaagcggat	gccgggagca	gacaagcccg	tcagggcgcg	tcagcgggtg	120
ttggcgggtg	tcggggctgg	cttaactatg	cggcatcaga	gcagattgta	ctgagagtgc	180
accatatgcg	gtgtgaaata	ccgcacagat	gcgtaaggag	aaaataccgc	atcaggcgcc	240
attcgccatt	caggctgcgc	aactgttggg	aaggcgatc	ggtgcgggcc	tcttcgctat	300
tacgccagct	ggcgaaagg	ggatgtgctg	caaggcgatt	aagttgggta	acgccagggt	360
tttcccagtc	acgacgttgt	aaaacgacgg	ccagtgccaa	gcttcagaa	ggtaattatc	420
caagatgtag	catcaagaat	ccaatgttta	cgggaaaaac	tatggaagta	ttatgtgagc	480
tcagcaagaa	gcagatcaat	atgccgcaca	tatgcaacct	atgttcaaaa	atgaagaatg	540
tacagataca	agatcctata	ctgccagaat	acgaagaaga	atacgtagaa	attgaaaaag	600
aagaaccagg	cgaagaaaag	aatcttgaag	acgtaagcac	tgacgacaac	aatgaaaaga	660
agaagataag	gtcggtgatt	gtgaaagaga	catagaggac	acatgtaagg	tggaaaatgt	720
aaggcgcgaa	agtaacctta	tcacaaagga	atcttatccc	ccactactta	tccttttata	780
tttttccgtg	tcattttttg	ccttgagttt	tcctatataa	ggaaccaagt	tcggcatttg	840
tgaaaacaag	aaaaaatttg	gtgtaagcta	ttttctttga	agtactgagg	atacaacttc	900
agagaaattt	gtaagtittg	agatctcgat	tctagaatgt	ctacagtcgg	aacagggaag	960
ttaactcgtg	cacaacgaag	ggctgcggcc	cgtaagaaca	agcggaacac	tcgtgtggtc	1020
caacctgtta	ttgtagaacc	catcgcttca	ggccaaggca	aggctattaa	agcatggacc	1080
ggttacagcg	tatcgaagtg	gaccgcctct	tgtgcggctg	ccgaagctaa	agtaacctcg	1140
gctataacta	tctctctccc	taatgagcta	tcgtccgaaa	ggaacaagca	gctcaaggta	1200
ggtagagttt	tattatggct	tgggttgctt	cccagtgtta	gtggcacagt	gaaatcctgt	1260
gttacagaga	cgcagactac	tgctgctgcc	tcctttcagg	tggcattagc	tgtggccgac	1320
aacgggatcc	ttagtaaatc	tcgtaagaaa	cgttctacct	ctgctggccc	taccgtgcct	1380
gatcgtaata	atgatggcat	tcctgatggg	atcctgtcga	aagatgttgt	cgtgctatg	1440
taccccgagg	cgtttaagg	tataaccctt	gaacaactca	ccgcggattt	aacgatctac	1500
ttgtacagca	gtgcggctct	cactgagggc	gacgtcatcg	tgcatttggg	ggttgagcat	1560
gtcagacctt	cgtttgacga	ctctttcact	ccggtgtatt	agtgcccgct	gaagagcggt	1620
acactagtgt	ggcctacttg	aaggctagtt	ataaccgttt	ctttaaacgg	taatcgttgt	1680
tgaaacgtct	tccttttaca	agaggattga	gctgcccttg	ggttttactc	cttgaaccct	1740
tcggaagaac	tcttttgaat	tcgagctcgg	taccggatcc	aattcccgat	cgttcaaaaa	1800
tttggcaata	aagtttctta	agattgaatc	ctgttgccgg	tcttgcatg	attatcatat	1860
aatttctggt	gaattacggt	aagcatgtaa	taattaacat	gtaatgcatg	acgttatatta	1920

-continued

tgagatgggt ttttatgatt agagtccgc aattatacat ttaatacgcg atagaaaaca	1980
aaatatagcg cgcaactag gataaattat cgcgcgcggt gtcactctatg ttactagatc	2040
ggggatcgat ccccaattcg taatcatggt catagctggt tcctgtgtga aattgttatc	2100
cgctcacaat tccacacaac atacgagccg gaagcataaa gtgtaaagcc tggggtgcct	2160
aatgctgtag ctaactcaca ttaattgcgt tgcgctcact gcccgctttc cagtcgggaa	2220
acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc ggggagaggc ggtttgcgta	2280
ttgggcgctc ttccgcttcc tcgctcactg actcgctgcg ctcggtcggt cggtgcggc	2340
gagcggatc agctcactca aaggcggtaa tacggttatc cacagaatca ggggataacg	2400
caggaaagaa catgtgagca aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt	2460
tgctggcggt tttccatagc ctccgcccc ctgacgagca tcacaaaaat cgagctcaa	2520
gtcagagggt gcgaacccg acaggactat aaagatacca ggcgtttccc cctggaagct	2580
ccctcgtgcg ctctcctggt ccgaccctgc cgcttacgg atacctgtcc gcctttctcc	2640
cttcgggaag cgtggcgctt tctcaatgct cagcgtgtag gtatctcagt tcggtgtagg	2700
tcgttcgctc caagctgggg tgtgtgcacg aacccccgt tcagcccgac cgtgcgcct	2760
tatccggtaa ctatcgtctt gagtccaacc cggtaagaca cgacttatcg ccactggcag	2820
cagccactgg taacaggatt agcagagcga ggtatgtagg cggtgctaca gatttcttga	2880
agtggtggcc taactacgac tacactagaa ggacagtatt tggatatctgc gctctgctga	2940
agccagttac cttcgaaaa agagttggt gctcttgatc cggcaaaaa accaccgctg	3000
gtagcgggtg tttttttgtt tgcaagcagc agattacgcg cagaaaaaaa ggtatctcaag	3060
aagatccttt gatcttttct acggggtctg acgctcagtg gaacgaaaac tcacgttaag	3120
ggattttgg catgagatta tcaaaaagga tcttcaccta gatcctttta aattaaat	3180
gaagtttta atcaatctaa agtatatatg agtaaaactg gtctgacagt taccaatgct	3240
taatcagtga ggcacctatc tcagcgatct gtctatttgc ttcattccata gttgcctgac	3300
tccccgtcgt gtagataact acgatacggg agggcttacc atctggcccc agtgctgcaa	3360
tgataccgcg agaccacgc tcaccggctc cagatttatc agcaataaac cagccagccg	3420
gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc ctccatccag tctattaatt	3480
gttgccggga agctagagta agtagttcgc cagttaatag tttgcgcaac gttgttgcca	3540
ttgtacagc catcgtggtg tcacgctcgt cgtttggtat ggcttcattc agtccgggt	3600
cccaacgac aaggcgaggt acatgatccc ccatgttggtg caaaaaagcg gttagctcct	3660
tcggtcctcc gatcgttgct agaagtaagt tggccgcagt gttatcactc atggttatgg	3720
cagcactgca taattctctt actgtcatgc catccgtaag atgcttttct gtgactggtg	3780
agtactcaac caagtcattc tgagaatagt gtatgcggcg accgagttgc tcttgcccgg	3840
cgtaaatag ggataatacc gcgccacata gcagaacttt aaaagtgtc atcattggaa	3900
aacgttcttc gggcgaaaa ctctcaagga tcttaggcgt gttgagatcc agttcgatgt	3960
aaccactcgc tgcacccaac tgatcttcag catcttttac tttcaccagc gtttctgggt	4020
gagcaaaaac aggaaggcaa aatgcgcgaa aaaagggaat aaggcgacga cggaaatgtt	4080
gaatactcat actcttcctt tttcaatatt attgaagcat ttatcagggt tattgtctca	4140
tgagcggata catatttgaa tgtatttaga aaaataaaca aataggggtt ccgcgcacat	4200

-continued

ttccccgaaa agtgccacct gacgtctaag aaaccattat tatcatgaca ttaacctata	4260
aaaataggcg tatcacgagg ccctttcgtc	4290

<210> SEQ ID NO 10
 <211> LENGTH: 4290
 <212> TYPE: DNA
 <213> ORGANISM: artificial
 <220> FEATURE:
 <223> OTHER INFORMATION: pDOW2170

<400> SEQUENCE: 10

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca	60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg	120
ttggcgggtg tcggggcttg cttaactatg cggcatcaga gcagattgta ctgagagtgc	180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc	240
attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat	300
tacgccagct gccgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt	360
tttcccagtc acgacgttgt aaaacgacgg ccagtgccaa gcttccagaa ggtaattatc	420
caagatgtag catcaagaat ccaatgttta cgggaaaaac tatggaagta ttatgtgagc	480
tcagcaagaa gcagatcaat atgcccagca tatgcaacct atgttcaaaa atgaagaatg	540
tacagatata agatcctata ctgccagaat acgaagaaga atacgtagaa attgaaaaag	600
aagaaccagg cgaagaaaag aatcttgaag acgtaagcac tgacgacaac aatgaaaaa	660
agaagataag gtcggtgatt gtgaaagaga catagaggac acatgtaagg tggaaaatgt	720
aagggcggaa agtaacctta tcacaaagga atcttatccc ccactactta tccttttata	780
tttttccgtg tcatttttgc ccttgagttt tcctatataa ggaaccaagt tcggcatttg	840
tgaaaacaag aaaaaatttg gtgtaagcta ttttctttga agtactgagg atacaacttc	900
agagaaattt gtaagtttgt agatctcgat tctagaatgt ctacagtcgg aacagggaag	960
ttaactcgtg cacaacgaag ggctgcggcc cgtaagaaca agcggaaacac tcgtgtggtc	1020
caacctgtta ttgtagaacc catcgcttca ggccaaggca aggcatttaa agcatggacc	1080
ggttacagcg tatcgaagtg gaccgcctct tgtgcggctg ccgaagctaa agtaacctcg	1140
gctataacta tctctctccc taatgagcta tcgtccgaaa ggaacaagca gctcaaggta	1200
ggtagagttt tattatggct tgggttgctt cccagtgtta gtggcacagt gaaatcctgt	1260
gttacagaga cgcagactac tgctgctgcc tcctttcagg tggcattagc tgtggccgac	1320
aacgggatcc ttagtcctga agctcgtcat cctctcgtgg ctgcgtatcc tattgtgcat	1380
gttgatatgg aaaatattat cctctctggg atcctgtcga aagatgttgt cgctgctatg	1440
taccccgagg cgtttaaggg tataaccctt gaacaactca ccgcggatth aacgatctac	1500
ttgtacagca gtgcggctct cactgagggc gacgtcatcg tgcatttgga ggttgagcat	1560
gtcagaccta cgtttgacga ctctttcact ccggtgtatt agtgcccgtc gaagagcggt	1620
acactagtgt ggcctacttg aaggctagtt ataaccgttt ctttaaacgg taatcgttgt	1680
tgaacgtctc tccttttaca agaggattga gctgcccttg ggttttactc cttgaaccct	1740
tcggaagaac tctttggaat tcgagctcgg tacccgatcc aattcccgat cgttcaaca	1800
tttggaata aagtttctta agattgaatc ctgttgccgg tcttgcgatg attatcatat	1860

-continued

aatttctgtt gaattacgtt aagcatgtaa taattaacat gtaatgcatg acgttatatta	1920
tgagatgggt ttttatgatt agagtccgc aattatacat ttaatacgcg atagaaaaca	1980
aaatatagcg cgcaactag gataaattat cgcgcgcggt gtcacttatg ttactagatc	2040
ggggatcgat ccccaattcg taatcatggt catagctgtt tcctgtgtga aattgttatac	2100
cgctcacaat tccacacaac atacgagccg gaagcataaa gtgtaaagcc tggggcgctt	2160
aatgctgag ctaactcaca ttaattgcgt tgcgctcact gcccgctttc cagtcgggaa	2220
acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc ggggagaggc ggtttgcgta	2280
ttgggcgctc ttccgcttcc tcgctcactg actcgctgcg ctcggtcggt cggtgcggc	2340
gagcggatc agctcactca aaggcggtaa tacggttatc cacagaatca ggggataacg	2400
caggaaagaa catgtgagca aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt	2460
tgctggcggt tttccatagg ctccgcccc ctgacgagca tcacaaaaat cgacgctcaa	2520
gtcagagggt gcgaaaccg acaggactat aaagatacca ggcgtttccc cctggaagct	2580
ccctcgctgc ctctcctgtt ccgaccctgc cgcttaccg atacctgtcc gcctttctcc	2640
cttcgggaag cgtggcgctt tctcaatgct cacgctgtag gtatctcagt tcggtgtagg	2700
tcgttcgctc caagctggg tgtgtgcacg aacccccgt tcagcccgac cgctgcgcct	2760
tatccggtaa ctatcgtctt gagtccaacc cggtagaca cgacttatcg ccaactggcag	2820
cagccactgg taacaggatt agcagagcga ggtatgtagg cgggtgtaca gagttcttga	2880
agtgggtgcc taactacggc tacactagaa ggacagtatt tggtatctgc gctctgctga	2940
agccagttac cttcggaaaa agagttggta gctcttgatc cggcaaaaa accaccgctg	3000
gtagcgggtg tttttttgtt tgcaagcagc agattacgcg cagaaaaaa ggatctcaag	3060
aaagatcctt gatcttttct acggggtctg acgctcagtg gaacgaaaac tcacgttaag	3120
ggattttgg catgagatta tcaaaaagga tcttcaccta gatcctttta aattaaat	3180
gaagtttta atcaatctaa agtatatatg agtaaacttg gtctgacagt taccaatgct	3240
taatcagtga ggcacctatc tcagcgatct gtctatctcg ttcacccata gttgcctgac	3300
tcccgcgtg gtagataact acgatacggg agggcttacc atctggcccc agtgcgtcaa	3360
tgataccgcg agaccacgc tcaccggctc cagatttatc agcaataaac cagccagccg	3420
gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc ctccatccag tctattaatt	3480
gttgccggga agctagagta agtagttcgc cagttaatag tttgcgcaac gttgttgcca	3540
ttgctacagg catcgtggtg tcacgctcgt cgtttggtat ggcttcattc agctccggtt	3600
cccaacgata aaggcgagtt acatgatccc ccatgttggtg caaaaaagcg gttagctcct	3660
tcggtcctcc gatcgttgtc agaagtaagt tggccgcagt gttatcactc atggttatgg	3720
cagcactgca taattctctt actgtcatgc catccgtaag atgcttttct gtgactggtg	3780
agtactcaac caagtcattc tgagaatagt gtatgcggcg accgagttgc tcttgcccg	3840
cgtaatacgc ggataatacc gcgccacata gcagaacttt aaaagtgtc atcattggaa	3900
aacgttcttc ggggcgaaaa ctctcaagga tcttaggcgt gttgagatcc agttcgatgt	3960
aaccactcgc tgcacccaac tgatcttcag catcttttac tttaccagc gtttctgggt	4020
gagcaaaaac aggaaggcaa aatgcgcgaa aaaagggaat aaggcgacga cggaatgtt	4080
gaatactcat actcttcctt tttcaatatt attgaagcat ttatcagggg tattgtctca	4140

-continued

tgagcggata catatttgaa tgtatttaga aaaataaaca aataggggtt ccgcgcacat	4200
ttccccgaaa agtgccacct gacgtctaag aaaccattat tatcatgaca ttaacctata	4260
aaaataggcg tatcacgagg ccccttcgtc	4290

<210> SEQ ID NO 11
 <211> LENGTH: 4287
 <212> TYPE: DNA
 <213> ORGANISM: artificial
 <220> FEATURE:
 <223> OTHER INFORMATION: pDOW2171

<400> SEQUENCE: 11

tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca	60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg	120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc	180
accatatgcy gtgtgaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc	240
attcgccatt caggctcgcy aactgttggg aagggcgatc ggtcggggcc tcttcgctat	300
tacgccagct ggcgaaagg ggatgtgctg caaggcgatt aagttgggta acgccagggt	360
tttcccagtc acgacgttgt aaaacgacgg ccagtgcga gcttcagaa ggtaattatc	420
caagatgtag catcaagaat ccaatgttta cgggaaaaac tatggaagta ttatgtgagc	480
tcagcaagaa gcagatcaat atgcggcaca tatgcaacct atgttcaaaa atgaagaatg	540
tacagatata agatctata ctgccagaat acgaagaaga atacgtagaa attgaaaaag	600
aagaaccagg cgaagaaaag aatcttgaag acgtaagcac tgacgacaac aatgaaaaga	660
agaagataag gtcgggtgatt gtgaaagaga catagaggac acatgtaagg tggaaaatgt	720
aaggcgcgaa agtaacctta tcacaaagga atcttatccc ccactactta tccttttata	780
ttttcccggt tcatttttgc ccttgagttt tcctatataa ggaaccaagt tcggcatttg	840
tgaaaacaag aaaaaatttg gtgtaagcta ttttctttga agtactgagg atacaacttc	900
agagaaatgt gtaagttgt agatctcgat tctagaatgt ctacagtcgg aacagggaag	960
ttaactcggt cacaacgaag ggctcgggcc cgtaagaaca agcgggaacac tcgtgtggtc	1020
caacctgtta ttgtagaacc catcgcttca ggccaaggca aggtatttaa agcatggacc	1080
ggttacagcy tatcgaagt gaccgcctct tgtcggtg cggaagctaa agtaacctcg	1140
gctataacta tctctctccc taatgagcta tcgtccgaaa ggaacaagca gctcaaggta	1200
ggtagagttt tattatggct tgggttgctt ccagtggtta gtggcacagt gaaatcctgt	1260
gttacagaga cgcagactac tgctgctgcc tcctttcagg tggcattagc tgtggccgac	1320
aacgggatcc gtattatttt caatggcaaa gatctcaatc tcgtggaacg tcgtattgct	1380
gctgtgaatc cttctgatcc tctcgggac ctgtcgaaag atgttgtcgc tgctatgtac	1440
cccaggcgt ttaagggat aacccttgaa caactcaccg cggatttaac gatctacttg	1500
tacagcagtg cggtctcac tgaggcgac gtcacgtgc atttgagggt tgagcatgtc	1560
agacctacgt ttgacgactc tttactccg gtgtattagt gcccgctgaa gagcggtaca	1620
ctagtgtggc ctacttgaag gctagttata accgtttctt taaacggtaa tcgttgttga	1680
aacgtcttcc ttttacaaga ggattgagct gcccttgggt tttactcctt gaaccttcg	1740
gaagaactct ttggaattcg agctcggtag cggtatccat tcccgatcgt tcaaacttt	1800

-continued

ggcaataaaag tttcttaaga ttgaatcctg ttgccggtct tgcgatgatt atcatataat	1860
ttctgttgaa ttacgttaag catgtaataa ttaacatgta atgcatgacg ttatttatga	1920
gatggggttt tatgattaga gtcccgaat tatacattta atacgcgata gaaaacaaaa	1980
tatagcgcgc aaactagat aaattatcgc gcgcggtgtc atctatgtta ctagatcggg	2040
gatcgatccc caattcgtaa tcatggatcat agctgtttcc tgtgtgaaat tgttatccgc	2100
tcacaattcc acacaacata cgagccgaa gcataaagt taaagcctgg ggtgccta	2160
gcgtgagcta actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc	2220
tgtcgtgcca gctgcattaa tgaatcgcc aacgcgcggg gagaggcggg ttgcgtattg	2280
ggcgtcttc cgcttcctcg ctcaactgact cgctgcgctc ggtcgttcgg ctgcgcgag	2340
cgttatcagc tcaactcaag gcggaatac ggttatccac agaatacagg gataacgcag	2400
gaaagaacat gtgagcaaaa gccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc	2460
tggcgttttt ccataggctc gccccctcg acgagcatca caaaaatcga cgctcaagtc	2520
agagggtggc aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc	2580
tcgtgcgctc tcctgttccg accctgccgc ttaccggata cctgtccgcc tttctccctt	2640
cgggaagcgt ggcgctttct caatgctcac gctgtaggta tctcagttcg gtgtaggtcg	2700
ttcgtccaa gctggggtgt gtgcacgaac ccccggttca gcccgaccgc tgcgccttat	2760
ccggttaacta tcgtcttag tccaaccccg taagacacga cttatcgcca ctggcagcag	2820
ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt	2880
ggtggcctaa ctacggtac actagaagga cagtatttgg tatctgcgct ctgctgaagc	2940
cagttacctt cgaaaaaga gttggtagct cttgatccgg caaacaacc accgctggta	3000
gcggtggttt tttgtttgc aagcagcaga ttacgcgcag aaaaaagga tctcaagaag	3060
atcctttgat cttttctacg gggctgacg ctcaagtggaa cgaaaactca cgttaaggga	3120
ttttggtcat gagattatca aaaaggatct tcacctagat ctttttaaat taaaaatgaa	3180
gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttaac caatgcttaa	3240
tcagttaggc acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc	3300
ccgtcgtgta gataactacg atacgggagg gcttaccatc tggcccaggt gctgcaatga	3360
taccgcgaga cccacgctca ccggctccag atttatcagc aataaaccag ccagccggaa	3420
gggcgagcgc cagaagtgtt cctgcaactt tatccgcctc catccagtct attaatgtt	3480
gccgggaagc tagagtaagt agttcgccag ttaatagttt gcgcaacgtt gttgccattg	3540
ctacaggcat cgtggtgtca cgctcgtcgt ttggtatggc ttcatcagc tccggttccc	3600
aacgatcaag gcgagttaca tgatccccc tgtgtgtgcaa aaaagcgggt agctccttcg	3660
gtcctccgat cgttgtcaga agtaagtgg ccgcagtggt atcactcatg gttatggcag	3720
cactgcataa ttctcttact gtcattgccat ccgtaagatg cttttctgtg actggtgagt	3780
actcaaccaa gtcattctga gaatagtgt tgcggcgacc gagttgctct tgcccgcgct	3840
caatacggga taataccgcg ccacatagca gaactttaaa agtgctcatc attggaaaac	3900
gttcttcggg gcgaaaaact tcaaggatct taggcgtgtt gagatccagt tcgatgtaac	3960
ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag	4020
caaaaacagg aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa	4080

-continued

tactcatact cttccttttt caatattatt gaagcattta tcagggttat tgtctcatga	4140
gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc	4200
cccgaagagt gccacgtgac gtctaagaaa ccattattat catgacatta acctataaaa	4260
ataggcgtat cacgaggccc ttctgtc	4287

<210> SEQ ID NO 12
 <211> LENGTH: 4287
 <212> TYPE: DNA
 <213> ORGANISM: artificial
 <220> FEATURE:
 <223> OTHER INFORMATION: pDOW2172

<400> SEQUENCE: 12

tcgcgcggtt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca	60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggcgcg tcagcgggtg	120
ttggcgggtg tcggggcttg cttaactatg cggcatcaga gcagattgta ctgagagtgc	180
accatatgag gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc	240
attcgccatt caggctgcgc aactgttggg aaggcgatc ggtgcggg cccttcgtat	300
tacgccagct gccgaaagg ggatgtgtg caaggcgatt aagttgggt acgccagggt	360
tttcccagtc acgacgttgt aaaacgacgc ccagtgcga gcttcagaa ggtaattatc	420
caagatgtag catcaagaat ccaatgttta cgggaaaaac tatggaagta ttatgtgagc	480
tcagcaagaa gcagatcaat atgcccagca tatgcaacct atgttcaaaa atgaagaatg	540
tacagataca agatcctata ctgccagaat acgaagaaga atacgtagaa attgaaaaag	600
aagaaccagg cgaagaaaag aatcttgaag acgtaagcac tgacgacaac aatgaaaaga	660
agaagataag gtcgggtgatt gtgaaagaga catagaggac acatgtaagg tggaaaatgt	720
aaggcggaag agtaacctta tcacaaagga atcttatccc ccactactta tccttttata	780
ttttccgtg tcatttttgc ccttgagttt tcctatataa ggaaccaagt tcggcatttg	840
tgaaaaaag aaaaaattg gtgtaagcta ttttcttga agtactgagg atacaacttc	900
agagaaattt gtaagtgtgt agatctcgat tctagaatgt ctacagtcgg aacagggaag	960
ttaactcgtg cacaacgaag ggctgcggcc cgtaagaaca agcggaacac tcgtgtggtc	1020
caacctgtta ttgtagaacc catcgcttca ggccaaggca aggtatttaa agcatggacc	1080
ggttacagcg tatcgaagt gaccgcctct tgtgcggctg ccgaagctaa agtaacctcg	1140
gctataacta tctctctccc taatgagcta tcgtccgaaa ggaacaagca gctcaaggta	1200
ggtgaggtt tattatggct tgggttgctt ccagtggtta gtggcacagt gaaatcctgt	1260
gttacagaga cgcagactac tgctgctgcc tcctttcagg tggcattagc tgtggccgac	1320
aacgggatcc gtattatttt caatggcaaa gatctcaatc tcgtggaacg tcgtattgct	1380
gctgtgaatc cttctgatcc tctcgggac ctgtcgaaag atgttgtgc tgcattgtac	1440
cccaggcggt ttaagggtat aacccttgaa caactcacg cggatttaac gatctacttg	1500
tacagcagtg cggctctcac tgaggcgac gtcacgtgc atttgagggt tgagcatgtc	1560
agacctactg ttgacgactc tttactccg gtgtattagt gcccgctgaa gagcggtaca	1620
ctagtgtggc ctacttgaag gctagttata accgtttctt taaacggtaa tcgttgttga	1680
aacgtcttcc ttttacaaga ggattgagct gcccttgggt tttactcctt gaacccttcg	1740

-continued

gaagaactct ttggaattcg agctcgggtac cggatccaat tcccgatcgt tcaaacattt	1800
ggcaataaag tttcttaaga ttgaatcctg ttgccggtct tgcgatgatt atcatataat	1860
ttctgttgaa ttacgttaag catgtaataa ttaacatgta atgcatgacg ttatttatga	1920
gatgggtttt tatgattaga gtcccgcgaat tatacattta atacgcgata gaaaacaaaa	1980
tatagcgcgc aaactaggat aaattatcgc gcgcgggtgc atctatgtta ctagatcggg	2040
gatcgatccc caattcgtaa tcatgggtcat agctgtttcc tgtgtgaaat tgttatccgc	2100
tcacaattcc acacaacata cgagccggaa gcataaagt taaagcctgg ggtgccta	2160
gcgtgagcta actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc	2220
tgtcgtgccca gctgcattaa tgaatcggcc aacgcgcggg gagaggcggg ttgcgtattg	2280
ggcgctcttc cgcttcctcg ctcaactgact cgctgcgctc ggtcgttcgg ctgcggcgag	2340
cggtatcagc tcaactcaaag gcggtaatc gggtatccac agaatacagg gataacgcag	2400
gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc	2460
tggcggtttt ccataggctc cgccccctg acgagcatca caaaaatcga cgctcaagtc	2520
agaggtggcg aaaccgcaca ggactataaa gataccaggc gtttccccct ggaagctccc	2580
tcgtgcgctc tcctgttccg accctgccgc ttaccggata cctgtccgcc tttctccctt	2640
cgggaagcgt ggcgctttct caatgctcac gctgtaggta tctcagttcg gtgtaggctg	2700
ttcgctccaa gctgggggtg gtgcacgaac ccccgttca gcccgaccgc tgcgccttat	2760
ccggtacta tcgtcttgag tccaaccgg taagacacga cttatcgcca ctggcagcag	2820
ccactggtaa caggattagc agagcgagg atgtaggcgg tgctacagag ttcttgaagt	2880
ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgct ctgctgaagc	2940
cagttacctt cggaaaaaga gttggtagct cttgatccgg caaacaacc accgctggta	3000
gcggtgggtt ttttgtttgc aagcagcaga ttacgcgcag aaaaaagga tctcaagaag	3060
atcctttgat cttttctacg gggcttgacg ctcaagtggaa cgaaaactca cgtaaggga	3120
ttttggtcat gagattatca aaaaggatct tcacctagat ctttttaaat taaaaatgaa	3180
gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa	3240
tcagtggagc acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc	3300
ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga	3360
taccgcgaga cccacgctca ccggctccag atttatcagc aataaaccag ccagccggaa	3420
gggcccagcg cagaagtggg cctgcaactt tatccgcctc catccagtct attaatgtt	3480
gccgggaagc tagagtaagt agttcgccag ttaatagttt gcgcaacgtt gttgccattg	3540
ctacaggcat cgtggtgtca cgctcgtcgt ttggtatggc ttcattcagc tccggttccc	3600
aacgatcaag gcgagttaca tgatcccca tgttgtgcaa aaaagcgggt agctccttcg	3660
gtcctccgat cgttgtcaga agtaagttgg ccgcagtggt atcactcatg gttatggcag	3720
cactgcataa ttctcttact gtcatgccat ccgtaagatg cttttctgtg actggtgagt	3780
actcaaccaa gtcattctga gaatagtgt tgcggcgacc gagttgctct tgcccggcgt	3840
caatacggga taataccgcg ccacatagca gaactttaaa agtgctcatc attggaaaac	3900
gttcttcggg gcgaaaactc tcaaggatct taggcgtgtt gagatccagt tcgatgtaac	3960
ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag	4020

-continued

```

caaaaacagg aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa 4080
tactcatact cttccttttt caatattatt gaagcattta tcagggttat tgtctcatga 4140
gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttcgc cgcacatttc 4200
cccgaaaagt gccacctgac gtctaagaaa ccattattat catgacatta acctataaaa 4260
ataggcgtat cacgaggccc ttctgctc 4287

```

```

<210> SEQ ID NO 13
<211> LENGTH: 190
<212> TYPE: PRT
<213> ORGANISM: Cowpea Chlorotic Mottle Virus

```

```

<400> SEQUENCE: 13

```

```

Met Ser Thr Val Gly Thr Gly Lys Leu Thr Arg Ala Gln Arg Arg Ala
1           5           10           15
Ala Ala Arg Lys Asn Lys Arg Lys Thr Arg Val Val Gln Pro Val Ile
20          25          30
Val Glu Pro Ile Ala Ser Gly Gln Gly Lys Ala Ile Lys Ala Trp Thr
35          40          45
Gly Tyr Ser Val Ser Lys Trp Thr Ala Ser Cys Ala Ala Ala Glu Ala
50          55          60
Lys Val Thr Ser Ala Ile Thr Ile Ser Leu Pro Asn Glu Leu Ser Ser
65          70          75          80
Glu Arg Asn Lys Gln Leu Lys Val Gly Arg Val Leu Leu Trp Leu Gly
85          90          95
Leu Leu Pro Ser Val Ser Gly Thr Val Lys Ser Cys Val Thr Glu Thr
100         105         110
Gln Thr Thr Ala Ala Ala Ser Phe Gln Val Ala Leu Ala Val Ala Asp
115         120         125
Asn Ser Lys Asp Val Val Ala Ala Met Tyr Pro Glu Ala Phe Lys Gly
130         135         140
Ile Thr Leu Glu Gln Leu Thr Ala Asp Leu Thr Ile Tyr Leu Tyr Ser
145         150         155         160
Ser Ala Ala Leu Thr Glu Gly Asp Val Ile Val His Leu Glu Val Glu
165         170         175
His Val Arg Pro Thr Phe Asp Asp Ser Phe Thr Pro Val Tyr
180         185         190

```

```

<210> SEQ ID NO 14
<211> LENGTH: 573
<212> TYPE: DNA
<213> ORGANISM: Cowpea Chlorotic Mottle Virus

```

```

<400> SEQUENCE: 14

```

```

atgtctacag tcggaacagg gaagttaact cgtgcacaac gaagggtctgc ggcccgttaag 60
aacaagcgga agactcgtgt ggtccaacct gttattgtag aaccatcgcc ttcaggccaa 120
ggcaaggcta ttaaagcatg gaccgggttac agcgtatcga agtggaccgc ctcttgctgcg 180
gctgccgaag ctaaagtaac ctcggctata actatctctc tccctaatga gctatcgtcc 240
gaaaggaaca agcagctcaa ggtaggtaga gttttattat ggcttggggt gcttcccaat 300
gttagtgga cagtgaatc ctgtgttaca gagacgcaga ctactgctgc tgcctccttt 360
cagggtggcat tagctgtggc cgacaactcg aaagatgttg tcgctgctat gtaccccgag 420

```

-continued

gcgtttaagg gtataaccct tgaacaactc accgcggatt taacgatcta ctgtacagc	480
agtgcggctc tcaactgagg cgacgtcatc gtgcatttgg aggttgagca tgcagacct	540
acgtttgacg actctttcac tccggtgtat tag	573

<210> SEQ ID NO 15
 <211> LENGTH: 75
 <212> TYPE: DNA
 <213> ORGANISM: Bacillus anthracis

<400> SEQUENCE: 15

agtaattctc gtaagaaacg ttctacctct gctggcccta ccgtgcctga tcgtgataat	60
gatggcattc ctgat	75

<210> SEQ ID NO 16
 <211> LENGTH: 75
 <212> TYPE: DNA
 <213> ORGANISM: Bacillus anthracis

<400> SEQUENCE: 16

agtcctgaag ctgcgtcatcc tctcgtggct gcgtatccta ttgtgcatgt tgatatggaa	60
aatattatcc tctct	75

<210> SEQ ID NO 17
 <211> LENGTH: 75
 <212> TYPE: DNA
 <213> ORGANISM: Bacillus anthracis

<400> SEQUENCE: 17

cgtattatatt tcaatggcaa agatctcaat ctcgtggaac gtcgtattgc tgctgtgaat	60
ccttctgatc ctctc	75

<210> SEQ ID NO 18
 <211> LENGTH: 75
 <212> TYPE: DNA
 <213> ORGANISM: bacillus anthracis

<400> SEQUENCE: 18

cgtaaatgatg gcaaaacctt cattgatttc aaaaagtata atgataaact ccctctctat	60
atttctaata ctaat	75

<210> SEQ ID NO 19
 <211> LENGTH: 26
 <212> TYPE: DNA
 <213> ORGANISM: artificial
 <220> FEATURE:
 <223> OTHER INFORMATION: CCMVRNA4.primers

<400> SEQUENCE: 19

ctcgagctaa tacaccggag tgaaag	26
------------------------------	----

<210> SEQ ID NO 20
 <211> LENGTH: 26
 <212> TYPE: DNA
 <213> ORGANISM: artificial
 <220> FEATURE:
 <223> OTHER INFORMATION: CCMVRNA4.F primer

<400> SEQUENCE: 20

-continued

ctgcagatgt ctacagtcgg aacagg 26

<210> SEQ ID NO 21
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: CCMV-CP-F1 primer

<400> SEQUENCE: 21

aacccatcgc ttcaggccaa 20

<210> SEQ ID NO 22
<211> LENGTH: 29
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: CCMV-CP-XbaI primer

<400> SEQUENCE: 22

gctctagaat gtctacagtc ggaacaggg 29

<210> SEQ ID NO 23
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: CCMV-CP-EcoRI primer

<400> SEQUENCE: 23

cggaattctt actaatcagg agtgaaagag 30

<210> SEQ ID NO 24
<211> LENGTH: 28
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: CCMV-CP-EcoRI-3'UTR

<400> SEQUENCE: 24

cggaatcctg gtctccttag agatcacc 28

<210> SEQ ID NO 25
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: CCMV.F primer

<400> SEQUENCE: 25

ccagaaggta attatccaag atgtag 26

<210> SEQ ID NO 26
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: artificial
<220> FEATURE:
<223> OTHER INFORMATION: NOS TERM R primer

<400> SEQUENCE: 26

cgatccccga tctagtaaca tagat 25

We claim:

1) A process for producing a heterologous peptide comprising:

- a) transforming a plant cell with a non-infectious nucleic acid comprising a nucleic acid sequence encoding a fusion peptide, wherein the fusion peptide is comprised of at least one heterologous peptide of interest and at least one viral capsid protein, and wherein the nucleic acid is integrated into the genome of the plant cell;
- b) expressing the nucleic acid in the plant cell, wherein the plant cells are grown in a suspension plant cell culture in a fermentation process;
- c) wherein the expression in the plant cell provides for in vivo assembly of the fusion peptide into a virus like particle; and
- d) isolating the virus like particle.

2) The process of claim 1, wherein the nucleic acid encoding the fusion peptide is operably linked to a promoter sequence and a terminator sequence.

3) The process of claim 2, wherein the nucleic acid further comprises a viral 3'UTR.

4) The process of claim 3, wherein the 3'UTR comprises an encapsidation signal.

5) The process of claim 2, wherein the nucleic acid further comprises an encapsidation signal.

6) The process of claim 1, wherein the viral capsid protein is derived from a virus that does not display native tropism for the plant cell host.

7) The process of claim 1, wherein the viral capsid protein is derived from a virus that displays native tropism for the plant cell host.

8) The process of claim 1, wherein the viral capsid protein is derived from an icosahedral plant virus.

9) The process of claim 8, wherein the icosahedral virus is selected from the group of viruses consisting of Cowpea Mosaic Virus, Cowpea Chlorotic Mottle Virus, and Alfalfa Mosaic Virus.

10) The process of claim 1, wherein the heterologous peptide is an antigenic peptide.

11) The process of claim 10, wherein the antigenic peptide is a *Bacillus anthracis* peptide.

12) The process of claim 11, wherein the antigenic peptide is selected from the group consisting of Seq. ID. Nos. 1-4.

13) The process of claim 1, wherein the nucleic acid is selected from the group consisting of Seq. ID. Nos. 5-12.

14) The process of claim 1, wherein the virus like particle is administered to an animal as a vaccine.

15) The process of claim 1, wherein the plant cell is a dicot.

16) The process of claim 15, wherein the plant cell is *Nicotiana tabacum*.

17) The process of claim 1, wherein the plant cell is a monocot.

18) The process of claim 17, wherein the plant cell is *Oryza sativa*.

19) The process of claim 1, wherein the nucleic acid does not contain viral sequences encoding viral proteins other than the capsid protein.

20) A plant cell comprising a non-infectious nucleic acid comprising a nucleic acid sequence encoding a fusion peptide, wherein the fusion peptide is comprised of at least one heterologous peptide of interest and at least one viral capsid protein, and wherein the nucleic acid is integrated into the genome of the plant cell.

21) The plant cell of claim 20, wherein the plant cell expresses the fusion peptide.

22) The plant cell of claim 21, wherein the fusion peptide assembles in vivo into a virus like particle.

23) The plant cell of claim 20, wherein the antigenic peptide is selected from the group consisting of Seq. ID. Nos. 1-4.

24) The plant cell of claim 20, wherein the nucleic acid is selected from the group consisting of Seq. ID. Nos. 5-12.

25) The plant cell of claim 20, wherein the plant cell is *Nicotiana tabacum*.

26) The plant cell of claim 20, wherein the plant cell is *Oryza sativa*.

27) The plant cell of claim 20, wherein the nucleic acid further comprises a 3'UTR

28) The plant cell of claim 20, wherein the nucleic acid further comprises a viral 3'UTR.

29) The plant cell of claim 28, wherein the 3'UTR comprises an encapsidation signal.

30) The plant cell of claim 20, wherein the nucleic acid further comprises an encapsidation signal.

31) The plant cell of claim 20, wherein the viral capsid protein is derived from a virus that does not display native tropism for the plant cell host.

32) The plant cell of claim 20, wherein the viral capsid protein is derived from a virus that displays native tropism for the plant cell host.

33) The plant cell of claim 20, wherein the viral capsid protein is derived from an icosahedral plant virus.

34) The plant cell of claim 33, wherein the icosahedral virus is selected from the group of viruses consisting of Cowpea Mosaic Virus, Cowpea Chlorotic Mottle Virus, and Alfalfa Mosaic Virus.

35) A nucleic acid construct comprising the sequences selected from the group consisting of Seq. ID. Nos. 5-12.

* * * * *