

(21) Application No: 1503218.8

(22) Date of Filing: 26.02.2015

(71) Applicant(s):
The Native Antigen Company
Cherwell Innovation Centre, Unit 77 Heyford Park,
UPPER HEYFORD, Oxfordshire, OX25 5HD,
United Kingdom

(72) Inventor(s):
Nicholas Alexander Roesen
Holger Hannemann
Geoffrey Hale

(74) Agent and/or Address for Service:
The Native Antigen Company
Cherwell Innovation Centre, Unit 77 Heyford Park,
UPPER HEYFORD, Oxfordshire, OX25 5HD,
United Kingdom

(51) INT CL:
C12N 7/04 (2006.01) C07K 14/005 (2006.01)
C12N 7/00 (2006.01)

(56) Documents Cited:
WO 2013/122262 A1 WO 2011/116226 A2
WO 2010/062757 A1 WO 2005/042695 A2
US 5175099 A1 US 20120052082 A1
US 20020183247 A1

(58) Field of Search:
INT CL C12N
Other: WPI, EPODOC, BIOSIS, MEDLINE

(54) Title of the Invention: **Particles comprising fusion proteins**
Abstract Title: **Virus-like particles comprising transmembrane fusion proteins**

(57) Provided is a virus-like particle comprising a membrane and a fusion polypeptide, wherein the membrane is a lipid bilayer 5 and the fusion polypeptide comprises: (a) a core polypeptide enveloped by the membrane 4; and (b) a membrane-spanning polypeptide which crosses said membrane at least once 2. The core polypeptide may be a protein such as ferritin or lipoamide acetyltransferase, or a viral capsid such as that of hepatitis E, Norwalk virus, Parvovirus B19, Rabbit haemorrhagic disease virus, murine polyoma virus or human papilloma virus. The fusion polypeptide may also comprise a linker polypeptide 3 and/or a polypeptide tag 1. Also provided are nucleic acids and vectors coding for the said particle and methods for producing the said particle. Also provided are methods for using the said particle or said nucleic acids as an immunogen or as a vaccine or for the production of a pharmaceutical composition for use in therapy.

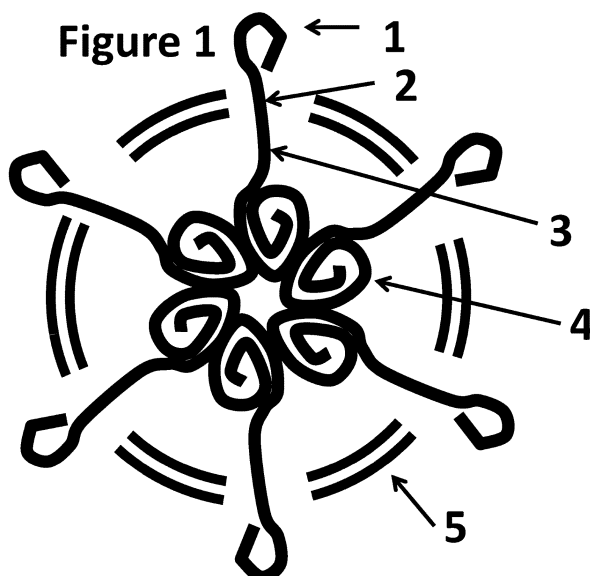
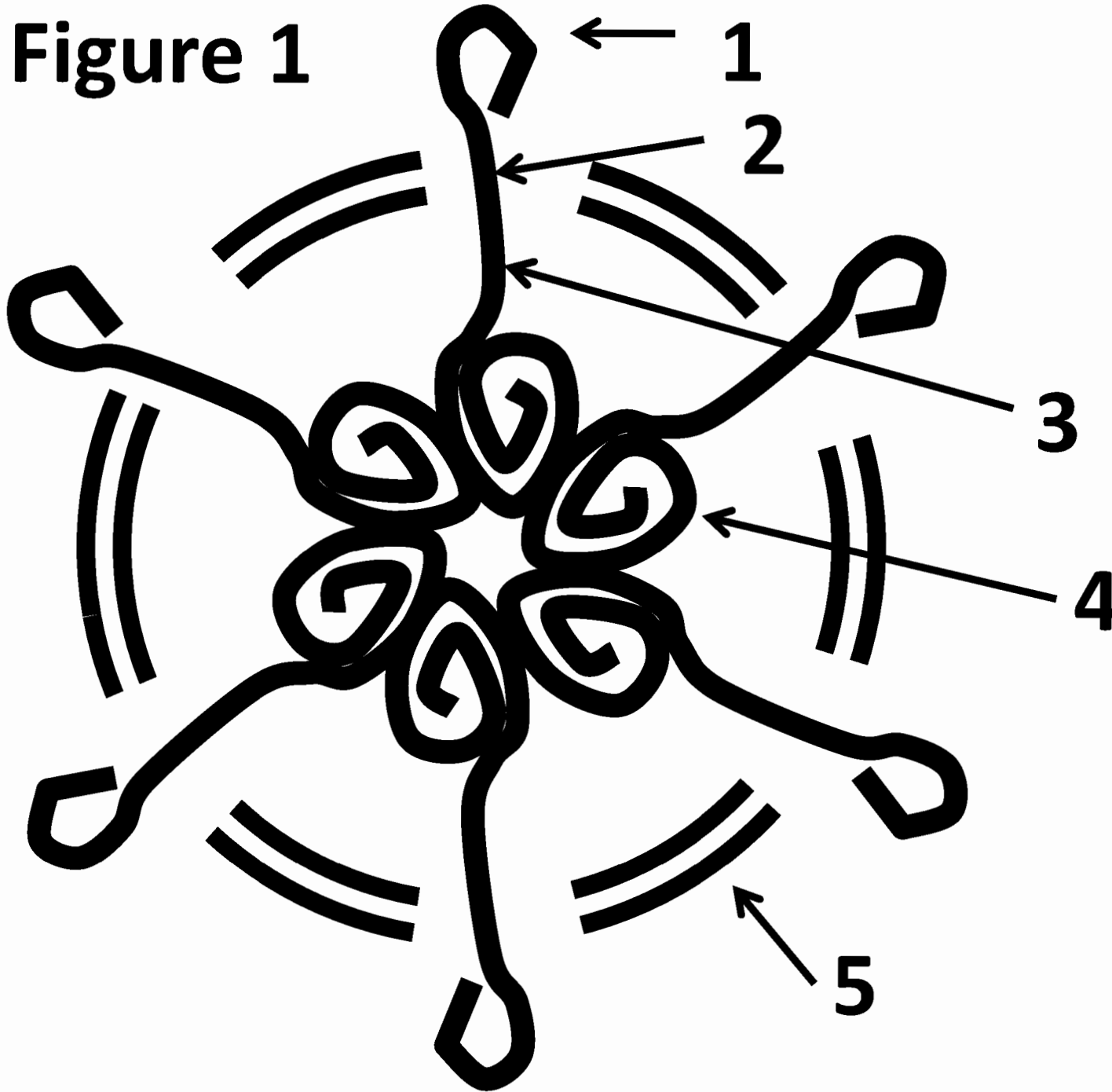


Figure 1



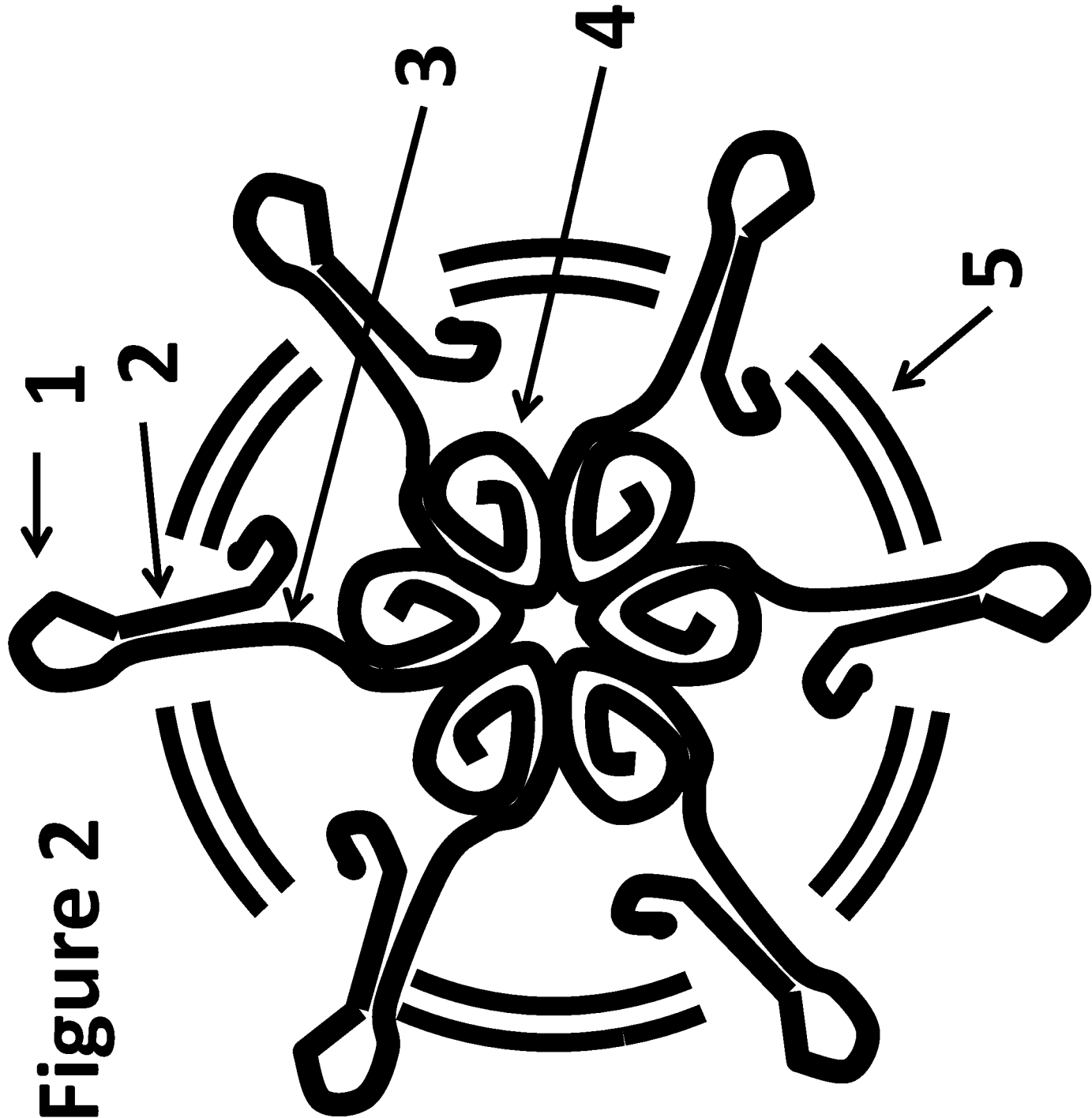


Figure 2

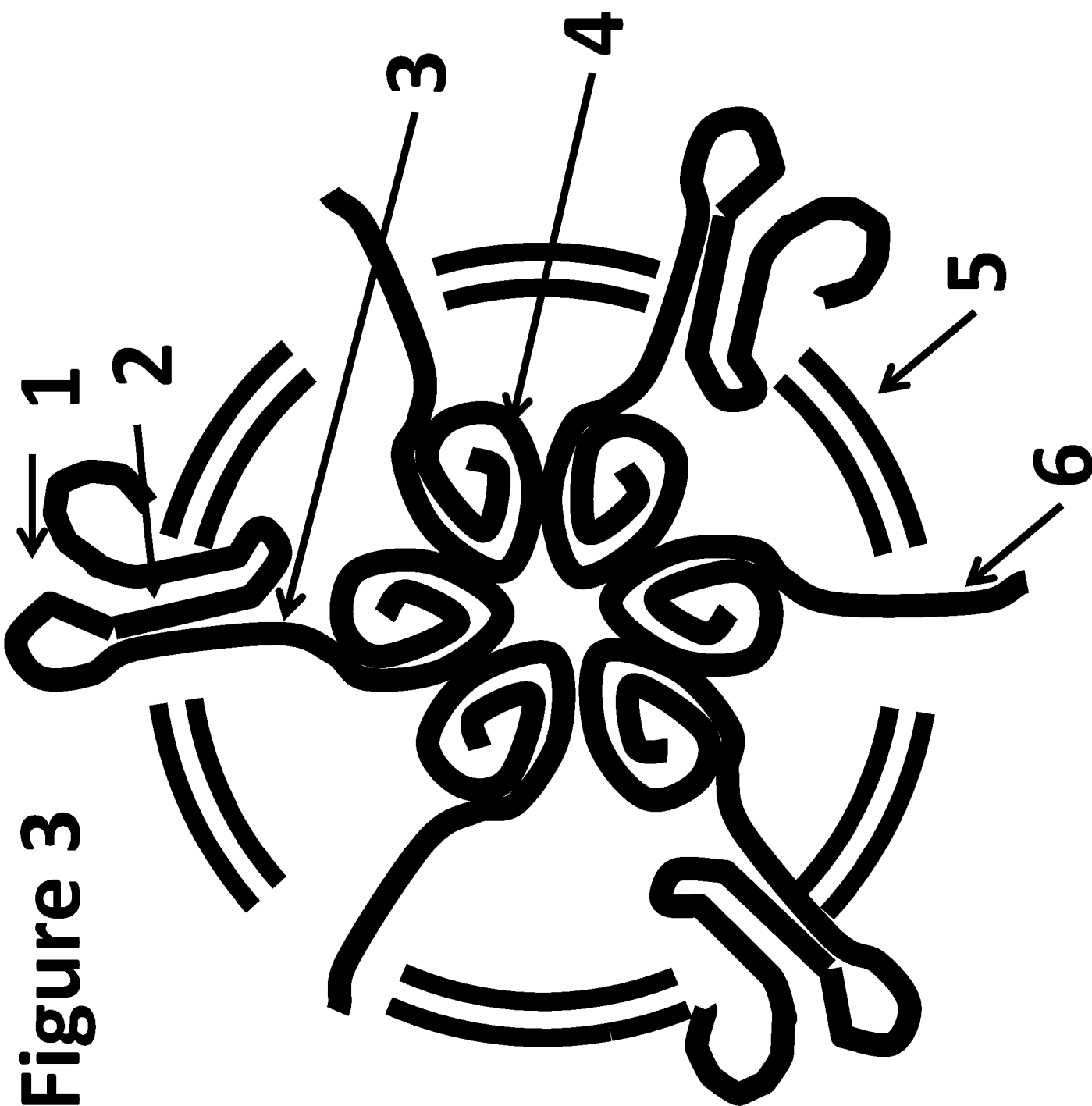


Figure 3

Figure 4

Type I



Type II



Type III



Type IV-A



NH_3^+



Type IV-B



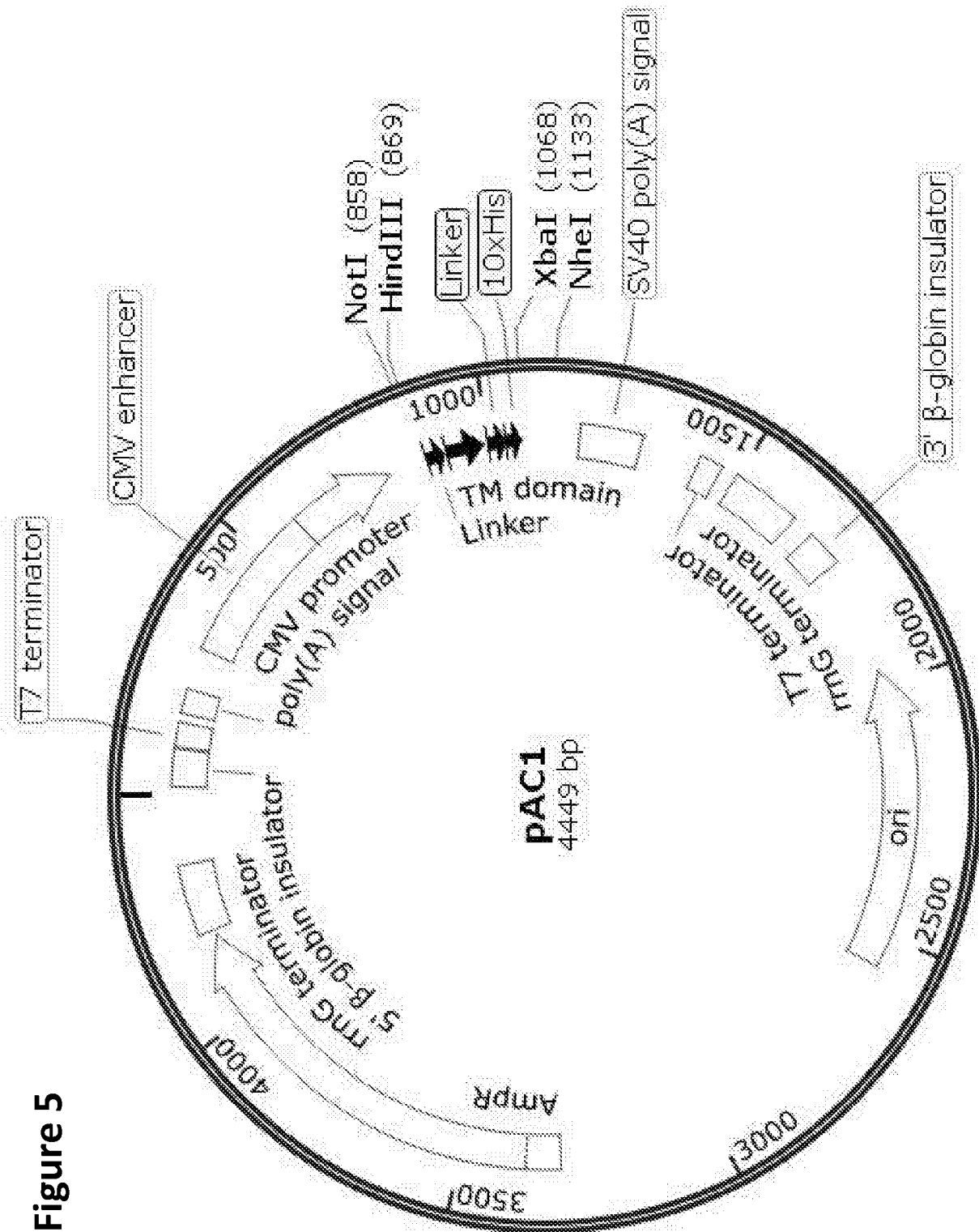
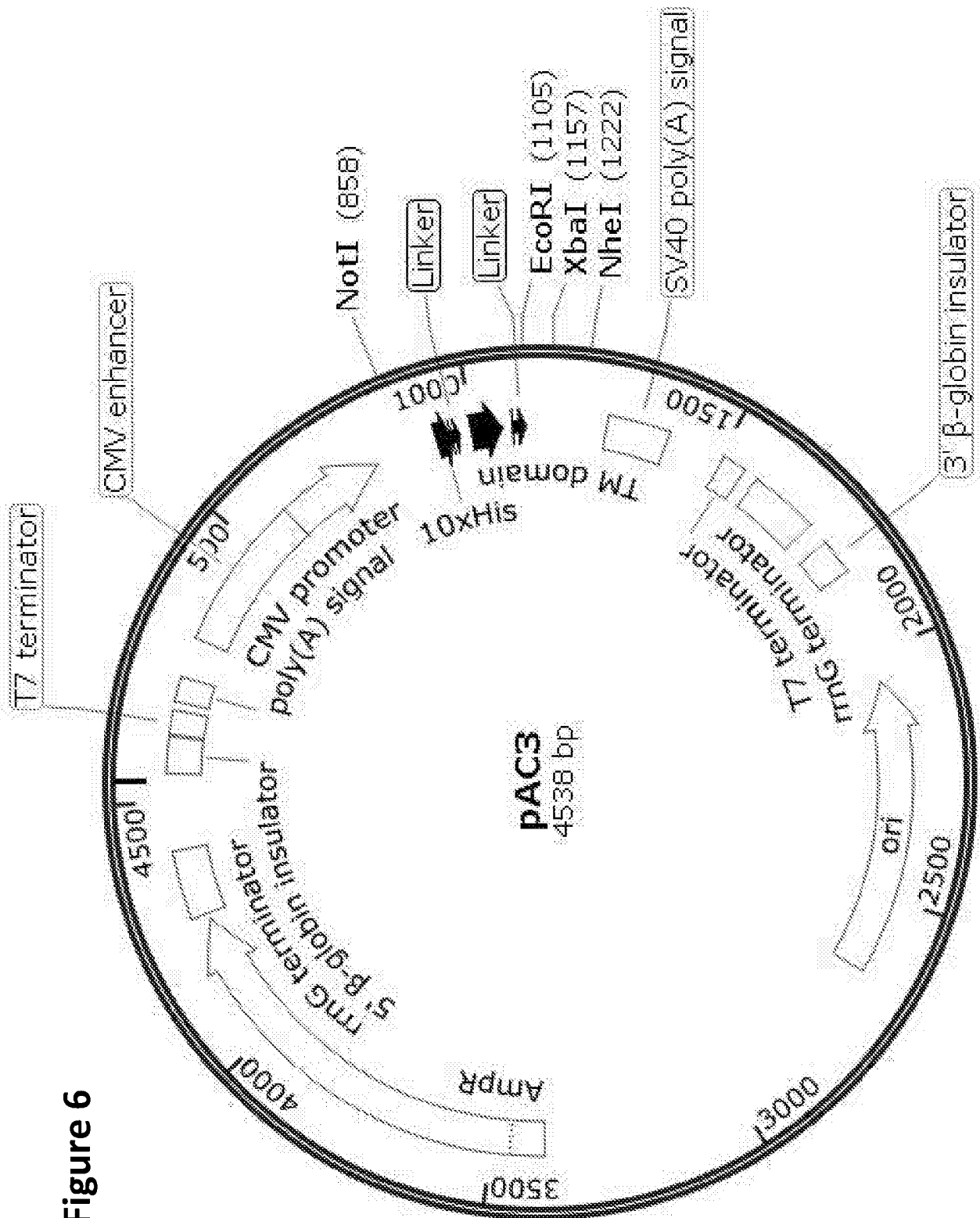


Figure 5

Figure 6



PARTICLES COMPRISING FUSION PROTEINS

BACKGROUND

Transmembrane proteins are of great importance in biology and medicine. They make up 27% of the total human proteome (Almen 2009) but are targets for more than 50% of modern drugs (Overington 2006). However, the study of transmembrane proteins presents formidable technical obstacles because they are frequently expressed at low levels and they are usually very difficult to purify (Smith 2011). As a result, transmembrane proteins represent fewer than 2% of the three-dimensional protein structures available in the protein data bank.

Integral membrane proteins are permanently attached to the membrane and can only be separated and purified using detergents, nonpolar solvents, or denaturing agents.

Transmembrane proteins are integral membrane proteins that span across the membrane at least once. The membrane-spanning peptides are most commonly alpha-helix bundles (found in all types of membrane). Beta-barrel proteins are found only in bacteria and outer membranes of mitochondria and chloroplasts.

It would be highly desirable to have a method for the efficient production of purified transmembrane proteins which did not require extraction with detergents, nonpolar solvents or denaturing agents. Avoidance of such reagents would greatly reduce the risk of damage or loss of functional integrity. The availability of purified transmembrane proteins in a native state would be greatly beneficial to biomedical research and development of pharmaceuticals. For example, they could facilitate structural and functional studies and provide candidate vaccines or antigens for raising antibodies.

Particles containing multiple copies of protein subunits are often highly immunogenic, probably because the immune system is attuned to recognition of pathogens with repetitive structures and because particles of about 40 to 50 nm are optimally taken up by dendritic cells (Fifis 2004).

Virus-like particles (VLPs) are particles which resemble viruses but do not contain nucleic acid and are therefore non-infectious. They commonly contain one or more virus capsid or envelope proteins which are capable of self-assembly to form the VLP. VLPs composed of the small hepatitis B derived surface antigen (HBsAg) occur naturally in patient sera (Bayer 1968). More recently, VLPs have been produced from components of a wide variety of virus families (Noad 2003, Grgacic 2006). Some VLPs have been approved as therapeutic vaccines, for example

Engerix-B (for hepatitis B), Cervarix and Gardasil (for human papilloma viruses). VLPs can be engineered to display heterologous antigens on their surface by creation of fusion proteins between a self-assembling VLP protein and a foreign protein. For example, the extracellular domain of the influenza A M2 ion channel protein has been fused with the hepatitis virus core protein (HBc) to create a VLP which gave mice protective immunity against a lethal challenge with influenza A virus (Neirynck 1999). Or again, a fragment of gonadotropin releasing hormone has been attached to bacteriophage AP205 coat protein to create a VLP which produced neutralising antibodies and concomitant reduction of testosterone levels in male mice (Tissot 2010). Many other examples are described in the literature (Grgacic 2006, Ludwig 2007).

In a few instances, particles have been created based on soluble self-assembling protein subunits not derived from viruses. For example the lipoyl acetyltransferase (E2) core of the pyruvate dehydrogenase multienzyme complex assembles to form an icosahedral structure which can be used as a scaffold to display a variety of soluble proteins and polypeptides (Domingo 2001). Artificial self-assembling proteins have also been created for similar purposes (Raman 2009).

However, VLPs based on non-enveloped viruses or on soluble self-assembling proteins are capable of displaying only foreign antigens which are soluble proteins. In contrast, display of a transmembrane protein in its native configuration requires the presence of a lipid membrane.

Enveloped viruses contain a lipid membrane derived from the host cell and might in principle be capable of displaying heterologous transmembrane proteins. Virus-like particles have been created based on capsid or envelope proteins of such enveloped viruses and used to display epitopes of heterologous proteins (for example Deml 2005, Grgacic 2006, Ludwig 2007, WO2013/122262). However, as explained below, substantial technical obstacles prevent the use of such VLPs for display of heterologous transmembrane proteins to meet the needs of biomedical scientists.

WO2013/122262 describes the use of VLPs based on an envelope protein of an alphavirus or a flavivirus. Certain sufficiently small heterologous peptide epitopes could be displayed by insertion of the desired sequence into an exposed loop of the envelope protein. However, as described in WO2013/122262, the attempt to incorporate a full length heterologous transmembrane protein was unsuccessful. It can be readily understood that such a construct would not be able to self-assemble due to steric interference between the heterologous transmembrane protein and the virus envelope protein.

Retroviruses such as human immunodeficiency virus (HIV) are a different class of enveloped virus which have been used as the basis for construction of VLPs. In this case the gag polyprotein complex can provide a self-assembling capsid around which the lipid membrane is organised (Deml 2005). Two strategies have been used for the incorporation of heterologous polypeptides. In so-called Type 1 VLPs sufficiently small peptide epitopes are fused within the gag polyprotein. This is analogous to the case of the alphavirus or flavivirus described above except that the epitope is now inside the membrane rather than outside. Such VLPs are suited for invoking cellular rather than humoral immunity. So-called Type 2 VLPs approach the problem of expression of transmembrane proteins by expression of the transmembrane protein separately from the gag protein. For example, US7763258 describes a VLP with an enveloped virus core and a heterologous multiple membrane spanning protein. However, the heterologous transmembrane protein is not specifically associated with the VLP; it is simply expressed on the cell membrane and is instead non-specifically gathered into the membrane of the VLP along with other cell membrane proteins. In order for the VLP to have a useful quantity of the desired heterologous protein, the latter must be expressed at a high level on the cell surface. Such VLPs do not contain a single type of transmembrane protein, as they also contain a sample of other cell-derived transmembrane proteins.

Surprisingly, we have found that it is possible to produce purified transmembrane proteins in their native state. In particular, we have shown that self-assembling fusion polypeptides, which are processed in the endoplasmic reticulum of a host cell, can be used to produce particles that contain a transmembrane protein embedded in a lipid membrane.

STATEMENT OF INVENTION

In the first aspect, the invention provides a virus-like particle comprising a membrane and a fusion polypeptide, wherein the membrane is a lipid bilayer and the fusion polypeptide comprises: (a) a core polypeptide enveloped by the membrane; and (b) a membrane-spanning polypeptide which crosses said membrane at least once.

In some embodiments the core polypeptide is capable of self-assembly to form an oligomeric structure greater than 400 kDa in size. The oligomer may be a hetero-oligomer, of two or more different subunits, or a homo-oligomer of identical subunits.

In some embodiments, the core polypeptide may be selected from the group consisting of: ferritin, encapsulin, dihydrolipoylsuccinyltransferase (EC 2.3.1.61), dihydrolipoamide branched chain

transacylase (EC 2.3.1.168), lipoyl acetyltransferase (EC 2.3.1.12) or may be a polypeptide fragment or a polypeptide homologue of a member of the said group.

In some embodiments, the core polypeptide is ferritin. In some embodiments, the core polypeptide is lipoyl acetyltransferase.

Alternatively, in some embodiments, the core polypeptide may be a capsid protein or a polypeptide fragment or a polypeptide homologue of the capsid protein wherein the N-terminus or the C-terminus of the core polypeptide is accessible on the outside of the assembled capsid.

In some embodiments the core polypeptide may be selected from a group consisting of: Newton's T=3 capsid protein, Bacteriophage GA capsid protein, Bacteriophage PP7 capsid protein, Bacteriophage AP205 capsid protein, Ryegrass mottle virus capsid protein, Turnip yellow mosaic virus capsid protein, Melon necrotic spot virus capsid protein, Pyrococcus furiosus virus-like protein, Norwalk virus capsid protein, Hepatitis E virus capsid protein, Rabbit haemorrhagic disease virus capsid protein or may be a polypeptide fragment or a polypeptide homologue of a member of the said group,

In some embodiments, the membrane-spanning polypeptide may be any transmembrane polypeptide. In some embodiments, it may be one or more selected from a group consisting of: Type I membrane proteins, Type II membrane proteins, Type III membrane proteins, Type IV-A membrane proteins, Type IV-B membrane proteins or is a polypeptide fragment or a polypeptide homologue of a member of the said group.

In some embodiments, the fusion polypeptide may also comprise one or more linker polypeptides. For example, a linker polypeptide may be fused at its N-terminus to the C-terminus of the membrane-spanning polypeptide and at its C-terminus to the N-terminus of the core polypeptide. Alternatively a linker polypeptide may be fused at its N-terminus to the C-terminus of the core polypeptide and at its C-terminus to the N-terminus of the membrane-spanning polypeptide.

In some embodiments, the fusion polypeptide also comprises a polypeptide tag, including an affinity, epitope or fluorescent tag. The tag may incorporate a protease recognition site.

In a further aspect, the invention provides a polynucleotide coding for the fusion polypeptide described above. The polynucleotide may be DNA or RNA.

In a further aspect, the invention provides a vector comprising the said polynucleotide.

In a further aspect, the invention provides a host cell comprising the said polynucleotide, wherein said host cell is capable of expressing the particle described above. The host cell may have been transfected with the polynucleotide or the polynucleotide may have been inherited upon cell division.

In a further aspect, the invention provides a method of preparing the particle described above comprising: culturing the said host cell with a culture medium; and purifying the particle from the culture medium or from the host cell. The method optionally includes biosynthesis of the particle in the endoplasmic reticulum of the host cell. This reduces the possibility of non-fused (that is non-specific) membrane proteins transecting the membrane of the particle.

In further aspects, the invention provides pharmaceutical and vaccine compositions comprising the particle described above and DNA vaccine compositions comprising the polynucleotide described above.

In further aspects, the invention provides a method of producing an antibody, comprising contacting the particle and/or the polynucleotide described above to a bird or mammal and a method of selecting an antibody, comprising screening an antibody library for binding to the particle described above.

In a further aspect, the invention provides an antibody so produced or selected, as well as a pharmaceutical composition comprising the antibody produced or selected as described above.

In a further aspect, the invention provides a medicament for use in therapy comprising any of the pharmaceutical, vaccine or DNA vaccine compositions described above.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention will now be described with reference to the following, non-limiting figures, wherein:

Figure 1 shows an example of one embodiment of the particle of the invention wherein an oligomerised core polypeptide (4) has been genetically fused to a single linker polypeptide (3), genetically fused to a single transmembrane protein (1+2). The core is surrounded by a lipid bilayer membrane (5), the membrane is transected by a single membrane-spanning domain (2). The membrane-spanning polypeptide displays an exodomain (1) externally. For simplicity, six

copies of the fusion protein have been represented, though it is envisaged that there will typically be more copies in total, for example 24, 60 or a multiple of 60 such as 120, 180, 240 or 420.

Figure 2 shows an example of one embodiment of the particle of the invention wherein an oligomerised core polypeptide (4) has been genetically fused to a single linker polypeptide (3), genetically fused to a single transmembrane protein (1+2). The core is surrounded by a lipid bilayer membrane (5), the membrane is transected by two membrane-spanning domains (2). The membrane-spanning polypeptide displays a transmembrane protein exodomain (1) externally and a cytoplasmic domain internally. For simplicity, six copies of the fusion protein have been represented, though it is envisaged that there will typically be more copies in total, for example 24, 60 or a multiple of 60 such as 120, 180, 240 or 420.

Figure 3 shows an example of one embodiment of the particle of the invention wherein an oligomerised core polypeptide (4) has been genetically fused to a single linker polypeptide (3), genetically fused to a single transmembrane protein (1+2). The core is surrounded by a lipid bilayer membrane (5), the membrane is transected by three membrane-spanning domains (2). The membrane-spanning polypeptide displays two exodomains (1) externally and a cytoplasmic domain internally. Some of the fusion proteins are truncated (6) and have only a single membrane-spanning domain. For simplicity, six copies of the fusion proteins have been represented, though it is envisaged that there will typically be more copies in total, for example 24, 60 or a multiple of 60 such as 120, 180, 240 or 420.

Figure 4 is a diagrammatic representation of the topology of some examples of fusion proteins which contain a core polypeptide (Core) linked by a linker polypeptide (Linker) to a transmembrane protein. Five different types of transmembrane proteins are illustrated. They contain domains which are exposed to the lumen of the endoplasmic reticulum (and eventually the outside of the particle) or the cytosol (and eventually the inside of the particle). These are linked by a membrane-spanning domain, either an internal stop-transfer anchor sequence (STA) or an internal signal-anchor sequence (SA). Type I transmembrane proteins have no SA domain and instead have a signal sequence at the N-terminus which is often cleaved off during processing in the endoplasmic reticulum. Note that the core polypeptide must be linked to a cytosolic domain of the transmembrane protein. Type IV-A transmembrane proteins have such a domain at each end of the polypeptide chain, so the core polypeptide may be attached to the N-terminus or C-terminus.

Figure 5 is a diagram of plasmid pAC1, an acceptor vector for linking core polypeptides to the N-terminus of a membrane polypeptide.

Figure 6 is a diagram of plasmid pAC3, an acceptor vector for linking core polypeptides to the C-terminus of a membrane polypeptide.

DETAILED DESCRIPTION OF THE INVENTION

Transmembrane proteins are of great importance in biology and medicine but are very difficult to purify in a native form. This invention provides biosynthetically-produced, self-assembling membrane-bound virus-like particles which are capable of displaying specific transmembrane proteins. These virus-like particles can be isolated and used in a range of applications, including for example as antigens, vaccines or as a source of transmembrane protein for biochemical or structural studies.

Virus-like particles based on non-enveloped viruses or on soluble self-assembling proteins are capable of displaying only foreign antigens which are soluble proteins, or soluble domains of membrane proteins. In contrast, display of a transmembrane protein in its native configuration requires the presence of a lipid membrane.

As the art shows, it is not trivial to provide a lipid membrane without incorporating large number of unwanted cellular membrane proteins. These can interfere with use of the VLP as a means to generate immune responses (for example by displaying unwanted epitopes or antigens). The presence of additional or un-specified cellular membrane proteins can also take up much needed space on the membrane itself and can even lead to steric hindrance. This is significant in nano-sized particles. Furthermore, the art does not provide a core polypeptide *fused* to the transmembrane polypeptide. Accordingly, until this invention, the problem of producing purified transmembrane proteins for structural or functional studies or for use as antigens or vaccines had not been solved.

The present invention allows for the efficient production of purified transmembrane proteins without extraction with detergents, nonpolar solvents or denaturing agents, thus reducing the risk of damage or loss of functional integrity. We provide purified transmembrane proteins which are in a native state. This is of great use in biomedical research and the development of pharmaceuticals. These native trans-membrane antigens can be used, for example, in structural and functional studies and can provide candidate vaccines or antigens for raising antibodies. The native state of the transmembrane protein is typically that found in a wild-type or un-transfected cell for the particular host or transformant species (e.g. bird or mammal, in particular chickens,

mice, hamsters, rats and other rodents, goats, rabbits, llamas, sheep or primates including humans, and so forth). This native state is achieved by using self-assembling fusion polypeptides which are processed in the endoplasmic reticulum (ER).

The transmembrane protein embedded in a lipid membrane may be a single transmembrane protein in the sense that just one copy of a particular transmembrane protein is provided. However, multiple copies of the same transmembrane protein may also be provided, provided that they are fused with the core. Indeed, two or more different transmembrane proteins may also be provided, with the same proviso. What distinguishes the present selected or specified transmembrane proteins from the art, where numerous cellular proteins are found in the lipid membrane of the particle, is that the present transmembrane proteins are fused with the core during formation of the particle as it buds off the endoplasmic reticulum (ER). The fusion typically occurs via a linker as discussed elsewhere. It is envisaged that the linker may be cleaved after particle formation (budding off from the ER), but it is preferred that this does not occur and that the core remains fused with the transmembrane protein (i.e. that the linker remains intact).

Five different types of transmembrane proteins are exemplified as illustrated in Figure 4. In some embodiments, the membrane-spanning polypeptide comprises one or more domains which are exposed to the lumen of the endoplasmic reticulum. These will eventually present on the outside of the particle. Alternatively, in some embodiments, the membrane-spanning polypeptide comprises one or more domains which are exposed to the cytosol (and will eventually present on the inside of the particle). In either case, these may be linked, in some embodiments, by a membrane-spanning domain, either an internal stop-transfer anchor sequence (STA) or an internal signal-anchor sequence (SA).

Certain embodiments of the invention can be visualised in Figures 1-3. Any or all of the following features are preferred and in any combination. A core polypeptide which is capable of forming oligomers has been genetically fused to a single linker polypeptide. The linker has been genetically fused to a single membrane-spanning polypeptide. The core is surrounded, encompassed or enveloped by or within a lipid membrane. The core may, therefore, be inside, not through, the membrane.

The membrane is transected at least once by the membrane-spanning polypeptide so that the membrane-spanning polypeptide has at least one single membrane-spanning domain. In some embodiments, the membrane-spanning polypeptide is encoded with more than one membrane-spanning domain and the exposed domains on the ER lumen or cytosolic sides are cleaved post translationally or after insertion in the membrane.

The membrane-spanning polypeptide may, in some embodiments, displays transmembrane protein exodomains externally.

The core is, in some embodiments, an oligomerised polypeptide. It may be fused to a linker polypeptide. The linker may also be fused to the membrane-spanning polypeptide. In either or both cases, this fusion is, in some embodiments, genetic, i.e. encoded by a single polynucleotide.

The membrane-spanning polypeptide may comprise or be further linked or fused to a polypeptide tag.

The particle is preferably in the nanometre range, i.e. typically in the range of 1 nm to 1000 nm. This distinguishes it greatly from a cell (aside from the absence of non-specified cellular membrane proteins). Preferably, the particle may be 8 nm to 500 nm. In some embodiments, the particle may be 10 nm to 400 nm. In some embodiments, the particle may be 15 nm to 200 nm. In some embodiments, the particle may be 10 nm to 150 nm. In some embodiments, the particle may be 10 nm to 100 nm. In some embodiments, the particle may be 10 nm to 40 nm. In some embodiments, the particle may be 15 nm to 30 nm. In some embodiments, the particle may be 20 nm to 40 nm. The particles measured in the Examples, for instance, were about 18 nm in diameter which is consistent with the size expected for a ferritin core (11 nm) surrounded by a lipid membrane.

This invention is not limited to specific compositions or process steps, as such may vary.

Whenever appropriate in this specification and the appended claims, terms used in the singular form will also include the plural and vice versa.

Unless defined otherwise, all technical and scientific terms have the same meaning as commonly understood by one of ordinary skill in the art to which this invention is related. For example, the Concise Dictionary of Biomedicine and Molecular Biology, Juo, Pei-Show, 2nd ed., 2002, CRC Press; The Dictionary of Cell and Molecular Biology, 3rd ed., 1999, Academic Press; and the Oxford Dictionary Of Biochemistry And Molecular Biology, Revised, 2000, Oxford University Press, provide general dictionaries of many of the terms used in this invention.

Amino acids may be referred to by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

The meaning of certain terms used in this invention may be different from the meanings of similar terms used in the literature. Definitions of such terms, in the context of this invention, are provided below.

Accessible. The N-terminus or C-terminus of a core polypeptide is considered to be "accessible" if it can form a peptide bond with another polypeptide, for example a linker polypeptide or a membrane-spanning polypeptide, and the resulting fusion protein is still capable of capable of self-assembly to form an oligomeric structure.

Antibody. An "antibody" is a molecule with one or more protein chains containing at least one domain which can potentially bind to an antigen (the "antigen-binding domain"). It may be an intact immunoglobulin of any isotype (for example IgG, IgE, IgM, IgD, IgA or IgY), subtype (for example IgG1, IgG2, IgG2a, IgG2b, IgG2c, IgG3, IgG4, IgA1 and IgA2) or allotype. It may be a fragment of an immunoglobulin (for example Fab, F(ab')₂, scFv, dAb). It may be of any species (for example human, monkey, camel, llama, goat, sheep, rabbit, rat, mouse, hamster or chicken) or it may be a hybrid derived from more than one species. It may be a polyclonal antibody or a monoclonal antibody. It may be naturally occurring or may be created by genetic engineering or phage display (for example a chimeric antibody, humanised antibody, camelised antibody, intrabody, domain antibody, single-chain variable region).

Antibody Library. An "antibody library" is a collection of different antibodies. The antibodies may be produced by any means, for example, from animal sera, from hybrid myeloma cell lines, from isolated lymphocytes, or by bacteriophage display.

Capsid Protein. A "capsid protein" is a protein which comprises the shell (capsid) of a virus or a virus-like particle or a peptidic nanoparticle which resembles a virus-like particle. It comprises multiple identical subunits which are capable of self-assembly (possibly together with other capsid proteins) to form a regular structure, for example an icosahedron or a helical cylinder.

Core Polypeptide. A "core polypeptide" is a polypeptide which lacks a membrane-spanning domain or a signal peptide and is capable of self-assembly to form an oligomeric structure. It may therefore form the core or central part of a particle of this invention.

Culture medium. "Culture medium" is an aqueous mixture which is capable of supporting the growth of a host cell and permitting the expression of a product encoded by a gene contained within its transfected nucleic acid.

Fusion polypeptide. A "fusion polypeptide" (sometimes known as a fusion protein or a chimeric protein) is a polypeptide created through the joining of two or more genes that originally or in the wild type coded for separate polypeptides. Transcription and translation of this fusion gene results in a single or multiple polypeptides with properties derived from each of the original proteins. Naturally occurring fusion polypeptides can be found in cancer cells. In the context of this invention, the term "fusion polypeptide" is taken to mean a fusion polypeptide created by genetic engineering, and which may include genes or gene segments which are altered (compared with an original gene which coded for a separate polypeptide). A fusion polypeptide may consist of single or multiple polypeptides, any of which may be post-translationally modified.

Host Cell. A "host cell" is a cell which is capable of being transfected with nucleic acid to produce a cell which expresses the product of a gene contained within the nucleic acid.

Icosahedral Virus. An "icosahedral virus" possesses one or more capsid proteins which assemble with 5:3:2 rotational symmetry having 20 equilateral triangular faces and 12 regular pentagon faces. Examples of icosahedral viruses are available at the VIPER database: <http://viperd.b.scripps.edu/> (Natarajan 2005)

Immunomodulation. "Immunomodulation" is the adjustment of an immune response to a desired level. Various forms of immunomodulation may be used as therapeutic interventions for treatment of disease. They may include: (a) immunopotential, induction or enhancement of an immune response, for example for the prevention or treatment of infectious disease or cancer, (b) immunosuppression, reduction or abrogation of an immune response, for example for the treatment of autoimmune disease, (c) induction of immunological tolerance, for example for the prevention or treatment of transplant rejection or autoimmune disease.

Linker Polypeptide. A "linker polypeptide" (sometimes known as a spacer) is a sequence of amino acids inserted into a fusion polypeptide between two other polypeptide segments for the purpose of ensuring or improving on the folding, activity, or spacing between the said segments. Various methods are well known in the art for design of flexible, rigid or cleavable linker polypeptides (for example Arai 2001, Wriggers 2005, Chen 2013). Non-limiting examples that we used include Gly-Ser linkers (SEQ ID NO: 33 to SEQ ID NO: 37), lipoyl domain from lipoamide acetyltransferase (SEQ ID NO: 45), EAAAK linker (SEQ ID NO: 46), and modified Moloney murine leukaemia virus gag protein (SEQ ID NO: 47).

Lipid membrane. A "lipid membrane" (sometimes known as lipid bilayer) is a thin polar membrane made of two layers of lipid molecules. Natural lipid membranes are found in all living

organisms and usually comprise phospholipids with a hydrophilic head and two hydrophobic fatty acid tails and may include other lipids, for example cholesterol. Such lipid membranes comprise part of a particle of this invention.

Membrane Protein. A "membrane protein" is a protein which interacts with a lipid membrane. There are many types and classifications of membrane proteins. Integral membrane proteins are permanently attached to the membrane. Transmembrane proteins (otherwise known as polytopic integral membrane proteins) span the membrane. Monotopic integral membrane proteins are attached to one side. Peripheral membrane proteins are attached to the membrane or to integral membrane proteins by electrostatic or hydrophobic forces. Transmembrane proteins may be further classified in relation to the number of membrane-spanning domains and the position of the N- and C- terminal domains as follows (Lodish et al. 2007). Types I, II, and III are single-pass molecules. Type I transmembrane proteins are anchored to the lipid membrane with a stop-transfer anchor (STA) sequence and have their N-terminal domains targeted to the endoplasmic reticulum (ER) lumen during synthesis (and the extracellular space, if mature forms are located on plasmalemma). Type II and III are anchored with a signal-anchor (SA) sequence, with type II being targeted to the ER lumen with its C-terminal domain, while type III have their N-terminal domains targeted to the ER lumen. Type IV-A are multi-pass molecules with an N-terminal domain targeted to the cytosol and Type IV-B are multi-pass molecules with an N-terminal domain targeted to the lumen.

Membrane-spanning domain. A "membrane-spanning domain" (sometimes known as a transmembrane domain) is a peptide which can potentially cross a lipid membrane once and make up part of a transmembrane protein. It typically consists of an alpha helix with between 5 and 25 mostly hydrophobic amino acids (the core) flanked by 1 to 5 mostly charged amino acids (the caps) on either side (Krogh 2001). Non-limiting examples that we have used include membrane-spanning domains from human transferrin receptor 1 (SEQ ID NO: 38) and influenza haemagglutinin (SEQ ID NO: 42).

Membrane-spanning polypeptide. A "membrane-spanning polypeptide" is a transmembrane protein or a fragment thereof which includes at least one membrane-spanning domain.

Passive Immunization. "Passive immunization" is the administration of an antibody to a human or animal subject with the intention of altering the immunity of the subject towards one or more antigens. Passive immunization may be used for the treatment or prophylaxis of disease, for example infectious disease, autoimmune disease, cancer, transplant rejection, cardiovascular disease.

Polypeptide Fragment. A "polypeptide fragment" is a polypeptide which has fewer amino acids than a similar naturally-occurring polypeptide and still retains a desirable characteristic of the naturally-occurring polypeptide (for example the ability to self-assemble into oligomers, exhibit an antigenic epitope or express a physiological activity). Compared with the similar naturally-occurring polypeptide, the polypeptide fragment may have one or more amino acid residues deleted from the N-terminus, the C-terminus, elsewhere in the sequence, or any combination of such deletions.

Polypeptide Homologue. A "polypeptide homologue" is a polypeptide which has statistically-significant sequence similarity to a naturally-occurring polypeptide and shares a desirable characteristic with the naturally-occurring polypeptide (for example the ability to self-assemble into oligomers, exhibit an antigenic epitope or express a physiological activity)

Polypeptide tag. A "polypeptide tag" is a peptide sequence which is genetically grafted onto a recombinant polypeptide in order to provide a novel property. An affinity tag (for example, polyhistidine, chitin-binding protein, maltose binding protein, glutathione-S-transferase) may allow purification using an affinity binding technique. An epitope tag (for example, FLAG-tag, V5-tag, Myc-tag, and HA-tag) may allow binding of an antibody. A fluorescent tag (for example, green-fluorescent protein or its variants) may allow visualisation of the recombinant polypeptide. A protein tag may allow post-translational modification (for example biotin ligase target sequence) or chemical modification (for example, fluorescein-arsenical helix binder FIAsh). Optionally, a tag may incorporate a protease recognition site for cleavage from the recombinant protein. Non-limiting examples that we have used include polyhistidine tag (SEQ ID NO: 39), FLAG tag (SEQ ID NO: 12) and V5 tag (SEQ ID NO: 20).

Post-translational modification. "Post-translational modification" is a step in protein biosynthesis whereby a polypeptide chain is modified in any of a number of ways including for example folding, cutting, cross-linking, modification of amino acid residues or addition of other molecules. Many different types of post-translational modification are known to exist in nature, some of the most common being N-linked glycosylation, phosphorylation, acetylation, methylation, palmitoylation, citrullination, sulfation, O-linked glycosylation and amidation (Khoury 2011) as well as formation of disulphide bonds, folding, specific proteolytic cleavages and assembly of multimeric proteins.

Protease cleavage site. A "protease cleavage site" is a peptide sequence which is amenable to specific cleavage by a protease.

Vector. In the context of this invention a "vector" is a polynucleotide used as a vehicle to artificially carry foreign genetic material into another cell, where it can be replicated and/or expressed. The host cell may be a prokaryote but is preferably a eukaryote. The vector may be a plasmid, a viral vector, a cosmid or an artificial chromosome. A vector for expression of a polypeptide in a eukaryotic cell typically comprises: (a) sequences necessary for replication in bacteria including an origin of replication and an antibiotic resistance gene (for example for resistance to kanamycin or ampicillin), (b) sequences necessary for expression of the foreign genetic material in eukaryotic cells including a multicloning site for insertion of the foreign genetic material, at least one promoter sequence to drive expression of the foreign genetic material, a polyadenylation signal, a ribosome recognition site eg a Kozak sequence and optionally (c) a sequence coding for a marker selectable in eukaryotic cells to confer, for resistance to an antibiotic such as G418, hygromycin or puromycin. A non-limiting example of an expression vector we have used is pSV-CMV (SEQ ID NO: 5).

Virus-like particle. A "virus-like particle" is a particle which resembles a virus but is not infectious, normally because it does not contain viral nucleic acid.

Overview

Particles of the invention ideally have any, and preferably all, of the following features:

1. They contain a core polypeptide which is capable of self-assembly to form a stable oligomeric structure
2. The core polypeptide is surrounded by a lipid membrane
3. A membrane-spanning polypeptide is embedded in the lipid membrane
4. The core polypeptide and the membrane-spanning polypeptide have been part of the same fusion polypeptide at some point.

In certain embodiments, the core polypeptide and the membrane-spanning polypeptide are derived from different sources. They are not parts of a single polypeptide which can be found in nature. Rather, they have been joined together in a fusion polypeptide.

In certain embodiments, the intact fusion polypeptide may be found in the particle of the invention or in other embodiments it may have been subjected to proteolytic cleavage (either during biosynthesis or subsequently) so that the membrane-spanning polypeptide and the core polypeptide need not be covalently connected in the particle.

In certain embodiments the fusion polypeptide may contain any number of other parts which facilitate the assembly and/or use of the particle. These other parts may, for example, be selected from the group including: linker polypeptide, membrane-spanning domain, polypeptide tag, protease cleavage site.

In certain embodiments, the particle of this invention is produced by a eukaryotic cell wherein the fusion polypeptide is subjected to processing within the endoplasmic reticulum. As a result, it may be subject to post-translational modification in a manner which is characteristic of the eukaryotic cell.

The core polypeptide: cellular proteins

In certain embodiments the core polypeptide is derived from a cellular protein. The protein is ideally capable of self-assembly. Examples of such proteins include ferritin, dihydrolipoylsuccinyltransferase (EC 2.3.1.61), dihydrolipoamide branched chain transacylase (EC 2.3.1.168), lipoyl acetyltransferase (EC 2.3.1.12). These proteins may be derived from any prokaryote or eukaryote, including for example bacteria, plants, invertebrates, vertebrates, mammals, humans. Other examples of potential core polypeptides include encapsulin and lumazine synthase (EC 2.5.1.78) which may be derived for example from bacteria. The core polypeptide may be the same as the natural form of the soluble protein or it may be altered in any number of the following ways: mutation of one or more amino acid residues, deletion of one or more amino acid residues, insertion of one or more amino acid residues, post-translational modification. Thus the core polypeptide may be a polypeptide fragment or a polypeptide homologue provided that it retains the capability of self-assembly into an oligomeric structure.

In preferred embodiments the core polypeptide is derived from lipoyl acetyltransferase. This enzyme (sometimes called dihydrolipoyl acetyltransferase or E2) forms the self-assembling core of the pyruvate dehydrogenase multienzyme complex (Perham 1991). In some species (for example *E. coli* and *Azotobacter vinelandii*) the core consists of 24 subunits with octahedral symmetry whereas in other species (for example *Bacillus stearothermophilus*, *Bacillus subtilis*, *Streptococcus faecalis*, *Saccharomyces cerevisiae* and mammals including humans) the core consists of 60 subunits with icosahedral symmetry. Lipoyl acetyltransferase which form cores of 60 subunits is preferred because such a particle will, in principle, be larger than a core consisting of 24 subunits and so may be better able to display a membrane-spanning polypeptide.

Lipoyl acetyltransferase monomers consist of one, two or three highly homologous lipoyl domains (each about 8kDa), linked together by long flexible peptides, followed by a binding domain of about 5 kDa which binds peripheral enzyme subunits and which is linked in turn to a C-terminal core domain of about 29 kDa. It is this core domain which forms the self-assembling core of the complex and contains the enzyme active site (Perham 1991). The E2 core enzymes (dihydrolipoylsuccinyltransferase and dihydrolipoamide branched chain transacylase) found in other 2-oxo acid dehydrogenase multienzyme complexes have similar structures, though there are differences in the number of lipoyl domains between enzymes of different species.

The domain structure of the E2 enzymes provides an opportunity to construct core polypeptides of different sizes which may be suitable for displaying different membrane-spanning polypeptides. In certain embodiments the core polypeptide consists of only the core domain and in other embodiments the core polypeptide consists of the binding domain linked to the core domain or one or more lipoyl domains linked to the binding domain linked to the core domain. In principle a larger core polypeptide will provide a greater surface area of lipid membrane to support larger membrane-spanning polypeptides, especially those which contain multiple membrane-spanning domains.

The core polypeptide: viral proteins

In certain embodiments the core polypeptide is derived from a viral capsid protein which is capable of self-assembly. There are many such proteins known in the literature. The most common self-assembled structures are icosahedra and helical cylinders. Not all viral capsid proteins may be suitable candidates for a core polypeptide. For example, some may have an N-terminus and C-terminus which is buried within the structure or exposed only on an inner surface and incapable of providing an attachment point for a fusion protein without disrupting the assembly. Ultimately it is necessary to test the suitability of a proposed capsid protein experimentally. This can be done, for example, using methods described in the Examples. But when the three-dimensional structure of the capsid protein (or a structurally related family member) is already known, it is possible to select potential candidates by examination of the structure to identify the position of the ends of the protein chain, for example as described in **Example 2** and **Example 3**.

By 2005 more than 70 three dimensional structures of icosahedral viral capsids were known and deposited in the VIPER database (Nataragan 2005). By 2008 the database contained 256 entries from 28 families of virus (Carillo-Tripp 2008). In January 2015 the database contained 544 entries including 271 unique structures from 38 families of viruses. As described in **Example 3**, it is

possible to interrogate the database to identify short-lists of viral capsids with properties that may make them preferable candidates for use as a core polypeptide. For example, one can identify a list of candidates which have amino acid residues at the N-terminus or the C-terminus or both the N-terminus and the C-terminus of the protein chain which are likely to be exposed on the outside of the oligomeric particle. This approach has some limitations. For example: (1) The ends of the protein chain may not be defined in the crystal structure. However, even a truncated end can be informative, since it is unlikely that unresolved residues are buried or cross the capsid shell. (2) Not all structures have been solved to high resolution and therefore the structural data for chain ends may not be definitive. (3) Attachment of protein chains to the N-terminus or C-terminus to create fusion proteins might disrupt folding and/or assembly of the viral capsid. For these reasons, short-listed candidates need to be experimentally tested to determine whether they are suitable for use as a core polypeptide. Nevertheless, the process of creating a short-list by examination of structures reduces the amount of experimental work and may also facilitate the design of fusion proteins.

In preferable embodiments the oligomeric capsid protein has a diameter greater than 20 nm. The surface area of a core polypeptide made from a smaller capsid may be insufficient for the presentation of substantial membrane polypeptides without the addition of polypeptide spacers. In some embodiments the oligomeric capsid protein has a diameter in the range 20 nm to 100 nm. In some embodiments it has a diameter in the range 20 nm to 80 nm. In some embodiments it has a diameter in the range 20 nm to 60 nm. In some embodiments it has a diameter in the range 20 nm to 40 nm. Using data available in the VIPER database It is straightforward to refine a short-list to include only capsid proteins within a desirable range of diameters.

Following is a non-exclusive short list of viruses or virus-like particles which may provide capsid proteins that are candidates for consideration as core polypeptides by virtue of their size and potential exposure of N-terminus and/or C-terminus: Bacteriophage GA, Bacteriophage PP7, Bacteriophage AP205, Ryegrass mottle virus, Turnip yellow mosaic virus, Melon necrotic spot virus, *Pyrococcus furiosus* virus-like protein, Norwalk virus, Hepatitis E virus, Rabbit haemorrhagic disease virus,

The core polypeptide: purpose-designed proteins

In certain embodiments the core polypeptide is a protein which has been designed to self-assemble into a suitable oligomeric structure. For example, EP1594469 describes peptidic nanoparticles which consist of aggregates of a continuous chain comprising two peptidic oligomerization domains connected by a linker segment. If a trimer-forming domain is linked to a

pentamer-forming domain, the resulting aggregates may contain a significant proportion of icosahedral particles (Raman 2009) and thus may be suitable candidates for core polypeptides of the present invention, particularly because both the N- and C-termini of the polypeptide are both exposed near the outside of the particle. For example, the polypeptide identified as Newtons T=3 capsid (ID upana1503) in the VIPER database or P6c peptide (Yang 2012) is a candidate for consideration as a core polypeptide.

The polypeptide may be further engineered to introduce other desirable properties. For example it may be stabilised by introduction of Cys residues in suitable positions to form inter-chain disulphide bonds, the net charge may be manipulated by altering the content of acidic or basic amino acid residues, the overall size may be manipulated by altering the length of the oligomerization domains or other modifications may be made as further described below.

Modified core polypeptides

In some embodiments the sequence of a native protein providing a core polypeptide is altered, for example by deletion or mutation of amino acid residues, to create a polypeptide fragment or polypeptide homologue of the original core polypeptide. Such a modified core polypeptide may have superior properties compared with the original core polypeptide, for example: a higher level of expression in mammalian cells, improved stability with respect to thermal, oxidative and/or proteolytic degradation, the ability to form larger particles and/or expressible fusion proteins with a greater variety of transmembrane polypeptides.

In some embodiments, two native core polypeptides may be fused together with a suitable linker polypeptide to improve the stability of the particle, for example as described by Caldeira (2010).

Core polypeptides: unpredictability of formation of particles

As discussed above, candidate core polypeptides can be selected from published literature based on rational principles such as the accessibility of N- and C-termini in the three-dimensional structure, the size and stability of the particle and ease of expression in mammalian cells. As shown in **Example 13** and **Example 15** there is a reasonable correlation between the theoretical predictions and the measured expression of particles. Nevertheless, there are exceptions; surprisingly some theoretically plausible particles are not detected and others are detected even though they were considered to be theoretically implausible. Therefore it remains necessary to experimentally test a candidate core polypeptide to determine whether it is capable of forming fusion polypeptides and particles suitable for a particular application. We provide several

examples of core polypeptides suitable for different applications and straightforward methods for identification of others.

The membrane-spanning polypeptide

An especially useful feature of this invention is the ability for a range of different membrane-spanning polypeptides to be incorporated into the particles of this invention. Naturally occurring transmembrane proteins have a number of different structures and many of them can be incorporated into the particles of this invention with a suitable choice of core polypeptide, and, if necessary, the inclusion of suitable linker polypeptide(s) and/or suitable protease cleavage site(s).

Throughout the following description, it is to be understood that a membrane-spanning polypeptide may be a naturally-occurring transmembrane protein or a polypeptide fragment of polypeptide homologue of such a naturally-occurring transmembrane protein. Examples of different topologies which are discussed below are illustrated in Figure 4.

In certain embodiments the membrane-spanning polypeptide is a Type I transmembrane protein (for example low-density lipoprotein receptor, influenza haemagglutinin protein, insulin receptor, growth hormone receptor and many others well known in the art).

In certain embodiments the membrane-spanning polypeptide is a Type II transmembrane protein (for example asialoglycoprotein receptor, transferrin receptor, DC-SIGN and many others well known in the art)

In certain embodiments the membrane-spanning polypeptide is a Type III transmembrane protein (for example cytochrome P450 and many others well known in the art)

In certain embodiments the membrane-spanning polypeptide is a Type IV-A or Type IV-B transmembrane protein (for example a G-protein coupled receptor, glucose transporter, voltage-gated Ca^{2+} ion channels, ABC small molecule pump, and many others well known in the art)

In certain embodiments the membrane-spanning polypeptide has an odd number of membrane-spanning domains with an N-terminal domain which is targeted towards the lumen of the endoplasmic reticulum and a C-terminal domain which remains in the cytosol (for example when the membrane-spanning polypeptide is a Type I, Type III, or Type IV-B transmembrane protein). In such cases the fusion polypeptide of the invention is preferably constructed with the

membrane-spanning polypeptide linked, either directly or with a linker polypeptide, to the N-terminus of the core polypeptide.

In other embodiments the membrane-spanning polypeptide has an odd number of membrane-spanning domains with an N-terminal domain which remains in the cytosol while the C-terminal domain is targeted to the lumen of the endoplasmic reticulum (for example when the membrane-spanning polypeptide is a Type II transmembrane protein). In such cases the fusion polypeptide of the invention is preferably constructed with the membrane-spanning polypeptide linked, either directly or with a linker polypeptide, to the C-terminus of the core polypeptide.

In other embodiments the membrane-spanning polypeptide has an even number of membrane-spanning domains with both N-terminal and C-terminal domains which remain in the cytosol (for example when the membrane-spanning polypeptide is a Type IV-A transmembrane protein). In such cases the fusion polypeptide of the invention may be constructed with the membrane-spanning polypeptide linked, either directly or with a linker polypeptide, to either the N-terminus or the C-terminus of the core polypeptide.

In certain embodiments the topology of a transmembrane protein may be altered by addition or deletion of membrane-spanning domains so that the membrane-spanning polypeptide is a polypeptide fragment or polypeptide homologue with a different, and potentially more convenient topology than the transmembrane protein on which it is based. So for example, a Type II transmembrane protein might be modified by addition of a stop-transfer anchor membrane-spanning domain to its C-terminus in order to retain the C-terminal domain in the cytosol which could be linked to the N-terminus of the core polypeptide. Or for example, a Type I, Type III or Type IV-B transmembrane protein might be modified by additional of an internal signal anchor sequence to its N-terminus in order to retain the N-terminal domain in the cytosol which could be linked to the C-terminus of the core polypeptide.

Mixtures of fusion polypeptides containing different membrane-spanning polypeptides

It can be appreciated that there is likely to be an upper limit to the number of membrane-spanning domains which can be included within a given fusion polypeptide. This limit is likely to depend on the size of the core polypeptide and the diameter of the assembled particle which it can form. Too many membrane-spanning domains would create steric hindrance to the successful assembly of the particle. Therefore, in certain embodiments the particle consists of a mixture of at least two types of fusion protein where all types have a similar or identical core polypeptide but at least one type has the desired membrane-spanning polypeptide (with 2, 3, 4, 5, or more membrane-

spanning domains) and at least one type has a membrane-spanning polypeptide with a smaller number of membrane-spanning domains. An example of such a particle is illustrated in Figure 3. In preferred embodiments, the proportion of the different types of fusion polypeptide in the particle is adjusted to give optimal expression of the larger membrane-spanning polypeptide, for example as described in **Example 18**.

Linker polypeptides

In certain embodiments the particle of this invention may contain one or more linker polypeptides, each of which may have different properties and purposes.

In certain embodiments the core polypeptide and the membrane-spanning polypeptide are connected by means of a linker polypeptide. Such a linker polypeptide is located inside the lipid membrane and may facilitate the assembly of the particle of this invention. For example, it may provide flexibility so that the core polypeptide and the membrane-spanning polypeptide may each be aligned in a configuration which is favourable for self-assembly of the particle. Or for example, it may provide a relatively rigid spacer so that the core polypeptide and the membrane-spanning polypeptide may be separated by a sufficient distance to be favourable for self-assembly of the particle.

In certain embodiments a linker polypeptide may link two domains of the membrane-spanning polypeptide, or it may link the membrane-spanning polypeptide to a polypeptide tag. Such a linker polypeptide is located on the outside of the lipid membrane of the particle of this invention. It may facilitate the folding or display of the membrane-spanning polypeptide or the folding or display of a polypeptide tag.

It is well known in the art that the construction of fusion proteins involves the linking of two proteins or two domains of a protein by means of a peptide linker. Suitable selection of the linker is important for the correct folding and function of the fusion protein. There are several studies of linker design and selection (for example Argos 1990, Alfthan 1995, Robinson 1998, Crasto 2000, Arai 2001, George 2002, Wriggers 2005, Chen 2013). Two independent studies (Argos 1990, George 2002) examined databases of naturally-occurring peptide linkers to identify common properties and preferred structures. Other investigators have designed empirical linkers (for example flexible linkers, rigid linkers, cleavable linkers) for different applications (reviewed by Chen 2013). Flexible linkers are generally composed of small non-polar (eg Gly) or polar (eg Ser or Thr) amino acids. Incorporation of Ser or Thr maintains stability in aqueous solution. Commonly used flexible linkers consist of repeating units (GGGS)_n. However, the genetic

stability and expression of a fusion protein containing many repetitive units might be compromised and in such a case a flexible linker more closely based on a natural linker peptide would be preferred. Rigid linkers may be advantageous when it is desired to maintain a certain separation between the core polypeptide and the membrane-spanning polypeptide. Peptides which form an alpha-helical structure are preferred and are often stabilised by salt bridges. Commonly used rigid linkers include repeating units of (EAAAK)_n (Arai 2001). Another type of rigid linker includes repeating units (XP)_n where X may be any amino acid but preferably Ala, Lys or Glu (George 2003).

In certain embodiments the linker polypeptide may comprise one or more naturally occurring protein domains, for example the lipoyl domain of lipoamide acetyltransferase (SEQ ID NO: 45) or a polypeptide homologue of a naturally occurring protein domain, for example modified Moloney murine leukaemia virus gag protein (SEQ ID NO: 47).

Because of their relative accessibility compared with other parts of the fusion polypeptide, linker polypeptides may be susceptible to cleavage by proteases. This may or may not be desirable. For example, a stable linker polypeptide may be required for favourable assembly of the particle of this invention and display of the membrane-spanning polypeptide. In such a case, the sequence of the linker polypeptide is preferably selected to avoid potential protease cleavage sites. Or for example, a cleavable linker polypeptide may be required in order that the membrane-spanning polypeptide be released and displayed in a particular (for example native) conformation. In such a case, the sequence of the linker polypeptide is preferably selected to introduce a desirable protease cleavage site.

In certain embodiments the linker polypeptide includes a furin cleavage site based on the recognition sequence Arg-Xaa-Arg ▼ Lys-Arg, where Xaa represents any amino acid and ▼ indicates the cleavage site. Furin is mainly localised in the trans-Golgi network and thus may be able to cleave such a linker polypeptide as it is made within the cell.

In certain embodiments the linker polypeptide includes a recognition site for a protease which is not normally found within the host cell. This may be a broad-specificity protease for example chymotrypsin, papain, plasmin, or trypsin or more preferably a protease with more restricted sequence specificity for example: factor Xa protease (Ile-(Glu/Asp)-Gly-Arg ▼), enterokinase (enteropeptidase) light chain (Asp-Asp-Asp-Asp-Lys ▼), tobacco etch virus protease (Glu-Asn-Leu-Tyr-Phe-Gln ▼ Ser), thrombin (Leu-Val-Pro-Arg ▼ Gly-Ser) or HRV 3C protease (PreScission) (Leu-Glu-Val-Leu-Phe-Gln ▼ Gly-Pro).

Polypeptide tags

In certain embodiments the particle of this invention may contain one or more polypeptide tags, each of which may have different properties and purposes. Many such tags are well known in the art (Terpe 2003).

In certain embodiments the polypeptide tag is connected to the core polypeptide, either directly or by means of a linker polypeptide. Such a polypeptide tag is located within the lipid membrane and in the intact particle of this invention it will generally be inaccessible to polar substances. Such a polypeptide tag might be useful for example for targeting degradation products of the particle in vivo to a particular cellular compartment.

In certain embodiments the polypeptide tag is connected to or inserted within the membrane-spanning peptide, either directly or by means of a linker polypeptide in such a way that it is located on the outside of the lipid membrane of the particle of this invention. Such a polypeptide tag will be accessible to a range of polar substances, including antibodies, ligands and chemical reagents for protein modification and accordingly can be useful for different purposes as exemplified below.

In certain embodiments the polypeptide tag comprises an epitope that can be recognised by an antibody. This may be useful for affinity purification, detection or visualisation of the particles of the invention. Examples of antibody epitope tags include: CD52 mimotope (TSSPSAD), E-tag (GAPVPYPDPLEPR), Flag-tag (DYKDDDDK), HA-tag (YPYDVPDYA), His-tag (HHHHHH), Myc-tag (EQKLISEEDL), S-tag (KETAAAKFERQHMDs), Softag 1 (SLAELLNAGLGGS), Softag 3 (TQDPSRVG), Ty1-tag (EVHTNQDPLD), V5-tag (GKIPNPLLGLDST), VSV-tag (YTDIEMNRLGK), Xpress tag (DLYDDDDK). These are SEQ ID NO: 10 to SEQ ID NO: 22.

In certain embodiments the polypeptide tag comprises a peptide which binds to a non-antibody ligand. These may also be useful for purification, detection or visualisation of the particles of the invention. Examples of such peptide tags include: Arg-tag (RRRRR), calmodulin-binding peptide (KRRWKNFIAVSAANRFKKISSSGAL), Glu-tag (EEEEEE), His-tag (HHHHHH), SBP-tag (MDEKTTGWRGGHVVEGLAGELEQLRARLEHHPQGQREP) Strep-tag II (WSHPQFEK). These are SEQ ID NO: 23 to SEQ ID NO: 28.

In certain embodiments the polypeptide tag comprises a protein which confers a useful property on the particle of the invention. Examples of protein tags and some of their possible uses include: biotin carboxyl carrier protein (BCCP), a protein domain biotinylated by biotin ligase enabling

recognition by avidin or streptavidin; cellulose-binding domain for affinity purification; chitin-binding domain for affinity purification, glutathione-S-transferase, a protein which binds to immobilized glutathione allowing convenient purification by affinity chromatography; green fluorescent protein, a protein which is spontaneously fluorescent; maltose binding protein, a protein which binds to amylose agarose allowing convenient purification; Fc derived from immunoglobulin Fc domain allows dimerization and can be used for affinity purification.

In certain embodiments the polypeptide tag comprises a peptide which is amenable to post-translational modification, either in vivo or in vitro. Such tags may permit the addition of a useful chemical moiety to a specific site on the particle of the invention. Examples of tags which are amenable to post-translational modification and some of the possible modifications include: AviTag, a peptide allowing biotinylation by the enzyme biotin ligase enabling recognition by avidin or streptavidin (GLNDIFEAQKIEWHE); fluorescein-arsenical helix binder FIAsh for site-specific modification with a fluorochrome (CCXXCC); Isopeptag, a peptide which binds covalently to pilin-C protein (TDKDMTITFTNKKDAE), SpyTag, a peptide which binds covalently to SpyCatcher protein (AHIVMVDAYKPTK). These are SEQ ID NO: 29 to SEQ ID NO: 32.

In certain embodiments any of the above polypeptide tags may be linked to a linker polypeptide which contains a protease cleavage site (for example factor Xa protease, enterokinase, tobacco etch virus protease, thrombin or HRV 3C protease) in order to facilitate removal of the tag when it is no longer required.

Nucleic acids

In another aspect, this invention also provides nucleotide sequences encoding the fusion polypeptide of this invention. The polynucleotides of the invention may be obtained, and the nucleotide sequence of the polynucleotides determined, by any method known in the art. Preferably the codon usage in the nucleotide sequence is optimised to facilitate expression of the fusion polypeptide in the chosen host cell. Various methods for codon optimisation are known in the art (for example see US 8,825,411).

Vectors

In another aspect, this invention also provides vectors that contain a polynucleotide encoding a fusion polypeptide of the invention. In an exemplary embodiment, a polynucleotide that encodes a fusion polypeptide may be incorporated into a vector in order to express particles of the invention in a suitable host cell. A variety of expression vectors may be used. They may comprise self-

replicating extra-chromosomal vectors or vectors which integrate into a host genome. Expression vectors are constructed to be compatible with the host cell type. As is known in the art, a variety of expression vectors are available, commercially or otherwise, that may find use for expressing fusion polypeptides. Preferred expression vectors include those which enable expression of fusion polypeptides in eukaryotic cells, for example yeast, plant, insect or mammalian cells. Especially preferred expression vectors include those which enable expression of fusion polypeptides in mammalian cells, including human cells, for example the vector pSV-CMV (SEQ ID NO: 5). An example of expression vector pSV-CMV adapted for linking core polypeptides to the N-terminus of a membrane polypeptide is illustrated in Figure 7 (plasmid pAC1, SEQ ID NO: 1). An example of expression vector pSV-CMV adapted for linking core polypeptides to the C-terminus of a membrane polypeptide is illustrated in Figure 8 (plasmid pAC3, SEQ ID NO: 3).

Host cells

In another aspect, this invention also provides host cells comprising a vector as described above which can be used for expression of a fusion polypeptide of the invention. If the fusion polypeptide is not glycosylated and does not contain disulphide bonds which are critical for its correctly folded structure, then suitable host cells may include prokaryotes such as *Escherichia coli* or *Bacillus subtilis*. For expression of a glycosylated or extensively disulphide-bonded fusion polypeptide, a eukaryotic host cell is preferable. Suitable eukaryotes may include yeast (for example *Saccharomyces cerevisiae*, *Pichia pastoris*), plants (for example *Nicotiana species*), insects (for example *Spodoptera frugiperda*) or mammalian cell lines. Preferred host cells include mammalian cell lines. Suitable mammalian cell lines available as hosts for expression of recombinant polypeptides are well known in the art and include immortalized cell lines available from the American Type Culture Collection (ATCC), including but not limited to Chinese hamster ovary (CHO) cells, HeLa cells, baby hamster kidney (BHK) cells, monkey kidney cells (COS), human hepatocellular carcinoma cells (Hep G2), and human embryonic kidney (HEK) 293 cells.

The vector is introduced into the host cell by transfection or transduction using any suitable technique known in the art. An example of transfection of HEK293 cells is provided in **Example 6**.

In certain embodiments, fusion polypeptides and particles of the invention are expressed in a cell line with stable expression of the fusion polypeptide. In other embodiments, fusion polypeptide and particles of the invention are expressed in a cell line with transient expression of the fusion polypeptide. The cell line, either stable or transiently transfected, is maintained in cell culture medium and conditions well known in the art resulting in the expression and production of particles of the invention.

Preparation of particles of the invention

In another aspect, this invention also provides a method for preparation of particles of the invention. Host cells as described above are cultured to allow the expression of the fusion polypeptide under suitable conditions as are well known in the art (for examples see Vicente 2011, Zeltins 2013). General principles of particle purification are well known and include the following steps: (i) lysis of the cell to transfer the synthesized particles to the solution, (ii) clarification to remove cell debris and other large aggregates; (iii) concentration and (iv) purification to remove residual impurities. Preferably the particles of the invention are secreted into the culture medium in which case lysis of the cells is not necessary. In the event that the particles are retained within the cells, then a suitable method is required to obtain the cell-free extract. For example, eukaryotic (including mammalian and human) cells can be lysed with a suitable detergent. Bacterial, yeast or plant cells may require mechanical treatment such as mills, French press, ultrasonication or repeated freeze/thawing. A suitable method for a particular case may be determined by analysis of the yield and integrity of the particles in the lysate. To protect the particles from oxidation or proteolytic degradation during the purification process, the extraction buffer may be supplemented with reducing and chelating agents and protease inhibitors. Nucleases may be added at this stage to reduce the amount of nucleic acids in the extracts.

The cell extract or culture supernatant containing the particles may be clarified by low-speed centrifugation or filtration or by low-speed centrifugation followed by filtration. Preferably centrifugation conditions (time and g-force) are optimised to obtain a high yield of particles in the supernatant whilst pelleting a high proportion of cells and/or cellular debris. Filtration may be accomplished in one or more steps. A coarse (depth) filter may be used first to remove cells and large debris, followed by a filter with smaller pore size to remove cell fragments and small debris. A preferred pore size is selected with reference to the anticipated size of the particles to avoid their retention. A suitable pore size for a particular case may be determined by analysis of the yield and integrity of the particles in the filtrate.

The clarified cell extract or culture supernatant may be concentrated by any of a number of methods known in the art, for example ultrafiltration or precipitation with an agent such as ammonium sulphate or polyethylene glycol. Suitable conditions for a particular case may be determined by analysis of the yield and integrity of the particles in the concentrate.

Culture medium containing the product may be separated from host cells by centrifugation or by filtration or by centrifugation followed by filtration. Alternatively, if particles are retained within the cells they may be lysed and the product solubilised using methods known in the art.

The particles may be purified by methods known in the art for purification of viruses and virus-like particles, for example, by chromatography (for example ion exchange, affinity, and size exclusion chromatography), centrifugation, differential solubility, or by any other standard technique for the purification of proteins.

The preferred methods used for the final purification of the particles depend on their intended application and the scale of the process. For example, laboratory-scale purifications for experimental or research use may be accomplished using ultracentrifugation in density gradients of caesium chloride or sucrose. However, this may be relatively labour intensive. For industrial-scale processes other methods such as chromatography may be preferable. Depending on the particular properties of the particles, ion-exchange, size-exclusion or affinity chromatography may be useful. Size-exclusion chromatography is a generic method which may be used for the majority of different particles, but is less easy to scale than other methods. Incorporation of a suitable polypeptide tag into the particle may greatly facilitate the use of ion-exchange or affinity chromatography.

If the particles are to be used for the preparation of a pharmaceutical composition or a vaccine composition, then additional purification steps well known in the art may be required to remove actual or potential impurities or contaminants and to comply with regulatory requirements.

Compositions and medicaments for use in therapy

In another aspect, this invention provides for pharmaceutical or vaccine compositions. In certain embodiments pharmaceutical or vaccine compositions comprise the particle of this invention. In other embodiments DNA vaccine compositions comprise the polynucleotide of the invention. Any of these compositions may also comprise a pharmaceutically acceptable diluent, excipient or carrier. Vaccine or DNA vaccine compositions may also comprise a pharmaceutically acceptable adjuvant.

In certain embodiments this invention provides for medicaments comprising the particle of this invention for use in therapy. For example, such medicaments may be used for immunomodulation, for prophylaxis or treatment of infectious disease, autoimmune disease,

cancer or transplant rejection. Or for example, such medicaments may be used for vaccination, for prevention or treatment of infectious disease or cancer.

Particles of this invention may be particularly advantageous for the manufacture of pharmaceutical or vaccine compositions or medicaments for immunomodulation or vaccination compared with previously available substances because they are capable of displaying a single type of membrane-spanning polypeptide in an oligomeric particle of a suitable size for effective recognition by the immune system. Furthermore, depending on the particular host cell used for manufacture of the particles of this invention, it is possible to ensure that any post-translational modifications (for example glycosylation) are of a desirable (for example native) type or structure.

In certain other embodiments this invention provides for medicaments comprising the polynucleotide of this invention for use in therapy. Preferably the polynucleotide is comprised within a vector which is capable of causing expression of the fusion polypeptide and particles of the invention in vivo. Such medicaments may be used for vaccination, for example for the treatment of infectious disease or cancer.

In certain embodiments any of the core polypeptide, linker polypeptide or polypeptide tag comprised within the medicaments described above may in turn comprise one or more peptide sequences which can be processed by antigen-presenting cells and presented to CD4⁺ T cells to provide help for an antibody response against the membrane-spanning polypeptide comprised within the medicament.

Antibodies

In another aspect, this invention provides for a method for producing antibodies directed against a transmembrane protein. Such antibodies may be very desirable as reagents for research, diagnosis or therapy. However, it was previously difficult to produce them due to the difficulty of obtaining sufficiently pure or sufficient amounts of purified transmembrane protein. Particles of this invention are particularly advantageous for immunisation compared with previously available substances because they are capable of displaying a single type of membrane-spanning polypeptide in an oligomeric particle of a suitable size for effective uptake by antigen-presenting cells, for example macrophages or dendritic cells. In addition, a suitable core polypeptide could provide T-cell help for the generation of an antibody response against the membrane-spanning polypeptide. In addition, particles of this invention which display a single membrane-spanning polypeptide in a physiological format are likely to be especially useful reagents for screening, detection and analysis of antibodies which recognise the membrane-spanning polypeptide.

In certain embodiments animals (for example llama, goat, sheep, rabbit, rat, mouse, hamster or chicken) are immunised with particles of this invention to raise polyclonal antibodies. Methods for immunisation, collection of sera (or eggs) and purification of polyclonal antibodies are well known in the art (for example Hanly 1995; Howard 2001).

In other embodiments animals (for example llama, rabbit, rat, mouse, hamster) are immunised with particles of this invention and antibody-producing cells (for example B-lymphocytes from the spleen) are immortalised to produce monoclonal antibodies. Methods for immunisation, collection of antibody-producing cells, immortalisation and cloning of cells and production of monoclonal antibodies are well known in the art (for example Shepherd 2000; Howard 2001; Ossipow 2014).

In certain embodiments any of the core polypeptide, linker polypeptide or polypeptide tag may comprise one or more peptide sequences which can be processed by antigen-presenting cells and presented to CD4⁺ T cells to provide help for any of the antibody responses described above.

In certain embodiments antibodies are selected from phage display libraries (produced for example from human lymphocytes) using particles of this invention for selection of the desired antibodies. Methods for production of antibodies from phage display libraries are well known in the art (for example Winter 1994; Barbas 2001).

In certain embodiments polyclonal antibodies, or more preferably monoclonal antibodies raised and/or selected using particles of this invention may be used in the manufacture of pharmaceutical compositions and medicaments for the treatment of disease including for example, cancer, autoimmune disease, transplant rejection, infectious disease, cardiovascular disease, neurological disease.

EXAMPLES

The examples below are given so as to illustrate the practice of this invention. They are not intended to limit or define the entire scope of this invention.

Example 1 Selection of candidate core polypeptides by searching primary literature

The PubMed database (<http://www.ncbi.nlm.nih.gov/pubmed>) was used to search for primary scientific literature reporting proteins that self-assemble into multimeric structures greater than 20 monomers in size with evidence for self-assembly of the protein following expression in prokaryotic and/or eukaryotic cells. Protein sequences were retrieved using the UniProt database (<http://www.uniprot.org/>). Non-exclusive lists of candidates are shown in Table 1 and Table 2.

Table 1 Candidates for core polypeptides identified by searching primary literature.

Genbank Entry	UniProt Entry	Name	Reference
CAZ61323	F8J354	Aspergillus fumigatus partitivirus-1 coat protein	Ochoa 2008
CAA41364	Q65664	brome mosaic virus KU-1 coat protein	Kao 2011, Moon 2014
AFU48534.1	K0IB68	barley yellow dwarf virus PAV coat protein	na ¹
BAA02061.1	P18027	cucumber mosaic virus Y coat protein	Moon 2014
NP_056749.1	Q76QV9	potato leaf roll virus coat protein	Lamb 1996
223380	<u>P03606</u>	satellite tobacco necrosis virus coat protein	Lane 2011
NP_062899.1	<u>P11689</u>	tomato bushy stunt virus coat protein	Kumar 2009
WP_010880027.1	<u>O66529</u>	Aquifex aeolicus lumazine synthase	Ladenstein 2013
WP_000269761.1	<u>P0A998</u>	Escherichia coli ferritin	Kim 2011
NP_001922.2	P10515	human pyruvate dehydrogenase E2 (lipoamide acetyltransferase)	Zhou 2001

¹ selected because of its similarity to potato leaf roll virus

Table 2 Candidates for core polypeptides identified by searching primary literature.

Genbank Entry	UniProt Entry	Name	Reference
AAB47793	P07299	Human parvovirus B19	Rueda 1999
Q9IVZ8.1	Q9IVZ8	Hepatitis E	Li 2005
NP_056821.2	Q83884	Norwalk virus	Debbink 2013
AAM21587.1	Q8JZ14	Rabbit haemorrhagic disease virus	Luque 2012
P03090.3	P03090	Murine polyomavirus	Simon 2014
P04012.1	P040121	Human papillomavirus	Mossadegh 2014

Example 2 Selection of candidate core polypeptides by searching the Protein Data Bank

The Protein Data Bank (<http://www.rcsb.org/pdb>) was searched for homomeric multi-subunit proteins with octahedral or icosahedral symmetry, other than virus capsids. The three-dimensional models were retrieved as .pdb files and viewed using Discovery Studio Client 3.5 Visualizer software (Accelrys Software Inc). The position and accessibility of the N- and C-termini of the proteins was assessed. Where at least one terminus appeared to be accessible on the outside of the particle, the protein was considered as a potential candidate for a core polypeptide. A non-exclusive list of candidates is shown in Table 3.

Table 3 Candidates for core polypeptides identified by searching the Protein Data Bank

PDB entry	Name	Subunits	Diameter (nm)	Exposed N-terminus	Exposed C-terminus
1fha	Human ferritin	24	15	Thr-5	inner surface
1hqk	Aquifex aeolicus lumazine synthase	60	16	Met-1	Arg-154
1b5s	Bacillus stearothermophilus dihydrolipoyl transacetylase (E2), catalytic domain	60	24	Ala-184	Met-426
3dkt	Thermotoga maritima encapsulin	180	24	inner surface	Lys-267
4pt2	Myxococcus xanthus encapsulin	180	32	inner surface	Leu-238

Based on particle size and apparent accessibility of both the N- and C-termini, the E2 catalytic domain appeared to be a good potential candidate for a core polypeptide.

Example 3 Selection of candidate core polypeptides by searching the VIPER database

The VIPER database (<http://viperd.b.scripps.edu>) contains structural data determined by X-ray crystallography or cryo-electron microscopy for a large number of icosahedral viruses. Using the database browser provided on the VIPER website, a set of data was extracted which included the following parameters: PDB entry, virus name, genus, family, number of protein chains, minimum and maximum radius of the capsid, radius and solvent-accessible surface area in the intact

capsid (sasa_bound) of the endmost amino acid residues. Data for a total of 507 structures were extracted. (The endmost amino acid residues in the database may not always be the N- and C-termini of the protein chains since these are sometimes not visualised in the three-dimensional structure.) The dataset was loaded into a Microsoft Excel spreadsheet and the relative solvent-accessible surface area (sasa_bound%) was calculated for each amino acid residue as a percentage of the average solvent-accessible surface area for the corresponding reference amino acid (sasa_reference) in the VIPER database. The dataset was then filtered to identify viral capsid proteins which might provide candidate core polypeptides. The filter criteria were as follows:

- No more than one type of protein chain in the assembly
- sasa_bound% greater than 50%
- radius of residue within 30 Angstroms of the maximum radius of the capsid

It is possible to use more or less stringent filter criteria to select a smaller or larger list of candidates. For example, the terminal NH₂ or COOH of a protein chain might be sufficiently exposed even if sasa_bound% for the residue were less than or equal to 50%. Or an exposed terminal residue might be adequately accessible on the edge of a hole in the virus capsid even though it might be more than 30 Angstroms from the boundary sphere enclosing the capsid.

The three-dimensional models of each candidate were retrieved as .pdb files and viewed using Discovery Studio Client 3.5 Visualizer software (Accelrys Software Inc). The position and accessibility of the N- and C-termini of the proteins was assessed. Where there were several similar structures for a particular virus, or virus family, then a representative was selected which appeared to have the most accessible NH₂ and/or COOH group at the ends of the protein chain. From the reduced short-list, capsids were selected with a maximum radius in the range 125 to 250 Angstroms (equivalent to a diameter of 25 to 50 nm).

A non-exclusive list of the final set of candidates is shown in Table 4.

Table 4 Candidates for core polypeptides identified by searching the VIPER database

VDB entry	Name	Subunits	Diameter (nm)	Exposed N-terminus	Exposed C-terminus
upana1503	Newton's T=3 capsid	180	28	His-1	Arg-104
1gav	Bacteriophage GA	180	28	Ala-1	Ala-130
1dwn	Bacteriophage PP7	180	30	Ala-1	Arg-127

2izw	Ryegrass mottle virus	180	31	buried	Gln-235
1auy	Turnip yellow mosaic virus	180	31	buried	Thr-189
2zah	Melon necrotic spot virus	180	34	buried	Ala-390
2e0z	Pyrococcus furiosus virus-like protein	180	38	buried	Glu-345
1ihm	Norwalk virus	180	40	buried	Ser-520

Example 4 Selection of additional candidate core polypeptide

Among the candidates identified in the VIPER database (**Example 3**) were a number of bacteriophage from the levivirus family, including bacteriophage MS2, FR, GA, Qbeta, PRR1 and PP7. These all have homologous structures (Tars 1997) with N- and C-termini close to the outer surface of the capsid. Previous workers have been able to display short peptides inserted into an exposed loop of such capsids, including MS2 and PP7 (Caldeira 2010) although successful insertion of desired epitopes remained unpredictable (Tissot 2010) and creation of fusion proteins linked at the N- or C-terminus has not been reported. Bacteriophage AP205 is an RNA bacteriophage which infects *Acinetobacter* species and has little sequence similarity with other RNA phages. Its three-dimensional structure has not been reported but it has been shown that fusion proteins having up to 55 amino acid residues attached to either the N- or C-terminus of the coat protein can be expressed and assembled into VLPs (Tissot 2010; EP2351770). Therefore, despite the lack of structural information, bacteriophage AP205 may be considered as another candidate core polypeptide.

Example 5 Synthesis of vectors containing genes coding for model transmembrane proteins as a tool for screening candidate core polypeptides

Acceptor vectors containing inserts coding for model transmembrane polypeptides with a HIS tag

A set of acceptor vectors pAC1 to pAC4 was engineered by insertion of specific coding sequences into the mammalian expression vector pSF-CMV (Oxford Genetics) so that fusion proteins could be created between a desired core polypeptide and a membrane polypeptide (Table 5). Vector pSF-CMV (SEQ ID NO: 5) contains a multiple cloning site which includes unique sites for the restriction enzymes Not1 (857/859), HindIII (869/870), EcoR1 (885/886), Xba1 (937/938) and Nhe1 (1002/1003).

Table 5 Acceptor Vectors

Plasmid	For fusion to	Transmembrane domain	MMLV gag late domains	DNA sequence of insert
pAC1	C-terminus	SEQ ID NO: 38	n/a	SEQ ID NO: 1
pAC2	C-terminus	SEQ ID NO: 38	SEQ ID NO: 40	SEQ ID NO: 2
pAC3	N-terminus	SEQ ID NO: 42	No	SEQ ID NO: 3
pAC4	N-terminus	SEQ ID NO: 42	SEQ ID NO: 40	SEQ ID NO: 4

Acceptor vectors for joining the N-terminus of the membrane polypeptide to the C-terminus of a core polypeptide

Vectors pAC1 and pAC2 were designed for fusion of the N-terminus of a membrane polypeptide to the C-terminus of a core polypeptide. They contained DNA coding for: (a) a 12 amino-acid flexible Gly-Ser linker (SEQ ID NO: 33), (b) the human transferrin receptor 1 (Genbank entry NP_003225.2) transmembrane domain, amino acid residues 65-89 (SEQ ID NO: 38) flanked on both sides by the amino acid pair Lys-Glu (c) a second 12 amino-acid flexible Gly-Ser linker (SEQ ID NO: 34), (d) a 10 amino-acid His tag (SEQ ID NO: 39), (e) a stop codon. Vector pAC2 also contained an additional Gly-Ser linker (SEQ ID NO: 35) and a modified sequence derived from the late domains of the Moloney murine leukemia virus (MMLV) gag protein (Genbank entry NP_057934) (amino acids 98-170) upstream of the N-terminal flexible linker and the transmembrane domain (SEQ ID NO: 47).

DNA inserts for vectors pAC1 and pAC2 were codon optimized for expression in human cells (SEQ ID NO: 1 and SEQ ID NO: 2). The DNA was synthesized and provided in pBSK (pBSK-Sequence1 and pBSK-Sequence2). It was digested with HindIII and XbaI, and the resulting 0.2kb or 0.4kb fragment was ligated to the 4.2kb fragment of pSF-CMV digested with HindIII and XbaI, resulting in pAC1 (Figure 5) and pAC2 respectively.

Acceptor vectors for joining the C-terminus of the membrane polypeptide to the N-terminus of a core polypeptide

Vectors pAC3 and pAC4 were designed for fusion of the C-terminus of a membrane polypeptide to the N-terminus of a core polypeptide. They contained DNA coding for: (a) human serum albumin (Genbank entry AAQ89947.1) signal peptide, residues 1-18 (SEQ ID NO: 41), (b) a 10 amino-acid His tag (SEQ ID NO: 39), (c) a 7 amino-acid flexible Gly-Ser linker (SEQ ID NO: 36), (d) the negatively charged amino acid cluster Asp-Gly-Glu-Glu-Gly-Gly-Gly, (e) the human influenza virus H3N2 (Genbank AAA43200.1) haemagglutinin (HA) transmembrane domain, residues 163-183 (SEQ ID NO: 42), (f) a cluster of positively charged amino acids Gly-Gly-Gly-Arg-Gly-Arg-Lys and (g) a flexible 8 amino acid glycine-serine linker (SEQ ID NO: 37). Vector

pAC4 also contained an additional Gly-Ser linker (SEQ ID NO: 33) and a sequence derived from the late domains of the Moloney murine leukemia virus (MMLV) gag protein (Genbank entry NP_057934) (amino acids 98-170) including the PTAP/PSAP and PPxY motifs downstream of the transmembrane domain.

DNA inserts for vectors pAC3 and pAC4 were codon optimized for expression in human cells (SEQ ID NO: 3 and SEQ ID NO: 4). The DNA was synthesized and provided in pBSK (pBSK-Sequence3 and pBSK-Sequence4). It was digested with NotI and EcoRI, and the resulting 0.25kb or 0.45kb fragment was ligated to the 4.3kb fragment of NotI and EcoRI digested pSF-CMV, resulting in pAC3 (Figure 6) and pAC4 respectively.

DNA coding for core polypeptides

The coding sequence of selected core polypeptides (for example as listed in Table 1, Table 2, Table 3 Table 4) were codon optimized for expression in human cells. To the 5' end of each coding sequence (starting with ATG) the nucleotides GCGGCCGCGGAATTCTTCCACCC (SEQ ID NO: 8) were added to provide both a NotI and an EcoRI restriction enzyme site and a consensus Kozak sequence. At the 3' end of each coding sequence the naturally occurring stop codon was replaced by the nucleotides GGAAGCTTGTAGTGAGCTAGC (SEQ ID NO: 9) to provide a HindIII site, two stop codons and an NheI site. The complete sequence was synthesized and provided in the pBSK cloning vector (Agilent Technologies, catalogue number 212205).

Expression vectors

Expression vectors pAC1-E2 and pAC2-E2 were constructed by digesting acceptor vectors pAC1 or pAC2 and the DNA coding for the E2 core polypeptide (SEQ ID NO: 6) with the restriction enzymes NotI and HindIII and ligating the resulting coding fragments with the result that the sequence coding for the core polypeptide was inserted upstream of the sequence coding for the membrane polypeptide.

Expression vectors pAC3-E2 and pAC4-E2 were constructed by digesting acceptor vectors pAC3 or pAC4 and the DNA coding for the E2 core polypeptide (SEQ ID NO: 6) with the restriction enzymes EcoRI and NheI and ligating the resulting coding fragments with the result that the sequence coding for the core polypeptide was inserted downstream of the sequence coding for the membrane polypeptide.

Expression vectors for other core polypeptides were constructed in the same way.

Expression vectors for fusion proteins tagged with green fluorescence protein (GFP) were prepared in a similar way except that a gene coding for GFP (Genbank entry AAF62891.1) (SEQ ID NO: 44) was substituted for the HIS tag.

Transfection grade plasmid vectors were produced using Qiagen Plasmid Midi Kit. Purity was assessed by measurement of the ratio of absorbance at 260 and 280 nm (A_{260}/A_{280}).

Example 6 Transfection and expression of particles in HEK-293FT cells

Cell culture

HEK-293FT cells (Life Technologies) were grown in Freestyle293 Expression Medium (Life Technologies) in suspension on a platform shaker in a humidified 37°C CO₂ incubator with rotation at 140 rpm. Cells were maintained between 4×10^5 and 3×10^6 cells/mL in a volume not exceeding 20% of the total volume of the culture flask. Cultures were maintained for at least three passages to ensure stable growth prior to performing transfections.

Stock solutions of reagents

A stock solution of linear polyethyleneimine (PEI) molecular weight 25,000 (Polysciences) was prepared at 1 mg/mL in 25mM HEPES pH 7.5 containing 150 mM NaCl. The PEI was added to the buffer and vortex mixed until completely dissolved. Valproic acid (VPA) (Sigma-Aldrich) was diluted in water to a concentration of 0.5M. The reagents were sterile filtered, aliquotted and stored frozen at -20°C.

Transfection

HEK-293FT cells at a viability of >95% were pelleted by centrifugation at 300 x g for 10 min and resuspended in fresh Freestyle293 Expression Medium to give a density of 3×10^6 cells/mL. Purified DNA of the desired expression vector (with A_{260}/A_{280} ratio greater than 1.8) was added to the cells to give a final concentration of 3 µg/mL. The cells were incubated for 5 min under shaking, and then PEI was added to give a final concentration of 9 µg/mL. The cells were incubated for 24 h then diluted with an equal volume of fresh medium containing 8 mM VPA, 100 U/mL penicillin and 100 µg/mL streptomycin.

Expression and harvest:

Four days after the transfection, the cells were removed by centrifugation at 4000 x g. The culture supernatant was passed through a 0.22 µm filter and analysed by ELISA or Western blotting, or stored at -80°C. Optionally, the culture supernatant was concentrated up to 100 fold by tangential

flow filtration using a 300 kDa cut-off membrane. The concentrated culture supernatant was stored frozen at -80°C.

Example 7 Purification of particles containing a HIS tag

HEK-293FT cells were transfected with an expression vector (**Example 5**) and culture supernatant was collected as described in **Example 6**. The particles were purified by affinity chromatography using a cobalt affinity column. (HiTrap Talon crude from GE Healthcare cat no 28-9537-66). A 1 mL Talon column was equilibrated with phosphate buffered saline (PBS) pH 7.4 (Life Technologies cat no 10010-015), and 50 mL of concentrated culture supernatant was applied at a flowrate of 1 mL per min. The column was washed with 20 mL of PBS, and particles were eluted with 5 mL of 100 mM imidazole in PBS. The column was washed with 20 mL of PBS and the process was repeated 3 times. Eluates were pooled and analysed by SDS-polyacrylamide gel electrophoresis alongside a set of molecular weight markers. The gel was stained with Coomassie brilliant blue.

When this was done using an expression vector coding for the fusion protein [ferritin-linker-membrane-spanning domain-linker-his tag] (SEQ ID NO: 43) there was a single stained band in the eluate samples with an apparent molecular weight of about 25 kDa closely corresponding to the molecular weight of 26 kDa predicted for the fusion polypeptide.

Example 8 Purification of particles having a net positive surface charge

Particles having a net positive surface charge are purified by cation exchange chromatography on carboxymethyl Sepharose (HiTrap CM FF from GE Healthcare cat no 17-5056-01) optionally followed by size exclusion chromatography on hydroxylated methacrylic polymer (Toyopearl HW-65S from Tosoh Bioscience cat no 07467). A 1 mL CM FF column is equilibrated in binding buffer (eg 20 mM sodium phosphate, pH 7.4) and culture supernatant or concentrated culture supernatant is applied at a flow rate of 1 mL per min. The column is washed with binding buffer until the absorbance of the eluate at 280 nm is stable. The particles are eluted with a linear gradient of binding buffer containing from 0 to 1 M sodium chloride over approximately 20 mL. Fractions of approximately 1 mL are collected. Samples of each fraction are analysed by SDS polyacrylamide gel electrophoresis and the fractions containing protein bands of the predicted molecular weight are pooled. Optionally the particles may be further purified by size exclusion chromatography. An XK 16/100 column (GE Healthcare cat no 28-9889-47) is packed with approximately 189 mL of Toyopearl HW-65S and equilibrated with phosphate-buffered saline. The pooled eluate from the cation exchange column is applied and eluted with phosphate-

buffered saline at 1.5 mL/min. Fractions of approximately 1 mL are collected and analysed by SDS polyacrylamide gel electrophoresis as before. Fractions containing protein bands of the predicted molecular weight are pooled.

Example 9 Purification of particles having a net negative surface charge

Particles having a net negative surface charge are purified by anion exchange chromatography on diethylaminoethyl (DEAE) Sepharose (HiTrap DEAE FF from GE Healthcare cat no 17-5055-01) optionally followed by size exclusion chromatography on hydroxylated methacrylic polymer (Toyopearl HW-65S from Tosoh Bioscience cat no 07467). A 1 mL DEAE FF column is equilibrated in binding buffer (eg 20 mM triethanolamine hydrochloride, pH 7.4) and culture supernatant or concentrated culture supernatant is applied at a flow rate of 1 mL per min. The column is washed with binding buffer until the absorbance of the eluate at 280 nm is stable. The particles are eluted with a linear gradient of binding buffer containing from 0 to 1 M sodium chloride over approximately 20 mL. Fractions of approximately 1 mL are collected. Samples of each fraction are analysed by SDS polyacrylamide gel electrophoresis and the fractions containing protein bands of the predicted molecular weight are pooled. Optionally the particles may be further purified by size exclusion chromatography as described in **Example 8**.

Example 10 SDS-PAGE and Western blotting

HEK-293FT cells were transfected with vectors encoding various fusion polypeptides (**Example 5**) and culture supernatant was collected as described in **Example 6**.

Samples of the cells and culture supernatant were analysed by SDS polyacrylamide gel electrophoresis. 20 µL of cell suspension was centrifuged for 1 min at 10,000 × g and the pellet was resuspended in 50 µL of NuPAGE LDS sample buffer (Life Technologies cat no NP0007) containing 2.5 µL XT reducing agent (Biorad cat no 161-0792). 14 µL of supernatant was mixed with 5 µL of NuPAGE LDS sample buffer and 1 µL of XT reducing agent. The samples were incubated at 95°C for 5 min then loaded onto a Novex NuPAGE Bis-Tris 4-12% Midi polyacrylamide gel (Life Technologies cat no NP0321) and electrophoresed at 200V for approximately 45 min.

Proteins were blotted onto nitrocellulose using Transblot Turbo membranes (Bio-Rad cat no 170-4159) and a Transblot Turbo blotting system (Bio-Rad). Membranes were washed in PBS containing 0.1% Tween 20 (PBS-T) for 5 min. All steps were conducted under gentle rocking at room temperature. The membranes were blocked for 1 h in PBS-T containing 5% dried non-fat

milk (blocking buffer) and then incubated for 1 h in blocking buffer containing 1 µg/mL mouse anti-histidine-tag monoclonal antibody (Serotec cat no MCA 1396). The membrane was washed 3 times for 5 min with PBS-T then incubated for 1 h in blocking buffer containing 1 µg/mL peroxidase-labelled goat anti-mouse-IgG (Serotec cat no 103005). The membrane was washed 3 times for 5 min with PBS-T and developed using TMB membrane peroxidase substrate (Insight Biotechnology cat no 50-77-18 KPL). Results from a representative experiment are shown in Table 8.

The apparent molecular weight of the principal bands was compared with the theoretical molecular weight of the fusion polypeptide construct calculated from its amino acid sequence. A sample was considered positive if a band was clearly visible in the culture supernatant with an apparent molecular weight within ± 5 kDa of the theoretical molecular weight of the fusion polypeptide, except for polypeptides containing a lipoyl acetyltransferase (E2) core where, due to its unusual structure, the apparent molecular weight by SDS-PAGE exceeds the true molecular weight by approx. 15 kDa (Stephens 1983) and so an appropriate correction was made.

Example 11 ELISA

HEK-293FT cells were transfected with vectors encoding various fusion polypeptides (**Example 5**) and culture supernatant was collected as described in **Example 6**.

100 µL of filtered supernatant was mixed with 100 µL of a 200 mM sodium bicarbonate pH9.4 (coating buffer) on Maxisorp plates (Nunc) and incubated for 1 h. All incubations were carried out at room temperature. The plate was washed three times with PBS containing 0.1% Tween 20 (PBS-T) and blocked with 50 µL/well of coating buffer containing 2% bovine serum albumin (BSA) for 1 h. The plate was washed three times with PBS-T and incubated for 1 h with PBS-T containing 0.5 µg/ml mouse anti-histidine-tag monoclonal antibody. The plate was washed three times with PBS-T and incubated for 1 h with PBS-T containing a 1:2000 dilution of peroxidase-labelled goat anti-mouse-IgG. The plate was washed three times with PBS-T and developed using TMB microwell peroxidase substrate (Insight Biotechnology cat no 53-00-01). The absorbance at 450 nm (A450) was measured using an Infinite F50 microplate reader (Tecan).

Example 12 Analysis of particles by electron microscopy

Particles containing the fusion protein [ferritin-linker-membrane-spanning domain-linker-his tag] (SEQ ID NO: 43) purified as described in **Example 7** were immobilised on coated copper grids and stained with 2% uranyl acetate for 1 min and washed three times with ultrapure water. After

drying the grids for 10 min they were analysed using a A-7650 tomographic transmission electron microscope (Hitachi). Approximately spherical particles with diameters in the range 17.3 nm to 19.3 nm could be clearly seen, consistent with the size of a ferritin molecule (11 nm) surrounded by a lipid membrane.

Example 13 Screening candidate core polypeptides by ELISA

Candidate core polypeptides were identified by literature review as described in **Example 1**. An experiment was carried out to determine which of the candidates might be capable of supporting the expression of a VLP when fused either at the N-terminus or the C-terminus to a membrane-spanning polypeptide. For this experiment a model membrane-spanning polypeptide was used which contained a single membrane-spanning domain and a HIS polypeptide tag to facilitate detection, similar to that illustrated in Figure 1.

HEK-293FT cells were transfected with vectors encoding various fusion polypeptides (**Example 5**) containing candidate core polypeptides as listed in Table 1. Culture supernatant was collected as described in **Example 6** and tested by ELISA as described in **Example 11**. The results are shown in Table 6.

Table 6 Expression of HIS-tagged fusion proteins in culture supernatant of transfected cells

Plasmid	Core polypeptide	Orientation N to C	MMLV gag	A450
none	N/A	N/A	N/A	0.043
pFCK5	<i>Aquifex aeolicus</i> lumazine synthase	TM-CORE	-	0.012
pFCK6	<i>Aquifex aeolicus</i> lumazine synthase	TM-CORE	+	0.007
pFCK7	<i>Aquifex aeolicus</i> lumazine synthase	CORE-TM	-	0.011
pFCK8	<i>Aquifex aeolicus</i> lumazine synthase	CORE-TM	+	0.005
pFCK9	<i>Escherichia coli</i> ferritin	TM-CORE	-	0.005
pFCK10	<i>Escherichia coli</i> ferritin	TM-CORE	+	0.005
pFCK11	<i>Escherichia coli</i> ferritin	CORE-TM	-	0.463
pFCK12	<i>Escherichia coli</i> ferritin	CORE-TM	+	0.026
pFCK13	Human lipoamide acetyltransferase (E2)	TM-CORE	-	0.014
pFCK14	Human lipoamide acetyltransferase (E2)	TM-CORE	+	0.007
pFCK15	Human lipoamide acetyltransferase (E2)	CORE-TM	-	0.430
pFCK16	Human lipoamide acetyltransferase (E2)	CORE-TM	+	0.514

pFCK17	<i>Aspergillus fumigatus</i> partitivirus coat protein	TM-CORE	-	0.003
pFCK18	<i>Aspergillus fumigatus</i> partitivirus coat protein	TM-CORE	+	0.004
pFCK19	<i>Aspergillus fumigatus</i> partitivirus coat protein	CORE-TM	-	0.012
pFCK20	<i>Aspergillus fumigatus</i> partitivirus coat protein	CORE-TM	+	0.006
pFCK21	Brome mosaic virus coat protein	TM-CORE	-	0.007
pFCK22	Brome mosaic virus coat protein	TM-CORE	+	0.007
pFCK23	Satellite tobacco necrosis virus coat protein	TM-CORE	-	0.008
pFCK24	Satellite tobacco necrosis virus coat protein	TM-CORE	+	0.010
pFCK25	Tomato bushy stunt virus coat protein	TM-CORE	-	0.004
pFCK26	Tomato bushy stunt virus coat protein	TM-CORE	+	0.027
pFCK27	Barley yellow dwarf virus coat protein	CORE-TM	-	0.009
pFCK28	Barley yellow dwarf virus coat protein	CORE-TM	+	0.010
pFCK29	Potato leafroll virus coat protein	TM-CORE	-	0.010
pFCK30	Potato leafroll virus coat protein	TM-CORE	+	0.013
pFCK31	Potato leafroll virus coat protein	CORE-TM	-	0.008
pFCK32	Potato leafroll virus coat protein	CORE-TM	+	0.007
pSF6	GFP (negative control)	-	-	0.012
pSF28	Dengue virus NS1 (positive control)	-	-	0.776

An absorbance (A₄₅₀) of greater than 0.05 was taken to indicate the presence of HIS-tagged protein in the culture supernatant, indicating that the cell was secreting VLPs containing the fusion protein. Plasmid pSF6 coding for intracellular GFP served as a negative control and plasmid pSF28 coding for HIS-tagged Dengue virus NS1 served as a positive control. They gave the expected results. Negative results were obtained with constructs containing core polypeptides derived from virus capsids. Positive results were obtained with constructs containing core polypeptides derived from E coli ferritin and human lipoamide acetyltransferase (E2) when the fusion polypeptide had the orientation NH₂-Core-Transmembrane-COOH, but not with the opposite orientation NH₂- Transmembrane-Core-COOH. The presence of a spacer polypeptide

between the core and the membrane-spanning domain obliterated the signal with the ferritin core but not the E2 core.

Three-dimensional models of the assembled proteins were retrieved as .pdb files from the Protein Data Bank or the VIPER database as described in **Example 2** and **Example 3** and viewed using Discovery Studio Client 3.5 Visualizer software to assess the position and accessibility of the N- and C-termini. If a structure was not available then a homologue was examined if possible. (No structures were found for luteoviridae such as barley yellow dwarf virus or potato leafroll virus.) The results are shown in Table 7.

Table 7 Exposure of terminal residues in three-dimensional structures

PDB entry	Name	Subunits	Diameter (nm)	Exposed N-terminus	Exposed C-terminus
1hqk	<i>Aquifex aeolicus</i> lumazine synthase	60	16	Met-1	Arg-154
1eum	<i>Escherichia coli</i> ferritin	24	12	Leu-2	inner surface
1b5s	<i>Bacillus stearothermophilus</i> dihydrolipoyl transacetylase (E2), catalytic domain (homologue of human E2)	60	24	Ala-184	Met-426
3iym	<i>Penicillium stoloniferum</i> partitivirus coat protein (homologue of <i>Aspergillus fumigatus</i> partitivirus coat protein)	120	38	inner surface	inner surface
1js9	Brome mosaic virus coat protein	180	28	inner surface	buried
2buk	Satellite tobacco necrosis virus coat protein	60	20	inner surface	buried
1f15	Cucumber mosaic virus coat protein	180	30	inner surface	buried
2tbv	Tomato bushy stunt virus coat protein	180	35	inner surface	Leu-387

In the light of this information the negative results obtained with core polypeptides from *A. fumigatus* partitivirus, brome mosaic virus, satellite tobacco necrosis virus and cucumber mosaic virus could have been predictable. However, the negative results with fusions to the N-terminus of *E. coli* ferritin, human E2, the C-terminus of tomato bushy stunt virus and both termini of *A.*

aeolicus lumazine synthase could not have been predicted and the positive result obtained with a fusion to the C-terminus of *E coli* ferritin was highly surprising considering that it is found on the inside of the assembled protein shell.

Example 14 Comparison of different polypeptide tags

HEK-293FT cells were transfected with vectors encoding fusion polypeptides (**Example 5**) containing either ferritin or lipoamide acetyltransferase (E2) core polypeptide, membrane-spanning domains and either polyhistidine (HIS) or green fluorescent protein (GFP) polypeptide tag fused to the C-terminus of the core proteins. Culture supernatant was collected as described in **Example 6** and tested by SDS-gel electrophoresis and Western blotting as described in **Example 10**. The results are shown in Table 8.

Table 8 Visualisation of fusion proteins by SDS-PAGE and Western blotting

Plasmid	Sample type	Core polypeptide	Tag	Predicted molecular weight	Band seen?
pFCK11	lysate	ferritin	HIS	26 kDa	25 kDa
pFCK16	lysate	E2	HIS	73 kDa	80 kDa
pFCK11	supernatant	ferritin	HIS	26 kDa	25 kDa
pFCK16	supernatant	E2	HIS	73 kDa	80 kDa
pFCK35	lysate	ferritin	GFP	51 kDa	48 kDa
pFCK36	lysate	E2	GFP	99 kDa	115 kDa
pFCK35	supernatant	ferritin	GFP	51 kDa	48 kDa
pFCK36	supernatant	E2	GFP	99 kDa	115 kDa

Specific protein bands of the predicted molecular weights were detected in the supernatants of all transfectants, indicating that either HIS or GFP may be used as a polypeptide tag with either *E. coli* ferritin or human lipoamide acetyltransferase (E2) may be used as a core polypeptide.

Example 15 Screening candidate core polypeptides by Western blotting

A second set of candidate core polypeptides were identified by literature review as described in **Example 1**. They were screened as described in **Example 13** except that the supernatant samples were tested by SDS-PAGE and Western blotting instead of by ELISA.

HEK-293FT cells were transfected with vectors encoding various fusion polypeptides (**Example 5**) containing candidate core polypeptides as listed in Table 2. All constructs included a HIS tag. Culture supernatant was collected as described in **Example 6** and tested by SDS-gel electrophoresis and Western blotting as described in **Example 10**. The results for the supernatant samples are shown in Table 9.

Table 9 Visualisation of fusion proteins by SDS-PAGE and Western blotting

Plasmid	Core polypeptide	Orientation N to C	MMLV gag	Predicted molecular weight	Band seen?
pFCK139	human parvovirus B19 VP2	core-TM	-	67 kDa	(65 kDa) ¹
pFCK140	human parvovirus B19 VP2	core-TM	+	74 kDa	(74 kDa)
pFCK141	human parvovirus B19 VP2	TM-core	-	67 kDa	None
pFCK142	human parvovirus B19 VP2	TM-core	+	74 kDa	None
pFCK144	hepatitis E virus ORF2	core-TM	+	67 kDa	70 kDa
pFCK145	hepatitis E virus ORF2	TM-core	-	60 kDa	None
pFCK146	hepatitis E virus ORF2	TM-core	+	67 kDa	None
pFCK147	Norwalk virus VP1	core-TM	-	70 kDa	65 kDa
pFCK148	Norwalk virus VP1	core-TM	+	77 kDa	None
pFCK149	Norwalk virus VP1	TM-core	-	70 kDa	65 kDa
pFCK150	Norwalk virus VP1	TM-core	+	77 kDa	None
pFCK151	Rabbit haemorrhagic disease virus VP60	core-TM	-	67 kDa	70 kDa
pFCK152	Rabbit haemorrhagic disease virus VP60	core-TM	+	74 kDa	None
pFCK153	Rabbit haemorrhagic disease virus VP60	TM-core	-	67 kDa	70 kDa
pFCK155	murine polyoma virus VP1	core-TM	-	49 kDa	None
pFCK156	murine polyoma virus VP1	core-TM	+	56 kDa	(60 kDa)
pFCK157	murine polyoma virus VP1	TM-core	-	49 kDa	None
pFCK158	murine polyoma virus VP1	TM-core	+	56 kDa	None
pFCK160	human papilloma virus 11 L1	core-TM	+	69 kDa	None
pFCK161	human papilloma virus 11 L1	TM-core	-	62 kDa	None
pFCK162	human papilloma virus 11 L1	TM-core	+	69 kDa	None

¹ Results in brackets indicate that only a weak band was seen.

Very weak or negative results were obtained for core polypeptides derived from human parvovirus, murine polyoma virus and human papilloma virus. Positive results were obtained for hepatitis E virus in the orientation NH₂-Core-Spacer-Transmembrane-COOH and for Norwalk virus and rabbit haemorrhagic disease virus in either orientation provided that the MMLV gag spacer polypeptide was omitted. This confirmed the observation in **Example 13** that some core polypeptides (eg E2, hepatitis E) could tolerate the presence of a MMLV gag spacer but others (eg ferritin, Norwalk virus, rabbit haemorrhagic disease virus) could not.

Three-dimensional models of the assembled proteins were retrieved as .pdb files from the Protein Data Bank or the VIPER database as described in **Example 2** and **Example 3** and viewed using Discovery Studio Client 3.5 Visualizer software to assess the position and accessibility of the N- and C-termini. The results are shown in Table 10 .

Table 10 Exposure of terminal residues in three-dimensional structures

PDB entry	Name	Subunits	Diameter (nm)	Exposed N-terminus	Exposed C-terminus
1s58	human parvovirus B19 VP2	60	27	inner surface	buried
2ztn	hepatitis E virus ORF2	60	28	inner surface	Ala-606
3iyo	hepatitis E virus ORF2	180	41	inner surface	Ala-606
1ihm	Norwalk virus VP1	180	40	Asp-29 buried	Ser-520
3jip	Rabbit haemorrhagic disease virus VP60	180	55	Asn-45 inner surface	Leu-569
1sid	murine polyoma virus VP1	360	31	inner surface	Gly-382 ¹
1dzt	human papilloma virus 11 L1	60	32	Lys-20 ¹	inner surface
1l0t	human papilloma virus 11 L1	360	60	Met-1 ¹	inner surface

¹ Residues appear to be partly exposed on the edge of a hole in the capsid

In the light of this information the weak or negative results obtained with core polypeptides from human parvovirus, murine polyoma virus and human papilloma virus could have been predictable. Likewise, the positive results with fusions to the C-terminus of core polypeptides from hepatitis E, Norwalk virus and rabbit haemorrhagic disease virus were not unexpected.

However, the positive results obtained with a fusion to the N-terminus of core polypeptides from Norwalk virus and rabbit haemorrhagic disease virus could not have been predicted considering that the termini visible in the crystal structure were buried or on the inner surface of the capsid. However, it was noted that in both cases the true N-terminus of the protein was not visualized in the three-dimensional structure.

Example 16 Comparison of different linker polypeptides

HEK-293FT cells were transfected with vectors encoding various fusion polypeptides (**Example 5**) containing ferritin or lipoamide acetyltransferase (E2) core polypeptide, membrane-spanning domains and a HIS tag. A variety of alternative linkers were inserted between the core polypeptide and the membrane-spanning domain including: (a) 4, 5 or 6 copies of alternating fragments of the MMLV gag late domain (SEQ ID NO: 4549 and 50), (b) 10, 20, 30, 40 or 50 repeats of the EAAAK peptide (SEQ ID NO: 46), (c) 4 repeats of the E2 lipoyl domain (SEQ ID NO: 45). Culture supernatant was collected as described in **Example 6** and tested by SDS-gel electrophoresis and Western blotting as described in **Example 10**. The results for the supernatant samples are shown in Table 11.

Table 11 Visualisation of fusion proteins by SDS-PAGE and Western blotting

Plasmid	Core polypeptide	Linker	Copies	Predicted molecular weight	Band seen?
pFCK34	human lipoamide acetyltransferase (E2)	Gly-Ser	1	73 kDa	78 kDa
pFCK82		MMLV gag fragments	4	85 kDa	85 kDa
pFCK84		MMLV gag fragments	5	90 kDa	90 kDa
pFCK86		MMLV gag fragments	6	95 kDa	95 kDa
pFCK210		EAAAK	20	76 kDa	80 kDa
pFCK202		EAAAK	50	90 kDa	90 kDa
pFCK203		lipoyl domain	4	92 kDa	95 kDa
pFCK207	<i>E. coli</i> ferritin	EAAAK	10	31 kDa	30 kDa
pFCK206		EAAAK	20	35 kDa	35 kDa
pFCK205		EAAAK	30	40 kDa	40 kDa

pFCK204		EAAAK	40	44 kDa	45 kDa
pFCK200		EAAAK	50	49 kDa	55 kDa
pFCK201		lipoyl domain	4	51 kDa	70 kDa

Specific protein bands of the predicted molecular weights were detected in the supernatants of all transfectants, indicating that any of the tested combinations of core polypeptide and linker polypeptide may be used.

Example 17 Expression of Type IV-A membrane polypeptide with polypeptide tags

HEK-293FT cells were transfected with vectors encoding various fusion polypeptides (**Example 5**) containing lipoamide acetyltransferase (E2) core polypeptide linked at the C-terminus by linker SEQ ID NO: 35 to a modified MMLV gag sequence (SEQ ID NO: 47) linked by linker SEQ ID NO: 35 to a transmembrane polypeptide which consisted of (a) a signal anchor (SA) membrane-spanning domain (SEQ ID NO: 38), (b) a polypeptide tag, either FLAG (SEQ ID NO: 12) or V5 (SEQ ID NO: 20) linked by linker SEQ ID NO: 34 to (c) a stop transfer anchor (STA) membrane-spanning domain (SEQ ID NO: 48). A negative control vector pFCK16 contained the E2 core polypeptide with a single pass HIS-tagged transmembrane polypeptide as described in **Example 14**, Table 8. This provided a system to test the expression of a transmembrane polypeptide which spanned the membrane twice, displaying a peptide epitope on the outside and terminating inside the membrane as illustrated in Figure 2. Culture supernatant was collected as described in **Example 6** and tested by ELISA as described in **Example 11** and by SDS-gel electrophoresis and Western blotting as described in **Example 10**. In each case the anti-HIS tag antibody was substituted by an appropriate mouse anti-FLAG or mouse anti-V5 antibody. The results for the supernatant samples are shown in Table 12.

Table 12 Visualisation of fusion proteins by ELISA, SDS-PAGE and Western blotting

Plasmid	Polypeptide tag	Anti-FLAG A450	Anti-V5 A450	Predicted molecular weight	Band seen?
pFCK34	negative control	0.010	0.008	73 kDa	70 kDa
pFCK136	FLAG	0.548	0.009	77 kDa	75 kDa
pFCK137	V5	0.004	0.706	78 kDa	75 kDa

Bands of the expected size were seen for all three constructs and positive results were obtained by ELISA for the FLAG peptide or the V5 peptide when present in the construct. This indicated

that VLPs containing a Type IV-A membrane polypeptide could be expressed and that a variety of polypeptide tags could be incorporated in the fusion protein.

Example 18 Expression of a larger Type IV-A membrane polypeptide including three membrane-spanning domains and expression of VLPs containing a mixture of different of different fusion proteins

It can be appreciated that there is likely to be an upper limit to the number of membrane-spanning domains which can be included within the membrane-spanning polypeptide. This limit is likely to depend on the size of the core polypeptide and the diameter of the assembled particle which it can form. Too many membrane-spanning domains would create steric hindrance to the successful assembly of the VLP. An experiment was carried out to test whether a polypeptide containing two membrane spanning domains could be supported by the comparatively small *E. coli* ferritin core polypeptide and if not, to test the possibility that a mixture of fusion proteins containing either full-size or truncated membrane-spanning polypeptides might allow expression of a VLP similar to that illustrated in Figure 3.

HEK-293FT cells were co-transfected with different ratios of vectors pFCK73 and pFCK11. Vector pFCK73 encoded a fusion polypeptide containing the *E. coli* ferritin core polypeptide, followed by a fragment of human SCN9A protein containing two membrane spanning domains (SEQ ID NO: 51) followed by a six amino acid HIS tag (SEQ ID NO: 26). Vector pFCK11 encoded a fusion polypeptide containing the same *E. coli* ferritin core polypeptide, a single membrane spanning domain (SEQ ID NO: 38) or followed by a six amino acid HIS tag (SEQ ID NO: 26). Culture supernatant was collected as described in **Example 6** and tested by SDS-gel electrophoresis and Western blotting as described in **Example 10**. Relative intensities of the stained bands corresponding to the expected molecular weight of the fusion protein were estimated visually. The results are shown in Table 13.

Table 13 Visualisation of fusion proteins by Western blotting following transfection with mixtures of plasmids

	Transfection 1		Transfection 2		Transfection 3		Transfection 4	
Plasmid vector	pFCK73	pFCK11	pFCK73	pFCK11	pFCK73	pFCK11	pFCK73	pFCK11
Proportion in the vector mix	100%	0%	50%	50%	10%	90%	5%	95%
Visible fusion protein in the cell pellet	Very high	none	high	high	medium	high	weak	high

Visible fusion protein in the supernatant	none	none	medium	medium	weak	high	none	high
--	------	------	--------	--------	------	------	------	------

Transfection with a single plasmid pFCK73 encoding a fusion protein with two transmembrane domains resulted in transcription of the protein (as indicated by the strong band in the cell pellet sample) but there was no evidence for formation and secretion of particles (as indicated by lack of a band in the supernatant sample). However, co-transfection with mixtures of the plasmids pFCK73 and pFCK11 in the proportions 50:50 or 10:90 resulted in visible bands corresponding to both fusion proteins in the supernatant samples, indicated that VLPs were formed and secreted. Thus although a small particle based on a ferritin core polypeptide may not be able to accommodate 24 copies of a membrane polypeptide with two membrane-spanning domains, it can support approximately 12 copies of such a membrane polypeptide.

REFERENCES

Alfthan K et al. (1995) Properties of a single-chain antibody containing different linker peptides. *Protein Engineering* vol.8 pp.725-731.

Almen MS et al. (2009) Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin. *BMC Biology* 2009, vol.7 pp.50 doi:10.1186/1741-7007-7-50.

Arai R et al (2001) Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein Engineering* vol.14 pp.529-532.

Argos P. (1990) An Investigation of Oligopeptides Linking Domains in Protein Tertiary Structures and Possible Candidates for General Gene Fusion. *J. Mol. Biol.* vol.211 pp.943-958.

Barbas CF et al. (2001) *Phage Display: A Laboratory Manual*. pub. Cold Spring Harbor Laboratory Press.

Bayer ME et al. (1968) Particles associated with Australia antigen in the sera of patients with leukaemia, Down's Syndrome and hepatitis. *Nature* vol.218 pp.1057–1059.

Caldeira et al. (2010) Immunogenic display of diverse peptides, including a broadly cross-type neutralizing human papillomavirus L2 epitope, on virus-like particles of the RNA bacteriophage PP7. *Vaccine* vol. 28 pp. 4384–4393.

Carillo-Tripp M et al. (2008) VIPERdb²: an enhanced and web API enabled relational database for structural virology. *Nucleic Acids Research* vol.37 pp.D436-D442

Chen X et al. (2013) Fusion protein linkers: Property, design and functionality. *Advanced Drug Delivery Reviews* vol.65 pp.1357–1369

Crasto CJ and Feng J (2000) LINKER: a program to generate linker sequences for fusion proteins. *Protein Engineering* vol.13 pp.309-312.

Debbink K et al. (2013) Human norovirus detection and production, quantification, and storage of virus-like particles. *Curr. Protoc. Microbiol.* vol. 31 pp. 15K.1.1-15K.1.45.

Deml L et al. (2005) Recombinant HIV-1 Pr55gag virus-like particles: potent stimulators of innate and acquired immune responses. *Molecular Immunology* vol.42 pp.259–277

Fifis T et al (2004) Size-Dependent Immunogenicity: Therapeutic and Protective Properties of Nano-Vaccines against Tumours. *J. Immunol.* vol.173 pp.3148-3154.

George RA and Heringa J. (2003) An analysis of protein domain linkers: their classification and role in protein folding. *Protein Engineering* vol15 pp.871–879.

Grgacic EVL et al. (2006) Virus-like particles: Passport to immune recognition. *Methods* vol.40 pp.60–65.

Hanly WC et al. (1995) Review of Polyclonal Antibody Production. *Procedures in Mammals and Poultry. ILAR Journal* vol.37 pp.93-118.

Howard GC and Bethel DR (eds) (2001) Basic methods in antibody production and characterization. pub. CRC Press LLC.

Kao CC et al. (2011) The coat protein leads the way: an update on basic and applied studies with the Brome mosaic virus coat protein. *Mol. Plant. Pathol.* vol 12 pp 403-12.

Khoury GA et al. (2011) Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Scientific Reports* Vol.1, Article number: 90 pp.1-5. doi:10.1038/srep00090

Kim SE et al (2011) Fluorescent ferritin nanoparticles and application to the aptamer sensor. *Anal. Chem.* vol. 83 pp. 5834-43.

Krogh A et al. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* vol.305 pp.567-580.

Kumar S, et al (2009) Tomato bushy stunt virus (TBSV), a versatile platform for polyvalent display of antigenic epitopes and vaccine design. *Virology* vol. 388 pp 185-90.

Ladenstein R et al. (2013) The lumazine synthase/riboflavin synthase complex: shapes and functions of a highly variable enzyme system. *FEBS J.* vol. 280 pp. 2537-63.

Lamb JW et al. (1996). Assembly of virus-like particles in insect cells infected with a baculovirus containing a modified coat protein gene of potato leafroll luteovirus. J. Gen. Virol. vol. 77 pp. 1349-58.

Lane SW et al. (2011) Construction and crystal structure of recombinant STNV capsids. J. Mol. Biol. vol. 413 pp. 41-50.

Li TC et al. (2005) Essential elements of the capsid protein for self-assembly into empty virus-like particles of hepatitis E virus. J Virol. vol. 79 pp. 12999-3006.

Lodish et al. (2007) Molecular Cell Biology, pub MacMillan 6th edition, p 546.

Ludwig C and Wagner R (2007) Virus-like particles - universal molecular toolboxes. Current Opinion in Biotechnology vol.18 pp.537-545.

Luque D et al. (2012). Epitope insertion at the N-terminal molecular switch of the rabbit haemorrhagic disease virus T = 3 capsid protein leads to larger T = 4 capsids. J. Virol. vol. 86 pp. 6470-80.

Moon KB et al. (2014) Overexpression and self-assembly of virus-like particles in *Nicotiana benthamiana* by a single-vector DNA replicon system. Appl. Microbiol. Biotechnol. vol. 98 pp. 8281-90.

Mossadegh N et al. (2004) Codon optimization of the human papillomavirus 11 (HPV 11) L1 gene leads to increased gene expression and formation of virus-like particles in mammalian epithelial cells. Virology. vol. 326 pp. 57-66.

Natarajan P et al. (2005) Exploring icosahedral virus structures with VIPER. Nature Reviews Microbiology vol.3 pp.809-817.

Neirynck S et al. (1999) Universal influenza A vaccine based on the extracellular domain of the M2 protein. Nature Medicine vol.5 pp.1157-1163.

Noad R and Roy P (2003) Virus-like particles as immunogens. Trends in Microbiology vol.11 pp.438-444.

Ochoa WF et al. (2008) Partitivirus structure reveals a 120-subunit, helix-rich capsid with distinctive surface arches formed by quasisymmetric coat-protein dimers. *Structure*. vol. 16 pp 776-86.

Ossipow V and Fischer N (eds.) (2014) *Monoclonal Antibodies Methods and Protocols*. Series: *Methods in Molecular Biology* vol.1131 . 2nd ed. pub. Humana Press

Overington JP et al. (2006). How many drug targets are there? *Nat Rev Drug Discov* vol.5 pp.993–6.

Perham RN (1991) Domains, motifs and linkers in 2-oxo acid dehydrogenase multienzyme complexes: a paradigm in the design of a multifunctional protein. *Biochemistry* vol.30 pp.8501-8512.

Raman S et al. (2009) Design of Peptide Nanoparticles Using Simple Protein Oligomerization Domains. *The Open Nanomedicine Journal* vol. 2 pp15-26.

Robinson CR and Sauer RT (1998) Optimizing the stability of single-chain proteins by linker length and composition mutagenesis. *Proc. Natl. Acad. Sci. USA* vol.95 pp. 529–5934.

Rueda P et al. (1999) Minor displacements in the insertion site provoke major differences in the induction of antibody responses by chimeric parvovirus-like particles. *Virology*. vol. 263 pp 89-99.

Simon C et al. (2014) Disulfide linkage and structure of highly stable yeast-derived virus-like particles of murine polyomavirus. *J. Biol. Chem.* vol. 289 pp. 10411-8.

Shepherd P and Dean C. (2000) *Monoclonal Antibodies: A practical approach*. pub Oxford University Press.

Smith SM. (2011) Strategies for the purification of membrane proteins. *Methods Mol Biol*. vol.681 pp.485-96. doi: 10.1007/978-1-60761-913-0_29.

Stephens PE et al. (1983) The Pyruvate Dehydrogenase Complex of *Escherichia coli* K12 Nucleotide Sequence Encoding the Dihydrolipoamide Acetyltransferase Component. *Eur. J. Biochem.* vol 133 pp 481-489.

- Tars K et al. (2007) The crystal structure of bacteriophage GA and a comparison of bacteriophages belonging to the major groups of Escherichia coli leviviruses. *J. Mol. Biol.* vol 271 pp 759-773.
- Terpe K (2003) Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol* vol.60 pp.523–533.
- Tissot AC et al. (2010) Versatile virus-like particle carrier for epitope based vaccines. *Plos One* vol 5 article e9809. doi: 10.1371/journal.pone.0009809.
- Vicente T et al. (2011) Large-scale production and purification of VLP-based vaccines. *Journal of Invertebrate Pathology* vol.107 pp.S42–S48.
- Winter G (1994) Making Antibodies by Phage Display Technology. *Annual Review of Immunology* vol.12 pp.433-455.
- Wriggers W et al. (2005) Control of Protein Functional Dynamics by Linkers. *Peptide Science* vol.80 pp.736-746.
- Yang Y & Burkhard P (2012) Encapsulation of gold nanoparticles into self-assembling protein nanoparticles. *Journal of Nanobiotechnology* vol. 10 pp. 42-53.
- Zeltins A (2013) Construction and Characterization of Virus-Like Particles: A Review. *Mol Biotechnol* vol.53 pp.92–107.
- Zhou ZH et al (2001) The remarkable structural and functional organization of the eukaryotic pyruvate dehydrogenase complexes. *Proc. Natl Acad. Sci. U S A.* vol. 98 pp. 14802-7.

CLAIMS

We claim:

1. A virus-like particle comprising a membrane and a fusion polypeptide, wherein the membrane is a lipid bilayer and the fusion polypeptide comprises:
 - (a) a core polypeptide enveloped by the membrane; and
 - (b) a membrane-spanning polypeptide which crosses said membrane at least once.
2. The particle according to claim 1 wherein the core polypeptide is selected from a group consisting of: ferritin, encapsulin, dihydrolipoylsuccinyltransferase (EC 2.3.1.61), dihydrolipoamide branched chain acetyltransferase (EC 2.3.1.168), lipoyl acetyltransferase (EC 2.3.1.12) or is a polypeptide fragment or a polypeptide homologue of a member of the said group.
3. The particle according to claim 1 wherein: (a) the core polypeptide is a capsid protein or is a polypeptide fragment or a polypeptide homologue of a capsid protein and (b) the N-terminus or the C-terminus of the core polypeptide is accessible on the outside of the assembled capsid.
4. The particle according to claim 1 wherein the core polypeptide is selected from a group consisting of: Newtons T=3 capsid protein, Bacteriophage GA capsid protein, Bacteriophage PP7 capsid protein, Bacteriophage AP205 capsid protein, Ryegrass mottle virus capsid protein, Turnip yellow mosaic virus capsid protein, Melon necrotic spot virus capsid protein, Pyrococcus furiosus virus-like protein, Norwalk virus capsid protein, Hepatitis E virus capsid protein, Rabbit haemorrhagic disease virus capsid protein or is a polypeptide fragment or a polypeptide homologue of a member of the said group,
5. The particle according to any of claims 1 to 4 wherein the membrane-spanning polypeptide is selected from a group consisting of: Type I membrane proteins, Type II membrane proteins, Type III membrane proteins, Type IV-A membrane proteins, Type IV-B membrane proteins or is a polypeptide fragment or a polypeptide homologue of a member of the said group.
6. The particle according to any of claims 1 to 5 wherein the fusion polypeptide also comprises one or more linker polypeptides.

7. The particle according to any of claims 1 to 6 wherein the fusion polypeptide also comprises a polypeptide tag.
8. A polynucleotide coding for the fusion polypeptide according to any of the claims 1 to 7.
9. A vector comprising the polynucleotide according to claim 8.
10. A host cell comprising the polynucleotide according to claim 8, wherein said host cell is capable of expressing the particle according to any of claims 1 to 7.
11. A method of preparing the particle according to any of claims 1 to 7 comprising: culturing the host cell of claim 10 with a culture medium; and purifying the particle from the culture medium or from the host cell.
12. A pharmaceutical composition comprising the particle according to any of claims 1 to 7.
13. A vaccine composition comprising the particle according to any of claims 1 to 7.
14. A DNA vaccine composition comprising the polynucleotide according to claim 8.
15. A method of producing an antibody, comprising contacting the particle according to any of claims 1 to 7 and/or the polynucleotide according to claim 8 to a bird or mammal.
16. A method of selecting an antibody, comprising screening an antibody library for binding to the particle according to any of claims 1 to 7.
17. A pharmaceutical composition comprising the antibody produced according to claim 15 and/or selected according to claim 16.
18. A medicament for use in therapy comprising any composition according to claim 12, 13, 14 or 17.



Application No: GB1503218.8

Examiner: Dr Andrew Guy

Claims searched: 1-16 and 18

Date of search: 15 December 2015

Patents Act 1977: Search Report under Section 17

Documents considered to be relevant:

Category	Relevant to claims	Identity of document and passage or figure of particular relevance
X	1, 3-4, 6-16 and 18	WO 2013/122262 A1 (VLP THERAPEUTICS) See in particular example 1, page 49
X	1, 3, 5-16 and 18	WO 2005/042695 A2 (INTEGRAL MOLECULAR INC) See in particular example 3, pages 153-154
A	-	US 5175099 A1 (WILLS) See in particular column 11 line 43 - column 12 line 30 and example 9
A	-	US 2002/0183247 A1 (DOMS ET AL) See in particular paragraphs 0125-0126
A	-	WO 2010/062757 A1 (LIGOCYTE PHARMACEUTICALS) See in particular paragraph 0085
A	-	US 2012/0052082 A1 (COMPANS ET AL) See abstract
A	-	WO 2011/116226 A2 (STC.UNM) See abstract and figure 12

Categories:

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.

Field of Search:

Search of GB, EP, WO & US patent documents classified in the following areas of the UKC^X :

Worldwide search of patent documents classified in the following areas of the IPC

C12N

The following online and other databases have been used in the preparation of this search report



WPI, EPODOC, BIOSIS, MEDLINE

International Classification:

Subclass	Subgroup	Valid From
C12N	0007/04	01/01/2006
C07K	0014/005	01/01/2006
C12N	0007/00	01/01/2006