Innovation, Sciences et Développement économique Canada
Office de la Propriété Intellectuelle du Canada

Innovation, Science and Economic Development Canada
Canadian Intellectual Property Office

CA 2916827 C 2021/11/02

(11)(21) 2 916 827

(12) BREVET CANADIEN
CANADIAN PATENT

(13) C

(86) Date de dépôt PCT/PCT Filing Date: 2014/07/08
(87) Date publication PCT/PCT Publication Date: 2015/01/15
(45) Date de délivrance/Issue Date: 2021/11/02
(85) Entrée phase nationale/National Entry: 2015/12/23
(86) N° demande PCT/PCT Application No.: EP 2014/064652
(87) N° publication PCT/PCT Publication No.: 2015/004158
(30) Priorité/Priority: 2013/07/08 (EP13175637.1)

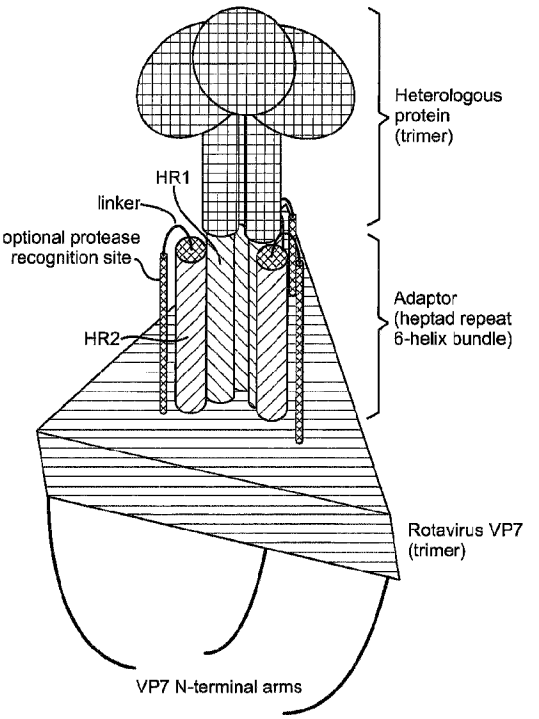
(51) Cl.Int./Int.Cl. *C07K 14/14* (2006.01)

(72) Inventeurs/Inventors:
DORMITZER, PHILIP R., US;
GRIGORIEFF, NIKOLAUS, US;
HARRISON, STEPHEN, US;
PAN, JUNHUA, US;
SETTEMBRE, ETHAN, US

(73) Propriétaires/Owners:
NOVARTIS AG, CH;
CHILDREN'S MEDICAL CENTER CORPORATION, US;
BRANDEIS UNIVERSITY, US

(74) Agent: BORDEN LADNER GERVAIS LLP

(54) Titre : PARTICULES DE ROTAVIRUS A PROTEINES DE SURFACE CHIMERES
(54) Title: ROTAVIRUS PARTICLES WITH CHIMERIC SURFACE PROTEINS



(57) Abrégé/Abstract:
The present invention relates to the use of rotavirus particles for displaying a heterologous protein, alone or in complex with another molecule. The invention further relates to methods that employ these modified rotavirus particles to rapidly determine the structure of the heterologous protein or the complex using cryo-electron microscopy (cryo-EM).The invention also relates to a method of immunising a patient, wherein said method comprises administering to the patient the modified rotavirus particles of the invention.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
15 January 2015 (15.01.2015)

WIPO | PCT

(10) International Publication Number
WO 2015/004158 A1

(51) International Patent Classification:
C07K 14/14 (2006.01)

(21) International Application Number:
PCT/EP2014/064652

(22) International Filing Date:
8 July 2014 (08.07.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
13175637.1 8 July 2013 (08.07.2013) EP

(71) Applicants: **NOVARTIS AG** [CH/CH]; Lichtstrasse 35, CH-4056 Basel (CH). **CHILDREN'S MEDICAL CENTER CORPORATION** [US/US]; 55 Shattuck Street, Boston, Massachusetts 02115 (US). **BRANDEIS UNIVERSITY** [US/US]; 415 South Street, Waltham, Massachusetts 02454 (US).

(72) Inventors: **DORMITZER, Philip R.**; c/o Novartis Vaccines, 350 Massachusetts Avenue, Cambridge, Massachusetts 02139 (US). **GRIGORIEFF, Nikolaus**; 19070 Coton Farm Court, Leesburg, Virginia 20176 (US). **HARRISON, Stephen**; 19R Sparhawk Street, Boston, Massachusetts 02135 (US). **PAN, Junhua**; 32 Prince Street, Brookline, Massachusetts 02445 (US). **SETTEMBRE, Ethan**; c/o Novartis Vaccines, 350 Massachusetts Avenue, Cambridge, Massachusetts 02139 (US).

(74) Agents: **MARSHALL, Cameron John** et al.; Carpmaels & Ransford LLP, One Southampton Row, London WC1B 5HA (GB).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,

[Continued on next page]

(54) Title: ROTAVIRUS PARTICLES WITH CHIMERIC SURFACE PROTEINS

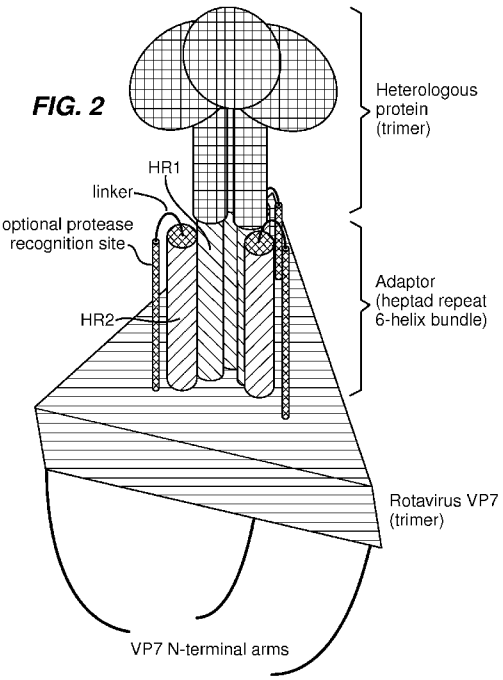


FIG. 2

(57) Abstract: The present invention relates to the use of rotavirus particles for displaying a heterologous protein, alone or in complex with another molecule. The invention further relates to methods that employ these modified rotavirus particles to rapidly determine the structure of the heterologous protein or the complex using cryo-electron microscopy (cryo-EM).The invention also relates to a method of immunising a patient, wherein said method comprises administering to the patient the modified rotavirus particles of the invention.

WO 2015/004158 A1

WO 2015/004158 A1

TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

— *with sequence listing part of description (Rule 5.2(a))*

Published:

— *with international search report (Art. 21(3))*

ROTAVIRUS PARTICLES WITH CHIMERIC SURFACE PROTEINS

SEQUENCE LISTING

An electronic copy of a sequence listing is filed herewith which is incorporated herein by reference and forms a part of the application as filed.

GOVERNMENT SUPPORT

This invention was made with U.S. government support under Grant Nos. P01-GM062580 and AI-89618 awarded by the National Institutes of Health. The U.S. government has certain rights in the invention.

FIELD OF THE INVENTION

The present invention relates to the use of rotavirus particles for displaying a heterologous protein, alone or in complex with another molecule. The invention further relates to methods that employ these modified rotavirus particles to rapidly determine the structure of the heterologous protein or the complex using cryo-electron microscopy (cryo-EM). The invention also relates to a method of immunising a patient, wherein said method comprises administering to the patient the modified rotavirus particles of the invention.

BACKGROUND

The family of *Reoviridae* comprises a structurally related group of viruses. Members of this family of viruses can cause infections of the gastrointestinal system and the respiratory tracts in mammals and birds. Some viruses of this family can also infect plants.

The *Reoviridae* are non-enveloped, double-stranded RNA viruses made up of an icosahedral capsid that is typically composed of an outer protein layer and one or more inner protein layers. The genus *Rotavirus* forms part of the *Reoviridae* family. Rotaviruses form triple-layered virus particles. The outer layer of a rotavirus particle consists of a shell protein, VP7, and a spike protein, VP4; the middle layer consists of VP6; and the inner layer is formed by VP2. The triple-layered virus particle is infectious, after activation by trypsin. During rotavirus cell entry, the outer layer of the virus particles is removed by a mechanism dubbed "uncoating," and the double-layered particle enters the cytoplasm. During cell entry, rotavirus penetrates a cellular membrane, possibly an endosomal membrane. The double-layered rotavirus particle is transcriptionally active. In the cytoplasm, it transcribes the genome, extruding mRNA through pores in its surface. The new transcripts are used to make rotavirus proteins and new double-stranded RNA genome segments, which are packaged into newly formed viral particles.

For some *Reoviridae*, the uncoating process can be simulated *in vitro* using purified viral particles. In the case of rotaviruses, calcium chelation or heat shock causes uncoating of viral particles *in vitro*. Once the outer protein layer has been removed, the resulting rotavirus double-layered particles can be "recoated" with recombinantly expressed versions of the proteins that form the outer layer. By recoating with recombinant VP4 and VP7, infectious recoated rotavirus particles that closely resemble mature virions

Date Reçue/Date Received 2020-10-01

can be formed. By recoating with VP7 alone, triple-layered, spike-less particles, which are minimally infectious, can be formed. By genetically modifying the coding regions of individual proteins of the outer layer and expressing them recombinantly, the properties of the proteins making up the outer layer can be studied in recoating experiments without requiring the production of fully recombinant viruses. To describe these types of experiments, the term "recoating genetics" has been coined.

In the case of reoviruses, the metastable state of the infectious subviral particle has allowed the use of cryo-electron microscopy (cryo-EM) in elucidating the structural basis for the priming mechanism that is used by reoviruses to enter their host cells. Similarly, *in vitro* recoated rotavirus particles have been used to study the molecular interactions in rotavirus assembly and uncoating using cryo-EM. The present invention builds on the technological advances that were brought about by these recoating and cryo-EM studies.

SUMMARY OF THE INVENTION

The combination of high-resolution cryo-EM in combination with recoating genetics has made it possible to advance the use of the rotavirus model beyond the study of rotaviruses themselves. The inventors have found that rotavirus particles can be used for rapid structure determination of heterologous trimeric proteins that are displayed on the surface of modified rotavirus particles. Rotavirus particles modified to display heterologous trimeric proteins on their surface are also useful in producing a wide range of novel vaccines.

The present invention provides rotavirus particles comprising a chimeric surface protein composed of a rotavirus surface protein linked to a heterologous protein. In one aspect of the invention, the rotavirus surface protein is linked to the heterologous protein via a linker sequence. In another aspect of the invention, the rotavirus surface protein is linked to the heterologous protein via an adapter system. In a preferred aspect of the invention, the heterologous protein is non-covalently bound to the rotavirus surface protein by a two-part adapter system, where one part of the adapter system is linked to the rotavirus surface protein and the other part of the adapter system is linked to the heterologous protein, whereby both parts of the adapter system form a stable complex thus non-covalently attaching the heterologous protein to the rotavirus surface protein. The chimeric surface protein can become part of the outer layer of the rotavirus particle by *in vitro* recoating double-layered rotavirus particles (DLPs) with it.

The invention further provides a first nucleic acid comprising an open reading frame encoding a modified rotavirus surface protein comprising a rotavirus surface protein, a first adapter polypeptide and, optionally, a linker sequence. The invention also provides a second nucleic acid comprising an open reading frame encoding a fusion protein comprising a heterologous protein, a second adapter polypeptide and, optionally, a linker sequence. The first adapter polypeptide and the second adapter polypeptide are able to form a stable complex. Typically, the open reading frame in the nucleic acids of the invention is operationally linked to a promoter sequence. In some embodiments, the adapter polypeptide comprises a heptad repeat sequence. A cell containing a nucleic acid of the invention is also provided. The nucleic acids and cells containing the same may be used to prepare the chimeric surface proteins of the invention.

The invention further relates to a kit comprising the first and second nucleic acid sequences of the invention. In some embodiments, a kit according to the invention may comprise a first nucleic acid encoding a modified rotavirus surface protein comprising a rotavirus surface protein and a first adapter

polypeptide, and a second nucleic acid comprising a nucleotide sequence encoding a second adapter polypeptide and a multiple cloning site, whereby insertion of a coding region for a heterologous protein in the multiple cloning site yields an open reading frame encoding a fusion protein comprising the heterologous protein and the second adapter polypeptide, wherein the first adapter polypeptide and the second adapter polypeptide are able to form a stable complex. The kits of the invention may further comprise a rotavirus particle. The rotavirus particle can be derived from the same species of rotavirus from which the rotavirus surface protein was derived or from a different species. For example, rhesus rotavirus VP7 can be used to recoat DLPs prepared from bovine rotavirus and *vice versa*.

Rotavirus particles comprising a chimeric surface protein can be prepared by various methods. A preferred method of the invention involves propagating a native rotavirus particle comprising an outer layer in a cell grown in a culture medium, purifying the particles from the culture medium, removing the outer layer from the particle to obtain a rotavirus DLP, and recoating the rotavirus DLP with a one or more chimeric surface proteins to form a rotavirus particle comprising the chimeric surface protein or proteins. A native rotavirus particle is a triple-layered particle, wherein the outermost third layer forms the outer shell of the rotavirus particle.

In a particular aspect, the invention relates to a first fusion protein comprising a trimer-forming rotavirus surface protein, a heptad repeat sequence, and optionally a linker sequence. The invention further provides a second fusion protein comprising a trimer-forming heterologous protein, a heptad repeat sequence, and optionally a linker sequence. The first fusion protein and the second fusion protein are able to form a stable complex via the heptad repeat sequences comprised in each of them. In some embodiments, the invention provides a chimeric surface protein formed by the first fusion protein and the second fusion protein. The chimeric surface protein may be displayed on the surface of a rotavirus particle.

In one aspect of the invention, rotavirus particles comprising a chimeric surface protein are used for determining the structure of a heterologous protein that forms a portion of the chimeric surface protein. For example, the invention provides a method for obtaining a three-dimensional model of a chimeric surface protein, wherein the method comprises the steps of (i) recoating a rotavirus DLP with a chimeric surface protein comprising all or part of a heterologous protein to yield a suspension of rotavirus particles displaying the chimeric surface protein, (ii) freezing the suspension, (iii) imaging the rotavirus particles using cryo-EM to obtain a plurality of micrographs, and (iv) analysing the plurality of micrographs to obtain a three-dimensional model of the chimeric surface protein. In some embodiments, step (i) may be subdivided into two steps, namely (a) a recoating step in which the rotavirus DLP is recoated with a rotavirus surface protein comprising a first adapter to form rotavirus particles and (b) a binding step in which the recoated rotavirus DLP is incubated in the presence of a heterologous protein comprising a second adapter whereby a complex is formed between the first adapter and the second adapter resulting in a chimeric surface protein being displayed on the rotavirus particles. In some instances, the method can be modified to determining the structure of a heterologous protein bound to a molecule that specifically binds to the heterologous protein. The modified method comprises the steps of (i) recoating a rotavirus DLP with a chimeric surface protein comprising all or part of a heterologous protein to yield a suspension of rotavirus particles displaying the chimeric surface protein, (ii) adding to the suspension a molecule that specifically binds to the heterologous protein, wherein the molecule forms a complex with the chimeric

surface protein, (iii) freezing the suspension, (iv) imaging the rotavirus particles using cryo-EM to obtain a plurality of micrographs, and (v) analysing the plurality of micrographs to obtain a three-dimensional model of the chimeric surface protein complexed to the molecule. The molecule may a proteinaceous molecule. For example, the proteinaceous molecule may be all or part of a receptor or antibody that specifically binds the heterologous protein. Alternatively, the proteinaceous molecule may be a polypeptide or peptide that specifically binds to the heterologous protein. In other embodiments, the molecule is a non-proteinaceous molecule. For instance, the non-proteinaceous molecule may be a nucleic acid.

In a further aspect, the rotavirus particles prepared in accordance with the invention are used as a medicament. In one embodiment, the invention relates to an immunogenic composition comprising a rotavirus particle that comprises the chimeric surface protein of the invention. The invention further relates to a method of treating a patient in need thereof comprising administering to said patient the rotavirus particle of the invention, *e.g.* in form of an immunogenic composition of the invention.

The underlying concept of the invention can be extended to other (trimeric or non-trimeric) heterologous proteins and other viruses (in particular reoviruses) or assemblies (*e.g.* virus-like particles, ferritin cages, *etc.*) by adapting the methods described herein.

DETAILED DESCRIPTION OF THE INVENTION

Viral particles

Rotavirus is a genus of double-stranded RNA viruses in the family *Reoviridae* comprising five known types (Rotaviruses A-E). Rotaviruses are a major cause of childhood gastroenteritis. Rotavirus is a non-enveloped, triple-layered icosahedral particle. An infectious triple-layered particle (TLP) or virion is formed from a non-infectious double layered particle (DLP) by coating the DLP with the shell protein VP7 and the spike protein VP4. VP7 is a trimer, and 260 such trimers decorate the outside of the coated particle. The DLP, which is composed of concentric VP2 and VP6 icosahedral protein layers, is about 700 Å in diameter and encapsidates 11 double-stranded RNA genome segments, the viral polymerase (VP1), and a capping enzyme (VP3). During the infection of a cell, the VP4 and VP7 proteins, which form the outer layer of the rotavirus particle, dissociate from the DLP in a low-calcium environment (probably in the endosomal compartment) – a process referred to as "uncoating" – and deliver the DLP containing the viral RNA transcription machinery into the cytoplasm. There, the DLP synthesizes, caps and releases copies of the 11 mRNA species.

Dissociation or "uncoating" of VP4 and VP7 from a rotavirus particle can be performed *in vitro* by incubating rotavirus particles in the presence of a calcium chelator such as EDTA or EGTA or by heat shock. The resulting DLPs can be recoated *in vitro* with recombinantly expressed VP4 and VP7 to form fully infectious rotavirus particles. Particles recoated in this way are very well ordered and give high-resolution cryo-EM images and density maps. Using *in vitro* reconstituted TLPs, cryo-EM has been used to study the molecular interactions in rotavirus assembly and uncoating at a resolution comparable with that of X-ray crystallography (see references 1 and 2).

The inventors demonstrate herein that a heterologous trimeric protein (such as influenza hemagglutinin) can be attached to the trimeric VP7 protein, thus forming a chimeric surface protein that can project from

the surface of a suitably recoated rotavirus DLP, therefore making it possible to determine the structure of the heterologous protein with the same cryo-EM methods that have previously provided near-atomic resolution structures of rotavirus and its subparticles. The method makes it possible for the first time to develop a high-throughput assay for the structure determination of antigen-antibody complexes. This was previously not possible due to the constraints of X-ray crystallography, which has been the method of choice to determine the structure of antigen-antibody complexes.

In principle, any non-enveloped icosahedral virus particle comprising an inner layer and an outer layer can be used to practise the methods of the invention. For example, displaying a chimeric surface protein comprising all or part of a heterologous protein is easily possible using any icosahedral virus for which a reverse genetics system for the production of virus particles has been established. A plasmid-based reverse genetics system consisting of ten reovirus cDNA constructs has been established for the mammalian reoviruses (see reference 3). The octameric symmetry of ferritin cages also has 3-fold axes of symmetry and is suitable for the practice of this invention. Other proteins that assemble into particulate, regular structures with three-fold axes of symmetry may be suitable for the practice of this invention.

Ideally, the outer layer of the virus particle of a non-enveloped icosahedral virus can be removed or stripped, *e.g.* by protease treatment or under low-calcium conditions, to yield subviral particles that comprise the inner layer(s) only and can be recoated with recombinantly produced outer layer proteins. *In vitro* reassembly or recoating of subviral particles to complete viral particles by addition of recombinantly expressed outer layer protein(s) is particularly advantageous in cases where the presence of the chimeric surface protein would interfere with the proper assembly of the virus propagated in cell culture using a reverse genetics approach. In addition, only expression vectors for the outer layer proteins need to be constructed, removing the need for an efficient plasmid-based reverse genetics system to be in place. Native virus can simply be propagated in tissue culture cells, and viral particles can be stripped of the outer layer and reassembled *in vitro* using recombinantly expressed outer layer proteins. Thus, using a method that relies on *in vitro* recoating of subviral particles to complete viral particles eliminates the need for transfecting large numbers of plasmids and removes the additional propagation step typically needed to yield large numbers of virus particles from plasmid-based reverse genetics system, making such a method more amiable for high-throughput applications.

According to the inventors' knowledge, all known non-enveloped icosahedral viruses where the outerlayer of the virus particle can be stripped and the subviral particle can be recoated with recombinantly expressed outer layer proteins to form infectious virus particles belong to the family of *Reoviridae*. This family is subdivided into two subfamilies, *Sedoreovirinae* and *Spinareovirinae*, which comprise six and nine genera, respectively. The subfamily *Sedoreovirinae* comprises the genera *Cardoreovirus, Mimoreovirus, Orbivirus, Phytoreovirus, Rotavirus,* and *Seadornavirus*. The subfamily *Spinareovirinae* comprises the genera *Aquareovirus, Coltivirus, Cypovirus, Dinovernavirus, Fijivirus, Idnoreovirus, Mycoreovirus, Orthoreovirus*, and *Oryzavirus*.

In addition to rotaviruses, mammalian orthoreoviruses may be useful in practising the invention. For these viruses, suitable conditions for the complete *in vitro* assembly of the outer capsid and the use of cryo-EM are well established (see references 1, 2, 4 and 5). Other orthoreoviruses (*e.g.* baboon or avaian

orthoreoviruses), oryzavirus (*e.g.* rice ragged stunt virus) and aquareoviruses for which cryo-EM has been used and for which structural information is already available (see references 6, 7 and 8) may also suitable for practising the invention.

Cypoviruses and dinovernaviruses have the equivalent of an inner capsid only and therefore are typically not considered suitable for practicing the invention and are therefore considered less preferable. In certain embodiments, the non-enveloped icosahedral viruses for use in the invention do not include cypoviruses and dinovernaviruses. In a preferred embodiment, the viruses of the invention can be manipulated at biosafety level 2 or lower (see reference 9).

Preferably, the virus particle used in practising the invention has three or fewer outer layer proteins. More preferably, the outer layer of the virus particle can be formed by a single outer layer protein, which is the outer surface protein used in preparation of the chimeric surface protein. A low number of outer layer proteins is advantageous because fewer proteins need to be expressed recombinantly to recoat the sub-viral particles after stripping.

For example, rotavirus has two outer layer proteins, VP4 and VP7, but only VP7 is a shell protein that is required to form the outer layer of a rotavirus particle. In most instances, only recoating of rotavirus DLPs with one outer layer protein, VP7, is sufficient in order to practice the invention.

Chimeric surface protein

In one aspect, the invention relates to a chimeric surface protein comprising a rotavirus surface protein covalently linked to a heterologous protein. The rotavirus surface protein may be linked to the heterologous protein via a linker sequence. In a specific embodiment, the heterologous protein is inserted in a flexible loop of the rotavirus surface protein which is the outer surface-exposed portion. In some instances, portions of the rotavirus surface protein are deleted to better accommodate the linker sequence and/or the heterologous protein. For examples, short N-terminal and C-terminal truncations ($\leq$ 10 amino acids) typically do not affect the ability of the rotavirus VP7 protein to recoat rotavirus DLPs. In addition, amino acid sequences that form surface loops that extend from away from the virus particle when the VP7 protein has been used to recoat DLPs are dispensable. Whether a deletion affect the ability of the VP7 protein to recoat DLPs can be assessed by incubating recombinantly expressed VP7 protein in the presence of DLPs and observe the formation of recoated virus particles.

In another aspect of the invention, the rotavirus surface protein is non-covalently linked to the heterologous protein via an adapter system. In a preferred aspect of the invention, the heterologous protein is non-covalently bound to the rotavirus surface protein by a two-part adapter system, where one part of the adapter system is linked to the rotavirus surface protein and the other part of the adapter system is linked to the heterologous protein, whereby both parts of the adapter system form a stable complex thus non-covalently attaching the heterologous protein to the rotavirus surface protein. The adapter system is typically composed of a first adapter polypeptide and a second adapter polypeptide. The first adapter polypeptide is fused to the rotavirus surface protein, optionally via linker sequence. The second adapter polypeptide is fused to the heterologous protein, optionally via a linker sequence. The first and second adapter polypeptides interact with each other to form a stable complex therefore non-covalently attaching the rotavirus surface protein to the heterologous protein, thus forming the chimeric surface protein.

Viral surface protein

The outer layer of a virus particle suitable for practising the invention typically comprises several different outer surface proteins. The main surface protein is particularly suitable for displaying a heterologous protein, as it covers most of the virus particle's surface. For example, the outer surface of rotavirus (excluding the spikes) is formed by 780 copies of the VP7 protein, which forms homotrimers. Choosing a viral particle with a major surface protein that forms homotrimers is particularly preferred for practising the invention.

In one specific embodiment, the viral surface protein is a rotavirus surface protein. In another specific embodiment, the viral surface protein is a glycoprotein.

Examples of other viruses with viral surface proteins that may be suitable for practising the invention include the orthoreoviruses. For example, the outer viral capsid of aquareovirus is formed by 200 trimers of a protein designated VP5. In mammalian orthoreoviruses, the outer layer of an infectious reoviral particle contains 600 copies of the trimeric membrane penetration protein μ1, which is studded with the chaperone protein σ3 (also present in 600 copies) thus forming a heterohexamer.

Heterologous Protein

Many heterologous proteins can be displayed using a trimeric rotavirus surface protein, provided a linker of appropriate length and/or an appropriate insertion site within the surface protein is chosen to avoid steric hindrance between monomers during assembly of the timer. No upper size limit for the heterologous protein exists as long as the heterologous protein does not form a volume that overlaps with the neighbouring volumes of other mounted heterologous proteins. Preferably an adapter system is used to display the heterologous protein on the trimeric rotavirus surface protein. Heterologous proteins that are not soluble, form higher-order oligomers or aggregate are typically not considered suitable in practising the invention.

Any size constraints due to steric hindrance may be overcome by choosing an appropriate linker or adapter system.

In the context of the heterologous protein, the term "heterologous" typically means that the protein is not a rotavirus protein. In some embodiments, the expression "heterologous protein" may mean that the protein is not derived from the same rotavirus strain that is used to display the protein.

Trimeric surface proteins are particularly suitable for displaying trimeric heterologous proteins. Examples of trimeric heterologous proteins include trimeric viral cell entry proteins and other trimeric viral surface proteins, in particular those that are targeted by neutralizing antibodies. Specific examples are the influenza haemagglutinin (HA), human immunodeficiency virus (HIV) gp140, the Ebola virus glycoprotein, rabies virus glycoprotein (RVG), the Env protein of caprine arthritis encephalitis virus, the F protein of respiratory syncytial virus (RSV), the gB protein and its complexes found in human herpes simplex viruses and human cytomegalovirus (HCMV) *etc.*

In some embodiments, the rotavirus surface protein and/or the heterologous protein include(s) a trimerization tag to aid in the assembly of the heterohexamer complex formed by the trimeric rotavirus surface protein and the trimeric heterologous protein. The trimerization tag, particularly coiled-coil based

trimerization tags (*e.g.* GCN4, [10]), can also serve as structural modules to extend the space available for a heterologous protein which is displayed on the surface of a rotavirus particle recoated with a modified rotavirus VP7 protein. Another suitable trimerization tag can be derived from bacteriophage T4 fibritin [11].

In certain embodiments, heterologous proteins that are not trimeric can interact with a trimerization tag present on the rotavirus surface proteins. As an example, if the heterologous protein has an accessible α-helix that could interact with α-helices of a coiled-coil-based trimerization tag, that heterologous protein could bind the trimerization tag. If there is no steric hindrance, up to three such heterologous proteins could bind the trimerization tag simultaneously, forming a six-helix bundle. If fewer than three such heterologous proteins bound, the six-helix bundle would be incomplete, containing the inner three helices contributed by the trimerization tag present on the rotavirus protein but only one or two of the outer helices of the bundle. The non-trimeric heterologous protein could have 1, 2, 4, or more subunits (*i.e.* be a monomer, dimer, tetramer, or other oligomer), provided that at least one of the subunits had an α-helix that could interact with the trimerization tag present on the rotavirus surface protein.

In certain embodiments, a monomeric heterologous protein can include multiple α-helices arranged such that more than one helix could interact with the α-helices of the trimerization tag present on the rotavirus surface protein, forming a complete or partial six-helix bundle. In certain embodiments, tetramers (or higher order structures) wherein at least three of the monomers contain suitably arranged α-helices could associate with the trimerization tag attached to the rotavirus surface protein to form a six-helix bundle. In certain embodiments, tetramers (or higher order structures), wherein at least two of the monomers contain suitably arranged α-helices, could associate with the trimerization tag associated with the rotavirus surface protein to form a partial six-helix bundle, missing one of the outer helices. The subunits of the heterologous protein may be identical (*i.e.* a homodimer, homotrimer, homotetramer, *etc.*) or may not be identical (*i.e.* a heterodimer, heterotrimer, heterotetramer, *etc.*).

If the trimerization tag present on the rotavirus surface protein is not an α-helical coiled-coil, the interacting structural element on the heterologous protein could have a secondary structure that binds the trimerization tag but is not α-helical.

Preferably, the heterologous protein is not an antibody which specifically binds to a rotavirus surface protein.

Adapter system

In some embodiments of the invention, an adapter system is used to display the heterologous protein on a rotavirus particle. The adapter system is typically composed of two adapter molecules. The first adapter molecule is covalently linked to the rotavirus surface protein that was chosen to display a heterologous protein. The second adapter molecule is covalently linked to the heterologous protein. The first adapter molecule and the second adapter molecule interact with each other to form a stable complex.

Adapter systems typically are composed of two adapter polypeptides. The first adapter polypeptide is fused to the rotavirus surface protein that was chosen to display a heterologous protein. The second adapter polypeptide is fused to the heterologous protein. The first adapter polypeptide and the second adapter polypeptide can interact with each other to form a stable complex.

Many polypeptides are known that associate which each other to form a stable complex. In some instances, these polypeptides will be derived from different proteins, *e.g.* a receptor and a ligand or an antibody and an antigen. In other instances, these polypeptides can be derived from the same protein such as the two heptad repeat sequences of HIV gp41.

Suitable adapter systems that do not rely on interactions between two adapter polypeptides can also be envisioned. For example, the rotavirus surface protein may be modified to include a specific glycosylation site, and the heterologous protein can be modified to include a lectin domain that specifically recognises the glycan at the glycosylation site. Examples of lectins that recognise specific glycan structures are well-known in the art.

Another suitable adapter system that does not rely on interactions between two adapter polypeptides is the streptavidin-biotin system. Monomeric streptavidin which has been mutated to prevent tetramer formation and to enhance solubility is preferably used [12]. The monomeric streptavidin can be fused to the rotavirus surface protein that was chosen to display a heterologous protein. For complex formation to occur, the heterologous protein is biotinylated. Enzymatic biotinylation is preferred as it allows biotin to be linked specifically to an amino acid residue present in the protein that is to be biotinylated. For example, the heterologous protein may be modified by insertion of an "AviTag" or "Acceptor Peptide" (AP), which can be specifically biotinylated by a biotin ligase (*e.g.*, BirA) in the presence of biotin and ATP (see reference 13 for details).

Another adapter system that does not rely on interactions between two adapter polypeptides can include an antibody bound to a hapten (*e.g.*, diethylene triamine pentaacetic acid (DTPA), 1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid (DOTA)). Alternatively, the chelators DTPA and DOTA can be coordinated with metal ions to form complexes.

Use of an adapter system is not essential for practising the invention. However, use of an adapter system may be preferred to optimally preserve the structural features of the heterologous protein, in particular of trimeric surface proteins of other viruses that are to be displayed on a rotavirus particle. The use of an adapter system is particularly advantageous because a rotavirus surface protein which has been modified to contain a first adapter needs to be prepared only once and can subsequently be used with any heterologous protein that has been modified to contain a compatible second adapter.

A further advantage of using an adapter system is that expression and purification can be optimised separately for each of the components of the chimeric surface protein. In most instances, the insertion of an adapter polypeptide sequence will not change those characteristics of the viral surface protein or of the heterologous protein which are relevant for their expression or purification. Thus existing expression systems and purification methods can typically be used without modification to prepare large quantities of the viral surface protein comprising the first adapter polypeptide and the heterologous protein comprising the second adapter polypeptide. In contrast, a chimeric surface protein composed of a viral surface protein covalently linked to a heterologous protein will likely have characteristics that differ substantially from the characteristics of the viral surface protein or the heterologous protein each by itself.

The rotavirus surface proteins and many reoviral surface proteins form trimers. Thus, in one aspect of the invention, the adapter system of choice preferably forms a trimeric complex. Particularly preferred are

adapter systems wherein the first adapter polypeptide and the second adapter polypeptide comprise heptad repeat sequences and form α-helical structures that allow the formation of six-helix bundles. Most preferably, three copies of the second adapter polypeptide form an interior coiled-coil trimer, while three copies of the first adapter polypeptide pack into grooves on the surface of this trimer completing the six-helix bundle. However, the inverse may also be possible, in which case three copies of the first adapter polypeptide form an interior coiled-coil trimer, while three copies of the second adapter polypeptide pack into grooves on the surface of this trimer. Various polypeptide pairs that comprise heptad repeat sequences and form six-helix bundles are known in the art. These are typically derived from viral fusion proteins. Because of the rotational symmetry and stability of their six-helix bundles, viral fusion proteins are the preferred source for adapter polypeptide sequences.

One of the most thoroughly studied viral fusion proteins is the envelope glycoprotein of human immunodeficiency virus 1 (HIV-1). The ectodomain of HIV-1 envelope glycoprotein consists of gp120 and gp41. gp41 mediates fusion between viral and cellular membranes. gp41 comprises two heptad repeat sequences, "helical region 1" (HR1) and "helical region 2" (HR2), which can form a six-helix bundle in the native protein [14]. In a preferred embodiment, the first adapter polypeptide and the second adapter polypeptide of the adapter system correspond to gp41 HR2 (SEQ ID NO: 1) and gp41 HR1 (SEQ ID NO: 2), respectively. In a further preferred embodiment, the first adapter polypeptide and the second adapter polypeptide of the adapter system correspond to Nipah virus F protein HR2 (SEQ ID NO: 3) and HR1 (SEQ ID NO: 4), respectively. However, various other viral fusion proteins are known that may be a source of interacting HR1 and HR2 polypeptides that can be used to prepare the first and second adapter polypeptides. These include the influenza hemagglutinin 2 (HA2), the transmembrane (TM) subunit of Moloney murine leukemia virus (Mo-MLV), the paramyxovirus F protein (including avian paramyxovirus F protein, *e.g.* Hendra virus F protein) *e.g.* from respiratory syncytial virus (RSV) and Newcastle disease virus (NDV), the spike protein of coronavirus (*e.g.* the spike proteins of mouse hepatitis virus and SARS-CoV), the ZEBRA protein of Epstein-Barr virus, simian virus 5 fusion protein *etc.* (see references 14, 15, 16 and 17).

Where the heterologous protein itself comprises heptad repeat sequences and these sequences are compatible with the heptad repeat sequence comprised in a modified rotavirus surface protein, the addition of heptad repeat sequences to the heterologous protein may not be necessary. In such a case, the native heptad repeat sequences of the heterologous protein may form a six-helix bundle with the heptad repeat sequence of the modified rotavirus surface protein and the heterologous protein may not need to be modified by the addition of a heptad repeat sequence that is foreign to the heterologous protein.

For example, any of the above listed viral fusion proteins contain interacting heptad repeat sequences, and these native heptad repeat sequences may be able to interact with the heptad repeat sequences in a rotavirus surface protein modified to contain heptad repeat sequences. In some embodiments, it may be preferred to rely on heptad repeat sequences heterologous to the viral fusion protein, in particular where the two-part adapter system is chosen to allow for a modular approach where the same modified rotavirus surface protein is used to display various different heterologous proteins.

Other protein domains comprising heptad repeat sequences that form six-helix bundles are known. These include the CARD domain of human Apaf-1 and RAIDD, death effector domain of FADD and the death

domain of p75 and Fas. Other sources of six-helix bundle-forming peptide sequences include the SNAREs and GCN4-pII. Any of these domains or peptide sequences may likewise be adapted for use in the present invention. For example, the three N-terminal helices and the three C-terminal helices of the CARD domain may be used for the first adapter polypeptide and the second adapter polypeptide, respectively, to form an adapter system suitable for monomeric or dimeric surface proteins.

Other peptide-based adapter systems can also be used. For example, an antibody or antigen binding portion thereof (e.g., a Fab fragment, an scFv, a domain antibody (DAb)) which binds the heterologous protein that one wishes to display on the virus surface can be inserted into the rotavirus surface protein. The antibody or antigen binding portion thereofmay recognise a specific epitope that either can be inserted into the heterologous protein (*e.g.*, in form of a peptide tag) or may naturally be present in the heterologous protein.

In some embodiments, a suitable two-part adapter system requires modification of the heterologous protein only. For example, in one embodiment of the invention, the rotavirus surface protein is a glycoprotein. In such a case, a heterologous protein may be modified to contain a lectin domain that specifically binds the glycan in the glycoprotein. For example, the rotavirus VP7 protein can be modified by the introduction of a single glycosylation sites in the surface-exposed portion of the protein so that the lectin domain of the modified heterologous protein specifically binds to the glycan in the modified VP7 protein.

Alternatively, the heterologous protein may be modified to contain an antigen-binding domain of an antibody that is specific for an epitope on the rotavirus surface antigen. For example, the heterologous protein may be fused to a Fab fragment of an antibody that binds the rotavirus surface antigen with high affinity.

Linker sequence

Linker sequences separate domains derived from different proteins and allow these domains to fold properly. Many linker sequences are known in the art (see reference 18). A linker sequence may be included to separate the rotavirus surface protein and/or the heterologous protein from the adapter sequence. Alternatively, where the rotavirus surface protein and the heterologous protein form two parts of the same protein, both parts may be separated from each other by a linker sequence.

Both rigid and flexible linkers are known. A typical sequence of a flexible linker is composed of repeats of the amino acids G and S. For instance, the linker may have the following sequence: GS, GSG, SGG, GGSGG (SEQ ID NO: 5) or GSGSGSGT (SEQ ID NO: 6). In some embodiments, the same sequence is repeated multiple times (*e.g.* two, three, four, five or six times) to create a longer linker. In other embodiments, a single amino acid such as S or G can be used as a linker. A rigid linker may be composed of several repeats of the amino acid sequence EAAAR (SEQ ID NO: 7).

When choosing a linker sequence, care should be taken to select a hydrophilic linker to avoid aggregation of the modified rotavirus surface protein or heterologous protein.

Typically, the linker is protease-insensitive but in some embodiments the linker contains a protease cleavage site. Protease cleavage sites may be useful to remove tags that are included for

detection/purification of *e.g.* a modified rotavirus surface protein of the invention. Protease cleavage sites may also be useful for exposing an amino acid sequence that has been inserted in a modified rotavirus surface protein so that the amino acid sequence becomes more accessible on the outer surface of the modified rotavirus surface protein after cleavage with a site-specific protease. For example, a protease cleavage site may be used to expose an adapter polypeptide for improving its binding to the corresponding adapter polypeptide of a two-part adapter system which has been fused to a heterologous protein that is to be displayed on the surface of a rotavirus particle. In some instances, the use of a protease cleavage to expose an adapter polypeptide which forms part of a modified rotavirus surface protein may obviate the need for additional linker sequences and thus reduce the number of additional amino acid sequences that need to be inserted into a modified rotavirus surface protein in order to form a stable complex with a heterologous protein.

Protease cleavage sites for a specific protease can be found in a number of proteins. For example, the blood-clotting cascade and the complementation cascade contains a number of very specific proteases that recognise cleavage sites in proteins further downstream in the cascade. Usually, the enzymes at the early stages of a cascade are more specific than are the later ones. For example, Factor X is more specific than thrombin. If thrombin is used, the most preferred thrombin-sensitive cleavage sites are those found in fibrinogen, Factor XIII, and prothrombin. Further examples of proteases of the blood clotting cascade, their target proteins and specific cleavage sites are listed in Table 1 below. The underlined portion of the sequence shown in Table 1 is the minimal cleavage site that needs to be included for the protease to recognise the target.

Table 1

Protease	**Target**	**Cleavage Site(s)**
Human factor XI	Human factor IX	QTSKLTRAEAVF (SEQ ID NO: 8) and SFNDFTRVVGGE (SEQ ID NO: 9)
Human kallikrein	Human factor XII	LFSSMTRVVGGLV (SEQ ID NO: 10)
Human facor XII	Human factor IX	KIKPRIVGGT (SEQ ID NO: 11)

Other proteases that have been used to cleave fusion proteins include enterokinase, collagenase, chymosin, urokinase, renin, Rhinovirus 3C protease, Tobacco Etch Virus (TEV) protease, factor Xa, thrombin, furin, and certain signal peptidases (see *e.g.* reference 19).

Preferably the cleavage site is positioned in such a way in the final construct that any tag that has been added to the rotavirus surface protein or the heterologous protein can easily be removed.

In some embodiments, the linker contains a tag for detection and/or purification. Many tags to facilitate the detection of proteins are known in the art. Frequently used peptide tags include FLAG-tag (DYKDDDDK; SEQ ID NO: 12), HA-tag (YPYDVPDYA; SEQ ID NO: 13), His-tag (*e.g.* HHHHHH; SEQ ID NO: 14), Myc-tag (EQKLISEEDL; SEQ ID NO: 15), Strep-tag I (AWRHPQFGG; SEQ ID

NO: 16), Strep-tag II (NWSHPQFEK; SEQ ID NO: 17), and protein C-tag (EDQVDPRLIDGK; SEQ ID NO: 18). His-tags are preferred as they allow easy detection by anti-His antibodies and permit purification of the tagged protein using a nickel-column. Strep-tag II allows simple and easy purification of recombinantly expressed proteins using streptavidin columns. In some cases, protein tags are used. For example, a glutathione-S-transferase-tag may be included to allow for the easy purification of a protein of the invention using a column comprising immobilized glutathione. A green fluorescent protein-tag can be used if easy detection by fluorescence microscopy is required.

In some embodiments of the invention, a tag (*e.g.* a protein C tag) is included as part of the linker sequence because tags typically used *e.g.* for protein detection and purification do not interfere with the function and folding of the tagged protein and are generally surface exposed. Therefore a tag may provide advantageous properties over other, artificially designed linker sequences.

In other embodiments, a linker sequence may comprise an epitope sequence. Including an epitope sequence can be useful for packing antibody fragments recognising said epitope to further stabilise the complex formed by a modified rotavirus surface protein and a heterologous protein via two-part adapter system. Stabilising the complex may be particularly important for achieving high resolution images for structural studies.

Signal peptide

In some embodiments of the invention, the rotavirus surface protein and the heterologous protein are further modified to comprise a heterologous signal peptide sequence, preferably replacing the native signal peptide sequence. The use of a heterologous signal peptide may be advantageous for achieving higher expression levels in the expression system used to prepare large amounts of the modified rotavirus surface protein and the heterologous protein for recoating rotavirus DLPs. Accordingly, the heterologous signal peptide will be derived from a protein that is known to be expressed in high levels in the chosen expression system. For example, the HIV consensus signal sequence or signal peptide of human tissue plasminogen activator (tPA) are particular suitable for expression in human cells. For expression in insect cells, the Baculovirus gp64 signal peptide or the honeybee melittin signal sequence may be used. Typically, a linker is placed after the heterologous signal peptide sequence in order to guarantee efficient signal peptide cleavage. Generally, the signal peptide is removed by a signal peptidase endogenous to the chosen expression system so that it is not present in the final protein (*i.e.* rotavirus surface protein and the heterologous protein recovered from the expression system).

Nucleic acids

The invention also relates to nucleic acids comprising an open reading frame encoding a chimeric surface protein of the invention operationally linked to a promoter sequence such that the chimeric surface protein is expressed in large amounts in an expression system.

The invention also relates to a nucleic acid construct that encodes a modified rotavirus surface protein, that comprises all or a portion of a rotavirus surface protein, a first adapter polypeptide and, optionally, a linker sequence. The invention further relates to a nucleic acid comprising an open reading frame encoding a fusion protein comprising a heterologous protein, a second adapter polypeptide, and, optionally, a linker sequence, wherein the open reading frame is operationally linked to a promoter

sequence such that the fusion protein is expressed in large amounts in an expression system. In one embodiment, the invention relates to a nucleic acid construct that comprises a nucleotide sequence encoding a second adapter polypeptide, optionally, a linker sequence and a multiple cloning site, wherein insertion of a coding region for a heterologous protein in the multiple cloning site yields an open reading frame encoding a fusion protein comprising the heterologous protein and the second adapter polypeptide. The nucleic acid construct further comprises a promoter sequence that can drive expression of the fusion protein comprising the heterologous protein and the second adapter polypeptide in an expression system.

The first adapter polypeptide and the second adapter polypeptide form part of a two-part adapter system so that the fusion protein comprising the heterologous protein and the second adapter polypeptide and the modified rotavirus surface protein comprising the first adapter polypeptide form a stable complex with each other.

Expression systems

The invention also relates to expression systems for expressing the proteins encoded by the nucleic acids of the invention.

In one embodiment, a first expression system is used to express a modified rotavirus surface protein comprising all or a portion of a rotavirus surface protein, a first adapter polypeptide and, optionally, a linker sequence. A second expression system is used to express a fusion protein comprising a heterologous protein, a second adapter polypeptide, and, optionally, a linker sequence. Optionally, a third expression system is used to express one or more rotavirus protein(s) that, together with the rotavirus surface protein, form(s) the outer layer of a rotavirus particle. The first and second adapter polypeptides interact with each other to form a stable complex. The first expression system comprises a first nucleic acid construct comprising an open reading frame encoding the modified rotavirus surface protein, wherein the open reading frame is operationally linked to a promoter sequence. The second expression system comprises a second nucleic acid construct comprising an open reading frame encoding the fusion protein, wherein the open reading frame is operationally linked to a promoter sequence. The third expression system comprising one or more expression vector(s) for the one or more rotavirus protein(s). The modified rotavirus surface protein and the fusion protein and optionally the one or more rotavirus protein(s) are then purified. The modified rotavirus surface protein and the fusion protein can be mixed in appropriate ratios to form a chimeric surface protein. The chimeric surface protein and optionally the one or more rotavirus protein(s) are then used to recoat rotavirus DLPs to form rotavirus particles displaying the heterologous protein on their surface. Alternatively, rotavirus DLPs are recoated with the modified rotavirus surface protein and optionally the one or more rotavirus protein(s) to form rotavirus particles. The rotavirus particles can then be mixed with the fusion protein to allow the formation of a complex between the first adapter polypeptide and the second adapter polypeptide yielding a chimeric surface protein so that the heterologous protein is displayed on the surface of the rotavirus particles.

In another embodiment, the invention relates to an expression system comprising (i) a first nucleic acid construct comprising an open reading frame encoding a modified rotavirus surface protein that comprises all or a portion of a rotavirus surface protein, a first adapter polypeptide and, optionally, a linker sequence, wherein the open reading frame is operationally linked to a promoter sequence, and (ii) a second nucleic acid construct comprising an open reading frame encoding a fusion protein comprising a

heterologous protein, a second adapter polypeptide, and, optionally, a linker sequence, wherein the open reading frame is operationally linked to a promoter sequence. In some instances, the expression system further comprises an expression vector for one or more rotavirus protein(s) that, together with the rotavirus surface protein, form(s) the outer layer of the rotavirus particle.

The expression system can be a bacterial cell, a yeast cell, a protozoan cell, an insect cell or a mammalian cell. The use of bacterial cells or yeast cells as expression systems is less preferred, in particular where proper glycosylation of the expressed proteins is desired.

Kits

The invention further provides kits comprising a first nucleic acid construct encoding a modified rotavirus surface protein that comprises all or a portion of a rotavirus surface protein and a first adapter polypeptide, and a second nucleic acid construct, wherein the second nucleic acid construct comprises a nucleotide sequence encoding a second adapter polypeptide and a multiple cloning site, and wherein insertion of a coding region for a heterologous protein in the multiple cloning site yields an open reading frame encoding a fusion protein comprising the heterologous protein and a second adapter polypeptide, wherein the first adapter polypeptide of the chimeric fusion protein and the second adapter polypeptide of the fusion protein interact with each other to form a stable complex.

The invention further relates to kits comprising a first nucleic acid construct encoding a modified rotavirus surface protein comprises all or a portion of rotavirus surface protein and a first adapter polypeptide, and a second nucleic acid construct encoding a fusion protein comprising all or a portion of a heterologous protein and a second adapter polypeptide, wherein the first adapter polypeptide of the modified rotavirus surface protein and the second adapter polypeptide of the fusion protein interact with each other to form a stable complex, thus yielding a chimeric surface protein.

Kits may further comprise a rotavirus particle. The rotavirus particle can be either from the same species from which the rotavirus surface protein was derived or from a different species. For example, rhesus rotavirus VP7 can be used to recoat DLPs prepared from bovine rotavirus and *vice versa*. The rotavirus can be uncoated and recoated with the chimeric surface protein. Alternatively, kits may comprise DLPs for recoating with the chimeric surface protein. Rotavirus DLPs can be prepared by uncoating native rotavirus particles or by recombinantly expressing the rotavirus inner shell proteins VP2 and VP6.

Recombinant expression of outer layer proteins

Recoating of rotavirus DLPs requires only one or two recombinant viral proteins, the outer layer protein VP7 or the outer layer protein VP7 together with the outer layer spike protein VP4. The expression and purification of VP4 and VP7 are described in detail in references 20 and 21, respectively.

The outer layer proteins including the chimeric surface protein of the invention can be produced using conventional expression systems known to the skilled person. In order to guarantee correct folding, and in some instances proper glycosylation, expression systems other than prokaryotic or yeast expression systems are preferred. For example, mammalian cells such as CHO cells or 293 cells may be used to overexpress the outer layer proteins needed for recoating of the rotavirus DLPs. Alternatively, the protozoan *Leishmania tarentolae* may be used to express the outer layer proteins. Insect cell systems are also suitable for the expression of the outer layer proteins. For example, the insect cell lines Sf9, Sf21 and

Hi-5 are suitable for the overexpression of glycosylated proteins. In some instances, baculovirus-based insect cell systems are preferred. The expression systems described in references 2, 4 and 5 are particularly suitable in practising the invention.

Various ways of recovering and purifying the overexpressed outer layer proteins are known in the art. Typically a series of chromatographic steps is used to purify the overexpressed proteins from cytoplasmic extracts or the supernatant of the cells which were used as the expression system. For example lectin affinity, immunoaffinity and size exclusion chromatography may be used. If the outer shell protein has been tagged with a peptide- or protein tag, this tag may advantageously be used for purification. If a protease cleavage site is present in the sequence preceding the tag, the tag may be removed after purification using a protease that specifically recognises the protease cleavage site.

In some instances, a crude preparation of the recombinantly expressed outer layer proteins can be used for the recoating reaction. For example, lysates of cells used for expressing the outer layer proteins can be prepared using a lysis buffer and/or mechanical disruption of the cells (*e.g.*, by scraping or sonicating the cells). Any cell debris is removed by centrifugation, and the supernatant containing a crude preparation of the recombinant outer layer proteins can be used in a recoating reaction. The crude preparation may be concentrated using ultrafiltration prior to being used in a recoating reaction.

In those aspects of the invention where a heterologous protein is non-covalently bound to a rotavirus surface protein by a two-part adapter system comprising a first adapter polypeptide and a second adapter polypeptide, the rotavirus surface protein fused to the first adapter polypeptide and the heterologous protein fused to the second adapter polypeptide are expressed separately in different cells. Separate expression may be preferable because both proteins can be purified separately using known purification protocols for each of the proteins. After purification, the proteins can be mixed in the appropriate ratios required for recoating of rotavirus DLPs. For example, the molar ratio of VP7 and VP4 in the outer layer of rotavirus is 13:1. The molar ratio of the modified VP7 protein comprising the first adapter polypeptide and the heterologous protein comprising the second adapter polypeptide is typically 1:1. Alternatively, DLPs are recoated with the rotavirus outer layer protein to form rotavirus particles. The rotavirus particles are then mixed with the heterologous protein to allow the formation of a complex between the first adapter polypeptide and the second adapter polypeptide so that the heterologous protein is displayed on the surface of the rotavirus particles.

Alternatively, both the rotavirus surface protein and the heterologous protein may be expressed in the same cell. Co-expression results in formation of a stable complex of the rotavirus surface protein and the heterologous protein mediated by the two-part adapter system (see above for details). Where the linker sequence connecting the adapter sequence to the rotavirus surface protein or to the heterologous protein contains a tag, this tag may be used to purify the complex from the cells/cell supernatant of the expression system.

To be able to recoat a reoviral core particle with a chimeric surface protein comprising all or part of a heterologous protein, not all proteins need to be provided to form an outer layer. For example, the outer layer of a mammalian reovirus is formed of μ1, σ1 and σ3. In order to recoat reoviral core particles for the purposes of the invention, typically providing a recombinant version of μ1 is sufficient to form viral particles. Optionally, recombinant versions of σ1 and σ3 can also be provided. The recombinant proteins

can be prepared with one of the expression systems mentioned above and can then be purified using known protocols.

Propagation of viral particles

Rotaviruses have been found in a large number of animal species including cattle, pigs, horses, rabbits, mice, dogs, cats, birds and exotic animal species such as addax, saiga, white-tailed gnu, grizzly bear, and red kangaroo. Accordingly, the suitability of a certain cell type for their propagation depends on the host-range of the chosen virus. Preferably, a virus is chosen that allows high-yield propagation in established cell culture systems.

For example, the cell line MA104 can be used to propagate rhesus rotavirus. Rotavirus can be recovered by lysing infected cells, *e.g.* by freeze-thawing infected MA104 cells in medium. Cell debris is cleared from the lysates by low-speed centrifugation, and the viral particles are concentrated by pelleting them using ultracentrifugation or by ultrafiltration. The concentrated suspension of viral particles can be further purified using CsCl gradient centrifugation [2]. A suitable buffer for preparation of the CsCl gradient is TNC (20 mM Tris pH 8.0, 100 mM NaCl, 1 mM $CaCl_2$).

Preferably a well-characterised cell is used for the propagation of rotavirus. Human, bovine and rhesus monkey rotaviruses can be propagated in Vero cells. A suitable Vero cell line (CCL81) can be obtained from the American Type Culture Collection (ATCC).

Similarly, members of the family *Reoviridae* have been isolated from a wide range of mammals, birds, reptiles, fish, crustaceans, insects, ticks, arachnids, plants and fungi. Culture systems have also been established for other viruses in the family *Reoviridae*. For instance, *Rice ragged stunt virus* (RRSV) can be prepared in sufficient amounts from infected rice leaves [7]. C6/36 cells can be used to propagate *Banna virus* (BAV) [22]. Mouse L-cells, in particular the cell line L929, and murine erythroleukemia (MEL) cells can be used to propagate mammalian reovirus strains type 1 Lang (T1L) and type 3 Dearing (T3D) [5]. Grass carp aquareovirus (GCRV) can be propagated in *Ctenopharyngodon idellus* kidney (CIK) cell culture [8].

Virus particles can typically be recovered from cell supernatants or lysed cells. For examples, any cell debris can be removed by low-speed centrifugation, and viral particles are pelleted from the supernatant using ultracentrifugation. The pellets contain virus particles that can be further purified if needed.

DLP preparation

Rotavirus DLPs are prepared by removing the outer layer proteins from purified rotavirus particles. Rotavirus particles can be uncoated by incubating virus particles in the presence of a calcium chelator such as EDTA or EGTA. A suitable buffer for uncoating rotavirus particles contains 20 mM Tris, pH8, 100 mM NaCl and 1 mM EDTA [2]. Rotavirus particles may also be uncoated by heat shock. The resulting DLPs can be purified by banding on two successive preformed CsCl gradients (ρ = 1.25 to 1.50 g/cm^3).

The method of choice for removing the outer layer proteins depends on the chosen viral particle. For example, mammalian reovirus particles can be stripped to the core by incubating the particle with α-chymotrypsin (CHT) for two hours at 37°C [4].

The above methods may be adapted to other viruses in the family *Reoviridae* depending on the sensitivity of their outer layer proteins to protease treatment or low-calcium conditions.

In some instances, subviral particle can be recovered by lysing the cells that were used to propagate the virus. The subviral particle can then be recovered from the cell lysates, *e.g.*, by a CsCl gradient as described for *Banna* virus [22].

Alternatively, DLPs or subviral particles may be prepared by recombinantly expressing the relevant viral proteins in one of the expression systems described above. For example, the rotavirus inner shell proteins VP2 and VP6 can be expressed recombinantly to from virus-like particles that resemble DLPs.

Recoating

Recoating of rotavirus DLPs occurs at a pH range including pH 4.5 and 6.5, but is most efficient between pH 5 and 5.5 (preferably pH 5.2). Recoating occurs at temperatures including 4°C and 37°C, but is more efficient between 4°C and 30°C than at 37°C.

The recoating of reoviral core particles can typically be done under conditions that mimic the conditions under which the assembly of the viral particles occurs in the cell. For example, mammalian orthoreovirus core particles can be recoated at 37°C. However, individual parameters such as pH and temperature may require further optimisation for efficient reassembly of viral particles from purified viral cores and recombinantly expressed outer layer proteins.

Recoated virus particles can be recovered by CsCl gradient centrifugation. CsCl gradient purification can be repeated a second time to increase the purity of the recovered virus particles.

Where an adapter system is used to display a heterologous protein on the surface of a recoated rotavirus particle, the rotavirus DLPs can be recoated with a rotavirus surface protein comprising a first adapter polypeptide. The heterologous protein comprising a second adapter polypeptide, which forms a stable complex with the first adapter polypeptide, can subsequently be added to the recoated virus particles. This may be advantageous where recoating with the complex formed by the heterologous protein and the rotavirus surface protein would be less efficient, *e.g.* due the size of the heterologous protein. Alternatively, a complex between the heterologous protein and the rotavirus surface protein via the adapter polypeptides may be formed first, and this complex may be added to rotavirus DLPs for recoating.

Complex formation

In some aspects of the invention, the chimeric surface protein is used to study the structure of complexes formed between the heterologous protein which forms part of the chimeric surface protein (*e.g.* by being non-covalently bound to a rotavirus surface protein via a two-part adapter system) and a molecule that specifically binds to the heterologous protein.

In one aspect of the invention, the molecule is a proteinaceous molecule. The proteinaceous molecule typically is another protein such as a receptor or ligand that interacts with the heterologous protein or an antibody or antibody fragment that recognises an epitope found on the surface of the heterologous protein.

For example, a hexameric complex of a trimeric rotavirus surface protein bound to a trimeric viral cell entry protein (*e.g.*, HIV gp140 or RSV F protein) which is displayed on the surface of a rotavirus particle may be used to study interaction between the viral cell entry protein and its host cell surface receptor.

The optimal conditions for complex formation between the heterologous protein and the proteinaceous molecule depend on the nature of the interaction. Receptor-ligand interactions may require different conditions from antibody-antigen interactions. Typically, complex formation is performed at room temperature in a buffered solution (*e.g.*, phosphate-buffered saline).

In one embodiment, the proteinaceous molecule is added to a suspension of recoated rotavirus particles displaying the chimeric surface protein. The buffered solution may contain additional components such as Ca^{2+} to prevent uncoating of the rotavirus particles and optionally one or more protease inhibitors to block degradation of the proteins. After incubation of the proteinaceous molecule in the presence of the recoated rotavirus particles, the newly formed complexes of the proteinaceous molecules bound to the rotavirus particles can be separated from any unbound proteinaceous molecules by centrifugation or ultrafiltration. However, removal of unbound proteinaceous molecules may not be necessary if the recoated rotavirus particles are used for cryo-EM analysis.

Alternatively, the proteinaceous molecule and the chimeric surface protein are incubated together to allow complex formation to occur. First, the heterologous protein and the rotavirus surface antigen may be incubated together to form the chimeric surface protein. Once the chimeric surface protein has formed, the proteinaceous molecule is added. The complex of the proteinaceous molecule and the chimeric surface protein can then be added to rotavirus DLPs to form rotavirus particles, optionally in the presence of any additional rotavirus proteins that, together with the rotavirus surface protein, form the outer layer of a native rotavirus particle.

In a particular aspect of the invention, the chimeric surface protein of the invention is used to determine the structure of an antigen-antibody complex. In this aspect of the invention, the heterologous protein may be derived from a pathogen such as a virus or a bacterium. For example, an immunodominant antigen may be chosen as the heterologous protein to study which parts of the protein are targeted by antibodies during an immune response against the antigen. Preferably, the heterologous protein is a trimeric viral surface protein such as influenza virus haemagglutinin, respiratory syncytial virus F, or HIV gp140. To determine the structure of an antigen-antibody complex, the use of a Fab fragment in place of the full-length antibody is typically preferred to avoid steric hindrance between neighbouring chimeric surface proteins and to guarantee maximal occupancy of the epitope found on the heterologous protein.

In another aspect of the invention, complex formation between a heterologous protein and a non-proteinaceous molecule may be studied. A non-proteinaceous molecule may be a nucleic acid (*e.g.*, RNA or DNA), a polysaccharide or oligosaccharide (*e.g.*, a glycan). For example, the heterologous protein may be a transcription factor or other DNA-binding protein that forms a complex with a specific DNA sequence. Alternatively, the heterologous protein may be a lectin that forms a complex with a glycan.

Cryo-EM

Using cryo-EM for structure determination has several advantages over more traditional approaches such as X-ray crystallography. In particular, cryo-EM places less stringent requirements on the sample to be

analysed with regard to purity, homogeneity and quantity. Importantly, cryo-EM can be applied to targets that do not form suitable crystals for structure determination.

A suspension of purified or unpurified recoated rotavirus particles, either alone or in complex with a proteinaceous molecule such as an antibody or non-proteinaceous molecule such as a nucleic acid, can be applied to carbon grids for imaging by cryo-EM. The coated grids are flash-frozen, usually in liquid ethane, to preserve the particles in the suspension in a frozen-hydrated state. Larger particles can be vitrified by cryofixation. The vitrified sample can be cut in thin sections (typically 40 to 200 nm thick) in a cryo-ultramicrotome, and the sections can be placed on electron microscope grids for imaging.

The quality of the data obtained from images can be improved by using parallel illumination and better microscope alignment to obtain resolutions as high as ~3.3 Å. At such a high resolution, *ab initio* model building of full-atom structures is possible. However, lower resolution imaging might be sufficient where structural data at atomic resolution on the chosen or a closely related rotavirus particle and the selected heterologous protein or a close homologue are available for constrained comparative modelling (see below).

To further improve the data quality, the microscope can be carefully aligned to reveal visible contrast transfer function (CTF) rings beyond 1/3 $Å^{-1}$ in the Fourier transform of carbon film images recorded under the same conditions used for imaging. The defocus values for each micrograph can then be determined using software such as CTFFIND [23]. Final pixel size of the density map can be calibrated using, *e.g.*, Tobacco Mosaic Virus (TMV).

Useful descriptions of applying cryo-EM to structural studies of rotavirus particles are found in references 24 and 25.

Image analysis and structure determination

Images obtained by cryo-EM are analysed to identify micrographs of single particles. Single particle selection can be done with the help of software tools such as SIGNATURE [26]. The astigmatic defocus, specimen tilt axis, and tilt angle for each micrograph can be determined using the computer programme CTFTILT [23]. Obtaining separate defocus values for each particle according to its coordinate in the original image improves the data quality of the cryo-EM density map which is obtained by averaging single-particle micrographs of rotavirus particles.

Fitting of known atomic models within a cryo-EM density map is a common approach for building models of complex structures such as viral particles. A number of computational fitting tools are available which range from simple rigid-body localization of protein structures, such as Situs [27], Foldhunter [28] and Mod-EM [29], to complex and dynamic flexible fitting algorithms like NMFF [30], Flex-EM [31], MDFF [32] and DireX [33, 34], which morph known structures to a density map.

When an atomic model is not known, cryo-EM density maps can be used in building and/or evaluating structural models from a gallery of potential models that are constructed *in silico* (see references 29, 35, 36, 37 and 38). A related template structure must be known for constrained comparative modelling or, for constrained *ab initio* modelling, the fold to be modelled must be relatively small. For example, an initial structure may be obtained using IMIRS [39]. Further alignment and reconstruction can be performed with

FREALIGN [40] using a known rotavirus structure and a known structure of the heterologous protein or a close homologue as template.

Significant structural and functional information can be obtained directly from the density map itself. For example, at 5-10 Å resolutions, some secondary structure elements are visible in cryo-EM density maps: α-helices appear as cylinders, while β-sheets appear as thin, curved plates. These secondary structure elements can be reliably identified and quantified using feature recognition tools to describe a protein structure or infer the function of individual proteins. At near-atomic resolutions (3-5 Å), the pitch of α-helices, separation of β-strands, as well as the densities that connect them, can be visualized unambiguously (see, *e.g.*, references 41, 42, 43 and 44).

De novo model building in cryo-EM comprises feature recognition, sequence analysis, secondary structure element correspondence, Cα placement and model optimisation. Various software applications can be used, *e.g.*, EMAN for density map segmentation and manipulation [45], SSEHunter [46] to detect secondary structure elements, visualization in UCSF's Chimera [47] and atom manipulation in Coot [48,49].

Secondary structure identification programs like SSEHunter provide a semi-automated mechanism for detecting and displaying visually observable secondary structure elements in a density map [46]. Registration of secondary structure elements in the sequence and structure, combined with geometric and biophysical information, can be used to anchor the protein backbone in the density map [41, 43]. This sequence-to-structure correspondence relates the observed secondary structure elements in the density to those predicted in the sequence. The modelling toolkit GORGON couples sequence-based secondary structure prediction with feature detection and geometric modelling techniques to generate initial protein backbone models [50]. Automatic modelling methods such as EM-IMO (electron microscopy-iterative modular optimization) can be used for building, modifying and refining local structures of protein models using cryo-EM maps as a constraint [51].

Once a correspondence has been determined using secondary structure element, Cα atoms can be assigned to the density beginning with α-helices and followed by β-strands and loops. For example, by taking advantage of clear bumps for Cα atoms, Cα models can be built using the Baton_build utility in the crystallographic programs O [52] and/or Coot [48]. Cα positions can be interactively adjusted such that they fit the density optimally while maintaining reasonable geometries and eliminating clashes within the model. Coarse full-atom models can be refined in a pseudocrystallographic manner using CNS [53]. Models can be further optimized using computational modeling software such as Rosetta [36]. Full-atom models can also be built with the help of other computational tools such as REMO [54]. The quality of a model can be confirmed by visual comparison of the model with the density map. Pseudocrystallographic R factor/Rfree analysis [55] provides a measure of the agreement between observed and computed structure factor amplitudes and may be used to confirm that the obtained atomic model provides a good fit to the cryo-EM density maps. Protein model geometry can be checked by PROCHECK [56].

Screening methods for immunogen design

Using cryo-EM in place of X-ray crystallography facilitates rapid structure determination. In comparison to X-ray crystallography, cryo-EM places less stringent requirements on the purity, homogeneity and

quantity of the sample to be analysed. These characteristics make cryo-EM attractive in the context of immunogen design for vaccines, particularly against pathogens that are subject to antigenic drift.

Having more structural information especially about epitopes shared by variants of the same pathogen or closely related pathogens may allow the rational design of immunogens, which can be used in vaccines that provide broad protection against a large number of variants of the same pathogen or closely related pathogens. Such so-called "universal" vaccines are believed to be more cost-effective than traditional vaccines because they would make it superfluous to include various variants of the same pathogen or closely related pathogens in the same vaccine composition (as is the case, *e.g.*, for currently available polio vaccines as well as for streptococcal and meningococcal conjugate vaccines) or to provide new vaccine compositions each year to account for the antigenic drift that occurred in the pathogen population in the previous season (as is the case for influenza vaccines).

Rational design of an immunogen for vaccination against a diverse pathogen involves identification of those regions of an immunodominant protein that are conserved among various variants/subtypes of the same pathogen or among closely related pathogens. The use of rotavirus particles for displaying heterologous proteins on chimeric rotavirus surface proteins makes it possible to rapidly determine a large number of structures in a relatively short time. Therefore the methods of the invention may be particularly useful in the identification of conserved epitopes of immunogenic proteins.

By determining the structure of immunogens in complex with antibodies or corresponding antigen-binding fragments that have been found to be elicited against a number of variants/subtypes of the same pathogen and/or against closely related pathogens, conserved regions in the immunodominant immunogen of this pathogen and/or closely related pathogens can be identified. In one embodiment, the invention therefore relates to a method for obtaining a three-dimensional model of an immunogen complexed to an antibody wherein said method comprises the steps of (i) recoating a rotavirus DLP with a chimeric surface protein comprising the immunogen to yield a suspension of rotavirus particles displaying the chimeric surface protein, (ii) adding to the suspension an antibody or antibody fragment that specifically binds to the immunogen, wherein the antibody or fragment forms a complex with the chimeric surface protein, (iii) freezing the suspension, (iv) imaging the rotavirus particles using cryo-EM to obtain a plurality of micrographs, and (vi) analysing the plurality of micrographs to obtain a three-dimensional model of the immunogen complexed to the antibody. The immunogen is typically heterologous to the rotavirus particle. The three-dimensional model of the immunogen complexed to the antibody can be used to define the epitope on the surface of the immunogen which is recognised by the antibody.

This method can be used to identify the epitope bound by an antibody that has broadly neutralising activity against a number of related pathogens or variants of a pathogen which arise due to antigenic drift. Alternatively, this method can be used to identify the epitope bound by an antibody that may recognise an epitope that interferes *e.g.* with the function of a viral cell entry protein and may be used to prevent or inhibit virus propagation. An antibody that binds to a specific three-dimensional configuration of a protein may also be useful as a quality control reagent during manufacturing of the protein, *e.g.* to exclude lots of a manufactured protein that contain a large number of denatured or misfolded copies of this protein. This method may further be employed to identify epitopes that are particularly useful when

bound by an antibody that is used as diagnostic reagent (*e.g.* epitopes that are specific to a particular pathogen in a group of closely related pathogens).

By repeating the method for obtaining a three-dimensional model of an immunogen complexed to an antibody for a panel of antibodies or their corresponding antigen-binding fragments which have been found to be elicited in response to infection by a number of variants/subtypes of the same pathogen and/or against closely related pathogens, the epitope of each antibody on the surface of the immunogen can be identified. Once the epitope of each antibody in the panel has been identified, this information can be used for the rational design of an immunogen. For examples, the method may be used to identify the epitopes of a panel of antibodies which have been found to have neutralising activity against a number of variants/subtypes of the same pathogen and/or against closely related pathogens to design an immunogen that provides broad neutralising activity against a large number of variants/subtypes of the same pathogen and/or closely related pathogens.

Alternatively, the method described above can be employed to map the repertoire of epitopes recognised by antibodies that are elicited in response to immunisation with an existing vaccine. By mapping the eptitopes most commonly recognised by these antibodies, immunodominant epitopes can be identified. This information can be used advantageously to further optimise an existing vaccine. For example, where an existing vaccine is composed of an inactivated pathogen, an optimised version of the vaccine may only include the immunodominant portions of said pathogen (*e.g.*, in form of a subunit vaccine that comprises recombinant versions of the identified immunodominant antigens or epitopes). Understanding of the structural determinants of immunodominant antibody epitopes, obtained using these technique, can be applied to prepare antigens that have been engineered to make the most useful epitopes (such as broadly neutralizing epitopes) immunodominant.

Examples of epitope mapping to aid in immunogen design are known in the art but generally have been limited to a small number of epitopes on an immunogen because of the efforts involved in getting structural data for immunogen-antibody complexes by X-ray crystallography. Often only a single antibody is tested (see references 57, 58 and 59). The present invention for the first time allows obtaining structural information for immunogen-antibody complexes using a high-throughput approach. The invention is particularly useful for antibodies that only weakly associate with the immunogen and thus present major challenges for crystallization.

Various approaches to design immunogens that are broadly protective against a number of variants of the same pathogen or closely related pathogens are known in the art. In recent years, many efforts have been made to design "universal" vaccines against influenza virus and HIV (see references 60, 61 and 62) Rational design of a modified immunogen has been hampered by the lack of structural data to guide the process. The use of rotavirus particles for displaying heterologous proteins on chimeric rotavirus surface proteins makes it possible to rapidly determine how modifications of the amino acid sequence of a protein affect its three dimensional structure. Amino acid modifications that change the structure of a protein may affect its ability to form crystals. Thus, structural information for assessing the impact of these modifications may not easily be available because the modified protein may not crystalize under the same conditions as the unmodified protein. Since the use of cryo-EM does not rely on the formation of crystals for the determination of protein structures and, in addition, does not require any adaptation of the basic

experimental set-up for different types of proteins, the rapid structural characterisation of modified proteins is possible.

Once an epitope on an immunogenic protein of a pathogen has been identified that is present in a number of variants/subtypes of the same pathogen and/or in many closely related pathogens, this information can be used to design a universal vaccine. In some instances, this may require modification of the immunogen, particularly where a conserved region which comprises a shared epitope is not easily accessible to antibodies in the native protein, as poor accessibility of the epitope usually translates into poor immunogenicity. By modifying the immunogen in such a way that it retains the native epitope which can be recognised by an antibody, but making the epitope more accessible, the modified immunogen may yield an antibody response that is protective against a wide range of variants/subtypes of the same pathogen or of closely related pathogens. In additional to mapping the epitopes of affinity matured antibodies, the epitopes recognized by the un-mutated ancestors of these antibodies, once deduced from B-cell repertoire cloning experiments, can be mapped by this technique. This information could aid in the design of immunogens that selectively amplify a desired un-mutated ancestral antibody and then direct its affinity maturation to a broadly and potently neutralizing antibody.

Candidate immunogens that have been modified from a native immunogen (be it in order to prepare a universal vaccine or for other reasons) can be tested for the presence of a native epitope of interest using the methods of the invention. In a particular embodiment, the invention therefore further relates to a method for determining the structural differences between two variants of a heterologous protein, wherein the method comprises the steps of (i) recoating a rotavirus DLP with a first chimeric surface protein comprising a first variant of the heterologous protein to yield a suspension of a first rotavirus particle displaying the first chimeric surface protein, (ii) freezing the suspension, (iii) imaging the first rotavirus particle using cryo-EM to obtain a plurality of micrographs, (iv) analysing the plurality of micrographs to obtain a three-dimensional model of the first chimeric surface protein, (v) repeating steps (i)-(iv) with a second chimeric surface protein comprising a second variant of the heterologous protein, wherein the first and second chimeric surface protein are identical to each other apart from the difference between the first variant and the second variant, and (vi) comparing the three-dimensional model of the first chimeric surface protein to the three dimensional model of the second chimeric surface protein to determine the structural differences between the first variant and the second variant of the heterologous protein. This method can be repeated for further variants of the same heterologous protein to provide a screening assay for the identification of a variant with certain desired structural characteristics.

In some instances, the structures of the native immunogen and of a plurality of modified immunogens each in complex with an antibody recognising an epitope may be determined. Comparing the structure of the modified immunogens to the native immunogen will assist in selecting candidate immunogens that best preserve the native epitope for further testing *in vivo*.

Immunogenic compositions

In a further aspect, the present invention relates to the use of rotavirus particles comprising the chimeric surface protein of the invention in the preparation of a medicament. In particular, rotavirus particles comprising the chimeric surface protein of the invention may be used in immunogenic compositions

suitable for human vaccination. For example, the chimeric surface protein can be designed to display a heterologous protein that elicits a protective immune response when administered to a patient.

The use of chimeric viruses to elicit an immune response to a pathogen-derived antigen is well-known in the art. Traditionally, this approach requires the reengineering of the genome of the host virus that is selected as a vector for the pathogen-derived antigen. Such an approach has a number of limitations. Typically, the virus genome is modified to contain a pathogen-derived gene encoding the antigen. The genome size of the chosen virus therefore may limit the size of the pathogen-derived antigen which can be expressed. Similarly, the insertion of a gene encoding the pathogen-derived antigen may interfere with the propagation of the virus and limit the yields that can be achieved. Therefore preparation of chimeric viruses needs to be optimised for each pathogen-derived antigen.

Creating a chimeric virus with a new surface protein using recombinant genetics to alter the genome of the host virus by inserting a coding region for a pathogen-derived antigen could potentially change the host range of the chimeric virus and therefore have unforeseeable consequences on the pathogenicity of the newly created virus. Therefore additional steps have to be taken to provide a sufficiently attenuated chimeric virus that is not able to replicate in the subject to be vaccinated.

The present invention overcomes the problems associated with the traditional approach because it allows the preparation of any kind of chimeric surface protein that can be used to recoat rotavirus DLPs. The recoated rotavirus particles do not contain any genetic information for the pathogen-derived antigen. Thus after the initial infection of a host cell with the recoated rotavirus particles, no progeny viruses are formed that carry the pathogen-derived antigen.

The use of a two-part adapter system further reduces the number of optimisation steps. Once a modified rotavirus protein fused to a first adapter polypeptide has been prepared, any known pathogen-derived antigen can be fused to a second adapter polypeptide that forms a stable complex with the first adapter polypeptide. This eliminates any need for genetic engineering of the rotavirus. In most cases, expression and purification of the pathogen-derived antigen is not affected by the presence of the second adapter polypeptide in the pathogen-derived antigen, so that no modification of existing expression and purification methods for the antigen is required. Existing protocols to recoat rotavirus DLPs with a chimeric surface protein comprising the modified rotavirus protein linked to the pathogen-derived antigen via the two-part adapter system can be used to prepare rotavirus particles comprising the chimeric surface protein. These rotavirus particles can then be included in immunogenic compositions to elicit an immune response against the pathogen-derived antigen.

The vaccine platform is particularly useful preparing immunogenic compositions in which the rotavirus particles comprise a trimeric chimeric surface protein that contains a heterologous trimeric viral cell entry protein. Trimeric viral cell entry proteins are the immunodominant surface antigens of many viruses. With the exceptions of hepatitis B surface antigen and human papilloma virus L1 protein, both of which form virus-like particles, and influenza HA, which forms rosettes, subunit vaccines based on immunodominant viral surface antigens have typically failed to result in clinically effective vaccines. By displaying these viral surface antigens in their native conformation on rotavirus particles, immunogenic compositions can be prepared that elicit a protective immune response against the virus from which the surface antigen was originally derived.

Specific examples of trimeric viral cell entry protein include influenza HA, HIV gp140, the Ebola virus glycoprotein, rabies virus glycoprotein, the Env protein of caprine arthritis encephalitis virus, the RSV F protein, the gB and optionally its complex with other proteins of human herpes simplex viruses and of HCMV.

Rotavirus particles

Rotaviruses are particularly useful because various animal strains exist that have a defined host range. For example, some strains are specific for cows and others for monkeys. Most animal strains are antigenically distinct from those strains that typically infect humans and therefore are mostly unable to cause disease in humans. Therefore humans typically do not have any pre-existing immunity to these rotavirus strains that could interfere with an immune response against the chimeric surface protein when these animal strains are used as vectors in immunogenic compositions for human vaccination. In addition, these viruses are naturally attenuated in humans.

In addition, two licensed live, attenuated rotavirus vaccines are currently marketed: Rotarix™ and RotaTeq™. Rotarix™ contains the attenuated human rotavirus strain RIX4414, which passaged 26 times in Primary Green Monkey Kidney cells (AGMK) and is propagated in Vero cells. RotaTeq™ contains a combination of five human rotavirus (HRV)-bovine rotavirus (BRV) reassortant strains, designated as G1, G2, G3, G4, and P1, respectively, which are also propagated in Vero cells. All reassortants are composed of the BRV strain WC3 (G6P7[5]) genome background expressing human VP7 or VP4 proteins.

The previously licensed rhesus rotavirus (RRV)-based vaccine RotaShield™ consists of RRV (G3P5B[3]) and 3 RRV-HRV reassortant rotaviruses. Each reassortant derives 10 genes from RRV and a single HRV gene encoding a VP7 protein for G serotype 1, 2, or 4 specificity (G1P5B[3], G2P5B[3], and G4P5B[3]).

Furthermore, safety of monovalent bovine (NCDV RIT4237 G6P6) and simian (RRV strain MMU18006 G3P5B) rotavirus strains for use as vaccines has been established.

The immunisation, clinical safety and manufacturing experience with these vaccine strains can be directly applied to the rotavirus particles of the invention. For example, any one of the rotaviruses contained in the licensed rotavirus vaccines could be used to prepare DLPs for recoating with the chimeric surface protein of the invention.

Preparation

The modified rotavirus particles included in the immunogenic compositions can be prepared using any of the methods described above. Native rotavirus particles may be propagated and purified. The purified rotavirus particles may then be uncoated to prepare DLPs which can then be recoated with a chimeric surface protein comprising a pathogen-derived antigen.

Typically, a modified rotavirus protein fused to a first adapter polypeptide is added to the DLPs to form rotavirus particles. The addition of further rotavirus proteins that, together with the rotavirus surface protein, form the outer layer of a native rotavirus particle is optional. The rotavirus particles are then

incubated with a pathogen-derived antigen fused to a second adapter polypeptide which forms a stable complex with the first adapter polypeptide to provide the chimeric surface protein.

Formulations

The immunogenic composition of the invention may be provided in form of a kit comprising a first container comprising lyophilised rotavirus particles and a second container comprising a solution for extemporaneous resuspension of the lyophilised rotavirus particles. Lyophilised rotavirus particles may comprise one or more lyoprotectant such as sucrose, dextran, sorbitol and amino acids to stabilise the rotavirus particles during lyophilisation.

Alternatively, the immunogenic composition is provided in a single container comprising the rotavirus particles in suspension. Where the immunogenic composition is for injection, it may be provided in a syringe.

Either solution may contain one or more excipient(s).

The solutions are typically water-based. Therefore purified water may form the main excipient. For example, dilution of the rotavirus particles to give the desired final concentration will usually be performed with water for injection (WFI).

The solution typically contains a buffer. Therefore further excipients include buffering agents and pH regulators such as sodium citrate, sodium dihydrogen phosphate monohydrate, and sodium hydroxide. Antacids such as calcium carbonate may be added to prevent inactivation of the virus during passage through the stomach if the immunogenic composition is administered orally. An acidity regulator such as di-sodium adipate may also be included, preferably in place of the antacid.

In some instances, a thickening agent such as xanthan may be present as a further excipient.

A surfactant, in particular a non-ionic surfactant such as polysorbate 80, may also be present.

Other excipients include sucrose, sorbitol, inorganic salts, amino acids and vitamins.

Compositions will generally have an osmolality of between 200 mOsm/kg and 400 mOsm/kg, preferably between 240-360 mOsm/kg, and will more preferably fall within the range of 280-320 mOsm/kg.

The pH of a composition of the invention will generally be between 5.0 and 7.5, and more typically between 5.0 and 6.0 for optimum stability.

Compositions of the invention preferably contain <1 EU (endotoxin unit, a standard measure; 1 EU is equal to 0.2 ng FDA reference standard Endotoxin EC 2 'RSE') per dose, and preferably <0.1 EU per dose.

Compositions of the invention are preferably gluten-free. Furthermore, compositions of the invention are preferably sterile.

Administration

The immunogenic compositions may be used to stimulate a mucosal immune response. Therefore the immunogenic compositions may be administered orally or intratracheally. Other routes of administration

such as intramuscular injection may be chosen depending on the pathogen-derived antigen displayed on the surface of the modified rotavirus particle included in the immunogenic composition of the invention.

General

The term "comprising" encompasses "including" as well as "consisting" *e.g.* a composition "comprising" X may consist exclusively of X or may include something additional *e.g.* X + Y.

The word "substantially" does not exclude "completely" *e.g.* a composition which is "substantially free" from Y may be completely free from Y. Where necessary, the word "substantially" may be omitted from the definition of the invention.

The term "about" in relation to a numerical value *x* means, for example, $x \pm 10\%$ or *x* + two standard deviations of the value. In certain embodiments, "about" is understood as acceptable variation and tolerances within the specific art. Unless clear from context, all numerical terms herein are understood to be modified by about. .The term "antibody" includes antibody fragments such as antigen-binding fragments (Fabs), single-chain variable fragments (scFv), *etc.* The term "antigen-binding portion" of an antibody (or "antibody portion") includes fragments of an antibody that retain the ability to specifically bind to an antigen. It has been shown that the antigen-binding function of an antibody can be performed by fragments of a full-length antibody. Examples of binding fragments encompassed within the term "antigen-binding portion" of an antibody include (i) a Fab fragment, a monovalent fragment consisting of the VL, VH, CL and CH1 domains; (ii) a $F(ab')_2$ fragment, a bivalent fragment comprising two Fab fragments linked by a disulfide bridge at the hinge region; (iii) a Fd fragment consisting of the VH and CH1 domains; (iv) a Fv fragment consisting of the VL and VH domains of a single arm of an antibody, (v) a dAb fragment [63], which consists of a VH domain; and (vi) an isolated complementarity determining region (CDR). Furthermore, although the two domains of the Fv fragment, VL and VH, are coded for by separate genes, they can be joined, using recombinant methods, by a synthetic linker that enables them to be made as a single protein chain in which the VL and VH regions pair to form monovalent molecules (known as single chain Fv (scFv); see *e.g.* references 64 and 65,. Such single chain antibodies are also intended to be encompassed within the term "antigen-binding portion" of an antibody.

As use herein, "or" is understood to be inclusive and is interchangeable with "and/or" unless otherwise clearly indicated by context.

As used herein, "a" and "the" are understood to include both singular and plural unless otherwise clearly indicated by context.

Unless specifically stated, a process comprising a step of mixing two or more components does not require any specific order of mixing. Thus components can be mixed in any order. Where there are three components then two components can be combined with each other, and then the combination may be combined with the third component, *etc.*

Where animal (and particularly bovine) materials are used in the culture of cells for preparation of material for administration to animals, especially humans, they should be obtained from sources that are free from transmissible spongiform encaphalopathies (TSEs), and in particular free from bovine

spongiform encephalopathy (BSE). Overall, when products are prepared for administration to animals, and especially humans, it is preferred to culture cells in the total absence of animal-derived materials.

Where a cell substrate is used for reassortment or reverse genetics procedures, it is preferably one that has been approved for use in human vaccine production *e.g.* as in Ph Eur general chapter 5.2.3 [66].

Identity between polypeptide sequences is preferably determined by the Smith-Waterman homology search algorithm as implemented in the MPSRCH program (Oxford Molecular), using an affine gap search with parameters gap open penalty=12 and gap extension penalty=1.

DESCRIPTION OF THE DRAWINGS

Fig. 1: Schematic overview of modifications made to (A) the rotavirus VP7 protein and (B) the heterologous protein. HR1 and HR2 form part of a two-part adapter system for non-covalently binding the heterologous protein to the rotavirus VP7 protein. The heterologous signal peptide sequences serve the purpose of achieving high expression levels in the chosen expression systems. The affinity tags allow easy purification. Protease recognition site P_1 can be used to remove the affinity tag A_1 after purification. The second set of affinity and protease recognition sites (A_2 and P_2) can be replaced by suitable linker sequences and may serve as a spacer sequence that may be needed to display large heterologous protein. The trimerization tag is optional and can be used in aiding in trimer formation of some heterologous proteins which then facilitates binding to the trimeric rotavirus VP7 protein. The trimerization tag, particularly coiled-coil based trimerization tags (*e.g.* GCN4), can also serve as structural modules to extend the space available for a heterologous protein which is displayed on the surface of a rotavirus particle recoated with a modified rotavirus VP7 protein.

Fig. 2: Schematic overview of the complex formed by a modified trimeric rotavirus VP7 protein and a trimeric heterologous protein. The rotavirus VP7 protein is non-covalently bound to the heterologous protein via a two-part adapter system, where one part of the adapter system (HR2) is linked to the rotavirus VP7 protein and the other part of the adapter system (HR1) is linked to the heterologous protein. HR1 and HR2 form a six-helix bundle resulting in a stable complex for non-covalently attaching the heterologous protein to the rotavirus VP7 protein. The chimeric surface protein can become part of the outer layer of the rotavirus particle by *in vitro* recoating DLPs.

Fig. 3: Electron micrographs of (A) purified DLPs, (B) DLPs recoated with rotavirus VP7 protein and (C) DLPs recoated with VP7 protein displaying influenza virus HA as the heterologous protein. The particles were negatively stained prior to image acquisition.

Fig. 4: (A) Coomassie-stained acrylamide gel after SDS-PAGE. Lane 1 shows the molecular weight marker (MW). Lane 2 was loaded with the purified DLPs used in the recoating reaction (DLP). Lane 3 was loaded with the purified VP7-HA protein complex (pro) used for recoating the DLPs. Lane 4 was loaded with the input mixture of DLPs, modified VP7 protein and HA protein for the recoating reaction (inp). Lanes 5-7 was loaded with the bands observed after purification of the recoated particles on a CsCl gradient (bands 1 and 2 and top band). Lanes 8-14 are loaded with the same samples in the same order as lanes 1-7, but no reducing agent was added to the samples prior to SDS-PAGE. (B) Rotavirus DLPs recoated with the VP7-HA protein complex on a CsCl gradient. Bands 1 and 2 and top band correspond

to the samples loaded on lanes 5-7 in panel A. The positions where recoated particles and DLPs would typically migrate are indicated by dotted arrows.

Fig. 5: Cross-section through a three-dimensional reconstruction of rotavirus DLP recoated with a modified VP7 protein (A) displaying HA, (B) displaying HA with bound ScFv fragments of antibody CR6261, and (C) displaying HA with bound Fab fragments of antibody CR6261. CR6261 recognizes a highly conserved helical region in the membrane-proximal stem of HA1/HA2. The reconstruction is based on images acquired by performing cryo-EM on recoated DLPs.

Fig. 6: Superposition of a DLP (white) onto a three-dimensional reconstruction of a DLP recoated with modified rotavirusVP7 protein and influenza virus HA as the modified heterologous protein. The reconstruction is based on images acquired by performing cryo-EM on recoated DLPs.

Fig. 7: Detail of three-dimensional reconstruction of a rotavirus particle displaying influenza virus HA bound to Fab fragments of antibody CR6261. The rotavirus particles were prepared by recoating DLPs with a modified VP7 protein containing an adapter sequence (HR2). The HA protein was non-covalently bound to the modified VP7 protein via an C-terminally fused HR1 heptad repeat sequence which forms a six-helix bundle with the HR2 heptad repeat sequences of the VP7 protein. CR6261 binds at the membrane proximal end of each HA subunit (HA1 and HA2). The reconstruction is based on images acquired by performing cryo-EM on recoated DLPs.

MODES FOR CARRYING OUT THE INVENTION

Example 1: Construction of expression vector for modified recombinant VP7, HA and HIV gp140 proteins

The wild type VP7 gene of rhesus rotavirus (RRV) G3 strain was cloned into a pFastBacDual vector (Invitrogen™) between the BamH I and Not I restriction sites via standard procedures. A Kozak sequence (GCCACC ; SEQ ID NO: 19) was designed at the 5' end before the start codon of the VP7 coding sequence. Modifications of the VP7 coding sequence were carried out by inverted PCR. Primers of appropriate annealing temperatures were designed containing the relevant modifications at the ends of the primers. Figure 1A shows the various modifications that were added to the rotavirus coding sequence.

The final protein encoded by the modified VP7 coding sequence has the following features: The first 21 amino acids encode the honeybee melittin signal sequence for expression of the modified protein in insect cells. The VP7 protein signal peptide has been removed. Other signal sequences can be used as appropriate, such as the HIV consensus signal sequence or the signal peptide of human tissue plasminogen activator (htPA) for expression in human cells. (b) Amino acid residues 22 to 33 form an optional affinity tag (Strep-tag II plus a linker for more efficient signal peptide cleavage) for protein purification purposes and can be replaced by any other affinity tags, such as His-tag, HA-tag, FLAG-tag, *etc.* (c) Amino acid residues 34 to 40 form a TEV protease recognition sequence for cleavage of the affinity tag and can be replaced by the recognition sequences of any other proteases, such as PreScission™ protease (*i.e.* Rhinovirus 3C protease), factor Xa, enterokinase, thrombin, furin, *etc.* (d) Amino acid residues 41 to 136 are the sequences of rhesus rotavirus VP7 N-terminal portion (VP7 amino acid residues 51 to 146). (e) Amino acid residue 137 is a one-amino acid linker and could be replaced by another appropriate linker sequence. (f) Amino acid residues 138 to 165 are part of the C-heptad repeat of HIV gp41 HXB2 strain and could be replaced by a heptad repeat sequences from another retrovirus,

paramyxovirus, *etc.* as long as replacement sequence is compatible to the heptad repeat sequences used in the modified HA or gp140 protein constructs (see below). (g) Amino acid residues 166 to 187 are a factor Xa protease recognition sequence followed by a protein C-tag flanked by linkers; this modification is optional and could be replaced by other sequences, such as a designed epitope or simply a linker sequence such as GGSGGSGGSGGSGGS (SEQ ID NO: 20) or GGSGGSGGSGGSGGSGG (SEQ ID NO: 21). Presenting an epitope can be useful for packing antibody fragments recognising the epitope to further stabilise the displayed assemblies (including the heterologous protein) for the purpose of structural studies. Like I said, designing for structural studies is way more complex than designing for presenting antigens. (h) Amino acid residues 188 to 367 are the C-terminal portion of rhesus rotavirus VP7 (VP7 amino acid residues 147 to 326). The modified VP7 protein encoded by this modified coding sequence has the amino acid sequence of SEQ ID NO: 22.

Alternative constructs were prepared containing some of the modifications described in the preceding paragraph. For example, some of these constructs include an additional epitope tag which is recognised by anti-HIV antibody 2F5 (*e.g.* SEQ ID NO: 29 and 31). In other constructs, the C-heptad repeat of the HIV gp41 HXB2 strain was replaced with the C-heptad repeat of Nipah virus (*e.g.* SEQ ID NO: 32-41). The length of heptad repeat sequence was varied in some constructs (*e.g.* SEQ ID NO: 27 and 28). Similarly, the linker connecting the C-heptad repeat sequence and the remainder of the VP7 protein coding sequence was shortened in some of the constructs (see *e.g.* SEQ ID NO: 24-26) The different variants are summarised in Table 1. The modified VP7 protein sequence of SEQ ID NO: 22 is included as reference sequence. The sequences in Table 1 are shown from N-terminus to C-terminus.

Table 1

Description of construct	**Sequence**
signal (honeybee melittin)-linker1 (EDSA)-strep tag II-TEV-VP7 (51-146)-linker 2 (G)-HR2 (HIV, 28 a.a.)-linker 3 (Factor Xa-linker 3a (SGG)-Protein C tag-linker 3b (SGG))-VP7 (147-326)	SEQ ID NO: 22
signal (honeybee melittin)-linker1 (EDSA)-strep tag II-TEV-VP7 (51-146)-linker 2 (G)-HR2 (HIV, 28 a.a.)-linker 3 (Factor Xa-linker 3a (SGG)-Protein C tag-linker 3b (G))-VP7 (147-326)	SEQ ID NO: 23
signal (honeybee melittin)-linker1 (EDSA)-strep tag II-TEV-VP7 (51-146)-linker 2 (G)-HR2 (HIV, 28 a.a.)-linker 3 (linker 3a (SGG)-Protein C tag-linker 3b (SGG))-VP7 (147-326)	SEQ ID NO: 24
signal (honeybee melittin)-linker1 (EDSA)-strep tag II-TEV-VP7 (51-146)-linker 2 (G)-HR2 (HIV, 28 a.a.)-linker 3 (linker 3a (SGG)-Protein C tag-linker 3b (G))-VP7 (147-326)	SEQ ID NO: 25
signal (honeybee melittin)-linker1 (EDSA)-strep tag II-TEV-VP7 (51-146)-linker 2 (G)-HR2 (HIV, 28 a.a.)-linker 3 (linker 3a (G)-Protein C tag-linker 3b (G))-VP7 (147-326)	SEQ ID NO: 26

Description of construct	Sequence
signal (honeybee melittin)-linker1 (EDSA)-strep tag II-TEV-VP7 (51-146)-linker 2 (G)-HR2 (HIV, 42 a.a.)-linker 3 (linker 3a (SGG)-Protein C tag-linker 3b (SGG))-VP7 (147-326)	SEQ ID NO: 27
signal (honeybee melittin)-linker1 (EDSA)-strep tag II-TEV-VP7 (51-146)-linker 2 (G)-HR2 (HIV, 42 a.a.)-linker 3 (linker 3a (G)-Protein C tag-linker 3b (G))-VP7 (147-326)	SEQ ID NO: 28
signal (honeybee melittin)-linker1 (EDSA)-strep tag II-TEV-VP7 (51-146)-linker 2 (G)-HR2 (HIV, 28 a.a.)-linker 3 (linker 3a (G)-HIV antibody 2F5 epitope-linker 3b (SGG)-Protein C tag-linker 3c (SGG))-VP7 (147-326)	SEQ ID NO: 29
signal (honeybee melittin)-linker1 (EDSA)-strep tag II-TEV-VP7 (51-146)-linker 2 (G)-HR2 (HIV, 28 a.a.)-linker 3 (linker 3a (G)-HIV antibody 2F5 epitope-linker 3b (G)-Protein C tag-linker 3c (G))-VP7 (147-326)	SEQ ID NO: 30
signal (honeybee melittin)-linker1 (EDSA)-strep tag II-TEV-VP7 (51-146)-linker 2 (G)-HR2 (HIV, 28 a.a.)-linker 3 (linker 3a (G)-HIV antibody 2F5 epitope-linker 3b (GS)-Factor Xa-linker 3c (SGG)-Protein C tag-linker 3d (SGG))-VP7 (147-326)	SEQ ID NO: 31
signal (honeybee melittin)-linker1 (EDSA)-strep tag II-TEV-VP7 (51-146)-linker 2 (G)-HR2 (Nipah, 30 a.a.)-linker 3 (linker 3a (K)-Factor Xa-linker 3b (SGG)-Protein C tag-linker 3c (SGG))-VP7 (147-326)	SEQ ID NO: 32
signal (honeybee melittin)-linker1 (EDSA)-strep tag II-TEV-VP7 (51-146)-linker 2 (G)-HR2 (Nipah, 30 a.a.)-linker 3 (linker 3a (K)-Factor Xa-linker 3b (SGG)-Protein C tag-linker 3c (G))-VP7 (147-326)	SEQ ID NO: 33
signal (honeybee melittin)-linker1 (EDSA)-strep tag II-TEV-VP7 (51-146)-linker 2 (G)-HR2 (Nipah, 30 a.a.)-linker 3 (linker 3a (SGG)-Protein C tag-linker 3b (SGG))-VP7 (147-326)	SEQ ID NO: 34
signal (honeybee melittin)-linker1 (EDSA)-strep tag II-TEV-VP7 (51-146)-linker 2 (G)-HR2 (Nipah, 30 a.a.)-linker 3 (linker 3a (SGG)-Protein C tag-linker 3b (G))-VP7 (147-326)	SEQ ID NO: 35
signal (honeybee melittin)-linker1 (EDSA)-strep tag II-TEV-VP7 (51-146)-linker 2 (G)-HR2 (Nipah, 30 a.a.)-linker 3 (linker 3a (G)-Protein C tag-linker 3b (G))-VP7 (147-326)	SEQ ID NO: 36
signal (honeybee melittin)-linker1 (EDSA)-strep tag II-TEV-VP7 (51-146)-linker 2 (G)-HR2 (Nipah, 30 a.a.)-linker 3 (HIV antibody 2F5 epitope (14 a.a.)-linker 3a (SGG)-Protein C tag-linker 3b (SGG))-VP7 (147-326)	SEQ ID NO: 37

Description of construct	Sequence
signal (honeybee melittin)-linker1 (EDSA)-strep tag II-TEV-VP7 (51-146)-linker 2 (G)-HR2 (Nipah, 30 a.a.)-linker 3 (HIV antibody 2F5 epitope (14 a.a.)-linker 3a (G)-Protein C tag-linker 3b (G))-VP7 (147-326)	SEQ ID NO: 38
signal (honeybee melittin)-linker1 (EDSA)-strep tag II-TEV-VP7 (51-146)-linker 2 (G)-HR2 (Nipah, 30 a.a.)-linker 3 (linker 3a (G)-HIV antibody 2F5 epitope (9 a.a.)-linker 3b (SGG)-Protein C tag-linker 3c (SGG))-VP7 (147-326)	SEQ ID NO: 39
signal (honeybee melittin)-linker1 (EDSA)-strep tag II-TEV-VP7 (51-146)-linker 2 (G)-HR2 (Nipah, 30 a.a.)-linker 3 (linker 3a (G)-HIV antibody 2F5 epitope (9 a.a.)-linker 3b (G)-Protein C tag-linker 3c (G))-VP7 (147-326)	SEQ ID NO: 40
signal (honeybee melittin)-linker1 (EDSA)-strep tag II-TEV-VP7 (51-146)-linker 2 (G)-HR2 (Nipah, 30 a.a.)-linker 3 (linker 3a (G)-HIV antibody 2F5 epitope (9 a.a.)-linker 3b (GS)-Factor Xa-linker 3c (SGG)-Protein C tag-linker 3d (SGG))-VP7 (147-326)	SEQ ID NO: 41
signal (htPA)-linker1 (EDSA)-strep tag II-TEV-VP7 (51-146)-linker 2 (G)-HR2 (HIV, 28 a.a.)-linker 3 (Factor Xa-linker 3a (SGG)-Protein C tag-linker 3b (SGG))-VP7 (147-326)	SEQ ID NO: 42
signal (htPA)-linker1 (EDSA)-strep tag II-TEV-VP7 (51-146)-linker 2 (G)-HR2 (HIV, 28 a.a.)-linker 3 (linker 3a (G)-Protein C tag-linker 3b (G))-VP7 (147-326)	SEQ ID NO: 43
signal (htPA)-linker1 (EDSA)-strep tag II-TEV-VP7 (51-146)-linker 2 (G)-HR2 (Nipah, 30 a.a.)-linker 3 (linker 3a (K)-Factor Xa-linker 3b (SGG)-Protein C tag-linker 3c (SGG))-VP7 (147-326)	SEQ ID NO: 44
signal (htPA)-linker1 (EDSA)-strep tag II-TEV-VP7 (51-146)-linker 2 (G)-HR2 (Nipah, 30 a.a.)-linker 3 (linker 3a (G)-Protein C tag-linker 3b (G))-VP7 (147-326)	SEQ ID NO: 45

DNAs consisting of both the vector and the modified genes were generated by PCR. The PCR products were gel purified and subjected to T4 polynucleotide kinase treatment to generate phosphorylated ends. Blunt end ligation was then carried out for the resulting DNA by incubating the DNA with appropriate amounts of T4 ligase for 2 hours or overnight at room temperature. The ligation products were then treated with Dpn I for 30 minutes at 37°C before being used to transform DH5α cells via standard protocols.

Typically three to four colonies were picked and overnight cell cultures were grown to prepare plasmid DNA using the Miniprep Kit (Qiagen™). The plasmids were examined by agarose gel electrophoresis and the correct sequences were confirmed by DNA sequencing. Plasmids of correct sequences were used to transform DH10Bac competent cells via standard procedures. Two to three white colonies were selected for each construct and overnight cell cultures were grown for the extraction of recombinant

bacmid DNA by isopropanol/ethanol precipitation (Solution I: 15 mM Tris, pH 8.0, 10 mM EDTA, and 100 µg/mL RNase A; Solution II: 0.2 M NaOH and 1% SDA; and Solution III: 3 M potassium acetate, pH 5.5; all filter-sterilized). The purified bacmids were examined by PCR using the M13 primers and the correct DNAs were used to transfect monolayers of sf9 cells in 6-well plates, each well seeded with 1 million cells. P1 viruses were harvested 5 days post transfection and P2 viruses were produced by infecting sf9 cells (density of 1.5~2 million/mL) with 0.05~0.1% P1 viruses. P2 viruses were harvested 5~7 days post infection and were used for protein expression.

The modified VP7 protein constructs described above can be used to display trimer-forming heterologous proteins with correspondingly modified trimer-forming heterologous protein (see Figure 1B for a schematic overview of suitable modifications). As shown in Figure 2, the HR2 heptad repeat sequences of the modified VP7 protein and the HR1 heptad repeat sequences of the correspondingly modified trimer-forming heterologous protein form a stable complex via a six-helix bundle which surfaces as an adaptor for non-covalently mounting the heterologous protein on the rotavirus VP7 protein. Influenza A haemagglutinin (HA) protein and the gp140 fusion protein of HIV-1 were chosen as examples for trimer-forming heterologous proteins.

The HA gene of influenza A virus H1N1 Solomon Islands 2006 was cloned into a pFastBac LIC™ vector (Life Technologies™) by means of ligation independeng cloning (LIC) method. The pFastBac LIC™ vector was created by inserting a LIC site in a pFastBac1 vector (Invitrogen). A Kozak sequence (SEQ ID NO: 19) was designed at the 5' end of the start codon before the coding sequence. Modification in the coding sequence of the HA gene were introduced by inverted PCR.

The final protein encoded by the modified HA coding sequence has the following features: (a) The first 38 amino acids encode the Baculovirus gp64 signal peptide for expression of the modified protein in insect cells. The HA protein signal peptide has been removed. Other signal sequences can be used as appropriate, such as the HIV consensus signal sequence or the signal peptide of human tPA for expression in human cells. (b) Amino acid residues 42 to 47 form an optional affinity tag (His-tag plus a linker for more efficient signal peptide cleavage) for protein purification purposes and can be replaced by any other affinity tags, such as strep-tag, HA-tag, FLAG-tag, *etc.* (c) Amino acid residues 48 to 54 form a TEV protease recognition sequence for cleavage of the affinity tag. A linker sequence is also included. The TEV protease recognition sequence can be replaced by the recognition sequence of any other proteases, such as PreScission™ protease (*i.e.* Rhinovirus 3C protease, GE Healthcare™, Life Sciences), factor Xa, enterokinase, thrombin, furin, *etc.* (d) Amino acid residues 55 to 381 are the sequences of the HA1 portion of the hemagglutinin of influenza A Solomon Islands 2006 strain, which could be modified/mutated as appropriate or be replaced by that of any other influenza strains as shown below. (e) Amino acid residues 382 to 386 are the recognition sequence of enterokinase engineered between HA1 and HA2 for cleavage purposes. This site is optional and could be replaced by recognition sequences of Factor Xa, TEV protease, or others as appropriate. (f) Amino acid residues 387 to 565 are the sequence of HA2 portion of HA. (g) Amino acid residues 566 to 595 are a trimerization tag from bacteriophage T4 fibritin (Foldon), which is optional and can be replaced by any other trimerization tag or linker sequences. (h) Amino acid residues 596 to 629 are part of the N-heptad repeat of HIV gp41 HXB2 strain and could be replaced by other suitable heptad repeat sequences *e.g.* from any other retrovirus, paramyxovirus, *etc.* The numbering of the HA amino acid sequences used in the description of the modified HA coding

sequence is according to GenBank ID ABU50586.1. The amino acid sequence of this construct is shown in SEQ ID NO: 46.

Alternative constructs were prepared containing some of the modifications described in the preceding paragraph. For example, the trimerization domain was omitted from of the alternative constructs (*e.g.* SEQ ID NO: 48 and 49). In some constructs, the enterokinase recognition sequence was replaced by the Factor Xa recognition sequence (e.g. SEQ ID NO: 52). This construct also did not include a linker between the HA coding sequence and the C-heptad repeat sequence. In other constructs, both the trimerization domain and the enterokinase recognition sequence were omitted (SEQ ID NO: 51). The N-heptad repeat of Nipah virus was sometimes used in place of the N-heptad repeat of HIV gp41 HXB2 strain (see *e.g.* SEQ ID NO: 47 and 49). The length of heptad repeat sequence was varied in some constructs (*e.g.* SEQ ID NO: 54, 56 and 58). In other constructs, the length of the linker connecting the HA coding sequence to the N-heptad repeat sequence was varied (see *e.g.* SEQ ID NO: 56 and 58). Modified HA genes of H3 Wisconsin 2005 (see SEQ ID NOs: 60 and 61) and H5 Vietnam 2004 (see SEQ ID NOs: 62 and 63) were codon optimized and synthesized by GeneArt™ (Life Technologies™) before they were included in modified constructs. The numbering of the HA amino acid sequences used in the descriptions of these constructs is according to GenBank IDs AAT73274.1 and ACV49644.1, respectively. The structure of these modified HA coding sequences is summarised in Table 2. The modified HA protein sequence of SEQ ID NO: 46 is included as reference sequence. The sequences in Table 2 are shown from N-terminus to C-terminus.

Table 2

Description of construct	**Sequence**
signal (Baculovirus gp64 signal peptide)-linker 1 (ADP)-6xHis tag-TEV-linker 2 (GYLLE)-HA (18-339, i.e. HA1)-enterokinase (DDDDK)-HA (344-519, i.e. HA2)-linker 3 (RSL)-Foldon (T4 Fibritin C terminal bit)-HR1 (HIV, 34 a.a.)	SEQ ID NO: 46
signal (Baculovirus gp64 signal peptide)-linker 1 (ADP)-6xHis tag-TEV-linker 2 (GYLLE)-HA (18-339, i.e. HA1)-enterokinase (DDDDK)-HA (344-519, i.e. HA2)-linker 3 (RSL)-Foldon (T4 Fibritin C terminal bit)-HR1 (Nipah, 34 a.a.)	SEQ ID NO: 47
signal (Baculovirus gp64 signal peptide)-linker 1 (ADP)-6xHis tag-TEV-linker 2 (GYLLE)-HA (18-339, i.e. HA1)-enterokinase (DDDDK)-HA (344-519, i.e. HA2)-HR1 (HIV, 34 a.a.)	SEQ ID NO: 48
signal (Baculovirus gp64 signal peptide)-linker 1 (ADP)-6xHis tag-TEV-linker 2 (GYLLE)-HA (18-339, i.e. HA1)-enterokinase (DDDDK)-HA (344-519, i.e. HA2)-linker 3 (RS)-HR1 (Nipah, 34 a.a.)	SEQ ID NO: 49
signal (Baculovirus gp64 signal peptide)-linker 1 (ADP)-6xHis tag-TEV-linker 2 (GYLLE)-HA (18-519)-HR1 (HIV, 34 a.a.)	SEQ ID NO: 50

Date Reçue/Date Received 2020-10-01

Description of construct	Sequence
signal (Baculovirus gp64 signal peptide)-linker 1 (ADP)-6xHis tag-TEV-linker 2 (GYLLE)-HA (18-519)-linker 3 (RS)-HR1 (Nipah, 34 a.a.)	SEQ ID NO: 51
signal (Baculovirus gp64 signal peptide)-linker 1 (ADP)-6xHis tag-TEV-linker 2 (GYLLE)-HA (18-339, i.e. HA1)-Factor Xa (IEGR)-HA (344-519, i.e. HA2)-HR1 (HIV, 34 a.a.)	SEQ ID NO: 52
signal (Baculovirus gp64 signal peptide)-linker 1 (ADP)-6xHis tag-TEV-linker 2 (GYLLE)-HA (18-339, i.e. HA1)-Factor Xa (IEGR)-HA (344-519, i.e. HA2)-linker 3 (RS) HR1 (Nipah, 34 a.a.)	SEQ ID NO: 53
signal (Baculovirus gp64 signal peptide)-linker 1 (ADP)-6xHis tag-TEV-linker 2 (GYLLE)-HA (18-339, i.e. HA1)-Enterokinase (DDDDK)-HA (344-511, i.e. HA2)-linker 3 (IGE)-HR1 (HIV, 39 a.a.)	SEQ ID NO: 54
signal (Baculovirus gp64 signal peptide)-linker 1 (ADP)-6xHis tag-TEV-linker 2 (GYLLE)-HA (18-339, i.e. HA1)-Enterokinase (DDDDK)-HA (344-511, i.e. HA2)-linker 3 (IGEARQ)-HR1 (Nipah, 34 a.a.)	SEQ ID NO: 55
signal (Baculovirus gp64 signal peptide)-linker 1 (ADP)-6xHis tag-TEV-linker 2 (GYLLE)-HA (18-339, i.e. HA1)-Enterokinase (DDDDK)-HA (344-519, i.e. HA2)-linker 3 (RSI)-HR1 (HIV, 38 a.a.)	SEQ ID NO: 56
signal (Baculovirus gp64 signal peptide)-linker 1 (ADP)-6xHis tag-TEV-linker 2 (GYLLE)-HA (18-339, i.e. HA1)-Enterokinase (DDDDK)-HA (344-519, i.e. HA2)-linker 3 (RSIRQ)-HR1 (Nipah, 34 a.a.)	SEQ ID NO: 57
signal (Baculovirus gp64 signal peptide)-linker 1 (ADP)-6xHis tag-TEV-linker 2 (GYLLE)-HA (18-339, i.e. HA1)-Enterokinase (DDDDK)-HA (344-519, i.e. HA2)-linker 3 (RSIKKLIGE)-HR1 (HIV, 39 a.a.)	SEQ ID NO: 58
signal (Baculovirus gp64 signal peptide)-linker 1 (ADP)-6xHis tag-TEV-linker 2 (GYLLE)-HA (18-339, i.e. HA1)-Enterokinase (DDDDK)-HA (344-519, i.e. HA2)-linker 3 (RSIKKLIGEARQ)-HR1 (HIV, 34 a.a.)	SEQ ID NO: 59
signal (Baculovirus gp64 signal peptide)-linker 1 (ADP)-6xHis tag-TEV-linker 2 (GYLLE)-HA (17-341, i.e. HA1)-Enterokinase (DDDDK)-HA (347-522, i.e. HA2)-linker 3 (RSL)-HR1 (HIV, 38 a.a.)	SEQ ID NO: 60
signal (Baculovirus gp64 signal peptide)-linker 1 (ADP)-6xHis tag-TEV-linker 2 (GYLLE)-HA (17-341, i.e. HA1)-Enterokinase (DDDDK)-HA (347-522, i.e. HA2)-linker 3 (RSL)-Foldon (T4 fibritin C-terminal bit)-HR1 (HIV, 35 a.a.)	SEQ ID NO: 61

Description of construct	Sequence
signal (Baculovirus gp64 signal peptide)-HA (17-20)-6xHis tag-TEV-HA (21-345, i.e. HA1)-Enterokinase (DDDDK)-HA (346-521, i.e. HA2)-linker (RSL)-HR1 (HIV, 38 a.a.)	SEQ ID NO: 62
signal (Baculovirus gp64 signal peptide)-HA (17-20)-6xHis tag-TEV-HA (21-345, i.e. HA1)-Enterokinase (DDDDK)-HA (346-521, i.e. HA2)-linker (RSL)-Foldon (T4 fibritin C-terminal bit)-HR1 (HIV, 35 a.a.)	SEQ ID NO: 63

Each of the three modified HA genes were subcloned into pFastBacDual™ (Life Technologies™) between the Sal I and Not I sites. Recombinant baculoviruses bearing the modified HA genes were created following similar procedures as those described for VP7.

Modified gp140 genes of HIV-1 clade A 1992 Uganda 037.8 serotype and clade C 1997 were codon optimized and synthesized by GeneArt™ (Life Technologies™). The modified genes were subcloned into pFastBacDual™ (Life Technologies™) between the Sal I and Not I sites.

The final protein encoded by the modified gp140 coding sequence of the HIV-1 clade A strain has the following features: (a) The first 21 amino acids encode the honeybee melittin signal peptide for expression of the modified protein in insect cells. The gp140 signal peptide has been removed. Other signal sequences can be used as appropriate, such as the HIV consensus signal sequence or the signal peptide of human tPA for expression in mammalian/human cells. (b) Amino acid residues 22 to 29 form an optional affinity tag (His-tag plus a linker for more efficient signal peptide cleavage) for protein purification purposes and can be replaced by any other affinity tags, such as strep-tag, HA-tag, FLAG-tag, *etc.* (c) Amino acid residues 30 to 36 form a TEV protease recognition sequence for cleavage of the affinity tag and can be replaced by the recognition sequences of any other proteases, such as PreScission™ protease (*i.e.* Rhinovirus 3C protease, GE Healthcare™, Life Sciences), factor Xa, enterokinase, thrombin, furin, *etc.* (d) Amino acid residues 37 to 685 is the gp140 coding sequence of HIV strain 1992 Uganda 037.8 which could be modified/mutated as appropriate or be replaced by that of any other HIV/SIV strains. (e) Amino acid residues 686 to 694 are a linker and could be replaced by any appropriate other linker sequences. (f) Amino acid residues 695 to 721 are a trimerization tag from bacteriophage T4 fibritin (Foldon), which is optional or can be replaced by any other trimerization tag or linker sequences. (g) Amino acid residues 722 to 755 are part of the N-heptad repeat of HIV gp41 HXB2 strain and could be replaced by other suitable heptad repeat sequences *e.g.* from any other retrovirus, paramyxovirus, *etc.* The amino acid sequence of this construct is shown in SEQ ID NO: 64.

Alternative constructs were prepared containing some of the possible modifications indicated in the preceding paragraph. For example, some of these constructs do not include the trimerization tag, and the linker sequence between the gp140 coding region and the N-heptad repeat has been shortened (*e.g.* SEQ ID NO: 66 and 67). Other constructs are adapted for expression in mammalian cells by replacement of the signal peptide (*e.g.* SEQ ID NO: 65 and 67). In some constructs, the N-heptad repeat of the HIV gp41 HXB2 strain has been replaced with the N-heptad repeat of Nipah virus (SEQ ID NO: 68-71 and 76-79), in others, the gp140 coding sequence from HIV-1 clade A 1992 Uganda 037.8 serotype has been replaced

with the gp140 coding sequence from HIV-1 clade C 1997 (SEQ ID NO: 72-79). The numbering for the clade A and C gp140 is according to GenBank ID AAB05027.1 and AF286227.1, respectively.

The different variants are summarised in Table 3. The modified gp140 protein sequence of SEQ ID NO: 64 is also included as reference sequence. The sequences in Table 3 are shown from N-terminus to C-terminus.

Table 3

Description of construct	**Sequence**
signal (honeybee melittin signal peptide)-linker1 (ED)-6xHis tag-TEV-HIV clade A gp140 (27-675)-linker2 (linker 2a (SR)-Factor Xa-linker 2b (GSG))-Foldon (T4 Fibritin C-terminal bit)-HR1 (HIV, 34 a.a.)	SEQ ID NO: 64
signal (HIV consensus signal peptide)-linker1 (ED)-6xHis tag-TEV-HIV clade A gp140 (27-675)-linker2 (linker 2a (SR)-Factor Xa-linker 2b (GSG))-Foldon (T4 Fibritin C-terminal bit)-HR1 (HIV, 34 a.a.)	SEQ ID NO: 65
signal (honeybee melittin signal peptide)-linker1 (ED)-6xHis tag-TEV-HIV clade A gp140 (27-675)-linker2 (GSG)-HR1 (HIV, 34 a.a.)	SEQ ID NO: 66
signal (HIV consensus signal peptide)-linker1 (ED)-6xHis tag-TEV-HIV clade A gp140 (27-675)-linker2 (GSG)-HR1 (HIV, 34 a.a.)	SEQ ID NO: 67
signal (honeybee melittin signal peptide)-linker1 (ED)-6xHis tag-TEV-HIV clade A gp140 (27-675)-linker2 (linker 2a (SR)-Factor Xa-linker 2b (GSG))-Foldon (T4 Fibritin C-terminal bit)-HR1 (Nipah, 32 a.a.)	SEQ ID NO: 68
signal (HIV consensus signal peptide)-linker1 (ED)-6xHis tag-TEV-HIV clade A gp140 (27-675)-linker2 (linker 2a (SR)-Factor Xa-linker 2b (GSG))-Foldon (T4 Fibritin C-terminal bit)-HR1 (Nipah, 32 a.a.)	SEQ ID NO: 69
signal (honeybee melittin signal peptide)-linker1 (ED)-6xHis tag-TEV-HIV clade A gp140 (27-675)-linker2 (S)-HR1 (Nipah, 34 a.a.)	SEQ ID NO: 70
signal (HIV consensus signal peptide)-linker1 (ED)-6xHis tag-TEV-HIV clade A gp140 (27-675)-linker2 (S)-HR1 (Nipah, 34 a.a.)	SEQ ID NO: 71
signal (honeybee melittin signal peptide)-linker1 (ED)-6xHis tag-TEV-Linker2 (AENLWV)-HIV clade C gp140 (31-667)-linker3 (S)-Foldon (T4 Fibritin C-terminal bit)-HR1 (HIV, 34 a.a.)	SEQ ID NO: 72

Date Reçue/Date Received 2020-10-01

Description of construct	Sequence
signal (HIV consensus signal peptide)-linker1 (ED)-6xHis tag-TEV-Linker2 (AENLWV)-HIV clade C gp140 (31-667)-linker3 (S)-Foldon (T4 Fibritin C-terminal bit)-HR1 (HIV, 34 a.a.)	SEQ ID NO: 73
signal (honeybee melittin signal peptide)-linker1 (ED)-6xHis tag-TEV-HIV clade C gp140 (31-667)-linker2 (SGI)-HR1 (HIV, 34 a.a.)	SEQ ID NO: 74
signal (HIV consensus signal peptide)-linker1 (ED)-6xHis tag-TEV-HIV clade C gp140 (31-667)-linker2 (SGI)-HR1 (HIV, 34 a.a.)	SEQ ID NO: 75
signal (honeybee melittin signal peptide)-linker1 (ED)-6xHis tag-TEV-Linker2 (AENLWV)-HIV clade C gp140 (31-667)-linker3 (S)-Foldon (T4 Fibritin C-terminal bit)-HR1 (Nipah, 32 a.a.)	SEQ ID NO: 76
signal (HIV consensus signal peptide)-linker1 (ED)-6xHis tag-TEV-Linker2 (AENLWV)-HIV clade C gp140 (31-667)-linker3 (S)-Foldon (T4 Fibritin C-terminal bit)-HR1 (Nipah, 32 a.a.)	SEQ ID NO: 77
signal (honeybee melittin signal peptide)-linker1 (ED)-6xHis tag-TEV-Linker2 (AENLWV)-HIV clade C gp140 (31-667)-linker3 (S)-HR1 (Nipah, 34 a.a.)	SEQ ID NO: 78
signal (HIV consensus signal peptide)-linker1 (ED)-6xHis tag-TEV-linker2 (AENLWV)-HIV clade C gp140 (31-667)-linker3 (S)-HR1 (Nipah, 34 a.a.)	SEQ ID NO: 79

Example 2: Preparation of a stable VP7-HA complexes

To produce HA-VP7 complexes, Hi5 cells at the density of $2x10^6$/mL were infected with 0.5% VP7 baculovirus and 1.0% HA baculovirus in the presence of 5.0% heat inactivated FBS (Sigma-Aldrich). The medium was harvested 5 days post infection by spinning down the cells at 4000×g for 45 minutes. The supernatant was diafiltrated against 8 liters of 1x TNC (20 mM Tris, pH 8.0, 100 mM NaCl, and 1.0 mM $CaCl_2$; supplemented with 0.02% sodium azide) using a filter of 10 KDa cutoff in a Cogent M Tangential Flow Filtration System (Millipore). The buffer exchanged samples were supplemented with 1.0 mM PMSF and clarified by centrifugation for 1 hour at 10,000 RPM in a JA10 rotor (Beckman™). The protein complexes were purified from the supernatant by a StrepTactin® column (IBA™), followed by a NiNTA column (Qiagen™), and a Superose 6 column (Amersham™).

The purified HA-VP7 complexes were treated with 0.002% (w/w) enterokinase at 4°C for 48 hours. The cleavage of HA0 into HA1 and HA2 was examined by SDS-PAGE. Enterokinase was inactivated by adding 1x EDTA-free complete protease inhibitor tablet. The sample was supplemented with a redox buffer (10:1 molar ratio of GSH and GSSG) at a final concentration of 0.2~0.5 mM and treated with TEV protease (1/50~200, w/w) either at room temperature for 4 hours or at 4°C for overnight. Removal of the tags was also examined by SDS-PAGE.

Example 3: Preparation of a stable HIV gp140-VP7 complexes

The modified VP7 and gp140 genes encoding the proteins of SEQ ID NOs: 22, 26, 32 and 36 and SEQ ID NOs: 64, 66, 68 and 70, respectively were subcloned into pVRC8400 and pVRC-IRES-Puro vectors between Sal I and Not I for expression in mammalian cells. The signal peptide for the VP7 gene were changed to a human tPA signal peptide and the signal peptide for the gp140 gene were changed to the HIV consensus signal peptide sequence. The resulting modified VP7 proteins and modified gp140 proteins have the amino acid sequences of SEQ ID NOs: 42-45 and SEQ ID NOs: 65, 67, 69 and 71, respectively. Co-expression and complex formation were tested for the following combinations of modified VP7 protein and modified gp140 protein: SEQ ID NOs. 42 + 65, SEQ ID NOs. 42 + 67, SEQ ID NOs. 43 + 65, SEQ ID NOs. 43 + 67, SEQ ID NOs. 44 + 69, SEQ ID NOs. 44 + 71, SEQ ID NOs. 45 + 69, SEQ ID NOs. 45 + 71.

HIV gp140-VP7 complexes were expressed by co-transfecting 293T cells with a plasmid hosting the gp140-HR1 gene and another hosting the VP7-HR2 gene. Transient transfection was done by following standard protocols using polyethylenimine (PEI, 25KDa, linear or branched). For higher protein yield, stable cell lines were selected against puromycin. In brief, 70% confluent 293T cells were transfected with a mixture of Lipofectoamine® 2000 (Invitrogen®) and the pVRC-IRES-puro version of the two plasmids (2:1 mass ratio of lipofectoamine to total DNA) following standard procedures. At 20 hours post transfection, the medium was replaced with DMEM (Gibco®) supplemented with 10% Fetal Bovine Serum (Atlas), 1% GlutaMAX™ (Gibco®), 100 unites/mL penicillin/streptomycin (Gibco®), and 2µg/mL puromycin (Gibco®). The medium was kept fresh by changing it every three to four days until colonies form (visible to the naked eye). Individual colonies were picked and cultured in 24-well plates for small-scale expression tests. Colonies harbouring both genes were selected based on the detection of both proteins on western blots. The colony with the highest yield of the complex was scaled up and saved as stocks in liquid nitrogen. The cells were scaled up in puromycin-containing media and protein expression could be carried out using either adherent or suspension cell cultures in 293 Freestyle medium (Gibco®) supplemented with 5% heat inactivated FBS (Sigma-Aldrich®). The medium was harvested after 5 days by spinning down the cells at 4000×g for 45 minutes. Diafiltration and protein purification were carried out following the same procedures as described for the HA-VP7 complexes.

Example 4: Preparation of Fabs and Fvs

Monoclonal antibody CR6261 recognises a highly conserved helical region in the membrane-proximal stem of the influenza HA protein [i]. Structure factors and coordinates for the CR6261 sHgL Fab fragment were deposited with the Protein Data Bank under accession 4EVN and have been described previously.

The coding sequence for the antigen-binding fragment (Fab) and the single chain variable fragment (scFv) of monoclonal antibody CR6261 were cloned into the pVRC-IRES-Puro vector between Sal I and Not I following standard procedures as described above. The human tPA signal peptide was used for efficient secretion in these constructs in mammalian expression systems. The amino acid sequences of the final constructs are shown in SEQ ID NOs: 81-83. An scFv of monoclonal antibody CR6261 for expression in bacterial cells was also prepared (see SEQ ID NO: 80).

Date Reçue/Date Received 2020-10-01

Fabs were expressed by transient transfection of 293T cells in roller bottles. For each roller bottle, 125 µg each of the purified plasmid DNA hosting the heavy and light chain genes were mixed in 12.5 mL unsupplemented DMEM (Gibco®) and incubated in the hood for 15 minutes. PEI (500µg, linear or branced) was also diluted in 12.5 mL unsupplemented DMEM and incubated in the hood for 15 minutes before being added drop by drop and mixed well into the DNA mixture. The total of 25 mL mixture was incubated in the hood for an extra 15 minutes to allow a DNA-PEI complex to form before being added to cells at 50~70% confluence in roller bottles. The cells were incubated for over 5 hours with the DNA-PEI complex at 37°C, and the medium was replaced by 250 mL of 293 FreeStyle™ medium (Gibco®) supplemented with 100 units/mL penicillin/streptomycin. Medium was harvested 5 days post transfection, and Fabs were purified using a CaptureSelect® Lc Kappa or Lc Lambda affinity resin, followed by size exclusion chromatography on an S200 column (Amersham™).

ScFvs were expressed either in 293T cells as secreted proteins or in *E.coli* as inclusion bodies, from which protein was then extracted and refolded. The same transient transfection procedures as described above were followed for mammalian cell expression. For the *E.coli* expression, protein expression was induced for about 5 hours using 1 mM IPTG at cell density of 0.6~0.8 O.D.600nm. The cells were harvested, washed with 1x PBS, and lysed by sonication on ice. The inclusion bodies were extracted after removing the soluble fractions of the cells by centrifugation for 15 minutes at 20000×g. Inclusion bodies were dissolved in 100 mM Tris pH 7.5, 8 M urea, and 10 mM β-mercaptoethanol, clarified by centrifugation (20 minutes at 40000×g), and the supernatant was purified on a NiNTA column. The eluted sample, also in the denatured form, was diluted into the refolding buffer (100 mM Tris pH 7.5, 1 M arginine, 500 mM NaCl, 10% glycerol, and 1 mM EDTA) drop by drop at 4°C to a final protein concentration of lower than 0.1 mg/mL. The refolded sample was dialyzed against 20 mM Tris, pH 7.5, 100 mM NaCl four times, each time 20 volumes. The dialyzed sample was then clarified by centrifugation at 10,000 rpm in a JA10 rotor (Beckman™) and the supernatant was passed through a NiNTA column to concentrate the protein. The eluted fractions were then combined, concentrated, and further purified on an S200 size exclusion column (Amersham™).

Example 5: DLP preparation

Rhesus rotavirus serotype G3 was amplified by infecting MA-104 cells. Briefly, MA104 cells were grown in M199 medium (Gibco®) supplemented with 10% fetal bovine serum (HyClone™ Laboratories), 10 mM HEPES, pH 7.0, 2 mM L-glutamine, and 100 units/mL penicillin/streptomycin. The cells were maintained and scaled up to a 10-stack cell culture chamber (Corning™) or roller bottles. When the cells became confluent, the medium was replaced by serum-free MA199 supplemented with 1 µg/mL porcine pancreatic trypsin (Sigma-Aldrich®) for rotavirus inoculation and amplification. The medium of infected MA104 cells was harvested about 36~42 hours post infection and stored at -80°C for future use.

The frozen medium of infected MA104 cells was thawed at 4°C overnight. Cell debris was cleared by low-speed centrifugation at 3000 × g for about 10 minutes. The resulting supernatant was filtered through Whatman® Filter paper to remove residual cell debris before it was further passed through a 0.45 µm ExpressPlus filter unit (Millipore™) under vacuum. The virus particles were then pelleted at 45,000 rpm for 1 h at 4°C in a 45Ti rotor (Beckman™).

The pellet was resuspended in a total of 10 mL of 1x TNE buffer (20 mM Tris, pH 8.0, 100 mM NaCl, and 1 mM EDTA), briefly sonicated, and extracted twice with Freon 113 (Sigma-Aldrich®). The aqueous phase was recovered and concentrated into about 1 mL in an Amicon® centrifugal filtration device (100 KDa cutoff) at 3000×g for about 10 minutes. The concentrated sample was resuspended by pipetting up and down a few times and layered over a preformed CsCl gradient in 1x TNE (1.26 to 1.45 g/mL density as determined by refractometry). Samples were centrifuged at 55000 rpm in an SW 55Ti rotor (Beckman™) at 4°C for about 2 hours. The DLP band was collected and dialyzed overnight at 4°C against 1x TN buffer (20 mM Tris, pH 8.0, 100 mM NaCl) supplemented with 0.02% sodium azide.

The DLPs were negatively stained and visualised using electron microscopy (EM; Figure 3A).

Example 6: Recoating of DLPs with HA-VP7 complex

The DLPs obtained in Example 5 and the modified rotavirus VP7 protein prepared in Example 1 were mixed and incubated at 4°C for at least 1 hour in the presence of 5~10 mM $CaCl_2$. The resulting particles were negatively stained and visualised using EM. A single VP7 layer covering each particle was observed confirming that the modified VP7 protein was able to recoat DLPs (Figure 3B).

Recoating of DLPs with the HA-VP7 complex obtained in Example 2 was carried out at 4°C for at least 1 hour by mixing them at a molar ratio of 1:1.5 to 1:2 in the presence of 5~10 mM $CaCl_2$. The recoated rotavirus particles were negatively stained and visualised using EM. In addition to the single VP7 layer covering the DLP, a further layer surrounding the recoated particles was observed suggesting that the HA-VP7 complex correctly assembled on the surface of the DLPs (Figure 3C).

To confirm that the additional layer was formed by the HA protein, the recoated particles were loaded onto a preformed CsCl gradient in 1x TNC (density 1.25 g/mL to 1.45 g/mL) and centrifuged in a SW60 rotor (Beckman™) at 58,000 rpm for 2 hours. The resulting bands were collected, dialyzed against 1x TNC, and examined by SDS-PAGE. As a control, DLPs, the HA-VP7 complex and the mixture of both as used in the recoating reaction were also loaded on the gel in separate lanes. After separation, separate bands corresponding to the VP2 and VP6 proteins of the DLPs, the VP7 protein and the HA protein were observed at the correct stoichiometry in the sample that corresponded to the recoated particles (Figure 4). This confirmed that the further layer surrounding the recoated particles was composed of HA protein at full occupancy.

Example 7: Three-dimensional image reconstruction of HA-VP7-antibody complex

High-resolution cryo-EM has the potential to replace X-ray crystallography as the method of choice for analysing large molecular complexes, in particular in areas where rapid structural elucidation of a large number of molecular complexes is needed. One such area is influenza vaccine design, where antigenic drift leads to rapid changes in the sequence of the immunodominat surface antigen, the HA protein. Broadly protective monoclonal antibodies have been previously identified and many of these antibodies recognise conserved epitopes that are not subjected to rapid change. To investigate their suitability as protective agents against newly emerging influenza strains, we need to screen these broadly neutralising anti-HA antibodies against different variants of the HA protein.

The feasibility of such an approach using cryo-EM analysis was tested. DLPs were recoated with modified VP7 protein alone or with a HA-VP7 complex as described in Example 6. HA-VP7-antibody

recoated particles were obtained by mixing the purified HA-VP7 complexes with the Fabs or scFvs of monoclonal antibody CR6261 prepared in Example 4 at a molar ratio of 1:1.5~2.0. The mixture was incubated at 4°C for 30 minutes. The complexes were purified by loading the mixture after the end of the incubation period on a Superose 6 column. The purified complexes were then used to recoat DLPs following the same procedures as described in Example 6. Alternatively, the antibody fragments were added directly to particles recoated with HA-VP7 complexes at a molar ratio of 1:1.2~1.5 either before or after the CsCl gradient purification step.

Cryo-grids were prepared with a Vitrobot™ Automated plunger (FEI). Quantifoil Holey carbon grids were glow discharged and left at room temperature overnight before use. For each grid, 4 µL sample at a concentration equivalent to ~5 mg/mL rotavirus DLPs was applied to one side surface of the grid. During plunging, the chamber moisture was maintained close to 100% and the temperature at around 22°C. The grid was then blotted for 4 seconds from both sides with filter paper, immediately followed by plunging into liquid ethane for vitrification. The grids were then stored in liquid nitrogen before being used for data collection.

Data were collected on a Tecnai™ F30 electorn microscrope (FEI) operated at 300 kV. The optical system was aligned using standard procedures (beam shift, beam tilt, eucentric height, pivot points, rotational center, astigmatism, *etc.*). The image acquisition procedures involved finding the desired imaging area, adjusting defocus (between 0.6 and 3.0 µm), testing grid drift rate (less than 2 Å per second), and the final image exposure. These procedures were semi-automatically achieved using the program SerialEM (Cryo-electron Microscopy Facility, University of Colorado - Boulder). The data were recorded as movies on a K2 Summit™ direct detection camera (Gatan™) using super resolution mode (pixel size: 0.99 Å). The "movie" protocol of reference 24 was used. The dose rate on the sample was ~3 electrons/Å^2 per second. Each frame recorded 0.5 second exposure and the final movie consisted of 24 frames with an accumulated dose of ~36 electrons/Å2.

The frames of each movie were aligned using IMOD (Cryo-electron Microscopy Facility, University of Colorado - Boulder) scripts based on image cross correlation. The aligned frames were simply averaged for initial image processing. As the resolution improved later, however, the averaged images from the first 13 frames (or the best series of frames) of each movie, which would correspond to ~20 electron dose, were used. Particle images were picked using e2boxer.py of the EMAN2 image processing suite [69]. Images with obvious defects, aberrations, abnormal focus, contamination, overlap, or large sample drifts were manually excluded after visual inspection. The particle coordinates were used to excise image stacks of individual particles with 1600 × 1600 pixel dimensions using proc2d of the EMAN2 image processing suite [69]. Defocus values were determined using the program CTFFIND3 [23].

The structure refinement and reconstructions were carried out using the program FREALIGN [40]. The initial orientation search was performed using 4x binned data and a previously calculated map of the VP7 recoated particles (7RP) (see reference 1) as a 3-dimensional reference. The initial alignment parameters of the excised particles, determined by a systematic search (mode 3) at a 1° angular interval, were further refined against the latest reconstructions (mode 1) until there was no further improvement in resolution. During alignment, a radial shell mask between 220 and 400 Å was applied to retain the density corresponding to the rotavirus portion and to exclude density corresponding to the RNA and HA spikes.

The images were also low pass filtered (up to 15 Å) to avoid possible overfitting in the alignment process. The alignment parameters were then used to calculate the reconstructions of the entire particle (within the radial shell of 600 Å or between the radii of 220 and 600 Å for the protein contents).

Subsequently, 2x binned and later unbinned images were used to further refine the alignment parameters following similar strategies to those used for the 4x binned data, *i.e.* an initial systematic search (mode 3) followed by multiple cycles of angular and positional refinement (mode 1) until no further improvement in resolution was observed. During each stage of refinement, the shell mask between 220 and 400 Å was applied and the data were restricted to the most reliable resolution (up to 10 Å for 2x binned data and up to 8 Å for unbinned data). Reconstructions were calculated for the entire particle within the radial shell of 600 Å.

Figure 5A shows a cross-section through a three-dimensional reconstruction of rotavirus DLP recoated with modified VP7 protein only. Figure 5B shows a cross-section through a three-dimensional reconstruction of rotavirus DLP recoated with an HA-VP7 complex as described in Example 6. Figure 6 shows a superposition of a DLP onto the three-dimensional reconstruction of the recoated particle shown in Figure 5B. Figure 5C shows a cross-section through a three-dimensional reconstruction of rotavirus DLP recoated with modified VP7 protein displaying HA with bound ScFv fragments of antibody CR6261.

To assess the reliability of refinement, different sets of refinement were carried out by masking out different regions of the particle, and the alignment parameters were used to calculate reconstructions for the omitted regions that were not included during refinement. For example, alignment parameters obtained by refining the HA portion (by applying a shell mask between 425 and 600 Å) were used to calculate the reconstruction for the rotavirus portion (by applying a shell mask between 220 and 400 Å). The resulting map contained densities clearly representing VP2, VP6, and VP7, suggesting minimum model bias. The T=13 quasi-equivalent arrangement of the rotavirus VP6 and VP7 layers clearly extended to the adaptor and HA portion of the particle. Local 13-fold averaging was not applied but can be performed to further improve the resolution. Rigid body fitting of the rotavirus VP2, VP6, and VP7, the six-helix bundle, the HA, and the Fabs into the electron density maps was performed using the UCSF chimera [47].

An example of the level of detail that can be achieve by further refining the reconstructions obtained from cryo-EM as described in the preceding paragraph is shown in Figure 7. A detail of the three-dimensional reconstruction of a rotavirus particle displaying influenza virus HA bound to Fab fragments of antibody CR6261 is shown. Some secondary structures can be visualised in the calculated map. As expected, the variable domain of CR6261 is seen to interact with both HA1 and HA2.

Example 8: Immunisation studies with HA-VP7 protein complexes mounted on rotavirus DLPs

Female 6-8-week-old BALB/c mice are immunised intramuscularly (IM) three times at three-week intervals with one of three different formulations, each formulated in 100 µl of a calcium-containing buffer. Group 1 receives of formulated HA-VP7 protein complexes mounted on rotavirus DLPs, group 2 receives purified HA, and group 3 receives purified HA mixed with VP7-recoated DLPs. A fourth group of mice is mock-injected with 100 µl buffer. Experiments can include various doses for each of the groups. Equivalent amounts of antigen can be used to compare immune response to equivalent doses of

antigen in various formats. Alternatively, equivalent amounts of total protein can be used to compare immune responses to various amounts of total protein.

One or more groups can include an adjuvant. Adjuvants can be particularly useful in stimulating an immune response in the mice in group 2 receiving purified HA antigen.

Each vaccine group contains eight mice, and the "buffer alone" control group contains five mice. Blood is drawn at days -1, 20, 41, 56 and 84 from administration of the first injection.

Mice are monitored after immunisation for body weight change, injection site reactions, and other clinical observations. The antibody responses to the administered antigens are analysed using methods known in the art to analyse antibody specificity (*e.g.* ELISA assay and competition ELISA assay against HA and rotavirus proteins including VP7 protein) and neutralization activity (*e.g.* influenza haemagglutination inhibition assay, see *e.g.* references 70 and 71).

Results demonstrate a substantially higher immune response to the HA-VP7 protein complexes mounted on rotavirus DLPs as compared to the purified HA, including purified HA formulated with adjuvant. Competition ELISA demonstrate that antibody responses to the HA-VP7 protein complexes are predominantly to the HA portion of the complex rather than the VP7 or rotavirus portion of the complex. Results demonstrate substantially higher neutralization antibody titres as a result of immunization with the HA-VP7 protein complexes mounted on rotavirus DLPs as compared to the purified HA, including purified HA formulated with adjuvant. These results demonstrate the increased immunogenicity of flu antigens when mounted on rotavirus VP7-coated DLP as compared to flu antigens not mounted on VP7-coated DLPs.

All animal experiments are performed in accordance with Institutional Animal Care and Use Committee protocols.

Example 9: Immunisation studies with HA-VP7 protein complexes mounted on rotavirus DLPs with and without adjuvant

Female BALB/c mice are immunized by two bilateral 50 µl intramuscular injections in the rear quadriceps on days 0, 21, and 42 with 0.3-15 µg of purified antigen H1N1 hemagglutinin (HA) protein complexes as set out in Table 4 below.

The complexes are formulated with PBS or adjuvant 1-2 hours prior to immunization. The formulated subunit vaccines are kept on ice until administration.

Mice are monitored after immunisation for body weight change, injection site reactions, and other clinical observations. Serum samples are obtained from the animals by retro-orbital sinus bleeds on day 20 (3 weeks post first immunization) and day 41 (3 weeks post second immunization) and from bleed-outs of euthanized animals on day 63 (3 weeks post third immunization).

The antibody responses to the administered antigens are analysed using methods known in the art (*e.g.* ELISA assay, neutralization assay). Results demonstrate a substantially higher immune response to the HA-VP7 protein complexes mounted on rotavirus DLPs as compared to the purified HA, including

purified HA formulated with adjuvant or purified HA formulated with, but not mounted on VP7-coated DLPs.

All animal experiments are performed in accordance with Institutional Animal Care and Use Committee protocols.

Table 4

Group	**IM Vaccine (dose)**	**N**
1	A/Sol Monobulk (0.3μg)	5
2	7RP (0.3 μg)	5
3	7RP (3 μg)	5
4	7RP (15 μg)	5
5	HA + 7RP (0.3 μg)	5
6	HA + 7RP (3 μg)	5
7	HA + 7RP (15 μg)	5
8	HA-7RP (0.3 μg)	5
9	HA-7RP (3 μg)	5
10	HA-7RP (15 μg)	5
11	A/Sol HA (0.3 μg)	5
12	A/Sol HA (3 μg)	5
13	A/Sol HA (15 μg)	5

A/Sol monobulk is lipid-oligomerized trimerized, recombinant HA.
7RP is VP7 recoated particles without HA.
HA + 7RP is VP7 recoated particles on which HA cannot be mounted + trimerized recombinant HA.
HA-7RP is VP7 recoated particles mounted with trimerized, recombinant HA.
A/Sol HA is trimerized recombinant HA.

It will be understood that the invention has been described by way of example only and modifications may be made whilst remaining within the scope and spirit of the invention.

Date Reçue/Date Received 2020-10-01

REFERENCES

[1] Chen *et al.* (2009) *Proc Natl Acad Sci U S A* 106(26):10644-8
[2] Trask & Dormitzer (2006) *J Virol* 80(22):11293-304
[3] Kobayashi *et al.* (2007) *Cell Host Microbe* 1(2):147-157
[4] Chandran *et al.* (1999) *J Virol* 73(5):3941-50
[5] Chandran *et al.* (2001) *J Virol* 75(11):5335-42
[6] Yan *et al.* (2011) *J Virol* 85(15):7483-95
[7] Miyazaki *et al.* (2008) *J Virol* 82(22):11344-53
[8] Zhang *et al.* (2010) *Cell* 141(3):472-82
[9] Biosafety in Microbiological and Biomedical Laboratories (BMBL) 5th Edition, CDC, 2009
[10] Yang *et al.* (2000) *J Virol* 74(10):4746-54
[11] Meier *et al.* (2004) *J Mol Biol* 344(4):1051-69
[12] Wu & Wong (2005) *J Biol Chem* 280(24):23225-31
[13] Beckett et al. (1999) *Protein Sci* 8(4):921-9
[14] Chen *et al.* (1997) *Cell* 89(2):263-73
[15] Zhu *et al.* (2003) *Protein Eng* 16(5):373-9
[16] Yu *et al.* (2002) *J Gen Virol* 83(Pt 3):623-9
[17] Xu *et al.* (2004) *J Biol Chem* 279(29):30514-22
[18] http://partsregistry.org/Protein_domains/Linker
[19] US 4,769,326
[20] Dormitzer et al. (2001) *J Virol* 75(16):7339-50
[21] Dormitzer *et al.* (2000) *Virology* 277(2):420-8
[22] Mohd Jaafar *et al.* (2005) *J Gen Virol* 86(Pt 4):1147-57
[23] Mindell & Grigorieff (2003) *J Struct Biol* 142(3):334-47
[24] Campbell *et al.* (2012) *Structure* 20(11):1823-8
[25] Settembre *et al.* (2011) *EMBO J.* 30(2):408-16
[26] Chen & Grigorieff (2007) *J Struct Biol* 157(1):168-73
[27] Wriggers *et al.* (1999) *J Struct Biol* 125(2-3):185-95
[28] Jiang *et al.* (2001) *J Mol Biol* 308(5):1033-44
[29] Topf *et al.* (2005) *J Struct Biol* 149(2):191-203
[30] Tama *et al.* (2004) *J Struct Biol* 147(3):315-2
[31] Topf *et al.* (2008) *Structure* 16(2):295-307
[32] Trabuco *et al.* (2009) *Methods* 49(2):174-80
[33] Schröder *et al.* (2007) *Structure* 15(12):1630-41
[34] Zhang *et al.* (2010) *Nature* 463(7279):379-83
[35] Baker *et al.* (2006) *PLoS Comput Biol* 2(10):e146
[36] DiMaio *et al.* (2009) *J Mol Biol* 392(1):181-90
[37] Topf *et al.* (2006) *J Mol Biol* 357(5):1655-68
[38] Zhu *et al.* (2010) *J Mol Biol* 397(3):835-51
[39] Liang et al. (2002) *J Struct Biol* 137(3):292-304
[40] Grigorieff (2007) *J Struct Biol* 157(1):117-25
[41] Cheng *et al.* (2010) *J Mol Biol* 397(3):852-63
[42] Jiang *et al.* (2008) *Nature* 451(7182):1130-4
[43] Ludtke *et al.* (2008) *Structure* 16(3):441-8
[44] Yu *et al.* (2008) *Nature* 453(7193):415-9
[45] Ludtke *et al.* (1999) *J Struct Biol* 128(1):82-97
[46] Baker *et al.* (2007) *Structure* 15(1):7-19
[47] Pettersen *et al.* (2004) *J Comput Chem* 25(13):1605-12
[48] Emsley & Cowtan (2004) *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126-32
[49] Emsley *et al.* (2010) *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486-501

[50] Baker *et al.* (2011) *J Struct Biol* 174(2):360-73
[51] Zhu et al. (2010) *J Mol Biol* 397(3):835-51
[52] Jones *et al.* (1991) *Acta Crystallogr A* 47 (Pt 2):110-9
[53] Brünger *et al.* (1998) *Acta Crystallogr D Biol Crystallogr* 54(Pt 5):905-21
[54] Li & Zhang (2009) *Proteins* 76(3):665-76
[55] Brünger (1992) *Nature* 355(6359):472-5
[56] Laskowski *et al.* (1993) *J Appl Cryst* 26:283-91
[57] Whittle *et al.* (2011) *Proc Natl Acad Sci U S A* 108(34):14216-21
[58] Aoki *et al.* (2009) *Science* 324(5933):1444-7
[59] Krause *et al.* (2011) *J Virol* 85(20):10905-8
[60] Bommakanti *et al.* (2010) *Proc Natl Acad Sci U S A* 107(31):13701-6
[61] Chakraborty *et al.* (2006) *Biochem J* 399(3):483-91
[62] Bhattacharyya *et al.* (2010) *J Biol Chem* 285(35):27100-10
[63] Ward *et al.* (1989) *Nature* 341:544-546
[64] Bird *et al.* (1988) *Science* 242:423-426
[65] Huston *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85:5879-5883
[66] Council of Europe (2014) *European pharmacopoeia*
[67] Ekiert *et al.* (2009) *Science* 324(5924): 246–251
[68] Lingwood *et al.* (2012) *Nature* 489(7417):566-70
[69] Tang *et al.* (2007) *J Struct Biol* 157(1):38-46
[70] Salk *et al.* (1945) *Am J Hyg* 42:57–93
[71] de Jong *et al.*, Haemagglutination-inhibiting antibody to influenza virus. In Brown *et al.* (eds.) (2003) *Laboratory Correlates of Immunity to Influenza- A Reassessment*. Basel, Switzerland: Karger, 63–73

CLAIMS

1. A chimeric protein complex comprising a trimer-forming rotavirus VP7 surface protein linked to a heterologous protein, wherein the rotavirus VP7 surface protein is linked to the heterologous protein non-covalently by a two-part adapter system, where one part of the adapter system is formed by a first adapter polypeptide which is fused to the rotavirus VP7 surface protein, and the other part of the adapter is formed by a second adapter polypeptide which is fused to the heterologous protein, wherein both parts of the adapter system form a stable complex with each other and the chimeric protein complex is for integration into the outer layer of a rotavirus particle by *in vitro* recoating double-layered rotavirus particles with said chimeric protein complex.

2. The chimeric protein complex of claim 1, wherein the first adapter polypeptide is fused to the rotavirus VP7 surface protein via a linker sequence.

3. The chimeric protein complex of claim 1, wherein the second adapter polypeptide is fused to the heterologous protein via a linker sequence.

4. The chimeric protein complex of any one of claims 1 to 3, wherein the first adapter polypeptide and the second adapter polypeptide comprise a heptad repeat sequence.

5. A rotavirus particle comprising the chimeric protein complex of claim 1 or 2.

6. A composition comprising the rotavirus particle as defined in claim 5, together with an acceptable excipient, diluent, or carrier.

7. A use of the rotavirus particle of claim 5 an immunogen.

8. A use of the rotavirus particle of claim 5 for raising an immune response in a subject.

Date Reçue/Date Received 2020-10-01

9. A use of the rotavirus particle of claim 5 for preparation of a medicament for raising an immune response in a subject.

10. A nucleic acid composition comprising:

(a) a first open reading frame encoding a modified rotavirus surface protein comprising a rotavirus VP7 surface protein, a first adapter polypeptide; and

(b) a second open reading frame encoding a fusion protein comprising a heterologous protein, a second adapter polypeptide;

optionally wherein the open reading frame is operationally linked to a promoter sequence.wherein the first adapter polypeptide and the second adapter polypeptide are for forming a stable complex with each other

11. The nucleic acid composition according to claim 10, wherein:

- the first open reading frame additionally comprises a linker sequence,
- the second open reading frame additional comprises a linker sequence, or
- the first open reading frame additionally comprises a first linker sequence and the second open reading frame additional comprises a second linker sequence.

12. The nucleic acid composition of claim 10 or 11, wherein the first open reading frame is operationally linked to a promoter sequence for expressing said first open reading frame.

13. The nucleic acid composition of any one of claims 10 to 12, wherein the second open reading frame is operationally linked to a promoter sequence for expressing said second open reading frame.

14. The nucleic acid composition of any one of claims 10 to 13, wherein:

- the first adapter polypeptide comprises a heptad repeat sequence,
- the second adapter polypeptide comprises a heptad repeat sequence, or
- . the first adapter polypeptide comprises a first heptad repeat sequence and the second adapter polypeptide comprises a second heptad repeat sequence.

Date Reçue/Date Received 2020-10-01

15. A kit comprising:

(a)(i) a first nucleic acid encoding a modified rotavirus surface protein comprising a rotavirus VP7 surface protein and a first adapter polypeptide, and (ii) a second nucleic acid comprising a nucleotide sequence encoding a second adapter polypeptide and a multiple cloning site, and wherein insertion of a coding region for a heterologous protein in the multiple cloning site yields an open reading frame encoding a fusion protein comprising the heterologous protein and the second adapter polypeptide; or

(b)(i) a first nucleic acid encoding a modified rotavirus surface protein comprising a rotavirus surface protein and a first adapter polypeptide and (ii) a second nucleic acid encoding a fusion protein comprising a heterologous protein and a second adapter polypeptide;

wherein the first adapter polypeptide and the second adapter polypeptide are able to form a stable complex.

16. A kit according to claim 15, wherein the kit further comprises a rotavirus particle, wherein the particle is derived from the same species of rotavirus as the rotavirus surface protein or from a different rotavirus species.

17. A method for preparing the rotavirus particle of claim 5, wherein the method comprises propagating a rotavirus particle comprising an outer layer in a cell grown in a culture medium, purifying the rotavirus particle from the culture medium, removing the outer layer from the rotavirus particle to obtain a rotavirus double-layered particle (DLP), and recoating the rotavirus DLP with the chimeric protein complex of any one of claims 1 to 4 to yield the rotavirus particle of claim 5.

18. A method for preparing the rotavirus particle of claim 5, wherein the method comprises propagating a rotavirus particle comprising an outer layer in a cell grown in a culture medium, purifying the rotavirus particle from the culture medium, removing the outer layer from the rotavirus particle to obtain a rotavirus DLP, and recoating the rotavirus DLP with a first fusion protein comprising a trimer-forming rotavirus VP7 surface protein, a first adaptor polypeptide

Date Reçue/Date Received 2020-10-01

comprising a heptad repeat sequence, and mixing the recoated rotavirus DLP with a second fusion protein comprising a trimer-forming heterologous protein, a second adaptor polypeptide comprising a heptad repeat sequence, to yield the rotavirus particle of claim 5.

19. A method for preparing the rotavirus particle of claim 5, comprising mixing a rotavirus particle comprising a first fusion protein comprising a trimer-forming rotavirus VP7 surface protein, a first adaptor polypeptide comprising a heptad repeat sequence, with a second fusion protein comprising a trimer-forming heterologous protein, a second adaptor polypeptide comprising a heptad repeat sequence.

20. The method of claim 18 or 19, wherein:

- the first fusion protein additionally comprises a linker sequence,
- the second fusion protein additionally comprises a linker sequence, or
- the first fusion protein additionally comprises a first linker sequence and the second fusion protein additionally comprises a second linker sequence.

21. A method for determining the structure of a heterologous protein, wherein the method comprises the steps of (i) recoating a rotavirus DLP with a chimeric protein complex according to any one of claims 1 to 4 to yield a suspension of rotavirus particles displaying the chimeric surface protein, (ii) freezing the suspension, (iii) imaging the rotavirus particles using cryo-EM to obtain a plurality of micrographs, and (iv) analysing the plurality of micrographs to obtain a three-dimensional model of the chimeric surface protein.

22. A method for determining the structure of a heterologous protein in complex with a molecule, wherein the method comprises the steps of (i) recoating a rotavirus DLP with a chimeric protein complex according to any one of claims 1 to 4 to yield a suspension of rotavirus particles displaying the chimeric protein complex, (ii) adding to the suspension a molecule that specifically binds to the heterologous protein, wherein the molecule forms a complex with the chimeric protein complex, (iii) freezing the suspension, (iv) imaging the rotavirus particles using cryo-EM to obtain a plurality of micrographs, and (vi) analysing the plurality of

Date Reçue/Date Received 2020-10-01

micrographs to obtain a three-dimensional model of the chimeric protein complex complexed to the molecule.

23. The method of claim 22, wherein the molecule is (a) an antibody or fragment thereof, wherein the antibody or fragment specifically binds the heterologous protein; or (b) a cell surface receptor, wherein the heterologous protein is a viral cell entry protein and the proteinaceous molecule is bound by the viral cell entry protein.

24. A vaccine comprising the rotavirus particle as defined in claim 5, together with a pharmaceutically acceptable excipient, diluent, or carrier.

Date Reçue/Date Received 2020-10-01

FIG. 1

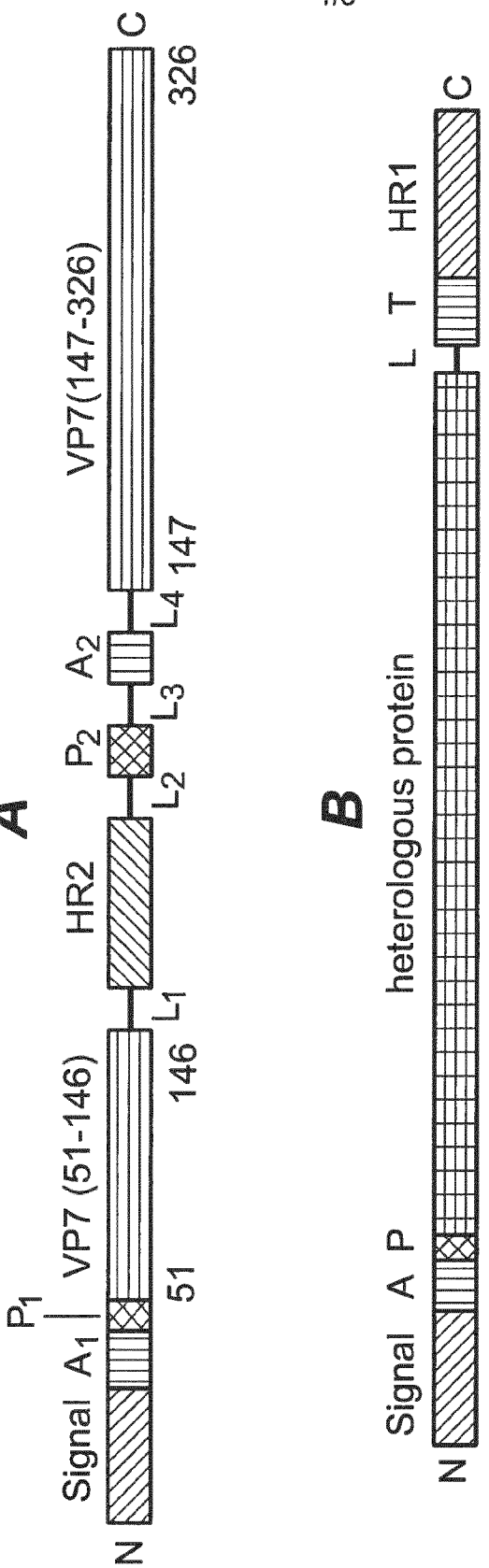

A
Signal
A1
P1
VP7 (51-146)
51
146
L1
HR2
L2
P2
L3
A2
L4
147
VP7(147-326)
326
N
C
B
Signal
A
P
heterologous protein
L
T
HR1
N
C
Signal = heterologous signal peptide sequence
A = affinity tag (optional)
P = protease recognition site (optional)
L = linker
T = trimerization tag (optional)
HR1, HR2 = heptad repeat sequences

FIG. 2

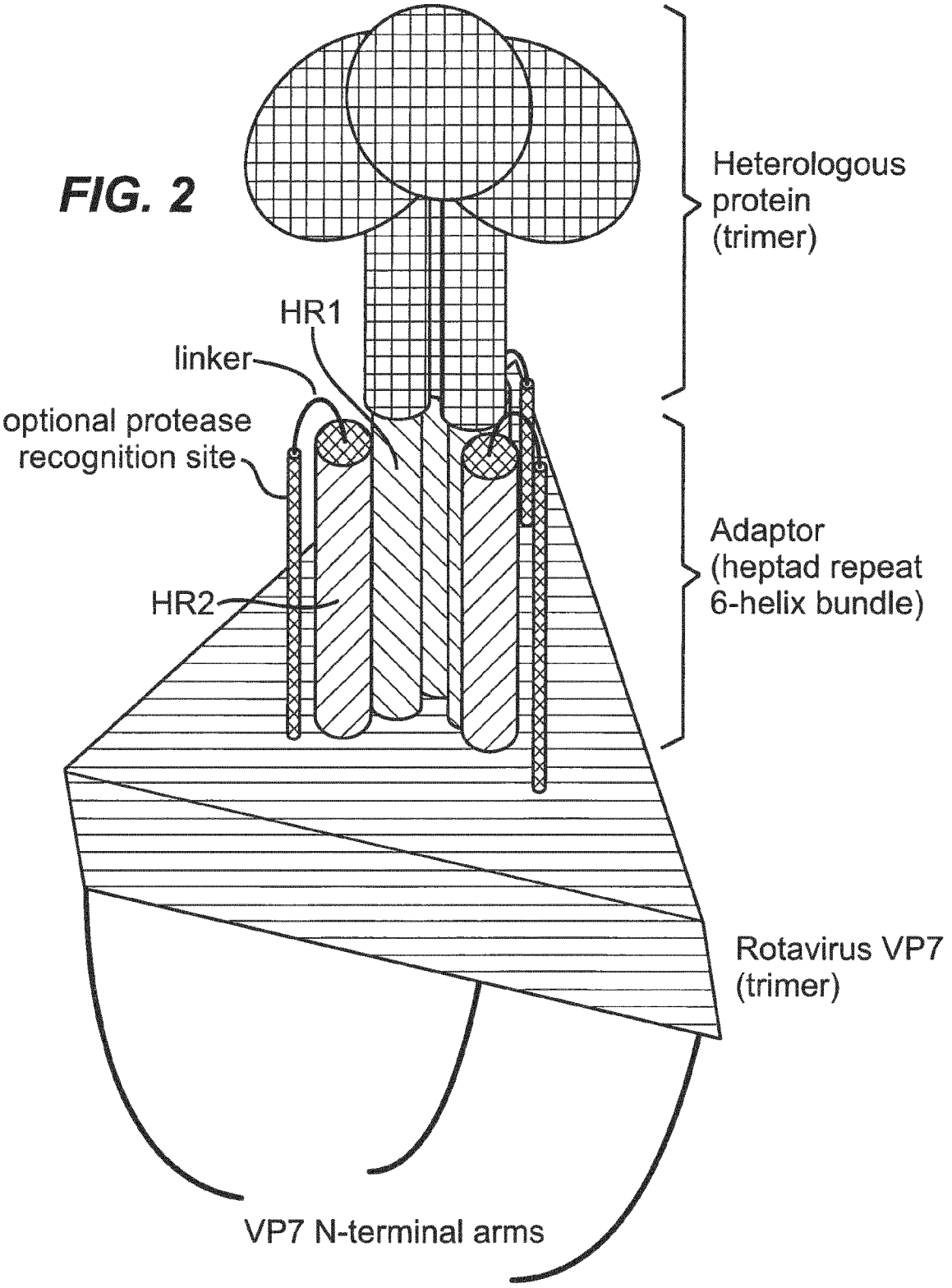
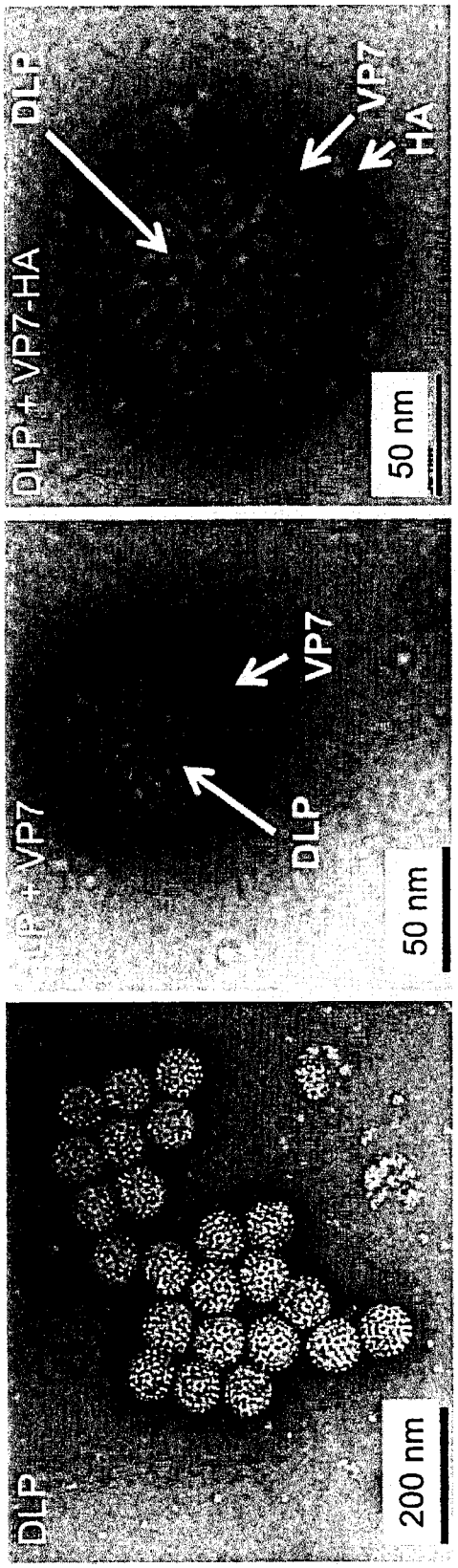

Heterologous protein (trimer)
HR1
linker
optional protease recognition site
HR2
Adaptor (heptad repeat 6-helix bundle)
Rotavirus VP7 (trimer)
VP7 N-terminal arms

FIG. 3



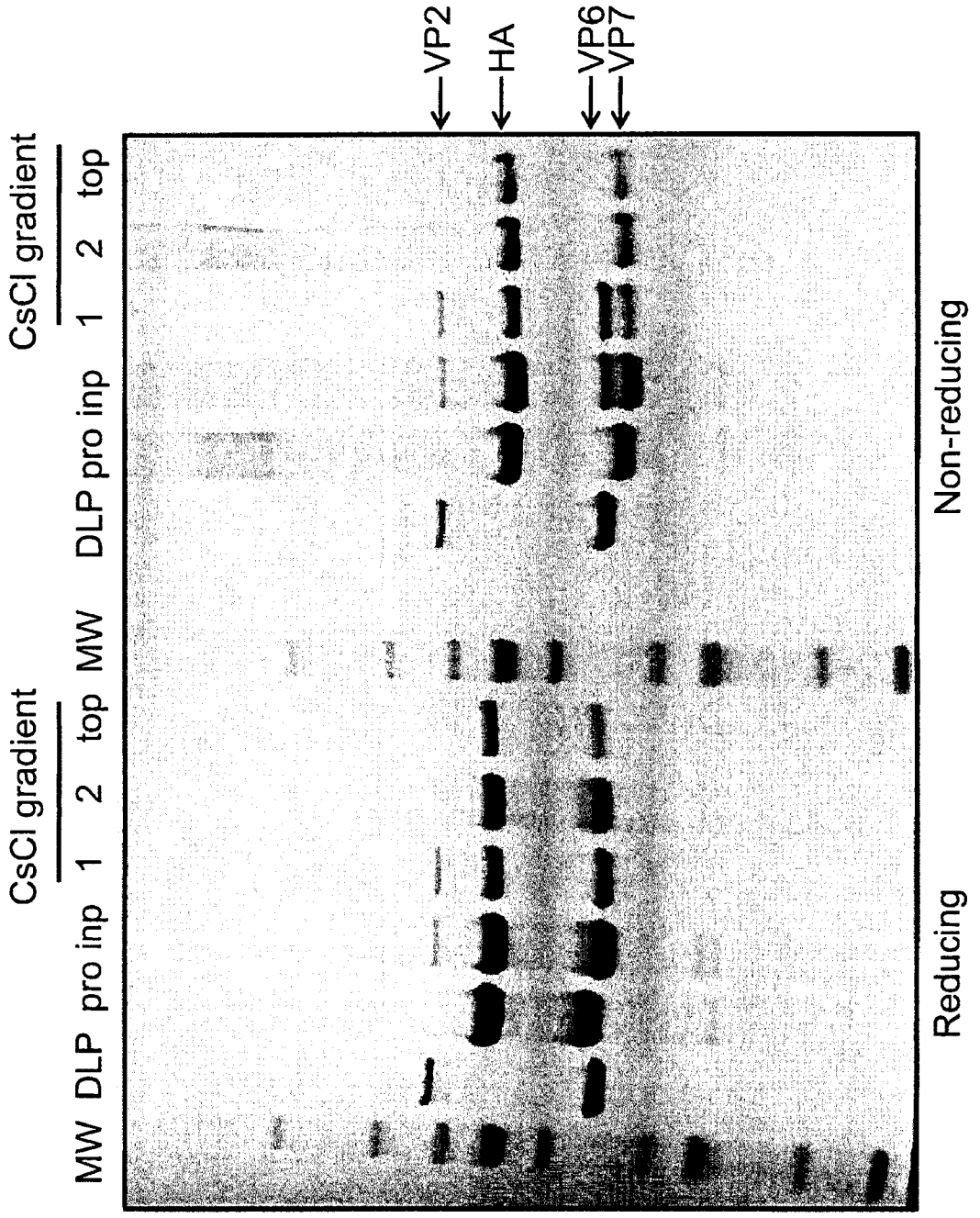

VP2
HA
VP6
VP7
CsCl gradient
MW DLP pro inp 1 2 top
CsCl gradient
MW DLP pro inp 1 2 top
Reducing
Non-reducing

FIG. 4A

FIG. 4B

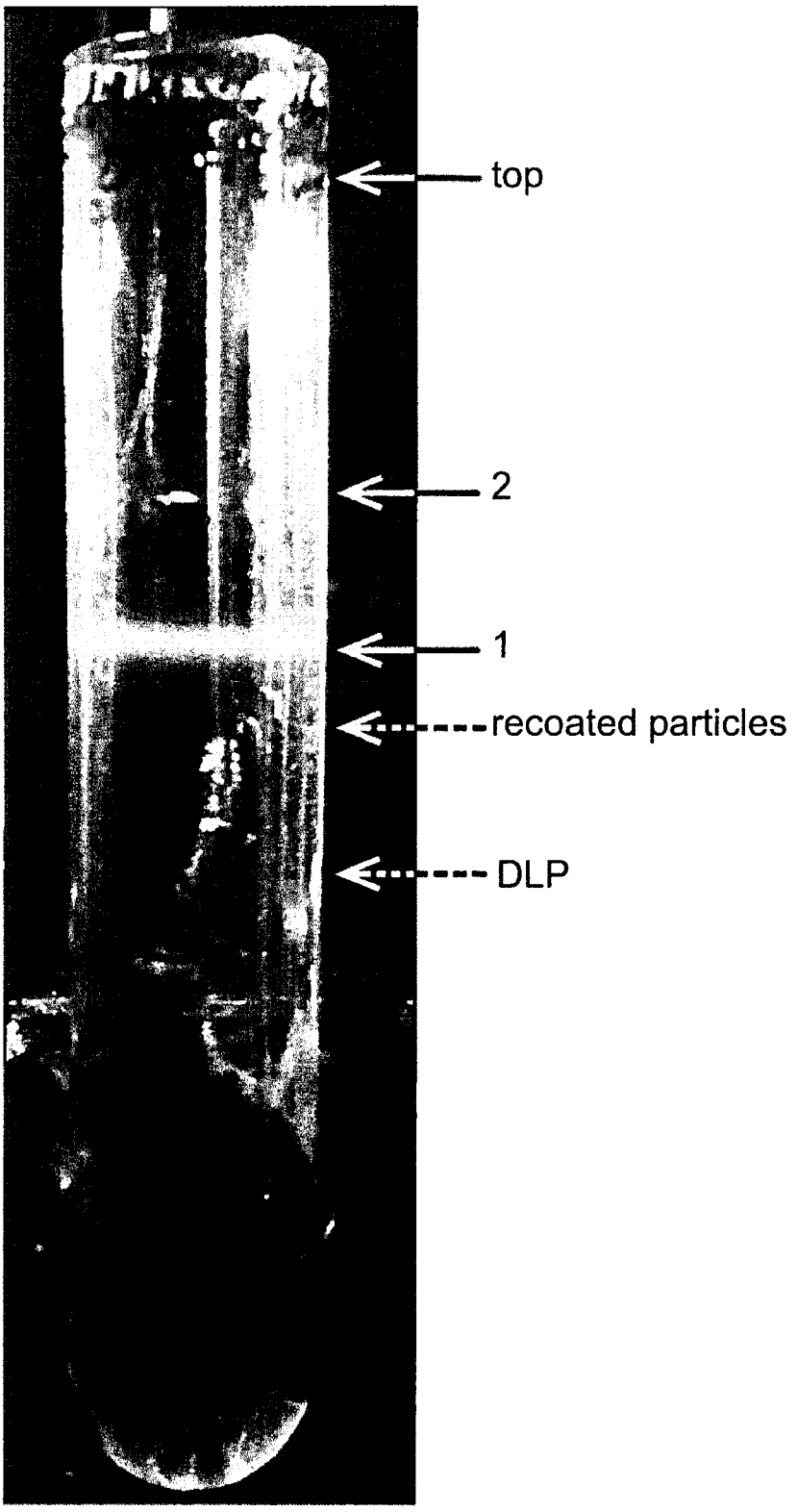
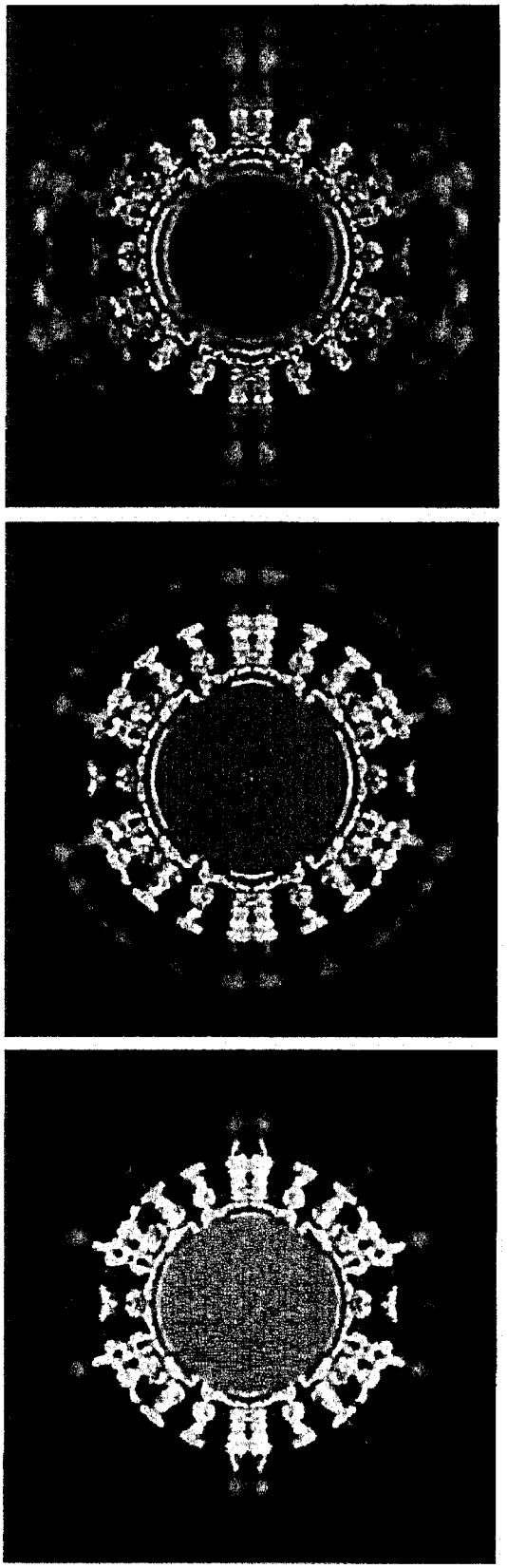


FIG. 5



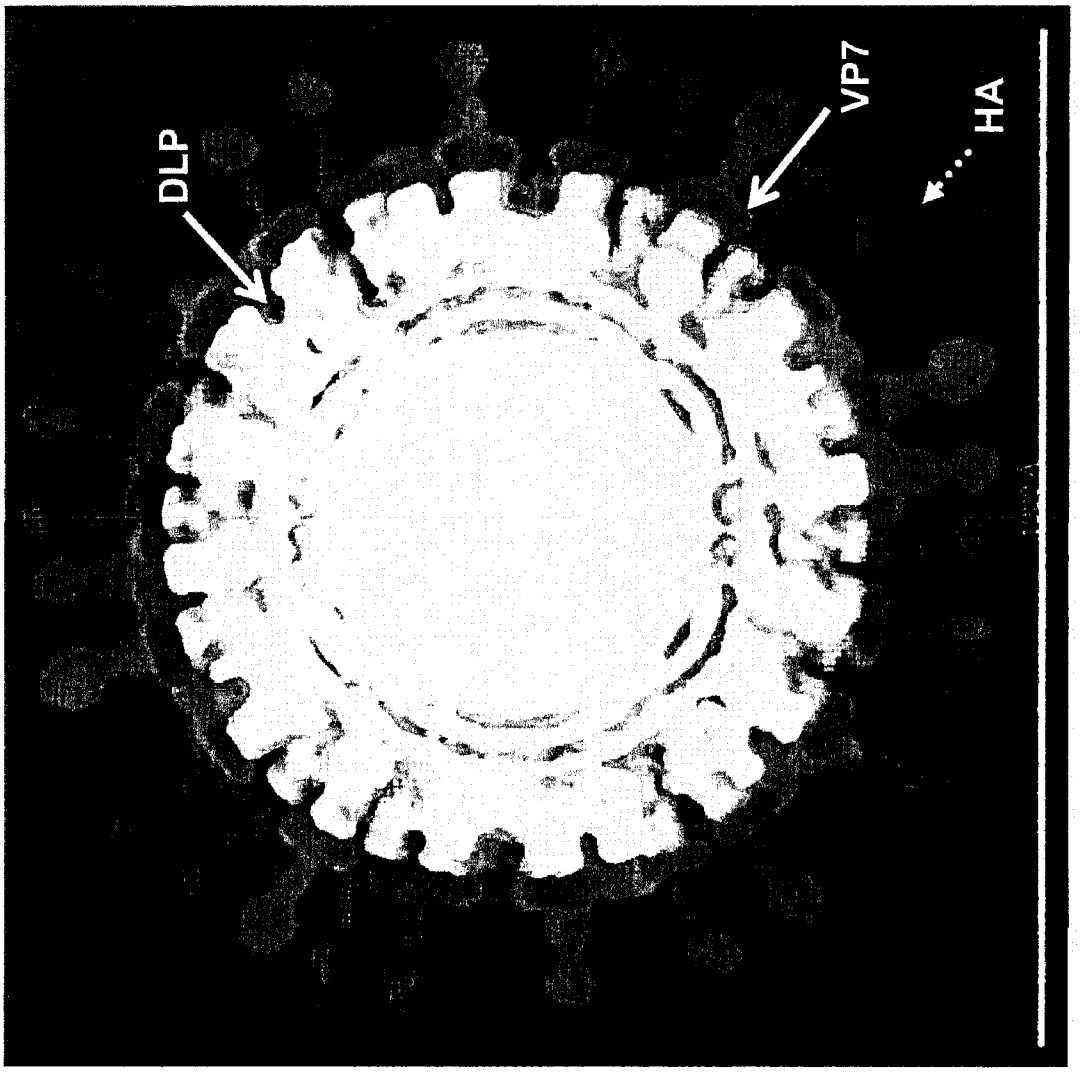


FIG. 6

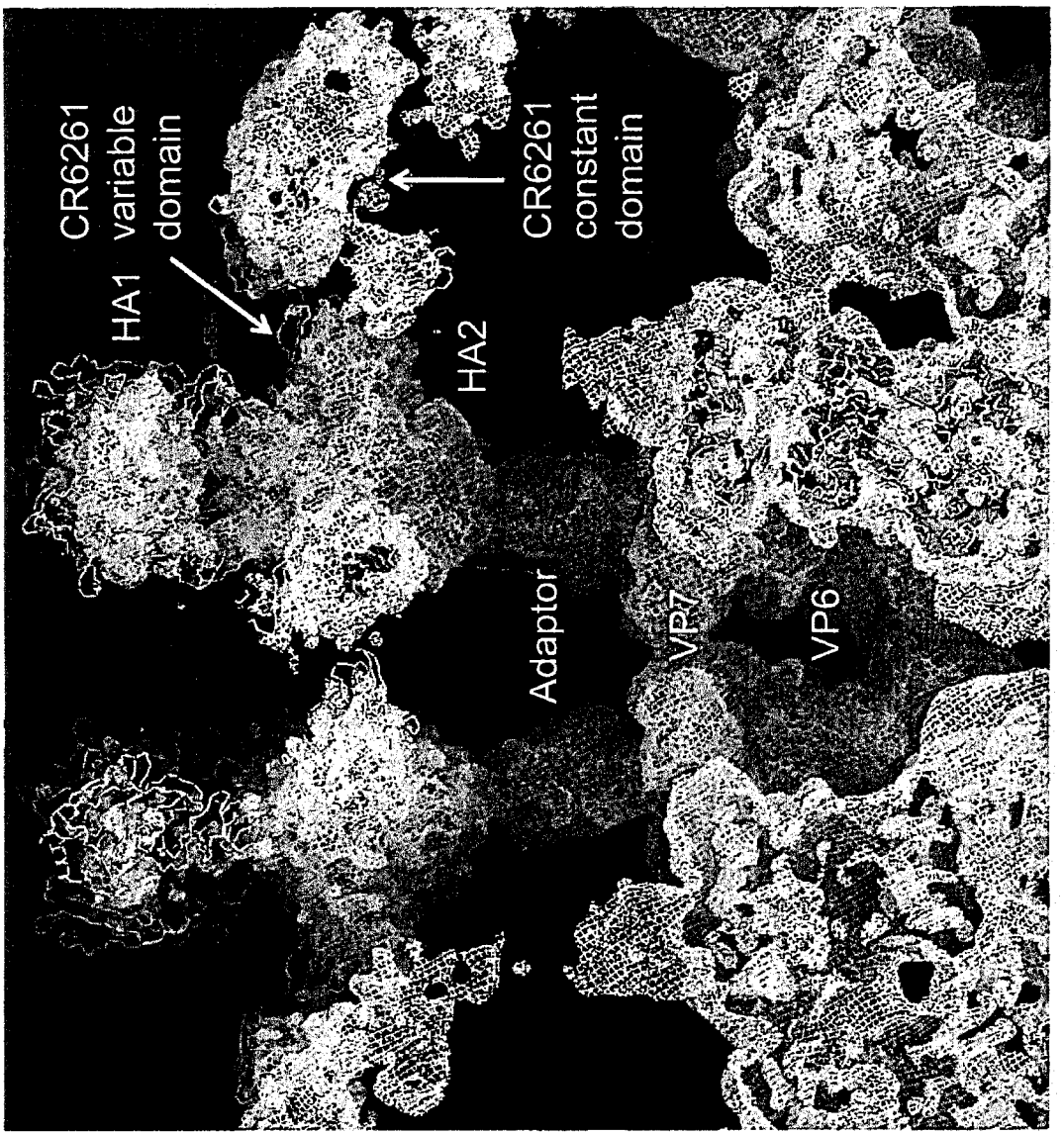
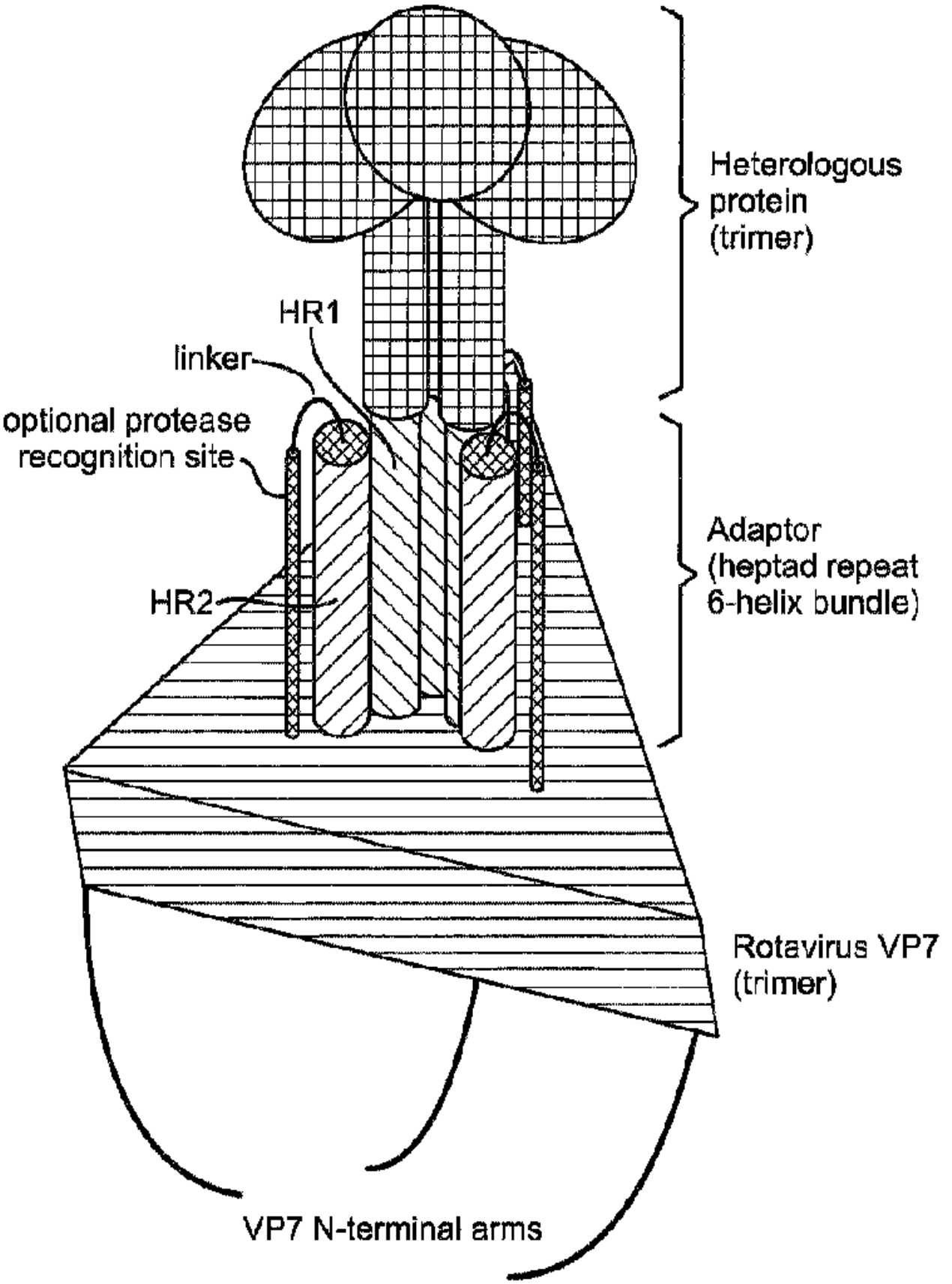

CR6261 variable domain
HA1
HA2
CR6261 constant domain
Adaptor
VP7
VP6

FIG. 7


Heterologous protein (trimer)
HR1
linker
optional protease recognition site
Adaptor (heptad repeat 6-helix bundle)
HR2
Rotavirus VP7 (trimer)
VP7 N-terminal arms