



(86) Date de dépôt PCT/PCT Filing Date: 2011/12/02
(87) Date publication PCT/PCT Publication Date: 2012/06/07
(45) Date de délivrance/Issue Date: 2019/04/30
(85) Entrée phase nationale/National Entry: 2013/05/30
(86) N° demande PCT/PCT Application No.: DK 2011/050460
(87) N° publication PCT/PCT Publication No.: 2012/072088
(30) Priorités/Priorities: 2010/12/02 (EP10193479.2);
2011/04/15 (US61/475,965)

(51) Cl.Int./Int.Cl. *C07K 14/005* (2006.01),
A61K 39/12 (2006.01), *C07K 14/025* (2006.01),
C07K 14/045 (2006.01), *C07K 14/075* (2006.01),
C07K 14/085 (2006.01), *C07K 14/11* (2006.01),
C07K 14/16 (2006.01), *C07K 14/18* (2006.01)

(72) Inventeurs/Inventors:
LANGE, EINAR TONNES, NO;
GRONVOLD, MAJA SOMMERFELT, NO;
HOLMBERG, JENS OLOF, SE;
SORENSEN, BIRGER, NO

(73) Propriétaire/Owner:
BIONOR IMMUNO AS, NO

(74) Agent: DEETH WILLIAMS WALL LLP

(54) Titre : CONCEPTION D'ECHAFAUDAGE PEPTIDIQUE
(54) Title: PEPTIDE SCAFFOLD DESIGN

$X^1 - X^2 - X^3 - X^4 - X^5$ (formula I)

(57) Abrégé/Abstract:

The present invention relates to novel peptides and methods for treatment, diagnosis and prognosis of virus infections including infections with HCV, HIV, CMV and Influenza. The invention further relates to methods for identifying and providing peptides useful for the treatment and diagnosis.

- (51) International Patent Classification:
C07K 14/005 (2006.01) C07K 14/045 (2006.01)
C07K 14/16 (2006.01) C07K 14/075 (2006.01)
C07K 14/18 (2006.01) C07K 14/085 (2006.01)
C07K 14/11 (2006.01) A61K 39/12 (2006.01)
C07K 14/025 (2006.01)

(21) International Application Number:
PCT/DK2011/050460

(22) International Filing Date:
2 December 2011 (02.12.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/475,965 2 December 2010 (02.12.2010) US
10193479.2 2 December 2010 (02.12.2010) EP

(71) Applicant (for all designated States except US): **BIONOR IMMUNO AS** [NO/NO]; Klostergata 33, P.O. Box 2870, N-3702 Skien (NO).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **LANGE, Einer Tønnes** [NO/NO]; Strømdalbakken 6, N-3744 Skien (NO). **GRØNVOLD, Maja Sommerfelt** [NO/NO]; Akland, N-4950 Risør (NO). **HOLMBERG, Jens Olof** [SE/SE]; Hjälmshultsgatan 15, S-254 41 Helsingborg (SE). **SØRENSEN, Birger** [NO/NO]; Meierlia 3, N-3744 Skien (NO).

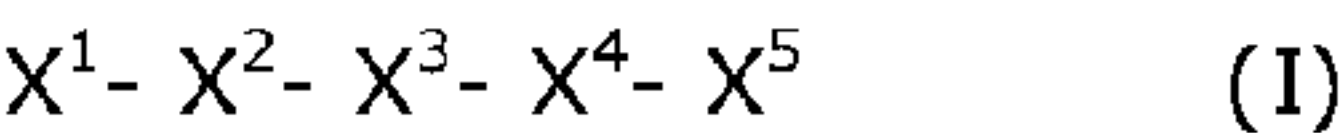
(74) Agent: **INSPICOS A/S**; Kogle Allé 2, P.O. Box 45, DK-2970 Hørsholm (DK).

(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, 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, IS, JP, KE, KG, KM, 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, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, 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, MD, 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, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— with international search report (Art. 21(3))
— with sequence listing part of description (Rule 5.2(a))

(54) Title: PEPTIDE SCAFFOLD DESIGN



(57) Abstract: The present invention relates to novel peptides and methods for treatment, diagnosis and prognosis of virus infections including infections with HCV, HIV, CMV and Influenza. The invention further relates to methods for identifying and providing peptides useful for the treatment and diagnosis.

PEPTIDE SCAFFOLD DESIGN

FIELD OF THE INVENTION

The present invention relates to novel peptides and methods for treatment, diagnosis and prognosis of virus infections including infections with HCV, HIV, CMV and Influenza. The invention further relates to methods for identifying and providing peptides useful for the treatment and diagnosis.

BACKGROUND OF THE INVENTION

Conventional approaches to vaccine development have implemented either whole replication competent virus which has been attenuated (e.g. Sabin polio vaccine, measles, mumps, rubella (MMR)) or inactivated virions that are not replication competent. On occasions, the inactivated virus vaccines may include split vaccines where the virus particles have been disrupted. Molecular techniques have also been used to develop the subunit vaccine (e.g. hepatitis B vaccine) that consists only of the surface glycoproteins of hepatitis B virus. The inactivated virus vaccines tend to induce primarily antibody responses to the viruses in question, whereas the live attenuated vaccines induce both cell-mediated immunity as well as an antibody response since the vaccine induces a transient infection.

The only disease which has been eliminated by virtue of a successful vaccination campaign is smallpox. A campaign is currently in progress to eradicate polio. Features of virus infections that can be eliminated by vaccination are infections caused by viruses with stable virus antigens (i.e. very low mutation frequency, few subtypes), that lack a reservoir in other animal species, viruses that do not persist in the body once the infection is over and where vaccination leads to long lasting immunity. Viruses such as polio and measles fulfill these criteria whereas viruses such as influenza virus (Flu), HCV, and HIV that vary their protein sequences do not. It is for this reason that new and alternate approaches are required to develop vaccines for these diseases.

Vaccination aims to stimulate the immune response to a specific pathogen in advance of infection. When an individual is exposed to that pathogen, a memory response is triggered which prevents the establishment of infection. Vaccines therefore stimulate the adaptive immune response which unlike innate immunity, is long lived and has memory. There are two major arms to the adaptive immune system. Humoral immunity which involves the development of antibodies that can bind virus particles and certain antibodies that can neutralize infection. Cell mediated immunity that leads to the development of cytotoxic T-

cells that kill infected cells exposing viral epitopes in the context of human leukocyte antigen (HLA) class I, in this way eliminating infected cells.

The challenge of providing vaccines suitable for stimulation of the adaptive immune system is that peptide epitopes need to be taken up by the antigen presenting cells.

- 5 Several peptides have been demonstrated to translocate across the plasma membrane of eukaryotic cells by a seemingly energy- independent pathway. These peptides are defined as cell-penetrating peptides (CPPs). Cellular delivery using these cell-penetrating peptides offers several advantages over conventional techniques. It is non-invasive, energy-independent, is efficient for a broad range of cell types and can be applied to cells en masse.
- 10 Hepatitis means inflammation of the liver which can be caused by a variety of factors including toxins, certain drugs, some diseases, heavy alcohol use, and bacterial and viral infections. Hepatitis is also the name of a family of viral infections that affect the liver; the most common types in the developed world are hepatitis A, hepatitis B, and hepatitis C.

- Hepatitis C is a liver disease that results from infection with the hepatitis C virus (HCV). It can range in severity from a mild illness lasting a few weeks to a serious, lifelong illness. Hepatitis C is spread via blood; the most common form of transmission is through sharing needles or other equipment used to inject drugs. The infection can be either "acute" or "chronic". Acute HCV infection is an asymptomatic, short-term illness that occurs within the first 6 months after someone is exposed to the hepatitis C virus. For most people, acute
- 15 infection leads to chronic infection, which can result in long-term complications and even death.
- 20

- HCV is an enveloped positive stranded ribonucleic acid (RNA) virus with a diameter of about 50nm, belonging to the genus Hepacivirus in the family Flaviviridae that replicate in the cytoplasm of infected cells. The only known reservoir for HCV is humans, although the virus
- 25 has experimentally been transmitted to chimpanzees. The natural targets of HCV are hepatocytes and possibly B-lymphocytes. As of 2008, six different genotypes and more than 100 subtypes of the virus are known. Replication occurs through an RNA-dependent RNA polymerase that lacks a proofreading function, which results in a very high rate of mutations. Rapid mutations in a hypervariable region of the HCV genome coding for the envelope
- 30 proteins enable the virus to escape immune surveillance by the host. As a consequence, most HCV-infected people proceed to chronic infection.

It is estimated that 170 million people are infected with HCV worldwide, equating to approximately 3% of the global population. There are also approximately 3-4 million people

who are infected every year; with an estimated 80% of these newly infected patients progressing to chronic infection.

The 6 genotypes of HCV have different geographical spread. The disease in the early stages is generally asymptomatic; the majority of patients with chronic infection eventually progress to complications such as liver fibrosis and cirrhosis, and, in 1-5% of cases, hepatocellular carcinoma.

HCV is the major cause of non-A, non-B hepatitis worldwide. Acute infection with HCV frequently leads to chronic hepatitis and end-stage cirrhosis. It is estimated that up to 20% of HCV chronic carriers may develop cirrhosis over a time period of about 20 years and that of those with cirrhosis between 1 to 4% is at risk to develop liver carcinoma.

The about 9.6 kb single-stranded RNA genome of the HCV virus comprises a 5'- and 3'-noncoding region (NCRs) and, in between these NCRs a single long open reading frame of about 9 kb encoding an HCV polyprotein of about 3000 amino acids.

HCV polypeptides are produced by translation from the open reading frame and cotranslational proteolytic processing. Structural proteins are derived from the amino-terminal one-fourth of the coding region and include the capsid or Core protein (about 21 kDa), the E1 envelope glycoprotein (about 35 kDa) and the E2 envelope glycoprotein (about 70 kDa, previously called NS1), and p7 (about 7kDa). The E2 protein can occur with or without a C-terminal fusion of the p7 protein (Shimotohno et al. 1995). An alternative open reading frame in the Core-region has been found which is encoding and expressing a protein of about 17 kDa called F (Frameshift) protein (Xu et al. 2001; Ou & Xu in US Patent Application Publication No. US2002/0076415). In the same region, ORFs for other 14-17 kDa ARFPs (Alternative Reading Frame Proteins), A1 to A4, were discovered and antibodies to at least A1, A2 and A3 were detected in sera of chronically infected patients (Walewski et al. 2001). From the remainder of the HCV coding region, the non-structural HCV proteins are derived which include NS2 (about 23 kDa), NS3 (about 70 kDa), NS4A (about 8 kDa), NS4B (about 27 kDa), NS5A (about 58 kDa) and NS5B (about 68 kDa) (Grakoui et al. 1993).

Influenza remains a significant cause of mortality and morbidity worldwide. The World Health Organisation (WHO) estimates that seasonal epidemics affect 3-5 million people with severe illness annually and result in 250,000 – 500,000 mortalities. Influenza is caused by viruses in the family Orthomyxoviridae which are negative stranded RNA viruses. The influenza virus exists as three types, A, B and C of which only A is associated with pandemics. Types A viruses are found in both humans and animals, particularly birds but also other mammals such as pigs. Type A viruses are further typed into subtypes according to different kinds and

combinations of virus surface proteins. Among many subtypes in 2009 influenza A (H1N1) and A (H3N2) subtypes were circulating among humans. Influenza A and B are included in the seasonal vaccine, whereas influenza C occurs only rarely, and so it is not included in the seasonal vaccine. Type B viruses are human specific and Type C viruses cause a very mild disease. The genomes of Orthomyxoviruses are segmented. Influenza viruses Types A and B have 8 segments whereas type C has seven. Pandemics may arise as a result of re-assortment of gene segments when two different type A viruses infect the same cell. There is no immunity in the population to this novel re-assorted virus. Three pandemics occurred in the twentieth century: "Spanish influenza" in 1918, "Asian influenza" in 1957, and "Hong Kong influenza" in 1968. The 1918 pandemic killed an estimated 40–50 million people worldwide. Subsequent pandemics were much milder, with an estimated 2 million deaths in 1957 and 1 million deaths in 1968. In June 2009 the WHO declared a pandemic from influenza virus H1N1 (swine Influenza) which was declared over in August 2010.

Human papillomaviruses are made up of a group of DNA viruses in the family Papillomaviridae which infect the skin and mucous membranes. Two groups which are derived from more than 100 different identified subtypes are the main cause for clinical concern: those causing warts (both benign and genital warts), and a group of 12 "high risk" subtypes that can result in cervical cancer. This latter group has been attributed as a contributory factor in the development of nearly all types of cervical cancer. Worldwide, cervical cancer remains the second most common malignancy in women, and is a leading cause of cancer-related death for females in developing countries. HPV 16 and 18 have been mainly associated with cervical cancer, however, the virus is also a cause of throat cancer in both men and women. HPV is transmitted through contact and enters the skin through abrasions. An abortive infection, where only the early proteins are expressed is associated with cancer development.

OBJECT OF THE INVENTION

It is an object of embodiments of the invention to provide peptides that may be used as immunogens to stimulate an adaptive immune response in a subject.

In particular, it is an object of embodiments of the invention to provide peptides that may be taken up by antigen presenting cells (macrophages and dendritic cells) such that epitopes within the peptides are correctly processed and presented to T-lymphocytes in order to stimulate an effective immune response.

Further, it is an object of embodiments of the invention to provide peptides that may be used as antigens, to provide immunogenic compositions and methods for inducing an immune response in a subject against an antigen.

SUMMARY OF THE INVENTION

- 5 The present invention pertains to a peptide design promoting uptake of peptide epitopes by antigen presenting cells (macrophages and dendritic cells) such that the epitopes can be correctly processed and presented in the context of HLA class I and II to stimulate both CD4+ and CD8+ T-lymphocytes. CD8+ T-lymphocytes with cytotoxic capacity will kill infected cells bearing the epitope of interest. CD4+ T-lymphocyte provide 'help' to sustain effective
10 CD8+ T-lymphocyte responses.

- It has been found by the present inventor(s) that peptide constructs - amino acid sequences with a particular pattern or scaffold design - have the ability to effectively penetrate the cell membrane. Accordingly, the peptide constructs according to the present invention may be used to load cells with an immunogenically effective amount of a peptide or fragments of this
15 peptide that can be presented by macrophages and dendritic cells. Accordingly these peptide constructs may elicit a Cytotoxic T-lymphocyte immune (CTL) response and/or a Humoral Immune Response.

So, in a first aspect the present invention relates to an isolated cell-penetrating peptide of not more than 60 amino acids with the following structure

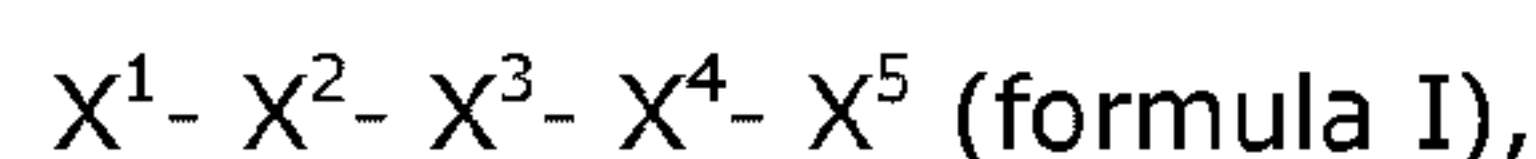
- 20 $X^1 - X^2 - X^3 - X^4 - X^5$ (formula I),

- wherein X^1 and X^3 independently defines a linear sequence of any 1, 2, 3 or 4 amino acid(s) independently selected from any basic amino acid, citrulline, tryptophan, or a derivative thereof; X^2 defines a linear sequence of 8-30 amino acids derived from an antigen; X^4 defines a linear sequence of 8-30 amino acids derived from said antigen, said sequence X^4 being
25 different from X^2 ; and wherein X^5 is any one optional amino acid selected from a basic amino acid, citrulline, tryptophan, or a derivative thereof.

- It is to be understood that the amino acid sequence of formula I refers to a peptide sequence in a standard N- to C-terminal direction, wherein the first amino acid mentioned is the N-terminal amino acid that may have an amino ($-NH_2$) group or alternatively an $-NH_3^+$ group.
30 The last amino acid mentioned is the C-terminal that may have a free carboxyl group ($-COOH$) or a carboxylate group. In some embodiments the N- and/or C-terminal amino acid is

modified, such as by N-terminal acetylation or C-terminal amidation. The symbol "-" used in formula I refers to a standard peptide bond, such as a standard peptide bond between X^1 and X^2 in " X^1 - X^2 ".

It is further to be understood that the peptides according to the invention primarily are intended for synthetic peptide synthesis. However, the peptides may be longer than 60 amino acids, if the peptides are produced by recombinant means. Accordingly, an alternative first aspect of the present invention relates to an isolated cell-penetrating peptide comprising a peptide sequence with the following structure



wherein X^1 and X^3 independently defines a linear sequence of any 1, 2, 3 or 4 amino acid independently selected from any basic amino acid, citrulline, tryptophan, or a derivative thereof; X^2 defines a linear sequence of 8-30 amino acids derived from an antigen; X^4 defines a linear sequence of 8-30 amino acids derived from said antigen, said sequence X^4 being different from X^2 ; and wherein X^5 is any one optional amino acid selected from a basic amino acid, citrulline, tryptophan, or a derivative thereof.

In a second aspect the present invention relates to an isolated peptide of not more than 60 amino acids comprising a sequence of X^2 and/or X^4 as independently defined in any one of table 1 or table 2.

In a third aspect the present invention relates to the use of a peptide comprising a sequence of X^2 or X^4 as independently defined in any one of table 1 or table 2 for inducing an immune response in a subject.

In a further aspect the present invention relates to an isolated peptide consisting of $X^1 - X^5$ as defined in any one of table 1 or table 2.

In a further aspect the present invention relates to a dimer peptide comprising two peptide monomers, wherein each peptide monomer is according to the invention.

In a further aspect the present invention relates to a peptide combination comprising two or more peptides according to the invention.

In a further aspect the present invention relates to an isolated nucleic acid or polynucleotide encoding a peptide according to the invention.

In a further aspect the present invention relates to a vector comprising the nucleic acid or polynucleotide encoding a peptide according to the invention.

In a further aspect the present invention relates to a host cell comprising the vector comprising the nucleic acid or polynucleotide encoding a peptide according to the invention.

- 5 In a further aspect the present invention relates to an immunogenic composition comprising at least one peptide according to the invention, a dimer peptide comprising two peptides monomers, wherein each peptide monomer is according to the invention, a peptide combination comprising two or more peptide according to the invention, the nucleic acid or polynucleotide encoding a peptide according to the invention, or the vector comprising the
10 nucleic acid or polynucleotide encoding a peptide according to the invention; in combination with a pharmaceutically acceptable diluent or vehicle and optionally an immunological adjuvant.

- In a further aspect the present invention relates to a method for inducing an immune response in a subject against an antigen which comprises administration of at least one
15 peptide according to the invention; a dimer peptide comprising two peptides monomers, wherein each peptide monomer is according to the invention; a peptide combination comprising two or more peptide according to the invention; the nucleic acid or polynucleotide encoding a peptide according to the invention; the vector comprising the nucleic acid or polynucleotide encoding a peptide according to the invention; or the composition according to
20 the invention.

- In a further aspect the present invention relates to a method for reducing and/or delaying the pathological effects of a virus in a subject infected with said virus, the method comprising administering an effective amount of at least one peptide according to the invention; a dimer peptide comprising two peptides monomers, wherein each peptide monomer is according to
25 the invention; a peptide combination comprising two or more peptide according to the invention; the nucleic acid or polynucleotide encoding a peptide according to the invention; the vector comprising the nucleic acid or polynucleotide encoding a peptide according to the invention; or the composition according to the invention.

- In a further aspect the present invention relates to a peptide according to the invention for
30 use as a medicament.

In a further aspect the present invention relates to a peptide according to the invention for treating the pathological effects of a virus in a subject infected with said virus.

DETAILED DISCLOSURE OF THE INVENTION

Definitions

When terms such as "one", "a" or "an" are used in this disclosure they mean "at least one", or "one or more" unless otherwise indicated. Further, the term "comprising" is intended to mean "including" and thus allows for the presence of other constituents, features, conditions, or steps than those explicitly recited.

"HIV" generally denotes human immunodeficiency virus I.

"HIV disease" is composed of several stages including the acute HIV infection which often manifests itself as an influenza-like infection and the early and medium stage symptomatic disease, which has several non-characteristic symptoms such as skin rashes, fatigue, night sweats, slight weight loss, mouth ulcers, and fungal skin and nail infections. Most HIV infected will experience mild symptoms such as these before developing more serious illnesses. It is generally believed that it takes five to seven years for the first mild symptoms to appear. As HIV disease progresses, some individuals may become quite ill even if they have not yet been diagnosed with AIDS (see below), the late stage of HIV disease. Typical problems include chronic oral or vaginal thrush (a fungal rash or spots), recurrent herpes blisters on the mouth (cold sores) or genitals, ongoing fevers, persistent diarrhea, and significant weight loss. "AIDS" is the late stage HIV disease and is a condition which progressively reduces the effectiveness of the immune system and leaves individuals susceptible to opportunistic infections and tumors.

The term "cell-penetrating peptide" as used herein refers to any peptide with the capability to translocate across the plasma membrane into either cytoplasmic and/or nuclear compartments of eukaryotic and/or prokaryotic cells, such as into cytoplasm, nucleus, lysosome, endoplasmic reticulum, golgi apparatus, mitochondria and/or chloroplast, seemingly energy-independently. This capability to translocate across the plasma membrane of a "cell-penetrating peptide" according to the invention may be non-invasive, energy-independent, non-saturable, and/or receptor independent. In one embodiment the term "cell-penetrating peptide" refers to a peptide, which is demonstrated to translocate across a plasma membrane as determined by the assay in example 1. It is to be understood that a cell-penetrating peptide according to the present invention may be translocated across the membrane with the sequence complete and intact, or alternatively partly degraded, but in a form where the antigens contained within this peptide is able to be presented within the cell to stimulate an immune response. Accordingly, a cell-penetrating peptide according to the

present invention is a peptide that may be demonstrated to translocate across a plasma membrane as determined by the assay in example 1 and be demonstrated to stimulate an effective immune response.

The term "derived from an antigen" when in reference to a peptide derived from a source (such as a virus etc.) as used herein is intended to refer to a peptide which has been obtained (e.g., isolated, purified, etc.) from the source. Preferably, the peptide may be genetically engineered and/or chemically synthesized to be essentially identical to the native peptide of the source. The term includes the use of variants of known native peptide sequences, such as peptide sequences, where 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 amino acids of the native peptide sequence have been substituted with any other amino acid, such as conservative substitutions. Alternatively, 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 amino acids have been removed or added to the native peptide sequence. Accordingly, in some embodiments, the peptides according to the present invention comprises the sequences X^2 and/or X^4 , that is defined as a sequence of 8-30 amino acids, such as 8-20 amino acids derived from an antigen, wherein the peptide sequence of the antigen comprises 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 substitutions, additions or deletions relative to the antigen, such as the addition of an arginine in the N- or C-terminal of the amino acid sequence of X^2 and/or X^4 . The amino acids used in the amino acid sequences according to the invention may be in both L- and/or D-form. It is to be understood that both L- and D-forms may be used for different amino acids within the same peptide sequence. In some embodiments the amino acids within the peptide sequence are in L-form, such as natural amino acids. It is to be understood that any known antigen may be used in the constructs according to the present invention.

In some specific embodiments, the first 1, 2, or 3 amino acids in the N-terminal of the amino acid sequences according to the invention are in the D-form. It is assumed that the N-terminal trimming and thereby degradation of the peptides are somewhat delayed by having amino acids of the D-form in the N-terminal of these cell-penetrating peptides. Alternatively and in some embodiments, the first 1, 2, or 3 amino acids in the N-terminal of the amino acid sequences according to the invention are amino acids in beta or gamma forms. Beta amino acids have their amino group bonded to the beta carbon rather than the alpha carbon as in the 20 standard natural amino acids.

Alternatively the first 1, 2, or 3 amino acids in the N-terminal of the amino acid sequences according to the invention may be modified by incorporation of fluorine, or alternatively cyclic amino acids or other suitable non-natural amino acids are used.

A "variant" or "analogue" of a peptide refers to a peptide having an amino acid sequence that is substantially identical to a reference peptide, typically a native or "parent" polypeptide. The

peptide variant may possess one or more amino acid substitutions, deletions, and/or insertions at certain positions within the native amino acid sequence.

"Conservative" amino acid substitutions are those in which an amino acid residue is replaced with an amino acid residue having a side chain with similar physicochemical properties.

- 5 Families of amino acid residues having similar side chains are known in the art, and include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine, tryptophan), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine), beta-branched side
- 10 chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). A particular form of conservative amino acid substitutions include those with amino acids, which are not among the normal 20 amino acids encoded by the genetic code. Since preferred embodiments of the present invention entail use of synthetic peptides, it is unproblematic to provide such "non-naturally occurring" amino
- 15 acid residues in the peptides disclosed herein, and thereby it is possible to exchange the natural saturated carbon chains in the side chains of amino acid residues with shorter or longer saturated carbon chains – for instance, lysine may be substituted with an amino acid having an the side chain $-(CH_2)_nNH_3$, where n is different from 4, and arginine may be substituted with an amino acid having the side chain $-(CH_2)_nNHC(=NH_2)NH_2$, where n is
- 20 different from 3, etc. Similarly, the acidic amino acids aspartic acid and glutamic acid may be substituted with amino acid residues having the side chains $-(CH_2)_nCOOH$, where $n > 2$.

The term "substantially identical" in the context of two amino acid sequences means that the sequences, when optimally aligned, such as by the programs GAP or BESTFIT using default gap weights, share at least about 50, at least about 60, at least about 70, at least about 80,

25 at least about 90, at least about 95, at least about 98, or at least about 99 percent sequence identity. In one embodiment, residue positions that are not identical differ by conservative amino acid substitutions. Sequence identity is typically measured using sequence analysis software. Protein analysis software matches similar sequences using measures of similarity assigned to various substitutions, deletions and other modifications, including conservative

30 amino acid substitutions. For instance, the publicly available GCG software contains programs such as "Gap" and "BestFit" which can be used with default parameters to determine sequence homology or sequence identity between closely related polypeptides, such as homologous polypeptides from different species of organisms or between a wild-type protein and a mutein thereof. See, e.g., GCG Version 6.1. Polypeptide sequences can also be

35 compared using FASTA or ClustalW, applying default or recommended parameters. A program in GCG Version 6.1., FASTA (e.g., FASTA2 and FASTA3) provides alignments and percent sequence identity of the regions of the best overlap between the query and search

sequences (Pearson, Methods Enzymol. 1990;183:63-98; Pearson, Methods Mol. Biol. 2000;132:185-219). Another preferred algorithm when comparing a sequence to a database containing a large number of sequences from various organisms, or when deducing the is the computer program BLAST, especially blastp, using default parameters. See, e.g., Altschul et al., J. Mol. Biol. 1990;215:403-410; Altschul et al., Nucleic Acids Res. 1997;25:3389-402 (1997). "Corresponding" amino acid positions in two substantially identical amino acid sequences are those aligned by any of the protein analysis software mentioned herein, typically using default parameters.

An "isolated" molecule is a molecule that is the predominant species in the composition wherein it is found with respect to the class of molecules to which it belongs (i.e., it makes up at least about 50% of the type of molecule in the composition and typically will make up at least about 70%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or more of the species of molecule, e.g., peptide, in the composition). Commonly, a composition of a peptide molecule will exhibit 98% - 99% homogeneity for peptide molecules in the context of all present peptide species in the composition or at least with respect to substantially active peptide species in the context of proposed use.

The term "linear sequence" as used herein refers to the specific sequence of amino acids connected by standard peptide bonds in standard N- to C-terminal direction. The peptide may contain only peptide bonds. However the term does not exclude that an amino acid within a sequence, such as within X^3 , may be connected, such as through the side chains, with another amino acid at a distant location within the peptide sequence, such as a distant location within X^3 .

In the context of the present invention, "treatment" or "treating" refers to preventing, alleviating, managing, curing or reducing one or more symptoms or clinically relevant manifestations of a disease or disorder, unless contradicted by context. For example, "treatment" of a patient in whom no symptoms or clinically relevant manifestations of a disease or disorder have been identified is preventive or prophylactic therapy, whereas "treatment" of a patient in whom symptoms or clinically relevant manifestations of a disease or disorder have been identified generally does not constitute preventive or prophylactic therapy.

The term "antigen" denotes a substance of matter which is recognized by the immune system's specifically recognizing components (antibodies, T-cells).

The term "immunogen" is in the present context intended to denote a substance of matter, which is capable of inducing an adaptive immune response in an individual, where said

adaptive immune response targets the immunogen. In relation to the present invention, an immunogen will induce a humoral and/or cell-mediated immune response. In other words, an immunogen is an antigen, which is capable of inducing immunity.

5 The terms "epitope", "antigenic determinant" and "antigenic site" are used interchangeably herein and denotes the region in an antigen or immunogen which is recognized by antibodies (in the case of antibody binding epitopes, also known as "B-cell epitopes") or by T-cell receptors when the epitope is complexed to a Major histocompatibility complex (MHC) molecule (in the case of T-cell receptor binding epitopes, i.e. "T-cell epitopes").

10 The term "immunogenically effective amount" has its usual meaning in the art, i.e. an amount of an immunogen, which is capable of inducing an immune response, which significantly engages pathogenic agents, which share immunological features with the immunogen.

15 The term "vaccine" is used for a composition comprising an immunogen and which is capable of inducing an immune response which is either capable of reducing the risk of developing a pathological condition or capable of inducing a therapeutically effective immune response which may aid in the cure of (or at least alleviate the symptoms of) a pathological condition.

20 The term "pharmaceutically acceptable" has its usual meaning in the art, i.e. it is used for a substance that can be accepted as part of a medicament for human use when treating the disease in question and thus the term effectively excludes the use of highly toxic substances that would worsen rather than improve the treated subject's condition.

25 A "T helper lymphocyte epitope" (a T_H epitope) is peptide, which binds an MHC Class II molecule and can be presented on the surface of an antigen presenting cell (APC) bound to the MHC Class II molecule. An "immunological carrier" is generally a substance of matter which includes one or many T_H epitopes, and which increase the immune response against an antigen to which it is coupled by ensuring that T-helper lymphocytes are activated and proliferate. Examples of known immunological carriers are the tetanus and diphtheria toxoids and keyhole limpet hemocyanin (KLH).

30 In the scaffold design according to the present invention, X^2 and X^4 defines a sequence of 8-30 amino acids, such as 8-20 amino acids derived from the antigen. This sequence of amino acids derived from an antigen may herein be referred to as an epitope.

Preferably the epitopes used in the scaffold according to the present invention are CTL epitopes. A "CTL inducing peptide" is a HLA Class I binding peptide that is capable of inducing

a CTL response. In other embodiments the epitopes used in the scaffold design according to the present invention are HTL inducing peptides. A "HTL inducing peptide" is a HLA Class II binding peptide that is capable of inducing a HTL response.

The term "basic amino acid" as used herein refers to any amino acid including both natural and non-natural amino acids that has an isoelectric point above 6.3 (such as above 7.4) as measured according to Kice & Marvell "Modern Principles of organic Chemistry" (Macmillan, 1974) or Matthews and van Holde "Biochemistry" Cummings Publishing Company, 1996. Included within this definition are Arginine, Lysine, Homoarginine (Har), and Histidine as well as derivatives thereof. Suitable non-natural basic amino acids are e.g. as described in US 6,858,396. Suitable positively charged amino acids includes non-natural alpha amino acids available from Bachem AG and includes alpha-amino-glycine, alpha,gamma-diaminobutyric acid, ornithine, alpha, beta-diaminopropionic acid, alpha-difluoromethyl-ornithine, 4-amino-piperidine-4-carboxylic acid, 2,6-diamino-4-hexynoic acid, beta-(1-piperazinyl)-alanine, 4,5-dehydro-lysine, delta-hydroxy-lysine, omega-hydroxy-norarginine, homoarginine, omega-amino-arginine, omega-methyl-arginine, alpha-methyl-histidine, 2,5-diiodo-histidine, 1-methyl-histidine, 3-methyl-histidine, beta-(2-pyridyl)-alanine, beta-(3-pyridyl)-alanine, beta-(2-quinolyl)-alanine, 3-amino-tyrosine, 4-amino-phenylalanine, and spinacine. Furthermore, any mono or dicarboxylic amino acid is a suitable positively charged amino acid.

The term "neutral amino acid" as used herein refers to an amino acid that has an isoelectric point above between 4.8 and 6.3 as measured according to Kice & Marvell "Modern Principles of organic Chemistry" (Macmillan, 1974). The term "acidic amino acid" as used herein refers to an amino acid that has an isoelectric point below 4.8 as measured according to Kice & Marvell "Modern Principles of organic Chemistry" (Macmillan, 1974).

Unless otherwise indicated amino acids are abbreviated and mentioned by their standard nomenclature known to the person skilled in the art, such as with reference to "nomenclature and symbolism for amino acids and peptides" by the international union of pure and applied chemistry (IUPAC) (www.iupac.org).

Citrulline (In this document referred to with the one-letter symbol "B") and/or homocitrulline are well known non-natural amino acids that in some embodiments may be used in the sequence defined by X¹, X³, or X⁵.

In other alternative embodiments, tryptophan or tryptophan derivatives are used in the sequence defined by X¹, X³, or X⁵. Any suitable tryptophan derivatives may be used. As used herein "tryptophan derivatives" means an unnatural modified tryptophan amino acid residue including those disclosed in US 7,232,803, such as tri tert.-butyltryptophan, di-tert-butyl

tryptophan, 7-benzyloxytryptophan, homotryptophan, 5'-aminoethyltryptophan (available as side chain Boc and N-alpha Fmoc derivative from RSP Amino Acids Analogues Inc, Boston, Mass., USA), N-Acetylhomotryptophan (Toronto Research), 7-Benzoyloxytryptophan (Toronto Research), Homotryptophan (Toronto Research), and tryptophan residues which have been substituted at the 1-, 2-, 5- and/or 7-position of the indole ring, positions 1- or 2- being preferred e.g. 5' hydroxy tryptophan.

The term "antibody response" refers to the production of antibodies (e.g., IgM, IgA, IgG) which bind to an antigen of interest, this response is measured for instance by assaying sera by antigen ELISA.

10 The term "adjuvant" as used herein refers to any compound which, when delivered together or simultaneously with an antigen, non-specifically enhances the immune response to that antigen. Exemplary adjuvants include but are not limited to oil in water and water in oil adjuvants, aluminum-based adjuvants (e.g., AIOH, AIPO₄, etc), and Montanide ISA 720.

15 The terms "patient" and "subject" refer to a mammal that may be treated using the methods of the present invention.

As used herein, the term "immune response" refers to the reactivity of an organism's immune system in response to an antigen. In vertebrates, this may involve antibody production, induction of cell-mediated immunity, and/or complement activation (e.g., phenomena associated with the vertebrate immune system's prevention and resolution of infection by microorganisms). In preferred embodiments, the term immune response encompasses but is not limited to one or more of a "lymphocyte proliferative response," a "cytokine response," and an "antibody response."

25 The term "net charge" as used herein with reference to a peptide sequence refers to the total electric charge of the peptide sequence represented by the sum of charges of each individual amino acid in the peptide sequence, wherein each basic amino acid are given a charge of +1, each acidic amino acid a charge of -1, and each neutral amino acid a charge of 0. Accordingly, the net charge will depend on the number and identities of charged amino acids.

Table 1 – Specific peptides according to the invention

	Reference ID	x1	x2	x3	x4	x5		Placement with

								reference to positions in SEQ ID NO:3; SEQ ID NO:6; SEQ ID NO:7, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:46, and SEQ ID NO:126.	
	P- biotin	R	QIKIWFQN	RR	MKWKK				
	N- Biotin		PVVHLTL	R	QAGDDFSR				
Antigen									
							Modified (m)	X2- seq	x4- seq
HCV	SP_2	RR	GYIPLVGAPLG	BGR	VARALAHGVRV			135- 145	147- 157
HCV	SP_3	R	GYIPLVGAPLG	RR	VARALAHGVRV			135- 145	147- 157
HCV	SP_4	R	GYIPLVGAPLG	RRR	VARALAHGVRV	R		135- 145	147- 157
HCV	SP_5	RR	GYIPLVGAPLG	RR	VARALAHGVRV			135- 145	147- 157
HCV	SP_6	RR	GYIPLVGAPLG	RRR	VARALAHGVRV			135- 145	147- 157
HCV	SP_7	BR	GYIPLVGAPLG	RR	VARALAHGVRV			135- 145	147- 157
HCV	SP_8	RRR	GYIPLVGAPLG	BR	VARALAHGVRV			135- 145	147- 157
HCV	SP_9	R	GYIPLVGAPLG	KKK	VARALAHGVRV			135- 145	147- 157
HCV	SP_10	R	GYIPLVGAPLG	RRR	VARALAHGVRV			135- 145	147- 157
HCV	SP_11	KK	GYIPLVGAPLG	KK	VARALAHGVRV			135- 145	147- 157
HCV	SP_12	W	GYIPLVGAPLG	RR	VARALAHGVRV			135- 145	147- 157
HCV	SP_13	WW	GYIPLVGAPLG	RR	VARALAHGVRV			135- 145	147- 157
HCV	SP_14	EE	GYIPLVGAPLG	EE	VARALAHGVRV			135- 145	147- 157
HCV	SP_15	GG	GYIPLVGAPLG	GG	VARALAHGVRV			135- 145	147- 157
HCV	SP_16	EE	GYIPLVGAPLG	RR	VARALAHGVRV			135- 145	147- 157
HCV	SP_17	RR	GYIPLVGAPLG	LRR	VARALAHGVRV			135- 145	147- 157
HCV	SP21:	WW	GYIPLVGAPLG	RR	VARALAHGVRV			135- 145	147- 157
HCV	SP22:	WW	GYIPLVGAPLG	RRR	VARALAHGVRV			135- 145	147- 157
HCV	SP23:	WW	GYIPLVGAPLG	R	VARALAHGVRV			135- 145	147- 157
HCV	SP24:	R	GYIPLVGAPLG	RR	VARALAHGVRV			135-	147-

								145	157
HCV	51_BIo tin	RR	GYLPAVGAPIG	BR	VIRVIAHGLRL		m	135- 144	147- 157
HCV	51b_BI otin	RR	GYIPLVGAPLG	BR	VARALAHGVRV			135- 145	147- 157
HCV	51_n		GYIPLVGAPLG	G	VARALAHGVRV			135- 145	147- 157
HCV	SP51_1 :	WW	GYLPAVGAPI	RR	VIRVIAHGLRL		m	135- 144	147- 157
HCV	SP1_C *		GYIPLVGAPLG	G	VARALAHGVRV			135- 145	147- 157
HCV	SP2_c	RR	GYIPLVGAPLG	BGR	VARALAHGVRV			135- 145	147- 157
HCV	SP3_c	R	GYIPLVGAPLG	RR	VARALAHGVRV			135- 145	147- 157
HCV	SP4_c	R	GYIPLVGAPLG	RRR	VARALAHGVRV			135- 145	147- 157
HCV	SP5_c	RR	GYIPLVGAPLG	RR	VARALAHGVRV			135- 145	147- 157
HCV	SP6_c	RR	GYIPLVGAPLG	RRR	VARALAHGVRV			135- 145	147- 157
HCV	SP7_c	BR	GYIPLVGAPLG	RR	VARALAHGVRV			135- 145	147- 157
HCV	SP8_c	RRR	GYIPLVGAPLG	BR	VARALAHGVRV			135- 145	147- 157
HCV	SP9_c	R	GYIPLVGAPLG	KKK	VARALAHGVRV			135- 145	147- 157
HCV	SP10_c	R	GYIPLVGAPLG	RRR	VARALAHGVRV			135- 145	147- 157
HCV	SP11_c	KK	GYIPLVGAPLG	KK	VARALAHGVRV			135- 145	147- 157
HCV	SP12_c	W	GYIPLVGAPLG	RR	VARALAHGVRV			135- 145	147- 157
HCV	SP13_c	WW	GYIPLVGAPLG	RR	VARALAHGVRV			135- 145	147- 157
HCV	SP17_c	RR	GYIPLVGAPLG	LRR	VARALAHGVRV			135- 145	147- 157
HCV	SP61_2	RR	NYVTGNIPG	BR	GITFSIFLIVS			163- 171	171- 181
HCV	SP61b_2	WW	NYATGNLPG	RR	CSFSIFLLAL		m	163- 171	171- 181
HCV	SP61_3	WW	NYVTGNIPG	BR	GITFSIFLIVS			163- 171	171- 181
HCV	SP61_4	WW	NYVTGNIPG	RR	GITFSIFLIVS			163- 171	171- 181
HCV	61b_BI otin	RR	NYATGNLPG	RR	GCSFSIFLLAL			163- 171	171- 181
HCV	SP25	RR	VTGNIPGSTYS	GBR	GITFSIYLIVS		m	165- 175	171- 181
HCV	42_BIo tin	RR	IRNLGRVIETLTG	BR	LZGYIPLIGA		m	116- 128	133- 142
HCV	42b_BI otin	RR	SRNLGKVIDTLT C	BR	LMGYIPLVGA			116- 128	133- 142
HCV	42n- BIOTIN		SRNLGKVIDTLT C	GFAD	LMGYIPLVGA			116- 129	133- 142
HCV	SP42_1	WW	IRNLGRVIETLT	RR	LZGYIPLIGA		m	116- 128	133- 142
HCV	SP42b_1	WW	SRNLGKVIDTLT C	RR	LMGYIPLVGA			116- 129	133- 142

HCV	BI310-11_Biotin	RR	GGGQIIGGNYLI P	RB	PBIGVRATB			26-38	42-50
HCV	BI310-11n_Biotin		GGGQIVGGVYLL P	RR	GPRLGVRATR			26-38	42-50
HCV	BI310-11n_sc_Biotin	RR	GGGQIVGGVYLL P	RR	GPRLGVRATR			26-38	42-50
HCV	SP11b-1-	WW	GGGQIVGGVYLL P	RR	GPRLGVRAT			26-38	42-50
FLU	BI100-12	BR	LIFLARSALIV		RGSVAHKS			256-266	267-274
FLU	BI100-22b	ED	LIFLARSALIL		RGSVAHKS			255-266	267-274
FLU	120b_Biotin	BR	LIFLARSALIL	BGR	SALILRGSVAHK			255-266	267-274
FLU	BI100-18b		SAYERMCNIL	KGK	FQTAAQRAMM			217-226	230-239
FLU	BI100-19		SAYERZVNIL	KGK	FQTAAQRAVZ			217-226	230-239
FLU	190_Biotin	BR	TAYERZCNIL	BRGR	FQTVVQBA			217-226	230-237
FLU	190b_Biotin	BR	IAYERMCNIL	LBRGK	FQTAAQRA			217-226	230-237
FLU	190n-BIOTIN		IAYERMCNIL	KGK	FQTAAQRA			217-226	230-237
FLU	BI100-24b		LFFKCIYRLFKHG L	KR	GPSTEGVPESM			46-59	62-72
FLU	BI100-26	BRR	LFFKTITRLFBHG L	RR	LLSTEGVPNSZ			46-59	62-72
FLU	260_Biotin	BR	GLEPLVIAGILA	RR	GSLVGLLHIVL			23-33	30-40
FLU	260b_Biotin	BR	GSDPLVVAASIV	RR	ASIVGILHLIL			23-33	30-40
CMV	BI 050-sc1	R	NLVPMVATV	RR	NLVPMVATV	B		485-493	485-493
CMV	BI 050-sc2	R	NLVPMVATV	BRR	NLVPMVATV	B		485-493	485-493
CMV	BI 050-sc5	R	NIVPZVVTA	RR	NIVPZVVTA	B	m	485-493	485-493
HIV	N10		PEVIPMFSALS	EGA	TPQDLNNTMLN				
HIV	V10	R	FIIPXFTALSG	GRR	ALLYGATPYAIG				
HIV	N13	K	ALGPAATL	EE	MMTACQGVG				
Neg c mod	SP_18	RR	GPVVHLTL	RRR	GQAGDDFS				
Neg c mod	SP_19	RR	GPVVHLTL	RRR	GQAGDDFS				
Neg c	SP_20	RR	GPVVHLTL	RGR	GQAGDDFS				

mod									
HPV		RR	LECVYCKQQL	RR	EVYDFAFRDLC			35-45	48-58
HPV		RR	GVYDFAFRDLC	RR	GFAFRDLCIVY	R		49-58	52-61
HPV		RR	GVFDYAFRDIN	RR	GFAYRDINLAY	R		49-58	52-61
CMV		RR	GATPVDLLGA	RR	GALNLCLPM	R		498-506	505-514
CMV		RR	GVTPAGLIGV	RR	GALQIBLPL	R		498-506	505-514
HPV		RR	VDIRTLEDLL	RR	GTLGIVCPIG	R		74-83	84-93

As used herein the one-letter-code 'Z' refers to the non-natural amino acid norleucine.

Table 2 – Specific peptides according to the invention

Antigen	X ¹	X ²	X ³	X ⁴	X ⁵
HCV	R	GYIPLVGAPLG	RRR	VARALAHGVRV	R
HCV	R	GYLPAVGAPIG	RRR	VIRVIAHGLRL	R
HCV	RR	GYIPLVGAPLG	RR	VARALAHGVRV	
HCV	RR	GYIPLVGAPLG	RRR	VARALAHGVRV	
HCV	RR	SRNLGKVIDTLTC	RR	LMGYIPLVGA	
HCV	RR	GGGQIVGGVYLLP	RR	GPRLGVRATR	
HCV	W	GYIPLVGAPLG	RR	VARALAHGVRV	
HCV	RR	IRNLGRVIETLTLZGYIPLIGA	RR	IRNLGRVIETLTLZGYIPLIGA	R
Flu	BR	TAYERZCNIL	BRGR	FQTVVQBA	
cmv	R	NLVPMVATV	BRR	NLVPMVATV	B

As used herein the one-letter-code 'Z' refers to the non-natural amino acid norleucine.

5

Antigens

Examples of viral antigens for use with the present invention include, but are not limited to, e.g., HIV, HCV, CMV, HPV, Influenza, adenoviruses, retroviruses, picornaviruses, etc. Non-limiting example of retroviral antigens such as retroviral antigens from the human immunodeficiency virus (HIV) antigens such as gene products of the gag, pol, and env genes, the Nef protein, reverse transcriptase, and other HIV components; hepatitis viral antigens such as the S, M, and L proteins of hepatitis B virus, the pre-S antigen of hepatitis B virus, and other hepatitis, e.g., hepatitis A, B, and C, viral components such as hepatitis C viral RNA; influenza viral antigens such as hemagglutinin and neuraminidase and other influenza viral components; measles viral antigens such as the measles virus fusion protein and other measles virus components; rubella viral antigens such as proteins E1 and E2 and other rubella virus components; rotaviral antigens such as VP7sc and other rotaviral components;

10

15

cytomegaloviral antigens such as envelope glycoprotein B and other cytomegaloviral antigen components; respiratory syncytial viral antigens such as the RSV fusion protein, the M2 protein and other respiratory syncytial viral antigen components; herpes simplex viral antigens such as immediate early proteins, glycoprotein D, and other herpes simplex viral antigen components; varicella zoster viral antigens such as gpl, gpII, and other varicella zoster viral antigen components; Japanese encephalitis viral antigens such as proteins E, M-E, M-E-NSI, NSI, NS1-NS2A, 80% E, and other Japanese encephalitis viral antigen components; rabies viral antigens such as rabies glycoprotein, rabies nucleoprotein and other rabies viral antigen components. See Fundamental Virology, Second Edition, eds. Fields, B. N. and Knipe, D. M. (Raven Press, New York, 1991) for additional examples of viral antigens.

The epitopes to be incorporated into the scaffold design according to the present invention may be derived from an adenovirus, retrovirus, picornavirus, herpesvirus, rotavirus, hantavirus, coronavirus, togavirus, flavivirus, rhabdovirus, paramyxovirus, orthomyxovirus, bunyavirus, arenavirus, reovirus, papillomavirus, parvovirus, poxvirus, hepadnavirus, or spongiform virus. In certain specific, non-limiting examples, the viral antigen are peptides obtained from at least one of HIV, CMV, hepatitis A, B, and C, influenza, measles, polio, smallpox, rubella; respiratory syncytial, herpes simplex, varicella zoster, Epstein-Barr, Japanese encephalitis, rabies, Influenza, and/or cold viruses.

HCV:

Peptides according to the present invention may comprise a known antigen. For antigens derived from HCV these antigens may be derived from the Core, E1, E2, P7, NS2, NS3, NS4 (NS4A and NS4B) and NS5 (NS5A and NS5B) protein of the Hepatitis C Virus (HCV). The epitopes are those which elicit a HLA class I and/or class II restricted T lymphocyte response in an immunized host. More specific, the HLA class I restricted peptides of the present invention may bind to at least one HLA molecule of the following HLA class I groups: HLA-A*01, HLA-A*02, HLA-A*03, HLA-A*11, HLA-A*24, HLA-B*07, HLA-B*08, HLA-B*35, HLA-B*40, HLA-B*44, HLA-Cw3, HLA-Cw4, HLA-Cw6 or HLA-Cw7. The HLA class II restricted peptides of the present invention may bind to at least one HLA molecule of the following HLA class II groups: HLA-DRB1, -DRB2, -DRB3, -DRB4, -DRB5, -DRB6, -DRB7, -DRB8 or -DRB9.

MHC binding HCV peptides that may be used according to the present invention as epitopes are disclosed in e.g. WO02/34770 (Imperial College Innovations Ltd), WO01/21189 and WO02/20035 (Epimmune), WO04/024182 (Intercell), WO95/25122 (The Scripps Research Institute), WO95/27733 (Government of the USA, Department of Health and Human Services), EP 0935662 (Chiron), WO02/26785 (Immunusystems GmbH), WO95/12677 (Innogenetics N.V.), WO97/34621 (Cytel Corp), and EP 1652858 (Innogenetics N.V.).

In other embodiments, the scaffold design according to the present invention comprises a PADRE peptide, such as the universal T cell epitope called PADRE as disclosed in WO95/07707 (Epimmune) the content of which are enclosed herein by reference. A 'PanDR binding peptide or PADRE peptide' is a member of a family of molecules that binds more than
 5 one HLA class II DR molecule. PADRE binds to most HLA-DR molecules and stimulates in vitro and in vivo human helper T lymphocyte (HTL) responses. Alternatively T-help epitopes can be used from universally used vaccines such as tetanus toxoid.

In a further embodiment, the peptides in the composition or polyepitopic peptide are characterized in that they are derived from a HCV protein, or more specifically from at least
 10 one of the following HCV regions selected from the group consisting of Core, E1, E2/NS1, NS2, NS3, NS4A, NS4B, NS5A and NS5B. Even more preferred is that peptides are characterized in that they are present in the HCV consensus sequence of genotype 1a, 1b and/or 3a.

Other HLA class I and II binding peptides that may be used according to the invention may
 15 be identified by the method as described in WO03/105058 -Algonomics, by the method as described by Epimmune in WO01/21189 and/or by three public database prediction servers, respectively Syfpeithi, BIMAS and nHLAPred. It is also an aspect of this present invention that each peptide may be used within the scaffold design of the invention in combination with the same peptide as multiple repeats, or with any other peptide(s) or epitope(s).

20 CMV:

The epitopes to be incorporated into the scaffold design according to the present invention may be derived from cytomegalovirus (CMV) including CMV glycoproteins gB and gH.

Influenza:

The epitopes to be incorporated into the scaffold design according to the present invention
 25 may be derived from fragments or portions of Influenza hemagglutinin (HA) or Influenza neuraminidase (NA), nucleoprotein (NP), M1, M2, NS1, NEP, PA, PB1, PB1-F2, PB2 for each of the subgroups, such as H1N1, H2N2 og H3N2.

Suitable epitopes may be derived from an HA protein of one, or more than one subtype,
 30 including H1, H2, H3, H4, H5, H6, H7, H8, H9, H10, H11, H12, H13, H14, H15 or H16 or fragment or portion thereof. Examples of subtypes comprising such HA proteins include A/New Caledonia/20/99 (H1N1) A/Indonesia/5/2006 (H5N1), A/chicken/New York/1995, A/herring gull/DE/677/88 (H2N8), A/Texas/32/2003, A/mallard/MN/33/00,

A/duck/Shanghai/1/2000, A/northern pintail/TX/828189/02, A/Turkey/Ontario/6118/68 (H8N4), A/shoveler/Iran/G54/03, A/chicken/Germany/N/1949 (H10N7), A/duck/England/56 (H11N6), A/duck/Alberta/60/76 (H12N5), A/Gull/Maryland/704/77 (H13N6), A/Mallard/Gurjev/263/82, A/duck/Australia/341/83 (H15N8), A/black-headed
 5 gull/Sweden/5/99 (H16N3), B/Lee/40, C/Johannesburg/66, A/PuertoRico/8/34 (H1N1), A/Brisbane/59/2007 (H1N1), A/Solomon Islands 3/2006 (H1N1), A/Brisbane 10/2007 (H3N2), A/Wisconsin/67/2005 (H3N2), B/Malaysia/2506/2004, B/Florida/4/2006, A/Singapore/1/57 (H2N2), A/Anhui/1/2005 (H5N1), A/Vietnam/1194/2004 (H5N1), A/Teal/HongKong/W312/97 (H6N1), A/Equine/Prague/56 (H7N7), A/HongKong/1073/99
 10 (H9N2)).

In some embodiments of the invention, the HA protein may be an H1, H2, H3, H5, H6, H7 or H9 subtype. In other embodiments, the H1 protein may be from the A/New Caledonia/20/99 (H1N1), A/PuertoRico/8/34 (H1N1), A/Brisbane/59/2007 (H1N1), or A/Solomon Islands
 15 3/2006 (H1N1) strain. The H3 protein may also be from the A/Brisbane 10/2007 (H3N2) or A/Wisconsin/67/2005 (H3N2) strain. In other embodiments, the H2 protein may be from the A/Singapore/1/57 (H2N2) strain. The H5 protein may be from the A/Anhui/1/2005 (H5N1), A/Vietnam/1194/2004 (H5N1), or A/Indonesia/5/2005 strain. In other embodiments, the H6 protein may be from the A/Teal/HongKong/W312/97 (H6N1) strain. The H7 protein may be
 20 from the A/Equine/Prague/56 (H7N7) strain. In other embodiments, the H9 protein is from the A/HongKong/1073/99 (H9N2) strain. In other embodiments, the HA protein may be from an influenza virus may be a type B virus, including B/Malaysia/2506/2004 or B/Florida/4/2006. The influenza virus HA protein may be H5 Indonesia.

25 Human immunodeficiency virus (HIV):

For HIV, the epitopes to be incorporated into the scaffold design according to the present invention may be derived from the group consisting of gp120, gp160, gp41, p24gag or p55gag derived from HIV, including members of the various genetic subtypes.

Human papillomavirus (HPV):

30 For HPV, the epitopes to be incorporated into the scaffold design according to the present invention may be derived from the group consisting E1, E2, E3, E4, E6 and E7, L1 and L2 proteins. The epitopes may be derived from any type including types 8, 11, 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, and 59.

Carriers, adjuvants and vehicles - delivery

- The isolated cell-penetrating peptides according to the invention may be delivered by various means and within various compositions, herein referred to as "compositions", "vaccine compositions" or "pharmaceutical compositions". The peptides of the present invention and pharmaceutical and vaccine compositions of the invention are useful for administration to mammals, particularly humans, to treat and/or prevent virus infection. Vaccine compositions containing the peptides of the invention are administered to a patient infected with the virus in question or to an individual susceptible to, or otherwise at risk for, virus infection to elicit an immune response against the specific antigens and thus enhance the patient's own immune response capabilities.
- Various art-recognized delivery systems may be used to deliver the peptides, into appropriate cells. The peptides can be delivered in a pharmaceutically acceptable carrier or as colloidal suspensions, or as powders, with or without diluents. They can be "naked" or associated with delivery vehicles and delivered using delivery systems known in the art.
- A "pharmaceutically acceptable carrier" or "pharmaceutically acceptable adjuvant" is any suitable excipient, diluent, carrier and/or adjuvant which, by themselves, do not induce the production of antibodies harmful to the individual receiving the composition nor do they elicit protection. Preferably, a pharmaceutically acceptable carrier or adjuvant enhances the immune response elicited by an antigen. Suitable carriers or adjuvant typically comprise one or more of the compounds included in the following non-exhaustive list: large slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers and inactive virus particles; aluminium hydroxide, aluminium phosphate (see International Patent Application Publication No. WO93/24148), alum ($KAl(SO_4)_2 \cdot 12H_2O$), or one of these in combination with 3-O-deacylated monophosphoryl lipid A (see International Patent Application Publication No. WO93/19780); N-acetyl-muramyl-L-threonyl-D-isoglutamine (see U.S. Patent No. 4,606,918), N-acetyl-normuramyl-L-alanyl-D-isoglutamine, N-acetylmuramyl-L-alanyl-D-isoglutamyl-L-alanine2-(1',2'-dipalmitoyl-sn-glycero-3-hydroxyphosphoryloxy) ethylamine; RIBI (ImmunoChem Research Inc., Hamilton, MT, USA) which contains monophosphoryl lipid A (i.e., a detoxified endotoxin), trehalose-6,6-dimycolate, and cell wall skeleton (MPL + TDM + CWS) in a 2% squalene/Tween 80 emulsion. Any of the three components MPL, TDM or CWS may also be used alone or combined 2 by 2; adjuvants such as Stimulon (Cambridge Bioscience, Worcester, MA, USA), SAF-1 (Syntex); adjuvants such as combinations between QS21 and 3-de-O-acetylated monophosphoryl lipid A (see International Application No. WO94/00153) which may be further supplemented with an oil-in-water emulsion (see, e.g., International Application Nos. WO95/17210, WO97/01640 and WO9856414) in which the oil-in-water emulsion comprises a metabolisable oil and a saponin, or a metabolisable oil, a saponin, and a sterol, or which may be further supplemented with a cytokine (see International Application

- No. WO98/57659); adjuvants such as MF-59 (Chiron), or poly[di(carboxylatophenoxy) phosphazene] based adjuvants (Virus Research Institute); blockcopolymer based adjuvants such as Optivax (Vaxcel, Cytrx) or inulin-based adjuvants, such as Algammulin and Gammalnulin (Anutech); Complete or Incomplete Freund's Adjuvant (CFA or IFA, respectively) or Gerbu preparations (Gerbu Biotechnik); a saponin such as QuilA, a purified saponin such as QS21, QS7 or QS17, -escin or digitonin; immunostimulatory oligonucleotides comprising unmethylated CpG dinucleotides such as [purine-purine-CG-pyrimidine-pyrimidine] oligonucleotides. These immunostimulatory oligonucleotides include CpG class A, B, and C molecules (Coley Pharmaceuticals), ISS (Dynavax), Immunomers (Hybridon).
- 10 Immunostimulatory oligonucleotides may also be combined with cationic peptides as described, e.g., by Riedl et al. (2002); Immune Stimulating Complexes comprising saponins, for example Quil A (ISCOMS); excipients and diluents, which are inherently non-toxic and non-therapeutic, such as water, saline, glycerol, ethanol, isopropyl alcohol, DMSO, wetting or emulsifying agents, pH buffering substances, preservatives, and the like; a biodegradable and/or biocompatible oil such as squalane, squalene, eicosane, tetratetracontane, glycerol, 15 peanut oil, vegetable oil, in a concentration of, e.g., 1 to 10% or 2,5 to 5%; vitamins such as vitamin C (ascorbic acid or its salts or esters), vitamin E (tocopherol), or vitamin A; carotenoids, or natural or synthetic flavanoids; trace elements, such as selenium; any Toll-like receptor ligand as reviewed in Barton and Medzhitov (2002).
- 20 Any of the afore-mentioned adjuvants comprising 3-de-O-acetylated monophosphoryl lipid A, said 3-de-O-acetylated monophosphoryl lipid A may be forming a small particle (see International Application No. WO94/21292).
- In any of the aforementioned adjuvants MPL or 3-de-O-acetylated monophosphoryl lipid A can be replaced by a synthetic analogue referred to as RC-529 or by any other amino-alkyl 25 glucosaminide 4-phosphate (Johnson et al. 1999, Persing et al. 2002). Alternatively it can be replaced by other lipid A analogues such as OM-197 (Byl et al. 2003).
- A "pharmaceutically acceptable vehicle" includes vehicles such as water, saline, physiological salt solutions, glycerol, ethanol, etc. Auxiliary substances such as wetting or emulsifying agents, pH buffering substances, preservatives may be included in such vehicles. Delivery 30 systems known in the art are e.g. lipopeptides, peptide compositions encapsulated in poly-DL-lactide-co-glycolide ("PLG"), microspheres, peptide compositions contained in immune stimulating complexes (ISCOMS), multiple antigen peptide systems (MAPs), viral delivery vectors, particles of viral or synthetic origin, adjuvants, liposomes, lipids, microparticles or microcapsules, gold particles, nanoparticles, polymers, condensing agents, polysaccharides, 35 polyamino acids, dendrimers, saponins, QS21, adsorption enhancing materials, fatty acids or, naked or particle absorbed cDNA.

Typically, a vaccine or vaccine composition is prepared as an injectable, either as a liquid solution or suspension. Injection may be subcutaneous, intramuscular, intravenous, intraperitoneal, intrathecal, intradermal, or intraepidermal. Other types of administration comprise electroporation, implantation, suppositories, oral ingestion, enteric application, inhalation, aerosolization or nasal spray or drops. Solid forms, suitable for dissolving in, or suspension in, liquid vehicles prior to injection may also be prepared. The preparation may also be emulsified or encapsulated in liposomes for enhancing adjuvant effect.

A liquid formulation may include oils, polymers, vitamins, carbohydrates, amino acids, salts, buffers, albumin, surfactants, or bulking agents. Preferably carbohydrates include sugar or sugar alcohols such as mono-, di-, tri-, oligo- or polysaccharides, or water-soluble glucans. The saccharides or glucans can include fructose, dextrose, lactose, glucose, mannose, sorbose, xylose, maltose, sucrose, dextran, pullulan, dextrin, alpha and beta cyclodextrin, soluble starch, hydroxethyl starch and carboxymethylcellulose, or mixtures thereof. Sucrose is most preferred. "Sugar alcohol" is defined as a C4 to C8 hydrocarbon having an -OH group and includes galactitol, inositol, mannitol, xylitol, sorbitol, glycerol, and arabitol. Mannitol is most preferred. These sugars or sugar alcohols mentioned above may be used individually or in combination. There is no fixed limit to the amount used as long as the sugar or sugar alcohol is soluble in the aqueous preparation. Preferably, the sugar or sugar alcohol concentration is between 1,0 % (w/v) and 7,0 % (w/v), more preferable between 2,0 and 6,0 % (w/v). Preferably amino acids include levorotary (L) forms of carnitine, arginine, and betaine; however, other amino acids may be added. Preferred polymers include polyvinylpyrrolidone (PVP) with an average molecular weight between 2,000 and 3,000, or polyethylene glycol (PEG) with an average molecular weight between 3,000 and 5,000. It is also preferred to use a buffer in the composition to minimize pH changes in the solution before lyophilization or after reconstitution. Any physiological buffer may be used, but citrate, phosphate, succinate, and glutamate buffers or mixtures thereof are preferred. Most preferred is a citrate buffer. Preferably, the concentration is from 0,01 to 0,3 molar. Surfactants that can be added to the formulation are shown in EP patent applications No. EP 0 270 799 and EP 0 268 110.

Additionally, polypeptides can be chemically modified by covalent conjugation to a polymer to increase their circulating half-life, for example. Preferred polymers, and methods to attach them to peptides, are shown in U.S. Patent Nos. 4,766,106; 4,179,337; 4,495,285; and 4,609,546. Preferred polymers are polyoxyethylated polyols and polyethylene glycol (PEG). PEG is soluble in water at room temperature and has the general formula:

$R(O-CH_2-CH_2)_nO-R$ where R can be hydrogen, or a protective group such as an alkyl or alkanol group. Preferably, the protective group has between 1 and 8 carbons, more

preferably it is methyl. The symbol n is a positive integer, preferably between 1 and 1.000, more preferably between 2 and 500. The PEG has a preferred average molecular weight between 1000 and 40.000, more preferably between 2000 and 20.000, most preferably between 3.000 and 12.000. Preferably, PEG has at least one hydroxy group, more preferably it is a terminal hydroxy group. It is this hydroxy group which is preferably activated. However, it will be understood that the type and amount of the reactive groups may be varied to achieve a covalently conjugated PEG/polypeptide of the present invention.

Water soluble polyoxyethylated polyols are also useful in the present invention. They include polyoxyethylated sorbitol, polyoxyethylated glucose, polyoxyethylated glycerol (POG), etc. POG is preferred. One reason is because the glycerol backbone of polyoxyethylated glycerol is the same backbone occurring naturally in, for example, animals and humans in mono-, di-, triglycerides. Therefore, this branching would not necessarily be seen as a foreign agent in the body. The POG has a preferred molecular weight in the same range as PEG. The structure for POG is shown in Knauf et al., 1988, and a discussion of POG/IL-2 conjugates is found in U.S. Patent No. 4,766,106.

Another drug delivery system for increasing circulatory half-life is the liposome. The peptides and nucleic acids of the invention may also be administered via liposomes, which serve to target a particular tissue, such as lymphoid tissue, or to target selectively infected cells, as well as to increase the half-life of the peptide and nucleic acids composition. Liposomes include emulsions, foams, micelles, insoluble monolayers, liquid crystals, phospholipid dispersions, lamellar layers and the like. In these preparations, the peptide or nucleic acids to be delivered is incorporated as part of a liposome or embedded, alone or in conjunction with a molecule which binds to a receptor prevalent among lymphoid cells, such as monoclonal antibodies which bind to the CD45 antigen, or with other therapeutic or immunogenic compositions. Thus, liposomes either filled or decorated with a desired peptide or nucleic acids of the invention can be directed to the site of lymphoid cells, where the liposomes then deliver the peptide and nucleic acids compositions. Liposomes for use in accordance with the invention are formed from standard vesicle-forming lipids, which generally include neutral and negatively charged phospholipids and a sterol, such as cholesterol. The selection of lipids is generally guided by consideration of, e.g., liposome size, acid lability and stability of the liposomes in the blood stream. A variety of methods are available for preparing liposomes, as described in, e.g., Szoka et al, 1980, and U.S. Patent Nos. 4,235,871, 4,501,728, 4,837,028, and 5,019,369.

For targeting cells of the immune system, a ligand to be incorporated into the liposome can include, e.g., antibodies or fragments thereof specific for cell surface determinants of the desired immune system cells. A liposome suspension containing a peptide may be

administered intravenously, locally, topically, etc. in a dose which varies according to, inter alia, the manner of administration, the peptide being delivered, and the stage of the disease being treated. For example, liposomes carrying either immunogenic polypeptides are known to elicit CTL responses in vivo (Reddy et al., 1992; Collins et al., 1992; Fries et al., 1992; 5 Nabel et al., 1992).

After the liquid pharmaceutical composition is prepared, it is preferably lyophilized to prevent degradation and to preserve sterility. Methods for lyophilizing liquid compositions are known to those of ordinary skill in the art. Just prior to use, the composition may be reconstituted with a sterile diluent (Ringer's solution, distilled water, or sterile saline, for example) which 10 may include additional ingredients. Upon reconstitution, the composition is preferably administered to subjects using those methods that are known to those skilled in the art.

Use of the peptides for evaluating immune responses.

The peptides according to the present invention may be used as diagnostic reagents. For example, a peptide of the invention may be used to determine the susceptibility of a 15 particular individual to a treatment regimen which employs the peptide or related peptides, and thus may be helpful in modifying an existing treatment protocol or in determining a prognosis for an affected individual. In addition, the peptides may also be used to predict which individuals will be at substantial risk for developing a chronic virus infection.

Accordingly, the present invention relates to a method of determining the outcome for a 20 subject exposed to a virus, comprising the steps of determining whether the subject has an immune response to one or more peptides according to the present invention.

In a preferred embodiment of the invention, the peptides as described herein can be used as reagents to evaluate an immune response. The immune response to be evaluated can be induced by using as an immunogen any agent that may result in the production of antigen-specific CTLs or HTLs that recognize and bind to the peptide(s) to be employed as the 25 reagent. The peptide reagent need not be used as the immunogen. Assay systems that can be used for such an analysis include relatively recent technical developments such as tetramers, staining for intracellular lymphokines and interferon release assays, or ELISPOT assays.

30 For example, a peptide of the invention may be used in a tetramer staining assay to assess peripheral blood mononuclear cells for the presence of antigen-specific CTLs following exposure to an antigen or an immunogen. The HLA- tetrameric complex is used to directly visualize antigen-specific CTLs (see, e.g., Ogg et al., 1998; and Altman et al., 1996) and

determine the frequency of the antigen-specific CTL population in a sample of peripheral blood mononuclear cells. A tetramer reagent using a peptide of the invention may be generated as follows: a peptide that binds to an HLA molecule is refolded in the presence of the corresponding HLA heavy chain and beta2-microglobulin to generate a trimolecular
5 complex. The complex is biotinylated at the carboxyl terminal end of the heavy chain at a site that was previously engineered into the protein. Tetramer formation is then induced by the addition of streptavidin. By means of fluorescently labeled streptavidin, the tetramer can be used to stain antigen-specific cells. The cells may then be identified, for example, by flow cytometry. Such an analysis may be used for diagnostic or prognostic purposes. Cells
10 identified by the procedure can also be used for therapeutic purposes. As an alternative to tetramers also pentamers or dimers can be used (Current Protocols in Immunology (2000) unit 17.2 supplement 35)

Peptides of the invention may also be used as reagents to evaluate immune recall responses. (see, e.g., Bertoni et al., 1997 and Perma et al., 1991.). For example, patient PBMC samples
15 from individuals with HCV infection may be analyzed for the presence of antigen-specific CTLs or HTLs using specific peptides. A blood sample containing mononuclear cells may be evaluated by cultivating the PBMCs and stimulating the cells with a peptide of the invention. After an appropriate cultivation period, the expanded cell population may be analyzed, for example, for cytotoxic activity (CTL) or for HTL activity.

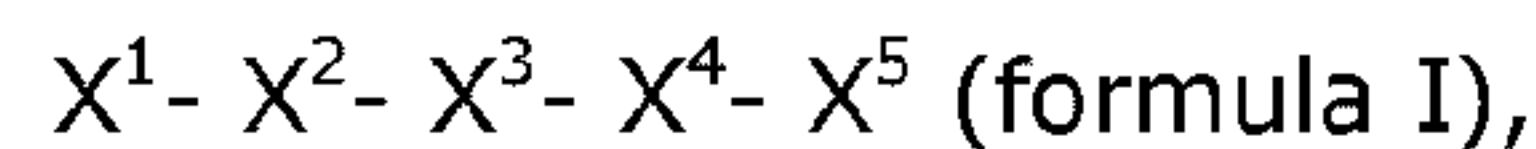
20 The peptides may also be used as reagents to evaluate the efficacy of a vaccine.

PBMCs obtained from a patient vaccinated with an immunogen may be analyzed using, for example, either of the methods described above. The patient is HLA typed, and peptide epitope reagents that recognize the allele-specific molecules present in that patient are selected for the analysis. The immunogenicity of the vaccine is indicated by the presence of
25 epitope-specific CTLs and/or HTLs in the PBMC sample.

The peptides of the invention may also be used to make antibodies, using techniques well known in the art (see, e.g. CURRENT PROTOCOLS IN IMMUNOLOGY, Wiley/Greene, NY; and Antibodies A Laboratory Manual, Harlow and Lane, Cold Spring Harbor Laboratory Press, 1989). Such antibodies include those that recognize a peptide in the context of an HLA
30 molecule, i.e., antibodies that bind to a peptide-MHC complex.

Specific embodiments of the invention

The present invention concerns a peptide an isolated cell-penetrating peptide comprising the following structure



- 5 wherein X^1 and X^3 independently defines a linear sequence of any 1, 2, 3 or 4 amino acid independently selected from any basic amino acid, citrulline, tryptophan, or a derivative thereof; X^2 defines a linear sequence of 8-30 amino acids derived from an antigen; X^4 defines a linear sequence of 8-30 amino acids derived from said antigen, said sequence X^4 being different from X^2 ; and wherein X^5 is any one optional amino acid selected from a basic amino
10 acid, citrulline, tryptophan, or a derivative thereof.

In some specific embodiments the sequence of X^4 is identical to the sequence of X^2 .

In some embodiments the isolated cell-penetrating peptide has a total of not more than 60 amino acids.

- 15 In some embodiments X^1 defines a linear sequence of any 1, 2, 3, or 4 amino acids independently selected from any basic amino acid, tryptophan, or a derivative thereof.

In some embodiments X^1 defines a linear sequence in any order of one citrulline and any 1, 2, or 3 amino acids independently selected from any basic amino acid, tryptophan, or a derivative thereof.

- 20 In some embodiments X^3 defines a linear sequence of any 1, 2, 3, or 4 amino acids independently selected from any basic amino acid.

In some embodiments X^1 defines a linear sequence of any 1, 2, 3, or 4 amino acids independently selected from any basic amino acid.

In some embodiments X^3 defines a linear sequence in any order of one citrulline and of any 1, 2, or 3 amino acids independently selected from any basic amino acid.

- 25 In some embodiments the sequence of amino acids defined by $X^2 - X^3 - X^4$ of formula I as defined in claim 1 is not found in the native sequence of said antigen.

In some embodiments the sequence of amino acids defined by X^1 - X^2 - X^3 - X^4 of formula I as defined in claim 1 is not found in the native sequence of said antigen.

In some embodiments the sequence of amino acids defined by X^1 - X^2 - X^3 - X^4 - X^5 of formula I as defined in claim 1 is not found in the native sequence of said antigen.

- 5 In some embodiments the peptide is demonstrated to translocate across a plasma membrane in the assay based on biotinylation of peptides as described in example 1.

In some embodiments the 1, 2, 3 or 4 amino acid independently selected from any basic amino acid, citrulline, tryptophan, or a derivative thereof, is a basic amino acid.

- 10 In some embodiments the 1, 2, 3 or 4 amino acid independently selected from any basic amino acid, citrulline, tryptophan, or a derivative thereof, is tryptophan, or a derivative thereof.

In some embodiments the 1, 2, 3 or 4 amino acid independently selected from any basic amino acid, citrulline, tryptophan, or a derivative thereof, is citrulline, or a derivative thereof.

- 15 In some embodiments the one optional amino acid selected from a basic amino acid, citrulline, tryptophan, or a derivative thereof is selected from Arg, Lys, and His.B).

In some embodiments the one optional amino acid selected from a basic amino acid, citrulline, tryptophan or a derivative thereof is tryptophan, or a derivative thereof.

In some embodiments the one optional amino acid selected from a basic amino acid, citrulline, tryptophan or a derivative thereof is citrulline, or a derivative thereof.

- 20 In some embodiments X^2 and/or X^4 defines a sequence identical to the native sequence of said antigen.

In some embodiments the peptide is capable of inducing a T-lymphocyte response.

In some embodiments the peptide is capable of inducing a CD4+ and/or a CD8+ T-lymphocyte response.

- 25 In some embodiments the antigen is a viral protein, such as a capsid protein.

In some embodiments the viral protein is selected from a protein of the Hepatitis C virus, such as a core protein; protein of influenza virus, such as an M2 protein.

In some embodiments the viral protein of Hepatitis C virus is selected from HCV consensus sequence of genotype 1, such as subtypes 1a and 1b, genotype 2 such as 2a and 2b and
5 genotype 3, such as 3a.

In some embodiments the cell-penetrating peptide is of 19-60 amino acids, such as of 20-60 amino acids, such as of 21-60 amino acids, such as of 22-60 amino acids, such as of 23-60 amino acids, such as of 24-60 amino acids, such as of 25-60 amino acids, such as of 26-60 amino acids, such as of 27-60 amino acids, such as of 28-60 amino acids, such as of 29-60
10 amino acids, such as of 30-60 amino acids, such as of 31-60 amino acids, such as of 32-60 amino acids, such as of 33-60 amino acids, such as of 34-60 amino acids, such as of 35-60 amino acids.

In some embodiments the cell-penetrating peptide is of 18-60 amino acids, such as 18-59 amino acids, such as 18-58 amino acids, such as 18-57 amino acids, such as 18-56 amino
15 acids, such as 18-55 amino acids, such as 18-54 amino acids, such as 18-53 amino acids, such as 18-52 amino acids, such as 18-51 amino acids, such as 18-50 amino acids, such as 18-49 amino acids, such as 18-48 amino acids, such as 18-47 amino acids, such as 18-46 amino acids, such as 18-45 amino acids, such as 18-44 amino acids, such as 18-43 amino acids, such as 18-42 amino acids, such as 18-41 amino acids, such as 18-40 amino acids,
20 such as 18-39 amino acids, such as 18-38 amino acids, such as 18-37 amino acids, such as 18-35 amino acids, such as of 18-34 amino acids, such as of 18-33 amino acids, such as of 18-32 amino acids, such as of 18-31 amino acids, such as of 18-30 amino acids, such as of 18-29 amino acids, such as of 18-28 amino acids, such as of 18-27 amino acids, such as of 18-26 amino acids, such as of 18-25 amino acids, such as of 18-24 amino acids, such as of
25 18-23 amino acids, such as of 18-22 amino acids, such as of 18-21 amino acids, such as of 18-20 amino acids, such as of 18-19 amino acids.

In some embodiments the net charge of X^2 is below or equal to 0.

In some embodiments the net charge of X^2 is below or equal to 0; and X^1 and X^3 defines a sequence of 2, 3 or 4 amino acid independently selected from any basic amino acid, citrulline,
30 tryptophan, or a derivative thereof.

In some embodiments the net charge of X^2 is below or equal to 0; and the net charge of X^4 is above or equal to 1.

In some embodiments the net charge of X^2 is below or equal to 0; the net charge of X^4 is above or equal to 1; and X^1 and X^3 defines a sequence of 2, 3 or 4 amino acid independently selected from any basic amino acid, citrulline, tryptophan, or a derivative thereof.

In some embodiments the net charge of X^2 and X^4 are below or equal to 0.

- 5 In some embodiments the net charge of X^2 and X^4 are below or equal to 0; and X^1 and X^3 defines a sequence of 2, 3 or 4 amino acid independently selected from any basic amino acid, citrulline, tryptophan, or a derivative thereof.

In some embodiments the net charge of X^2 and X^4 are above or equal to 1.

- 10 In some embodiments the net charge of X^2 and X^4 are above or equal to 1 and X^1 and X^3 defines a sequence of 1 or 2 amino acid independently selected from any basic amino acid, citrulline, tryptophan, or a derivative thereof.

In some embodiments the net charge of X^2 is above or equal to 1; and the net charge of X^4 is below or equal to 0.

- 15 In some embodiments the net charge of X^2 is above or equal to 1; the net charge of X^4 is below or equal to 0; X^1 defines a sequence of 1 or 2 amino acids with a positive charge; and X^3 defines a sequence of 2, 3 or 4 amino acid independently selected from any basic amino acid, citrulline, tryptophan, or a derivative thereof.

In some embodiments the basic amino acid is independently selected from Arg, Lys, and His.

- 20 In some embodiments X^2 and/or X^4 comprises the sequence of amino acids defined by position 135-157 of SEQ ID NO:3, or a fragment or variant thereof.

- 25 In some embodiments X^2 and/or X^4 consist of a sequence selected from GYIPLVGAPLG, GYLPAVGAPIG, GYLPAVGAPI, NYVTGNIPG, NYATGNLPG, NYATGNLPG, VTGNIPGSTYS, IRNLGRVIETLTG, SRNLGKVIDTLTC, IRNLGRVIETLT, GGGQIIGGNYLIP, GGGQIVGGVYLLP, LIFLARSALIV, LIFLARSALIL, LIFLARSALIL, SAYERM CNIL, SAYERZVNIL, TAYERZCNIL, IAYERMCNIL, IAYERMCNIL, LFFKCIYRLF KHGL, LFFKTITR LFBHGL, GLEPLVIAGILA, GSDPLVVAASIV, NLVPMVATV, NLVPMVATV, NIVPZVVT A, PEVIPMFSALS, FIIPXFTALSG, ALGPAATL, GPVVHLTL, LECVYCKQQLL, GVDYAFRDLC, GVFDYAFRDIN, GATPVDLLGA, GVTPAGLIGV, VARALAHGVRV, VIRVIAHGLRL, GITFSIFLIVS, CSFSIFLLAL, GCSFSIFLLAL, GITFSIYLIVS, LZGYIPLIGA, LMGYIPLVGA, LZGYIPLIGA, PBIGVRATB, GPRLGVRATR,

GPRLGVRAT, RGSVAHKS, SALILRGSVAHK, FQTAAQRAMM, FQTAAQRAVZ, FQTVVQBA, FQTAAQRA, GPSTEGVPESM, LLSTEGVPNSZ, GSVGLLHIVL, ASIVGILHLIL, NLVPMVATV, NIVPZVVTA, TPQDLNTMLN, ALLYGATPYAIG, MMTACQGVG, GQAGDDFS, EVYDFAFRDLC, GFAFRDLCIVY, GFAYRDINLAY, GALNLCLPM, and GALQIBLPL, IRNLGRVIETLTLZGYIPLIGA, or a
5 fragment or variant thereof.

In some embodiments X^2 and/or X^4 consist of a sequence derived from an amino acid sequence selected from SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:126, and SEQ ID NO:198, or a fragment or variant thereof.
10
15

In some embodiments the peptide according to the invention is selected from RRGYIPLVGAPLGBGRVARALAHGVRV, RGYIPLVGAPLGRRVARALAHGVRV, RGYIPLVGAPLGRRRVARALAHGVRV, RRGYIPLVGAPLGRRVARALAHGVRV, RRGYIPLVGAPLGRRRVARALAHGVRV, BRGYIPLVGAPLGRRVARALAHGVRV, RRRGYIPLVGAPLGBRVARALAHGVRV, RGYIPLVGAPLGKKKVARALAHGVRV, RGYIPLVGAPLGRRRVARALAHGVRV, KKGYIPLVGAPLGKKVARALAHGVRV, WGYIPLVGAPLGRRVARALAHGVRV, WWGYIPLVGAPLGRRVARALAHGVRV, EEGYIPLVGAPLGEEVARALAHGVRV, GGGYIPLVGAPLGGGVARALAHGVRV, EEGYIPLVGAPLGRRVARALAHGVRV, RRGYIPLVGAPLGLRRVARALAHGVRV, WWGYIPLVGAPLGRRVARALAHGVRV, WWGYIPLVGAPLGRRRVARALAHGVRV, WWGYIPLVGAPLGRVARALAHGVRV, RGYIPLVGAPLGRRVARALAHGVRV, RRGYIPLVGAPLGBRVARALAHGVRV, RRGYIPLVGAPLGGVARALAHGVRV, WWGYIPLVGAPLRRVIRVIAHGLRL, GYIPLVGAPLGGVARALAHGVRV, RRGYIPLVGAPLGBGRVARALAHGVRV, RGYIPLVGAPLGRRVARALAHGVRV, RGYIPLVGAPLGRRRVARALAHGVRV, RRGYIPLVGAPLGRRVARALAHGVRV, RRGYIPLVGAPLGRRRVARALAHGVRV, BRGYIPLVGAPLGRRVARALAHGVRV, RRRGYIPLVGAPLGBRVARALAHGVRV, RGYIPLVGAPLGKKKVARALAHGVRV, RGYIPLVGAPLGRRRVARALAHGVRV, KKGYIPLVGAPLGKKVARALAHGVRV, WGYIPLVGAPLGRRVARALAHGVRV, WWGYIPLVGAPLGRRVARALAHGVRV, RRGYIPLVGAPLGLRRVARALAHGVRV, RRNYVTGNIPGBRGITFSIFLIVS, WWNYATGNLPGRRCSFSIFLLAL,
20
25
30
35

WWNYVTGNIPGBRGITFSIFLIVS, WWNYVTGNIPGRRGITFSIFLIVS,
 RRNYATGNLPGRRGCSFSIFLLAL, RRVGTGNIPGSTYSGBRGITFSIYLIVS,
 RRIRNLGRVIETLTGBRLZGYIPLIGA, RRSRNLGKVIDTLTCBRLMGYIPLVGA,
 SRNLGKVIDTLTCGFADLMGYIPLVGA, WWIRNLGRVIETLTRRLZGYIPLIGA,
 5 WWSRNLGKVIDTLTCRRLMGYIPLVGA, RRGGGQIIGGNYLIPRBPBIGVRATB,
 GGGQIVGGVYLLPRRGPRLGVRATR, RRGGGQIVGGVYLLPRRGPRLGVRATR,
 WWGGGQIVGGVYLLPRRGPRLGVRAT, , BRLIFLARSALIVRGSAHKS,
 EDLIFLARSALILRGSAHKS, BRLIFLARSALILBGRSALILRGSAHKS,
 SAYERM CNILKGKFQTAAQRAMM, SAYERZVNILKGKFQTAAQRAVZ,
 10 BRTAYERZCNILBRGRFQTVVQBA, BRIAYERMCNILLBRGKFQTAAQRA,
 IAYERMCNILKGKFQTAAQRA, LFFKCIYRLFKHGLKRG PSTEGVPESM,
 BRRLFFKTITRLFBHGLRRLSTEGVPNSZ, BRGLEPLVIAGILARRGSLVGLLHIVL,
 BRGSDPLVVAASIVRRASIVGILHLIL, , RNLVPMVATVRRNLVPMVATVB,
 RNLVPMVATVBRRNLVPMVATVB, RNIVPZVVTARRNIVPZVVTAB, ,
 15 PEVIPMFSALESEGATPQDLNTMLN, RFIIPXFTALSGGRRALLYGATPYAIG,
 KALGPAATLEEMMTACQGVG, , RRGPVVHLTLRRRGQAGDDFS, RRGPVVHLTLRRRGQAGDDFS,
 RRGPVVHLTLRGRRGQAGDDFS, RRLECVYCKQQLLRREVYDFAFRDLC,
 RRGVYDFAFRDLCRRGFADLCIVYR, RRGVFDYAFRDINRRGFAYRDINLAYR,
 RRGATPVDLLGARRGALNLCLPMR, RRGVTPAGLIGVRRGALQIBLPLR,
 20 RGYLPAVGAPIGRRRVIRVIAHGLRLR, RRSRNLGKVIDTLTCRRLMGYIPLVGA,
 RRIRNLGRVIETLTZGYIPLIGARRIRNLGRVIETLTZGYIPLIGAR, or a fragment or variant
 thereof.

In some embodiments the peptide consist of a sequence selected from X¹-NYVTGNIPG-X³-
 GITFSIYLIVS; X¹-IRNLGRVIETLT-X³-LZGYIPLIGA; X¹-GYLPAVGAPI-X³-VIRVIAHGLRL; X¹-
 25 GGGQIIGGNYLIP-X³-PBIGVRATB; X¹-NYATGNLPG-X³-GCSFSIFLLAL; X¹-SRNLGKVIDTLTC-X³-
 LMGYIPLVGA; X¹-GYIPLVGAPL-X³-VARALAHGVRV; X¹-GGGQIVGGVYLLP-X³-PRLGVRATR; X¹-
 LTFLVRSVLLI-X³-GSVLIVRGSLVH; X¹-TAYERZCNIL-X³-GRFQTVVQBA; X¹-SDPLVVAASIV-X³-
 ASIVGILHLIL; X¹-LIFLARSALIL-X³-SALILRGSAH; X¹-IAYERMCNIL-X³-GKFQTAAQRA; and X¹-
 LEPLVIAGILA-X³-GSLVGLLHIVL; X¹-NLVPMVATV-X³-NLVPMATV; X¹-GYLPAVGAPIG-X³-
 30 VIRVIAHGLRL; X¹-IRNLGRVIETLTG-X³-LZGYIPLIGA; X¹-GVYDFAFRDLC-X³-GFAFRDLCIVYR, X¹-
 GVFDYAFRDIN-X³-GFAYRDINLAYR, X¹-GATPVDLLGA-X³-GALNLCLPMR, X¹-GVTPAGLIGV-X³-
 GALQIBLPLR, and X¹-IRNLGRVIETLTZGYIPLIGA-X³-IRNLGRVIETLTZGYIPLIGA; optionally
 with an X⁵ in the C-terminal of the peptide wherein X¹, X³ and X⁵ refers to X¹, X³, and X⁵ of
 formula I.

35 In some embodiments the peptide comprises one or more cysteine.

In some embodiments the peptide contain intramolecular bonds, such as intramolecular disulfide (S-S) bonds between two cys residues.

In other embodiments the peptide contains intramolecular bonds, such as in the form of a acylal moiety (COO-CH₂-OOC, COO-CHR-OOC or COO-CR₂-OOC).

- 5 In some embodiments the N- and/or C-terminal amino acid in X² is a hydrophilic or polar amino acid.

In some embodiments the N-terminal amino acid in X⁴ is a hydrophilic or polar amino acid.

- In some embodiments the peptide according to the invention is not more than 58 amino acids, such as not more than 56, 54, 52, 50, 48, 46, 44, 42, 40, 38, 36, 34, 32, 30, 28, 26,
10 24, 22, 20, 18 amino acid residues.

In some embodiments X¹ consist of 2 or 3 amino acids.

In some embodiments X¹ consist of WW, BR, or RR.

In some embodiments X¹ consist of R, or W.

In some embodiments X³ consist of 2 or 3 amino acids.

- 15 In some embodiments X³ consist of WW, BR, or RR.

In some embodiments X³ consist of RRR, BRR, or BRGR.

In some embodiments an isolated peptide according to the invention consists of a sequence of X² or X⁴ as defined in table 1 or in table 2.

- In some embodiments an isolated peptide according to the invention comprises a sequence of
20 X² and/or X⁴ as defined in table 1 or in table 2, or a fragment thereof.

In some embodiments X² defines a sequence of 8-25 amino acids, such as 8-20 amino acids, such as 8-15 amino acids, such as 8-11 amino acids.

In some embodiments X² defines a sequence of less than 25, such as less than 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7 or 6 amino acids.

In some embodiments X^2 defines a sequence of more than 8, such as more than 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 amino acids.

In some embodiments X^4 defines a sequence of 8-25 amino acids, such as 8-20 amino acids, such as 8-15 amino acids, such as 8-11 amino acids.

- 5 In some embodiments X^4 defines a sequence of less than 25, such as less than 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7 or 6 amino acids.

In some embodiments X^4 defines a sequence of more than 8, such as more than 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 amino acids.

- 10 In some embodiments the dimer peptide according to the invention consist of two identical peptide monomers.

In some embodiments the immunogenic composition according to the invention is in the form of a vaccine composition.

- 15 In some embodiments, the cell-penetrating peptide of the invention comprises at most 60, at most 59, at most 58, at most 57, at most 56, at most 55, at most 54, at most 53, at most 52, at most 51, at most 50, at most 49, at most 48, at most 47, at most 46, at most 45, at most 44, at most 43, at most 42, at most 41, at most 40, at most 39, at most 38, at most 37, at most 36, at most 35, at most 34, at most 33, at most 32, at most 31, at most 30, at most 29, at most 28, at most 27, at most 26, at most 25, at most 24, at most 23, at most 22, at most 21, at most 20, at most 19, at most 18 amino acids.

- 20 In some embodiments, the cell-penetrating peptide of the invention comprises at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, at least 30, at least 31, at least 32, at least 33, at least 34, at least 35 at least 36, at least 37, at least 38, at least 39, at least 40, at least 41, at least 42, at least 43, at least 44, at least 45, at least 46, at least 47, at least 48, at least 49, 25 at least 50, at least 51, at least 52, at least 53, at least 54, at least 55, at least 56, at least 57, at least 58, at least 59, at least 60 amino acid residues.

- 30 In some embodiments, the cell-penetrating peptide of the invention consists of 18 amino acid residues or 19 amino acid residues or 20 amino acid residues or 21 amino acid residues or 22 amino acid residues or 23 amino acid residues or 24 amino acid residues or 25 amino acid residues or 26 amino acid residues or 27 amino acid residues or 28 amino acid residues or 29 amino acid residues or 30 amino acid residues or 31 amino acid residues or 32 amino acid

residues or 33 amino acid residues or 34 amino acid residues or 35 amino acid residues or 36 amino acid residues or 37 amino acid residues or 38 amino acid residues or 39 amino acid residues or 40 amino acid residues or 41 amino acid residues or 42 amino acid residues or 43 amino acid residues or 44 amino acid residues or 45 amino acid residues or 46 amino acid residues or 47 amino acid residues or 48 amino acid residues or 49 amino acid residues or 50 amino acid residues or 51 amino acid residues or 52 amino acid residues or 53 amino acid residues or 54 amino acid residues or 55 amino acid residues or 56 amino acid residues or 57 amino acid residues or 58 amino acid residues or 59 amino acid residues or 60 amino acid residues.

10 In some embodiments the cell-penetrating peptide of the invention does not consist of the following sequence RFIIP[Nle]FTALSGGRRALLYGATPYAIG, where Nle denotes a nor-leucine.

In some embodiments X^2 and/or X^4 is not derived from HIV.

In some embodiments X^4 is a linear sequence of less than 12 amino acids.

In some embodiments X^2 is a linear sequence of less than 12 amino acids.

15 In some embodiments X^2 and/or X^4 do not contain nor-leucine.

In some embodiments X^2 do not contain nor-leucine.

In some embodiments X^2 and/or X^4 only contains natural amino acids.

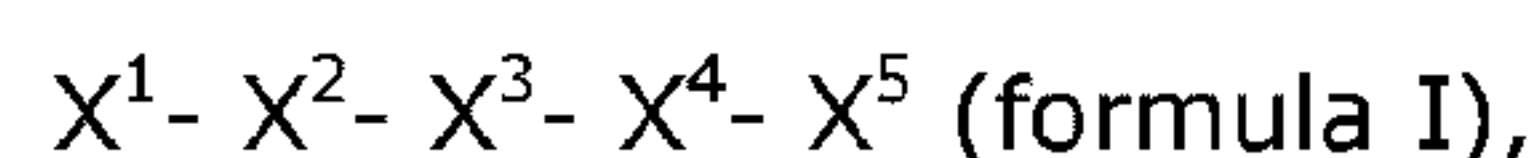
In some embodiments X^2 only contains natural amino acids.

In some embodiments X^2 only contains natural amino acids if derived from HIV.

20 In some embodiments X^2 and/or X^4 is derived from HCV, CMV, HPV, Influenza, adenoviruses, or picornaviruses.

Numbered embodiments according to the invention:

1. An isolated cell-penetrating peptide comprising the following structure



- wherein X^1 and X^3 independently defines a linear sequence of any 1, 2, 3 or 4 amino acid independently selected from any basic amino acid, citrulline, tryptophan, or a derivative thereof; X^2 defines a linear sequence of 8-30 amino acids derived from an antigen; X^4 defines a linear sequence of 8-30 amino acids derived from said antigen, said sequence X^4 being different from X^2 ; and wherein X^5 is any one optional amino acid selected from a basic amino acid, citrulline, tryptophan, or a derivative thereof.
2. The isolated peptide according to embodiment 1, wherein X^1 defines a linear sequence of any 1, 2, 3, or 4 amino acids independently selected from any basic amino acid, tryptophan, or a derivative thereof.
 - 10 3. The isolated peptide according to embodiment 1, wherein X^1 defines a linear sequence in any order of one citrulline and any 1, 2, or 3 amino acids independently selected from any basic amino acid, tryptophan, or a derivative thereof.
 - 15 4. The isolated peptide according to any one of embodiment 1-3, wherein X^3 defines a linear sequence of any 1, 2, 3, or 4 amino acids independently selected from any basic amino acid.
 5. The isolated peptide according to any one of embodiment 1-3, wherein X^3 defines a linear sequence in any order of one citrulline and of any 1, 2, or 3 amino acids independently selected from any basic amino acid.
 - 20 6. The isolated peptide according to any one of embodiment 1-5, wherein the sequence of amino acids defined by X^2 - X^3 - X^4 of formula I as defined in embodiment 1 is not found in the native sequence of said antigen.
 7. The isolated peptide according to any one of embodiment 1-6, wherein said peptide is demonstrated to translocate across a plasma membrane in the assay based on biotinylation of peptides as described in example 1.
 - 25 8. The isolated peptide according to any one of embodiments 1-7, wherein said one optional amino acid selected from a basic amino acid, citrulline, tryptophan, or a derivative thereof is selected from Arg, Lys, and His.
 9. The isolated peptide according to any one of embodiments 1-8, wherein X^2 and/or X^4 defines a sequence identical to the native sequence of said antigen.

10. The isolated peptide according to any one of embodiments 1-9, wherein said peptide is capable of inducing a T lymphocyte response.
11. The isolated peptide according to any one of embodiments 1-10, wherein said antigen is a viral protein, such as a capsid protein.
- 5 12. The isolated peptide according to embodiment 11, wherein said viral protein is selected from a protein of the Hepatitis C virus, such as a core protein; protein of influenza virus, such as an M2 protein.
13. The isolated peptide according to embodiment 11, wherein said viral protein of Hepatitis C virus is selected from HCV consensus sequence of genotype 1, such as subtypes
10 1a and 1b, genotype 2 such as 2a and 2b and genotype 3, such as 3a.
14. The isolated peptide according to any one of embodiments 1-13, which cell-penetrating peptide is of 19-60 amino acids, such as of 20-60 amino acids, such as of 21-60 amino acids, such as of 22-60 amino acids, such as of 23-60 amino acids, such as of 24-60 amino acids, such as of 25-60 amino acids, such as of 26-60 amino acids, such as of 27-60
15 amino acids, such as of 28-60 amino acids, such as of 29-60 amino acids, such as of 30-60 amino acids, such as of 31-60 amino acids, such as of 32-60 amino acids, such as of 33-60 amino acids, such as of 34-60 amino acids, such as of 35-60 amino acids.
15. The isolated peptide according to any one of embodiments 1-14, which cell-penetrating peptide is of 18-60, such as 18-59, such as 18-58, such as 18-57, such as 18-56,
20 such as 18-55, such as 18-54, such as 18-53, such as 18-52, such as 18-51, such as 18-50, such as 18-49, such as 18-48, such as 18-47, such as 18-46, such as 18-45, such as 18-44, such as 18-43, such as 18-42, such as 18-41, such as 18-40, such as 18-39, such as 18-38, such as 18-37, such as 18-35 amino acids, such as of 18-34 amino acids, such as of 18-33 amino acids, such as of 18-32 amino acids, such as of 18-31 amino acids, such as of 18-30
25 amino acids, such as of 18-29 amino acids, such as of 18-28 amino acids, such as of 18-27 amino acids, such as of 18-26 amino acids, such as of 18-25 amino acids, such as of 18-24 amino acids, such as of 18-23 amino acids, such as of 18-22 amino acids, such as of 18-21 amino acids, such as of 18-20 amino acids, such as of 18-19 amino acids.
16. The isolated peptide according to any one of embodiments 1-15, wherein the net
30 charge of X^2 is below or equal to 0; and wherein X^1 and X^3 defines a sequence of 2, 3 or 4 amino acid independently selected from any basic amino acid, citrulline, tryptophan, or a derivative thereof.

17. The isolated peptide according to any one of embodiments 1-16, wherein the net charge of X^2 is below or equal to 0; wherein the net charge of X^4 is above or equal to 1; and wherein X^1 and X^3 defines a sequence of 2, 3 or 4 amino acid independently selected from any basic amino acid, citrulline, tryptophan, or a derivative thereof.
- 5 18. The isolated peptide according to any one of embodiments 1-17, wherein the net charge of X^2 and X^4 are below or equal to 0 and wherein X^1 and X^3 defines a sequence of 2, 3 or 4 amino acid independently selected from any basic amino acid, citrulline, tryptophan, or a derivative thereof.
- 10 19. The isolated peptide according to any one of embodiments 1-18, wherein the net charge of X^2 and X^4 are above or equal to 1 and wherein X^1 and X^3 defines a sequence of 1 or 2 amino acid independently selected from any basic amino acid, citrulline, tryptophan, or a derivative thereof.
- 15 20. The isolated peptide according to any one of embodiments 1-19, wherein the net charge of X^2 is above or equal to 1; wherein the net charge of X^4 is below or equal to 0; wherein X^1 defines a sequence of 1 or 2 amino acids with a positive charge; and wherein X^3 defines a sequence of 2, 3 or 4 amino acid independently selected from any basic amino acid, citrulline, tryptophan, or a derivative thereof.
21. The isolated peptide according to any one of embodiments 1-20, wherein said basic amino acid is independently selected from Arg, Lys, and His.
- 20 22. The isolated peptide according to any one of embodiments 1-21, wherein X^2 and/or X^4 comprises the sequence of amino acids defined by position 135-157 of SEQ ID NO:3, or a fragment or variant thereof.
- 25 23. The isolated peptide according to any one of embodiments 1-22, wherein X^2 and/or X^4 consist of a sequence selected from GYIPLVGAPLG, GYLPAVGAPIG, GYLPAVGAPI, NYVTGNIPG, NYATGNLPG, NYATGNLPG, VTGNIPGSTYS, IRNLGRVIETLTG, SRNLGKVIDTLTC, IRNLGRVIETLT, GGGQIIGGNLIP, GGGQIVGGVYLLP, LIFLARSALIV, LIFLARSALIL, LIFLARSALIL, SAYERM CNIL, SAYERZVNIL, TAYERZCNIL, IAYERMCNIL, IAYERMCNIL, LFFKCIYRLF KHGL, LFFKTITRLFBHGL, GLEPLVIAGILA, GSDPLVVAASIV, NLVPMVATV, NLVPMVATV, NIVPZVVT A, PEVIPMFSALS, FIIPXFTALSG, ALGPAATL, GPVVHLTL, 30 LECVYCKQQLL, GVDYFAFRDLC, GVFDYAFRDIN, GATPVDLLGA, GVTPAGLIGV, VARALAHGVRV, VIRVIAHGLRL, GITFSIFLIVS, CSFSIFLLAL, GCSFSIFLLAL, GITFSIYLIVS, LZGYIPLIGA, LMGYIPLVGA, LZGYIPLIGA, PBIGVRATB, GPRLGVRATR, GPRLGVRAT, RGSVAHKS, SALILRGSVAHK, FQTAAQRAMM, FQTAAQRAVZ, FQTVVQBA, FQTAAQRA, GPSTEGVPESM,

LLSTEGVPNSZ, GSLVGLLHIVL, ASIVGILHLIL, NLVPMVATV, NIVPZVVT, TPQDLNTMLN, ALLYGATPYAIG, MMTACQGVG, GQAGDDFS, EVYDFAFRDLC, GFAFRDLCIVY, GFAYRDINLAY, GALNLCLPM, GALQIBLPL, and IRNLGRVIETLTZGYIPLIGA, or a fragment or variant thereof.

24. The isolated peptide according to any one of embodiments 1-23, wherein X^2 and/or X^4 consist of a sequence derived from an amino acid sequence selected from SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:23, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:28, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:35, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:42, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:126, and SEQ ID NO:198, or a fragment or variant thereof.
25. The isolated peptide according to any one of embodiments 1-24, wherein said peptide is selected from RRGYIPLVGAPLGBGRVARALAHGVRV (SEQ ID NO:47), RGYIPLVGAPLGRRVARALAHGVRV (SEQ ID NO:48), RGYIPLVGAPLGRRRVARALAHGVRV (SEQ ID NO:49), RRGYIPLVGAPLGRRVARALAHGVRV (SEQ ID NO:50), RRGYIPLVGAPLGRRRVARALAHGVRV (SEQ ID NO:51), BRGYIPLVGAPLGRRVARALAHGVRV (SEQ ID NO:52), RRRGYIPLVGAPLGBRVARALAHGVRV (SEQ ID NO:53), RGYIPLVGAPLGKKKVARALAHGVRV (SEQ ID NO:54), RGYIPLVGAPLGRRRVARALAHGVRV (SEQ ID NO:55), KKGYIPLVGAPLGKKVARALAHGVRV (SEQ ID NO:56), WGYIPLVGAPLGRRVARALAHGVRV (SEQ ID NO:57), WWGYIPLVGAPLGRRVARALAHGVRV (SEQ ID NO:58), EEGYIPLVGAPLGEEVARALAHGVRV (SEQ ID NO:59), GGGYIPLVGAPLGGGVARALAHGVRV (SEQ ID NO:60), EEGYIPLVGAPLGRRVARALAHGVRV (SEQ ID NO:61), RRGYIPLVGAPLGLRRVARALAHGVRV (SEQ ID NO:62), WWGYIPLVGAPLGRRVARALAHGVRV (SEQ ID NO:63), WWGYIPLVGAPLGRRRVARALAHGVRV (SEQ ID NO:64), WWGYIPLVGAPLGRVARALAHGVRV (SEQ ID NO:65), RGYIPLVGAPLGRRVARALAHGVRV (SEQ ID NO:66), RRGYLPVAVGAPIGBRVIRVIAHGLRL (SEQ ID NO:67), RRGYIPLVGAPLGBRVARALAHGVRV (SEQ ID NO:68), GYIPLVGAPLGGVARALAHGVRV (SEQ ID NO:69), WWGYLPVAVGAPIRRVIRVIAHGLRL (SEQ ID NO:70), GYIPLVGAPLGGVARALAHGVRV (SEQ ID NO:71), RRGYIPLVGAPLGBGRVARALAHGVRV (SEQ ID NO:72), RGYIPLVGAPLGRRVARALAHGVRV (SEQ ID NO:73), RGYIPLVGAPLGRRRVARALAHGVRV (SEQ ID NO:74), RRGYIPLVGAPLGRRVARALAHGVRV (SEQ ID NO:75), RRGYIPLVGAPLGRRRVARALAHGVRV (SEQ ID NO:76), BRGYIPLVGAPLGRRVARALAHGVRV (SEQ ID NO:77), RRRGYIPLVGAPLGBRVARALAHGVRV (SEQ ID NO:78), RGYIPLVGAPLGKKKVARALAHGVRV (SEQ ID NO:79),

RGYIPLVGAPLGRRRVARALAHGVRV (SEQ ID NO:80), KKGYIPLVGAPLGKKVARALAHGVRV (SEQ ID NO:81), WGYIPLVGAPLGRRRVARALAHGVRV (SEQ ID NO:82), WWGYIPLVGAPLGRRRVARALAHGVRV (SEQ ID NO:83), RRGYIPLVGAPLGLRRVARALAHGVRV (SEQ ID NO:84), RRNYVTGNIPGBRGITFSIFLIVS (SEQ ID NO:85),

5 WWNYATGNLPGRRCSFSIFLLAL (SEQ ID NO:86), WWNYVTGNIPGBRGITFSIFLIVS (SEQ ID NO:87), WWNYVTGNIPGRRGITFSIFLIVS (SEQ ID NO:88), RRNYATGNLPGRRGCSFSIFLLAL (SEQ ID NO:89), RRVVTGNIPGSTYSGBRGITFSIYLIVS (SEQ ID NO:90), RRIRNLGRVIETLTGBRLZGYIPLIGA (SEQ ID NO:91), RRSRNLGKVIDTLTCBRLMGYIPLVGA (SEQ ID NO:92), SRNLGKVIDTLTCGFADLMGYIPLVGA (SEQ ID NO:93),

10 WWIRNLGRVIETLTRRLZGYIPLIGA (SEQ ID NO:94), WWSRNLGKVIDTLTCRRLMGYIPLVGA (SEQ ID NO:95), RRGGGQIIGGNYLIPRBPBIGVRATB (SEQ ID NO:96), GGGQIVGGVYLLPRRGPRLGVRATR (SEQ ID NO:97), RRGGGQIVGGVYLLPRRGPRLGVRATR (SEQ ID NO:98), WWGGGQIVGGVYLLPRRGPRLGVRAT (SEQ ID NO:99), BRLIFLARSALIVRGSVAHKS (SEQ ID NO:100), EDLIFLARSALILRGSVAHKS (SEQ ID NO:101),

15 BRLIFLARSALILBGRSALILRGSVAHK (SEQ ID NO:102), SAYERM CNILKGKFQTAAQRAMM (SEQ ID NO:103), SAYERZVNILKGKFQTAAQRAVZ (SEQ ID NO:104), BRTAYERZCNILBRGRFQTVVQBA (SEQ ID NO:105), BRIAYERMCNILLBRGKFQTAAQRA (SEQ ID NO:106), IAYERMCNILKGKFQTAAQRA (SEQ ID NO:107), LFFKCIYRLFKHGLKRG PSTEGVPESM (SEQ ID NO:108), BRRLFFKTITRLFBHGLRRLSTEGVPNSZ (SEQ ID NO:109),

20 BRGLEPLVIAGILARRGSLVGLLHIVL (SEQ ID NO:110), BRGSDPLVVAASIVRRASIVGILHLIL (SEQ ID NO:111), RNLVPMVATVRRNLVPMVATVB (SEQ ID NO:112), RNLVPMVATVBRRNLVPMVATVB (SEQ ID NO:113), RNIVPZVVTARRNIVPZVVTAB (SEQ ID NO:114), PEVIPMFSALSEGATPQDLNTMLN (SEQ ID NO:115), RFIIPXFTALSGGRRALLYGATPYAIG (SEQ ID NO:116), KALGPAATLEEMMTACQGVG (SEQ ID NO:117), RRGPVVHLTLRRRGQAGDDFS (SEQ ID NO:118), RRGPVVHLTLRRRGQAGDDFS (SEQ ID NO:119), RRGPVVHLTLRGRRGQAGDDFS (SEQ ID NO:120), RRLECVYCKQQLLRREVYDFAFRDLC (SEQ ID NO:121), RRGVYDFAFRDLCRRGFAFRDLCIVYR (SEQ ID NO:122), RRGVFDYAFRDINRRGFAYRDINLAYR (SEQ ID NO:123), RRGATPVDLLGARRGALNLCLPMR (SEQ ID NO:124), RRGVTPAGLIGVRRGALQIBLPLR (SEQ ID NO:125), RGYLPAVGAPIGRRRVIRVIAHGLRLR (SEQ ID NO:196), RRSRNLGKVIDTLTCRRLMGYIPLVGA (SEQ ID NO:197), and

30 RRIRNLGRVIETLTZGYIPLIGARRIRNLGRVIETLTZGYIPLIGAR (SEQ ID NO:199), or a fragment or variant thereof.

26. The isolated peptide according to any one of embodiments 1-25, wherein the peptide consist of a sequence selected from X¹-NYVTGNIPG-X³-GITFSIYLIVS; X¹-IRNLGRVIETLT-X³-LZGYIPLIGA; X¹-GYLPAVGAPI-X³-VIRVIAHGLRL; X¹-GGGQIIGGNYLIP-X³-PBIGVRATB; X¹-NYATGNLPG-X³-GCSFSIFLLAL; X¹-SRNLGKVIDTLTC-X³-LMGYIPLVGA; X¹-GYIPLVGAPL-X³-VARALAHGVRV; X¹-GGGQIVGGVYLLP-X³-PRLGVRATR; X¹-LTFLVRSVLLI-X³-GSVLIVRGSLVH; X¹-TAYERZCNIL-X³-GRFQTVVQBA; X¹-SDPLVVAASIV-X³-ASIVGILHLIL; X¹-LIFLARSALIL-X³-

35

- SALILRGSVAH; X¹-IAYERMCNIL-X³-GKFQTAAQRA; and X¹-LEPLVIAGILA-X³-GSLVGLLHIVL; X¹-NLVPMVATV-X³-NLVPMATV; X¹-GYLPAVGAPIG-X³-VIRVIAHGLRL; X¹-IRNLGRVIETLTG-X³-LZGYIPLIGA; X¹-GVYDFAFRDLC-X³-GFAFRDLCIVYR, X¹-GVFDYAFRDIN-X³-GFAYRDINLAYR, X¹-GATPVDLLGA-X³-GALNLCLPMR, X¹-GVTPAGLIGV-X³-GALQIBLPLR, and X¹-
- 5 IRNLGRVIETLTZGYIPLIGA-X³-IRNLGRVIETLTZGYIPLIGA; optionally with an X⁵ in the C-terminal of the peptide, wherein X¹ and X³ and X⁵ refers to X¹, X³, and X⁵ of formula I.
27. The isolated peptide according to any one of embodiments 1-26, wherein the peptide comprises one or more cysteine.
28. The isolated peptide according to any one of embodiments 1-27, wherein the peptide
10 contain intramolecular bonds, such as intramolecular disulfide (S-S) bonds between two cysteines or acylals.
29. The isolated peptide according to any one of embodiments 1-28, wherein the N- and/or C-terminal amino acid in X² is a hydrophilic or polar amino acid.
30. The isolated peptide according to any one of embodiments 1-29, wherein the N-
15 terminal amino acid in X⁴ is a hydrophilic or polar amino acid.
31. The isolated peptide according to any one of embodiments 1-30, which is not more than 58 amino acids, such as not more than 56, 54, 52, 50, 48, 46, 44, 42, 40, 38, 36, 34, 32, 30, 28, 26, 24, 22, 20, 18 amino acid residues.
32. The isolated peptide according to any one of embodiments 1-31, wherein X¹ consist of
20 2 or 3 amino acids.
33. The isolated peptide according to embodiment 32, wherein X¹ consist of WW, BR, or RR.
34. The isolated peptide according to any one of embodiments 1-33, wherein X³ consist of 2 or 3 amino acids.
35. The isolated peptide according to embodiment 34, wherein X³ consist of WW, BR, or
25 RR.
36. Isolated peptide of not more than 60 amino acids comprising a sequence of X² and/or X⁴ as independently defined in any one of table 1 or table 2.

37. Isolated peptide according to embodiment 36, which peptide consists of a sequence of X^2 or X^4 as independently defined in any one of table 1 or table 2.
38. The isolated peptide according to any one of embodiments 1-37, wherein X^2 defines a sequence of 8-25 amino acids, such as 8-20 amino acids, such as 8-15 amino acids.
- 5 39. The isolated peptide according to any one of embodiments 1-38, wherein X^4 defines a sequence of 8-25 amino acids, such as 8-20 amino acids, such as 8-15 amino acids.
40. The isolated peptide according to any one of embodiments 1-39, wherein X^2 defines a sequence of less than 25, such as less than 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7 or 6 amino acids.
- 10 41. The isolated peptide according to any one of embodiments 1-40, wherein X^2 defines a sequence of more than 8, such as more than 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 amino acids.
42. The isolated peptide according to any one of embodiments 1-41, wherein X^4 defines a sequence of less than 25, such as less than 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7 or 6 amino acids.
- 15 43. The isolated peptide according to any one of embodiments 1-42, wherein X^4 defines a sequence of more than 8, such as more than 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 amino acids.
44. The isolated peptide according to any one of embodiments 1-43, wherein the cell-
20 penetrating peptide of the invention does not consist of the following sequence
RFIIP[Nle]FTALSGGRRALLYGATPYAIG, where Nle denotes a nor-leucine.
45. The isolated peptide according to any one of embodiments 1-44, wherein X^2 and/or X^4 is not derived from HIV.
46. The isolated peptide according to any one of embodiments 1-45, wherein X^4 is a linear
25 sequence of less than 12 amino acids.
47. The isolated peptide according to any one of embodiments 1-46, wherein X^2 is a linear sequence of less than 12 amino acids.

48. The isolated peptide according to any one of embodiments 1-47, wherein X^2 and/or X^4 do not contain nor-leucine.
49. The isolated peptide according to any one of embodiments 1-48, wherein X^2 do not contain nor-leucine.
- 5 50. The isolated peptide according to any one of embodiments 1-49, wherein X^2 and/or X^4 only contains natural amino acids.
51. The isolated peptide according to any one of embodiments 1-50, wherein X^2 only contains natural amino acids.
52. The isolated peptide according to any one of embodiments 1-51, wherein X^2 only
10 contains natural amino acids if derived from HIV.
53. The isolated peptide according to any one of embodiments 1-52, wherein X^2 and/or X^4 is derived from HCV, CMV, HPV, Influenza, adenoviruses, or picornaviruses.
54. Use of a peptide comprising a sequence of X^2 or X^4 as independently defined in any one of table 1 or table 2 for inducing an immune response in a subject.
- 15 55. A dimer peptide comprising two peptides monomers, wherein each peptide monomer is as defined in any one of embodiments 1-53.
56. A dimer peptide according to embodiment 55, wherein the two peptide monomers are identical.
57. Peptide combination comprising two or more peptide according to any one of
20 embodiments 1-53.
58. An isolated nucleic acid or polynucleotide encoding a peptide according to any one of embodiments 1-53.
59. A vector comprising the nucleic acid or polynucleotide according to embodiment 58.
60. A host cell comprising the vector according to embodiment 59.

61. An immunogenic composition comprising at least one peptide according to any one of embodiments 1-53, a dimer peptide according to any one of embodiments 55-56, a peptide combination according to embodiment 57, the nucleic acid or polynucleotide according to embodiment 58, or the vector according to embodiment 59; in combination with a
5 pharmaceutically acceptable diluent or vehicle and optionally an immunological adjuvant.

62. The immunogenic composition according to embodiment 61 in the form of a vaccine composition.

63. A method for inducing an immune response in a subject against an antigen which comprises administration of at least one peptide according to any one of embodiments 1-53,
10 a dimer peptide according to any one of embodiments 55-56, a peptide combination according to embodiment 57, the nucleic acid or polynucleotide according to embodiment 58, or the vector according to embodiment 59; or the composition according to any one of embodiments 61-62.

64. A method for reducing and/or delaying the pathological effects of a virus in a subject
15 infected with said virus, the method comprising administering an effective amount of at least one peptide according to any one of embodiments 1-53, a dimer peptide according to any one of embodiments 55-56, a peptide combination according to embodiment 57, the nucleic acid or polynucleotide according to embodiment 58, or the vector according to embodiment 59; or the composition according to any one of embodiments 61-62.

20 65. A peptide according to any one of embodiments 1-53 for use as a medicament.

66. A peptide according to any one of embodiments 1-53 for treating the pathological effects of a virus in a subject infected with said virus.

Sequence list (amino acids in bold represents suitable antigenic sequences that may be used as any of X² and/or X⁴ as defined in formula I of the present invention)

25 SEQ ID NO:1: Accession no AF009606; Hepatitis C virus subtype 1a polyprotein gene, complete cds.

30 MSTNPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRL
GVRATRKTSESRQPRGRRQPIPKARRPEGRTWAQPGYPWPLYGNEGCGWAGWLLSPRG
SRPSWGPDPFRRRSRNLGKVIDTLTCGFADLMGYI**PLVGAPLGGAARALAHGVRVLED**
GVNYATGNLPGCSFSIFLLALLSCLTVPASAYQVRNSSGLYHVTNDCPNSSIVYEAD
AILHTPGCVPCVREGNASRCWVAVTPTVATRDGKLPTTQLRRHIDLLVGSATLCSALY
VGDLCGSVFLVGQLFTFSPRRHWTQDCNCSTYPGHITGHRMAWDMMNWSPTAALVV
AQLLRIPQAIMDMIAGAHWGVLAGIAYFSMVGNWAKVLVVLVLLFAGVDAETHVTGGSA
35 GRTTAGLVGLLTPGAKQNIQLINTNGSWHINSTALNCNESLNTGWLGLFYQHKFNSS
GCPERLASCRRRLTDFAGWGPISYANGSGLDERPYCWHYPPRPCGIVPAKSVCGPVYC

FTPSPVVVGTTDRSGAPTYSWGANDTDVFVLNNTRPPLGNWFGCTWMNSTGFTKVCGA
 PPCVIGGVGNNTLLCPTDCFRKHPEATYSRCGSGPWITPRCMVDYPYRLWHYPCTINY
 TIFKVRMYVGGVEHRLEAACNWTRGERCDLEDRDRSELSPLLLSTTQWQVLPSCSFTTL
 PALSTGLIHLHQNIVDVQYLYGVGSSIASWAIKWEYVLLFLLLADARVCSCIWMMLL
 5 ISQAEAALENLVI LNAASLAGTHGLVSFLVFFCFAWYLKGRWVPGAVYAFYGMWPLLL
 LLLALPQRAYALDTEVAASC GGVLVGLMALTLSPYYKRYISWCMWWLQYFLTRVEAQ
 LHVWVPPLNVRGGRDAVILLMCVVHPTLVFDITKLLLAIFGPLWILQASLLKVPYFVR
 VQGLLRICALARKIAGGHYVQMAIIKLGALTGTYYVNHLTPLRDWAHNGLRDLAVAVE
 PVVFSRMETKLITWGADTAACGDIINGLPVSARRGQEILLGPADGMVSKGWRL LAPIT
 10 AYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTR
 TIASPKGPVIQMYTNVDQDLVGWPAPQGSRS LTPCTCGSSDLYLVTRHADVIPVRRRG
 DSRGSLSPRPISYLGSSGGP LCPAGHAVGLFRAAVCTRGVAKAVDFIPVENLETT
 MRSPVFTDNSSPPAVPQSFQVAHLHAPTGS GKSTKVPAAYAAQGYKVLVLPNSVAATL
 GFGAYMSKAHGVDPNIRTGVRTITTGSPITYSTY GKFLADGGCSGGAYDIIICDECHS
 15 TDATSILGIGTVLDQAETAGARLVVLATATPPGSVTVSHPNIEEVALSTTGEIPFYGK
 AIPLEVIKGGRH LIFCHSKKKCDELA AKLVALGINAVAYYRGLDVSVIPTSGDVVVVS
 TDALMTGFTGDFDSVIDCNTCVTQTVD FSLDPTFTIETTTLPQDAVSRTQRRGRTGRG
 KPGIYRFVAPGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNT PGLPVCQ
 DHLEFWEGVFTGLTHIDAHFLSQT KQSGENFPYLVAYQATVCARAQAPPPSWDQMWK
 20 LIRLKPTLHGPTPLLYRLGAVQNEVT LTHPITKYIMTCMSADLEVVTSTWVLVGGVLA
 ALAAYCLSTGCVVIVGRIVLSGKPAIIPDREVLYQEFDEMEEC SQHLPYIEQGMMLAE
 QFKQKALGLLQTASRQAEVITPAVQTNWQKLEVFWAKHMWNFISGIQYLAGLSTLPGN
 PAIASLMAFTA AVTSPLTTGQTLLENILGGWVAAQLAAPGAATAFVGAGLAGAAIGSV
 GLGKVLVDILAGYGAGVAGALVAFKIMSGEVPSTEDLVNLLPAILSPGALVVGVC AA
 25 I LRRHVGPGE GAVQWMNRLIAFASRGNHVSPTHYVPESDAAARVTAI LSSLTVTQLLR
 RLHQWISSECTTPCSGSWLRDIWDWICEVLSDFKTWLKAKLMPQLPGIPFVSCQRGYR
 GVWRGDGIMHTRCHCGAEITGHVKNGTMRIVGPRTCRNMWSGTFPINAYTTGPCTPLP
 APNYKFALWRVSAEEYVEIRRVGDFHYVSGMTDN LKCPQCIPSPFEFFTELDGVR LHR
 FAPPCKPL LREEVSFRVGLHEYFVGSQ L PCEPEPDVAVLTSMLTDP SHITAEAAGRRL
 30 ARGSPPSMASSSASQLSAPSLKATCTANHDS PDAELIEANLLWRQEMGGNITRVESEN
 KVVILDSFDPLVAEEDEREVSVP AEILRKSRRFARALPVWARPDYNPPLVETWKKPDY
 EPPVHVGCP LPPRSPVP PPRKRTVVLTESTLSTALAE LATSFGSSSTSGITGDN
 TTTSSEPAPSGC PPDSDVESYSSMPLEGE PGDPDLSDGSWSTVSSGADTEDVCCSM
 SYSWTGALVTPCA AEEQKL PINALSNSLLRHHNLVYSTTSRSACQRQKKVTFDRLQVL
 35 DSHYQDVLKEVKAAASKVKANLLSVEEACSLTPPHSAKSKFGYGAKDVRCHARKAVAH
 INSVWKDLLED SVTPI DTTIMAKNEVF CVQPEKGGRKPARLIVFPDLGVRVCEKMALY
 DVVSKLPLAVMGSSYGFQYSPGQRVEFLVQAWKSKKTPMGFSYDTRCFDSTVTESDIR
 TEEAIYQCCDLDPQARVAIKSLTERLYVGGPLTNSRGENCYRRCRASGVLTTS CGNT
 LTCYIKARAACRAAGLQDCTMLVCGDDL VVICESAGVQEDAASLRAFTEAMTRY SAPP
 40 GDPPQPEYDLELITSCSSNVSV AHDGAGKRVYYLTRDPTT PLARA AWETARHTPVNSW
 LGNIIMFAPTLWARMILMTHFFSVLIARDQLEQALNCEIYGACYSIEPLDLPPIIQRL
 HGLSAFSLHSYSPGEINRVAACLRKLGVPPLRAWRHRARSVRARLLSRGGRAAICGKY
 LFNWAVRTKLKLTPIAAAGRLDLSGWFTAGYSGGDIYHSVSHARPRWFWFC LLLLAAG
 VGIYLLPNR
 45
 SEQ ID NO:2: HCV core protein, H77, Accession AF009606
 Genbank number: 2316097
 >gi|2316098|gb|AAB66324.1| polyprotein [Hepatitis C virus subtype 1a]
 MSTNPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRG PRLGVRATRKT SERSQPRGRRQPIPKARR
 50 PEGRTWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSWGPTDPRRRSRNLGKVIDTLTCGFADLMGYIPLVGAPLGGAAR
 ALAHGVRVLEDGVNYATGNLPGCSFSIFLLALLSCLTVPASA

 SEQ ID NO:3:
 55 Hepatitis C virus mRNA, complete cds; ACCESSION M96362 M72423; Hepatitis C
 virus subtype 1b
 MSTNPKPQRKTKRNTNRRPQDIKFPGGGQIVGGVYLLPRRG PRL
 GVRATRKT SERSQPRGRRQPIPKARRPEGRAWAQPGYPWPLYGNEGLGWAGWLLSPRG
 SRPSWGPTDPRRK **SRNLGKVIDTLTCGFADLMGYIPLVGAPLGGVARALAHGVRVLED**
 60 **GVNYATGNLPGCSFSIFLLALLSCLTTPV**SAYEVNRNASGMYHVTNDCSNSSIVYEAAD
 MIMHTPGCVPCVREDNSSRCWVALTPTLAARNASVPTTT LRRHVDLLVGVAAFCSAMY
 VGDLCGSVFLVSQ LFTFSPRRHETVQDCNCSIYPGRVSGHRMAWDMMMNWSPTTALVV

5 S QLLRIPQAVVDMVTGSHWGILAGLAYYSMVGNWAKVLIAMLLFAGVDGTTHTVTGGAQ
 GRAASSLTSLFSPGPVQHLQLINTNGSWHINRTALSCNDSLNTGFVAALFYKYRFNAS
 GCPERLATCRPIDTFAQGWGPITYTEPHDLDQRPYCWYAPQPCGIVPTLQVCGPVYC
 FTPSPVAVGTTDRFGAPTYRWGANETDVLLNNAGPPQGNWFGCTWMNGTGFTKTCGG
 10 PPCNIGGVGNNTLTCPTDCFRKHGATYTKCGSGPWLTPRCLVDYPYRLWHYPCTVNF
 TIFKVRMYVGGAEHRLDAACNWTGRGERCDLEDRDRSELSPLLLSTTEWQVLPFSFTTL
 PALSTGLIHLHQNIVDIQYLYGIGSAVVSFAIKWEYIVLLFLLLADARVCACLWMMLL
 VAQAEAALENLVVNAASVAGAHGILSFIVFFCAAWYIKGRLVPGAAYALYGVWPLLL
 LLLALPPRAYAMDREMAASC GGAVFVGLVLLTSLPHYKVFLARFIWWLQYLITRTEAH
 15 LQVWVPLNVRGGRDAIILLTCVVHPELIFDITKYLLAIFGPLMVLQAGITRVYPYFVR
 AQGLIRACMLARKVVGGHYVQVMFMKLAALAGTYVYDHLTPLRDWAHTGLRDLAVAVE
 PVVFSDMETKVITWGADTAACGDIILALPASARRGKEILLGPADSLEGQGWRLAPIT
 AYSQQTRGLLGCIITSLTGRDKNQVEGEVQVVSTATQSFLATCINGVCWTVFHGAGSK
 TLAGPKGPITQMYTNVDQDLVGWPAPPGARSLTPCTCGSSDLYLVTRHADVIPVRRRG
 20 DGRGSLPPRPVSYLKGSSGGPLLCPSGHAVGILPAAVCTRGVAMAVEFIPVESMETT
 MRSPVFTDNPSPPAVPQTFQVAHLHAPTGS GKSTRVPAAYAAQGYKVLVNLNPSVAATL
 GFGAYMSKAHGIDPNLRTGVRTITTTGAPITYSTYGKFLADGGGSGGAYDIIMCDECHS
 TDSTTIYGIGTVLDQAETAGARLVVLSTATPPGSVTVPHLNIEEVALSNTGEIPFYGK
 AIPIEAIKGRHLIFCHSKKKCDELAAKLSGLGLNAVAYYRGLDVSVIPITSGDVVVVA
 25 TDALMTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTTETTTVPQDAVSRSQRRGRTGRG
 RAGIYRFVTPGERPSGMFDSVLCCECYDAGCAWYELTPAETSVRLRAYLNTPLPVCQ
 DHLEFSEGVFTGLTHIDAHFLSQTKQAGENFPYLVAYQATVCARAQAPPPSWDEMWRC
 LIRLKPTLHGPTPLLYRLGAVQNEVTLTHPITKFIIMTMSADLEVVTSTWVLVGGVLA
 ALAAYCLTTGSGVIVGRIILSGKPAIIPDREVLYQEFDEMEECASHLPYFEQGMQLAE
 30 QFKQKALGLLQTATKQAEAAAPVVESKWRALETFWAKHMWNFISGIQYLAGLSTLPGN
 PAIRSPMAFTASITSPLTTQHTLLFNILGGWVAAQLAPPSAASAFVGAGIAGAAVGTI
 GLGKVLVDILAGYGAGVAGALVAFKIMSGEMPSAEDMVNLLPAILSPGALVVGIVCAA
 ILRRHVGPGE GAVQWMNRLIAFASRGNHVSPRHYPPESEPAARVTQILSSLTITQLLK
 RLHQWINE DCSTPCSSSWLREIWDWICTVLTDFKTWLQSKLLPRLPGVPPFFSCQRGYK
 35 GVWRGDGIMHTTCPCGAQITGHVKN GMSMRIVGPKTCSNTWYGTFFPINAYTTGPCTPSP
 APNYSKALWRVAAEEYVEVTRVGD FHYVTGMTTDNVKCPQCQVPAPEFFTEVDGVR LHR
 YAPACRPLLREEVVFQVGLHQYLVGSQLPCEPEPDVAVLTSMLTDP SHITAETAKRRL
 ARGSPPSLASSSASQLSAPSLKATCTTHHDS PADLIEANLLWRQEMGGNITRVESEN
 KVVILDSFDPLRAEDDEGEISVPAEILRKSRKFPPALPIWAPPDYNPPLLESWKDPDY
 40 VPPVVHGCPLPPTKAPPIPPPRRKRTVVLTESTVSSALAE LATKTFGSSGSSAIDSGT
 ATAPPDQASGDGDRES DVESFSSMPPLEGE PGDPDLSDGSWSTVSEEASEDVVCCSMS
 YTWTGALITPCAAEESKL PINPLSNSLLRHHNMVYATT SRSAGLRQKKVTFDRLQVLD
 DHYRDVLKEMKAKASTVKAKLLSVEEACKLT PPHSAKSKFGYGAKDVRSLSRAVTHI
 RSVWKD LLED TETPISTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYD
 45 VVSTLPQAVMGSSYGFQYSPKQ RVEFLVNTWKS KCPMGFSYDTRCFDSTVTENDIRV
 EESIYQCCDLAPEAKLAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLT TSCGNTL
 TCYLKATAACRAAKLRDCTMLVNGDDL VICESAGTQEDAASLRVFTEAMTRY SAPPG
 DPPQPEYDLELITSCSSNVSVAH DASGKR VYYLTRDPTT PLARAAWETARHTPVNSWL
 GNIIMYAPTLWARMILMTHFFSILLAQE QLEKTLDCQIYGACYSIEPLDLPQIIERLH
 50 GLSAFSLHSYSPGEINRVASCLRLKLGVPPLRAWRHARSVRAKLLSQGGR AATCGKYL
 FNWAVRTKLKLTPIPAASRLDLSGW FVAGYSGGDIYHSLSRARPRWFMLCLLLLSVG
 GIYLLPNR

50

SEQ ID NO:4, nucleocapsid protein of influenza A virus

1 MASQGTKRSY EQMETS GERQ NATEIRASVG RMVGGIGRFY IQMCTELKLS DHEGRLIQNS
 61 ITIERMVL SA FDERRNKYLE EHPSAGKDPK KTGGPIYRRR DGKWMRELIL YDKEEIRRIW
 121 RQANNGEDAT AGLTHMMIWH SNLNDATYQR TRALVRTGMD PRMCSLMQGS TLPRRSGAAG
 55 181 AAVKGVGTMV MELIRMIKRG INDRNFWRGE NGRRTRIAYE RMCNILKGKF QTAAQRAMMD
 241 QVRESRNPGN AEIEDLIFLA RSALILRGSV AHKSCLPACV YGLAVASGYD FEREGYSLVG
 301 IDPFRLIQNS QVFS LIRPNE NPAHKSQLVW MACHSAAFED LRVSSFIRGT RVVPRGQLST
 361 RGVQIASNEN METMDSSTLE LRSRYWAIRT RSGGNTNQQR ASAGQISVQP TFSVQRNLPF
 421 ERATIMAAFT GNTEGRTSDM RTEIIRMMEN ARPEDVSFQG RGVFELSDEK ATNPVPSFD
 60 481 MSNEGS

SEQ ID NO:5

>gi|73919153|ref|YP_308840.1| matrix protein 2 [Influenza A virus (A/New York/392/2004 (H3N2))]

5 MSLLTEVETPIRNEWGCRCNDS**SDPLVVAASIIGILHLILWILDRLFFKCVYRLFKHGLKRG**PSTEGVPE 70
SMREEYRKEQQNAVDADDSHFVSIELE

SEQ ID NO:6

10 >gi|73919147|ref|YP_308843.1| nucleocapsid protein [Influenza A virus (A/New York/392/2004 (H3N2))]
MASQGTKRSYEQMETDGDGRQNATEIRASVGKMIDGIGRFYIQMCTELKLSDEGRLIQNSLTIEKMVLSA 70
FDERRNKYLEEHPSAGKDPKKTGGPIYRRVDGKWMRELVLVDKEEIRRIWRQANNGEDATAGLTHIMIWH 140
SNLNDATYQRTALVRTGMDPRMC SLMQGSTLPRRSGAAGAAVKGIGTMVMELIRMVKRGINDRNFWRGE 210
NGRKTR**SAYERMCN**ILKGK**FQTAAQ**RAMVDQVRESRNP GNAEIED**LI FLARSALI LRGSVAHKS**CLPACA 280
15 YGPAVSSGYDFEKEGYSLVGIDPFKLLQNSQIYSLIRPNENPAHKSQLVWMACHSAAFEDLRLLSFIRGT 350
KVSPRGKLSTRGVQIASNENMDNMGSSTLELRSGYWAI RTRSGGNTNQQRASAGQTSVQPTFSVQRNLPF 420
EKSTIMAAFTGNTEGRTSDMRAEI I RMMEGAKPEEVSFGRGVFELSDEKATNP IVP SFDM SNEGSYFFG 490
DNAEEYDN

20

SEQ ID NO:7

>gi|56583270|ref|NP_040979.2| matrix protein 2 [Influenza A virus (A/Puerto Rico/8/34 (H1N1))]
25 MSLLTEVETPIRNEWGCRCNGS**SDPLAIAANIIGILHLILWILDRLFFKCIYRRFKYGLKGG**PSTEGVPK
SMREEYRKEQQSAVDADDGHFVSIELE

SEQ ID NO:8

>gi|8486130|ref|NP_040982.1| nucleocapsid protein [Influenza A virus (A/Puerto Rico/8/34 (H1N1))]
30 MASQGTKRSYEQMETDGERQNATEIRASVGKMIGGIGRFYIQMCTELKLSDYEGRLIQNSLTIERMVLSA
FDERRNKYLEEHPSAGKDPKKTGGPIYRRVNGKWMRELILYDKEEIRRIWRQANNGDDATAGLTHMMIWH
SNLNDATYQRTALVRTGMDPRMC SLMQGSTLPRRSGAAGAAVKGVGTVMELVRMIKRGINDRNFWRGE
NGRKTR**IAYERMCN**ILKGK**FQTAAQ**KAMMDQVRESRDPGNAEFED**LT FLARSALI LRGSVAHKS**CLPACV
YGPAVASGYDFEREGYSLVGIDPFRLQLNSQVYSLIRPNENPAHKSQLVWMACHSAAFEDLRVLSFIKGT
35 KVVPRGKLSTRGVQIASNENMETMESSTLELR SRYWAI RTRSGGNTNQQRASAGQISIQPTFSVQRNLPF
DRTTVMAAFTGNTEGRTSDMRTEI I RMMESARPEDVSFQGRGVFELSDEKAASPIVP SFDM SNEGSYFFG
DNAEEYDN

40

SEQ ID NO:9

>gi|73912687|ref|YP_308853.1| membrane protein M2 [Influenza A virus (A/Korea/426/68 (H2N2))]
MSLLTEVETPIRNEWGCRCNDS**SDPLVVAASIIGILHFI**LWILDRL**FFKCIYRFFKHGLKRG**PSTEGVPE
45 **SMREEYRKEQQSAVDADDSHFVSIELE**

SEQ ID NO:10

>gi|73921307|ref|YP_308871.1| nucleoprotein [Influenza A virus (A/Korea/426/68 (H2N2))]
50 MASQGTKRSYEQMETDGERQNATEIRASVGKMIDGIGRFYIQMCTELKLSDYEGRLIQNSLTIERMVLSA
FDERRNKYLEEHPSAGKDPKKTGGPIYKRVDGKWMRELVLVDKEEIRRIWRQANNGDDATAGLTHMMIWH
SNLNDTTYQRTALVRTGMDPRMC SLMQGSTLPRRSGAAGAAVKGVGTVMELIRMIKRGINDRNFWRGE
NGRKTR**SAYERMCN**ILKGK**FQTAAQ**RAMMDQVRESRNP GNAEIED**LI FLARSALI LRGSVAHKS**CLPACV
YGPAIASGYNFEKEGYSLVGIDPFKLLQNSQVYSLIRPNENPAHKSQLVWMA CN SAAFEDLRVLSFIRGT
KVSPRGKLSTRGVQIASNENMDTMESSTLELR SRYWAI RTRSGGNTNQQRASAGQISVQPAFSVQRNLPF
55 DKPTIMAAFTGNTEGRTSDMRAEI I RMMEGAKPEEMS FQGRGVFELSDEKATNP IVP SFDM SNEGSYFFG
DNAEEYDN

SEQ ID NO:11
 >gi|330647|gb|AAA45994.1| pp65 [Human herpesvirus 5]
 MASVLGPISGHVLKAVFSRGDTPVLPHETRLQLTGIHVRVSQPSLILVSQYTPDSTPCHRGDNQLQVQHT 70
 YFTGSEVENVSVNVHNPTGRSICPSQEPMSIYVYALPLKMLNIPSINVHHYPSAAERKHRHLPVADAVIH 140
 5 ASGKQMWQARLTVSGLAWTRQQNQWKEPDVYYTSAFVFPTKDVALRHVVCAHELVCSEMENTRATKMQVIG 210
 DQYVKVYLESFCEDEVPSGKLFMHVTLGSDVEEDLTMTNRNPQPFMRPHERNGFTVLCPKNMIKPGKISHI 280
 MLDVAFTSHEHFGLLCPKSIPLGLSISGNLLMNGQQIFLEVQAIRETVELRQYDPVAALFFFDIDLLLQRG 350
 PQYSEHPTFTSQYRIQGKLEYRHTWDRHDEGAAQGDDVWTSGSDSDEELVTTERKTTPRVTTGGGAMAGAS 420
 TSAGRKRKSASSATACTAGVMTGRGLKAESTVAPEEDTDESDNEIHNPVFTWPPWQAGILARN**NLVPMV** 490
 10 **ATV**QGQNLKYQEFFWDANDIYRIFAELGVWQPAAPKRRRHRQDALPGPCIASTPKKHRG 541

SEQ ID NO:12
 >gi|33330937|gb|AAQ10712.1| putative transforming protein E6 [Human
 15 papillomavirus type 16]
 MHQKRTAMFQDPQERPGKLPQLCTELQTTIHDIILECVYCKQQLLRRE**VYDFAERDLCIVY**RDGNPYAVC 70
 DKCLKFYISKISEYRHYCYSVYGTTLQYQYNKPLCDLLIRCINCQKPLCPEEKQRHLDKKQRFHNIRGRWT 140
 GRCMSCCRSSRTRRETQL

20
 SEQ ID NO:13
 >gi|56583270|ref|NP_040979.2| matrix protein 2 [Influenza A virus (A/Puerto
 Rico/8/34(H1N1))]
 MSLLTEVETPIRNEWGCRCNGSSDPLAIAANIIGILHLILWILDRLEFFKCIYRRFKYGLKGGPSTEGVPK
 25 SMREEYRKEQQSAVDADDGHFVSIELE

SEQ ID NO:14
 >gi|8486139|ref|NP_040987.1| PB2 protein [Influenza A virus (A/Puerto
 Rico/8/34(H1N1))]
 30 MERIKELRNLSQSRTREILTCTTVDHMAIIKKYTSGRQEKNPALRMKWMAMKYPIITADKRITEMIPER
 NEQGQTLWSKMNDAGSDRVMVSLAVTWWRNGPMTNTVHYPKIKYTYFERVERLKHGTFGPVHFRNQVK
 IRRRVDINPGHADLSAKEAQDVIMEVVFNPVEVGARILTSSESQTLITKEKKEELQDCKISPLMVAYMLERE
 LVRKTRFLPVAGGTSSVYIEVLHLTQGTCEWQMYTPGGEVKNDDVDQSLIIAARNIVRRAAVSADPLASL
 LEMCHSTQIGGIRMVDILKQNPTEEQAUGICKAAMGLRISSEFSFGGFTEFKRTSGSSVKREEEVLTGNLQ
 35 TLKIRVHEGYEEFTMVGRRATAILRKATRRLIQLIVSGRDEQSIAEAIIVAMVFSQEDCMIKAVRGDLNF
 VNRRANQRLNPMHQLLRHFQKDAKVLFFQNWGVPEIDNVGMIGILPDMPSTIEMSMRGVRISKMGVDEYSS
 TERVVVSIDRFLRVRDQRGNVLLSPEEVSETQGTKEKLTITYSSSMWEINGPESVLVNTYQWIIIRNWETV
 KIQWSQNPTMLYNKMEFEPFQSLVPKAIKRGQYSGFVRTLFQQMRDVLGTFDFTAQIIKLLPFAAAPPKQSR
 MQFSSFTVNVRGSGMRILVRGNPVPFNYNKATKRLTVLGKDAGTLTEDPDEGTAGVESAVLRGFLILGKE
 40 DRRYGPALSINELSNLAKGEKANVLIQGDVVLVMKRKRDSILTDSQTATKRIRMAIN

SEQ ID NO:15
 >gi|8486137|ref|NP_040986.1| polymerase PA [Influenza A virus (A/Puerto
 Rico/8/34(H1N1))]
 45 MEDFVRQCFNPMIVELAEKTMKEYGEDLKIETNKFAAICTHLEVCFMYSDFHFINEQGESIIVELGDPNA
 LLKHRFEIIEGRDRTMAWTVVNSICNTTGAEKPKFLPDLYDYKENRFIEIGVTRREVHIYYLEKANKIKS
 EKTHIHIFSTGEEMATKADYTLDEESRARIKTRLFTRQEMASRGLWDSFRQSERGEETIEERFEITGT
 MRKLADQSLPPNFSSLENFRAYVDGFEPNGYIEGKLSQMSKEVNARIEPFLKTTTPRPLRLPNGPPCSQRS
 KFLMDALKLSIEDPSHEGEGIPLYDAIKCMRTFFGWKEPNVVKPHEKGINPNYLLSWKQVLAELQDIEN
 50 EEKIPKTKNMKKTSQLKWALGENMAPEKVDFFDDCKDVGLKQYDSDEPELRSLSASWIQNEFNKACELTDS
 SWIELDEIGEDVAPIEHIA SMRRNYFTSEVSHCRATEYIMKGVYINTALLNASCAAMDDFQLIPMISKCR
 TKEGRKTNLYGFIIKGRSHLRNDTDVVNFVSMEFSLTDPRLEPHKWEKYCVLEIGDMLLRSAIGQVSRP
 MFLYVRTNGTSKIKMKWGMEMRRCLLQSLQQIESMIEAESSVKEKDMTKEFFENKSETWPIGESPKGVEE
 SSIGKVCRTLLAKSVFNSLYASPQLEGFSAESRKLILLVQALRDNLPGTFDLGGLYEAIEECLINDPWV
 55 LLNASWFNSFLTHALS

SEQ ID NO:16
 >gi|8486133|ref|NP_040984.1| nonstructural protein NS1 [Influenza A virus
 (A/Puerto Rico/8/34(H1N1))]

MDPNTVSSFQVDCFLWHVRKRVADQELGDAPFLDRLRRDQKSLRGRGSTLGLDIETATRAGKQIVERILK
EESDEALKMTMASVPASRYLTDMTLEEMSREWSMLIPKQKVAGPLCIRMDQAIMDKNIILKANFSVIFDR
LETLILLRAFTEEGAIVGEISPLPSLPGHTAEDVKNAVGVVLIGGLEWNDNTVRVSETLQRFARSSNENG
RPPLTPKQKREMAGTIRSEV

5

SEQ ID NO:17

>gi|8486132|ref|NP_040983.1| nonstructural protein NS2 [Influenza A virus
(A/Puerto Rico/8/34(H1N1))]

MDPNTVSSFQDILLRMSKMQLSSSEDNLNGMITQFESLKLRYRDSLGEAVMRMGDLHSLQNRNEKWREQLG
10 QKFEEIRWLIEEVRHKLKVTENSFEQITFMQALHLLLEVEQEIRTFSSFQLI

SEQ ID NO:18

>gi|8486128|ref|NP_040981.1| neuraminidase [Influenza A virus (A/Puerto
Rico/8/34(H1N1))]

15 MNPNQKIITIGSICLVVGLISLILQIGNIISIWISHSIQTGSQNHTGICNQNIITYKNSTWVKDTTSVIL
TGNSSLCPIRGWAIYSKDNSIRIGSKGDFVIREPFISCSHLECRTFFLTQGALLNDRHSNGTVKDRSPY
RALMSCPVGEAPSPYNSRFESVAWSASACHDGMGWLITIGISGPDNGAVAVLKYNGIITETIKSWRKKILR
TQESECACVNGSCFTIMTDGPSDGLASYKIFKIEKGKVTKSIELNAPNSHYEECSYCPDTGKVMCVCARDN
WHGSNRPWVSFDQNLDYQIGYICSGVFGDNPRPKDGTGSCGPVYVDGANGVKGFSYRYGNGVWIGRTKSH
20 SSRHGFEMIWDPNGTETDSKFSVRQDVVAMTDWSGYSGSFVQHPELTGLDCIRPCFWVELIRGRPKEKT
IWTSASSISFCGVNSD TVDWSWPDGAELPFTIDK

SEQ ID NO:19

25 >gi|8486126|ref|NP_040980.1| haemagglutinin [Influenza A virus (A/Puerto
Rico/8/34(H1N1))]

MKANLLVLLCALAAADADTICIGYHANNSTD TVD TVLEKNVTVTHSVNLLLED SHNGKLCRLKGIAPLQLG
KCNIAGWLLGNPECDPLLPVRSWSYIVETPNSENGICYPGDFIDYEELREQLSSVSSFERFEIFPKESSW
PNHNTTKGVTAACSHAGKSSFYRNLLWLTEKEGSYPKLKNSYVNKKGKEVLVLWGIHHPNSNSKDQQNIYQ
NENAYVSVVTSNYNRRFTPEIAERP KVRDQAGRMNYYWTL LKPGDTIIFEANGNLIAPRYAFALSRGFGS
30 GIITSNASMHECNTKCQTPLGAINSSLPFQNIHPVTIGECPKYVRS AKLRMTGLRNIPSIQSRGLFGAI
AGFIEGGWTGMIDGWYGYHHQNEQGS GYAADQKSTQNAINGITNKVNSVIEKMNIQFTAVGKEFNKLEKR
MENLNKKVDDGFLDIWTYNAELLV LLENERTLDFHDSNVKNLYEKVKSQ LKNNAKEIGNGCFFEFYHKCDN
ECMESVRNGTYDYPKYSEESKLNREKVDGVKLESMGIYQILAIYSTVASSLVLLVSLGAISFWMC SNGSL
QCRICI

35

SEQ ID NO:20

>gi|8486123|ref|NP_040978.1| matrix protein 1 [Influenza A virus (A/Puerto
Rico/8/34(H1N1))]

MSLLTEVETYVLSIIPSGPLKAEIAQRLEDVFAGKNTDLEVLMEWLKTRPILSPLTKGILGFVFTLTVPS
40 ERGLQRRRFVQNALNGNDPNNMDKAVKLYRKLRKREITFHGAKEISLSYSAGALASCMGLIYNRMGAVTT
EVAFGLVCATCEQIADSQHRSHRQMVTTTNPLIRHENRMV LASTTAKAMEQ MAGSSEQAAEAMEVASQAR
QMVQAMRTIGTHPSSSAGLKNDDLLENLQAYQKRMGVQMQRFK

SEQ ID NO:21

45 >gi|83031685|ref|YP_418248.1| PB1-F2 protein [Influenza A virus (A/Puerto
Rico/8/34(H1N1))]

MGQEQDTPWILSTGHISTQKRQDGQQTPKLEHRNSTRLMGHCQKTMNQVVM PKQIVYWKQWLSLRNPILV
FLKTRVLKRWRLFSKHE

50

SEQ ID NO:22

>gi|8486135|ref|NP_040985.1| polymerase 1 PB1 [Influenza A virus (A/Puerto
Rico/8/34(H1N1))]

MDVNPTLLFLKVP AQNAISTTFPYTGDPPYSHGTGTGYTMDTVNRTHQYSEKARWTTNTETGAPQLNPID
GPLPEDNEPSGYAQ TDCVLEAMAFLEESHPIGIFENSCIETMEVVQQTRVDKLTQGRQTYDWT LN RNQPAA
55 TALANTIEVFRSNGLTANESGRLIDFLKDVME SMKKEEMGITTHFQRKRRVRDNMTKKMITQRTIGKRKQ
RLNKRSYLIRALT LNTMTKDAERGKLRRAIATPGMQIRGFVYFVETLARSICEKLEQSGLPVGGNEKKA
KLANVVRKMMTNSQDTELSLTITGDNTKWNENQNPRMFLAMITYMTRNQPEWFRNVLSIAPIMFSNKMAR
LGKGYMFESKSMKLRTQIPAEMLASIDLKYFNDSTRKKIEKIRPLLI EG TASLSPGMMMGMFNMLSTVLG
VSILNLGQKRYTKTTYWWDGLQSSDDFALIVNAPNHEGIQAGVDRFYRTCKLHGINMSKKKSYINRTGTF

- EFTSFFYRYGFVANFSMELPSFGVSGSNESADMSIGVTVIKNNMINNDLGPATAQMALQLFIKDYRYTYR
CHRGDTQIQTRRSFEIKKLWEQTRSKAGLLVSDGGPNLYNIRNLHIPEVCLKWELMDDEDYQGRLCNPLNP
FVSHKEIESMNNAVMMPAHGPAKNMEYDAVATTHSWIPKRNRSILNTSQRGVLEDEQMYQRCCNLFEKFF
PSSSYRRPVGISSMVEAMVSRARIDARIDFESGRIKKEEFTEIMKICSTIEELRRQK
- 5
SEQ ID NO:23
>gi|8486130|ref|NP_040982.1| nucleocapsid protein [Influenza A virus
(A/Puerto Rico/8/34(H1N1))]
MASQGTKRSYEQMETDGERQNA TEIRASVGKMIGGIGRFYIQMCTELKLSDYEGRLIQNSLTIERMVLSA
10 FDERRNKYLEEHPSAGKDPKKTGGPIYRRVNGKWMRELILYDKEEI RRIWRQANNGDDATAGLTHMMIWH
SNLNDATYQRTALVRTGMDPRMCSLMQGSTLPRRSGAAGAAVKGVGTMVMELVRMI KRGINDRNFWRGE
NGRKTRIAYERMCNILKGKFQTAAQKAMMDQVRESRDPGNAEFEDLTFLARSALILRGSVAHKSCLPACV
YGPASVAGSYDFEREGYSLVGIDPFRLLQNSQVYSLIRPNENPAHKSQLVWMACHSAAFEDLRVLSFIKGT
KVVPRGKLSTRGVQIASNENMETMESSTLELRSRYWAI RTRSGGNTNQQRASAGQISIQPTFSVQRNLPF
15 DRTTVMAAFTGNTEGRTSDMRTEIIRMMESARPEDVSFQGRGVFELSDEKAASPIVPSFDMSNEGSYFFG
DNAEEYDN
- 20
SEQ ID NO:24
>gi|73918826|ref|YP_308855.1| polymerase 2 [Influenza A virus
(A/Korea/426/1968(H2N2))]
MERIKELRNLSQSRTREILT KTTVDHMAIIKKYTSGRQEKNP SLRMKWMAMKYPI TADKRITEMVPER
NEQGQTLWSKMSDAGSDRVMVSPLAVTWWRNNGPMTSTVHYPKIYKTYFEKVERLKHGTFGPVHFRNQVK
IRRRVDINPGHADLSAKEAQDVIMEVVF PNEVGARILTS ESQLTITKEKKEELQDCKISPLMVAYMLERE
25 LVRKTRFLPVAGGTSSVYIEVLHLTQGT CWEQMYTPGGEVRNDDVDQSLIIAARNIVRRAAVSADPLASL
LEMCHSTQIGGTRMVDILRQNPTEEQA VDICKAAMGLRISSSFSGGFTEFKRTSGSSIKREEEVLTGNLQ
TLKIRVHEGYEEFTMVGKRATAILRKATRRLVQLIVSGRDEQSIAEAIIVAMVFSQEDCMIKAVRGDLNF
VNRRANQRLNPMHQLLRH FQKDAKVL FQNWGIEHIDNVMGMIGVLPDMTPSTEMSMRGIRVSKMGVDEYSS
TERVVV SIDRFLRVRDQ RGNVLLSPEEVSETQGT EKL TITYSSSMWEINGPESVLVNTYQWII RNWETV
30 KIQWSQNPTMLYNKMEFEPFQSLVPKAI RGQYSGFVRTLFQQMRDVLGTFD TTQIIKLLPFAAAPPKQSR
MQFSSSLTVNVRGSGMRILVRGN SPVFNYNKTTKRLTILGKDAGTLTEDPDEGTSGVESAVLRGFLILGKE
DRRYGPALSINELSTLAKGEKANVLIGQGDVVLVMKRKR DSSILTDSQTATKRIRMAIN
- 35
SEQ ID NO:25
>gi|73919145|ref|YP_308850.1| hemagglutinin [Influenza A virus
(A/Korea/426/68(H2N2))]
MAIIYLILLFTAVRGDQICIGYHANNSTEKVD TILERNVTVT HAKDILEKTHNGKLCKLNGIPPLELGDC
40 SIAGWLLGNPECDRLLSVPEWSYIMEKENPRYSLCYPGSFNDYEELKHLLSSVKHF EKVKILPKDRWTQH
TTTGGSWACAVSGKPSFFRN MVWLTRKGSNYPVAKGSYNNTSGEQMLIIWGVHHPNDEAEQRALYQNVGT
YVSVATSTLYKRSIPEIAARPKV NGLGRRMEFSWTL LDMWD TINFESTGNLVAPEYGFKISKRGSSGIMK
TEGTLENCETKCQTPLGAIN TTLPFHNVHPLTIGEC PKYVKSEKLVLATGLRNVPQIESRGLFGAIAGFI
EGGWQGMVDGWYGYHHSNDQ GSGYAADKESTQKAFNGITNKVNSVIEKMNTQFEAVGKEFSNLEKRLENL
45 NKKMEDGFLDVW TYNAELLVLMENERTLDFHDSNVKNLYDKVRMQLRDNVKELGNGCFEFYHKCDNECMD
SVKNGTYDYPKYEEESKLNRNEIKGVKLSSMGVYQILAIYATVAGSLSLA IMMAGISFWMCSNGSLQCRI
CI
- 50
SEQ ID NO:26
>gi|73912688|ref|YP_308854.1| membrane protein M1 [Influenza A virus
(A/Korea/426/68(H2N2))]
MSLLTEVETYVLSIVPSGPLKAEIAQRLEDVFAGKNTDLEALMEWLKTRPILSPLTKGILGFVFTLTVPS
ERGLQRRRFVQNALNGNDPNNMDRAVKLYRK LKREITFHGAKEVALSYSAGALASCMGLIYNRMGAVTT
EVAFAVVCATCEQIADSQHRSHRQMVTTTNPLIRHENRMVLASTTAKAMEQMAGSSEQAAEAMEVASQAR
55 QMVQAMRAIGTPPSSSAGLKD DDLLENLQAYQKRMGVQMQRFK
- SEQ ID NO:27
>gi|73912687|ref|YP_308853.1| membrane protein M2 [Influenza A virus
(A/Korea/426/68(H2N2))]

MSLLTEVETPIRNEWGCRCNDSSDPLVVAASIIGILHFILWILDRLEFFKCIYRFFKHGLKRGFSTEGVPE
SMREEYRKEQQSAVDADDSHFVSIELE

SEQ ID NO:28

5 >gi|73912685|ref|YP_308852.1| polymerase PA [Influenza A virus
(A/Korea/426/68 (H2N2))]
MEDFVRQCFNPMIVELAEKAMKEYGEDLKIETNKFAAICTHLEVCFMYSDFHFINEQGGSIMVELDDPNA
LLKHRFEIIEGRDRTMAWTVVNSICNTTGAEKPKFLPDLYDYKENRFIEIGVTRREVHIYYLEKANKIKS
ENTHIHIFSFTGEEMATKADYTLDEESRARIKTRLFTIRQEMANRGLWDSFRQSERGEETIEERFEITGT
10 MRRLADQSLPPNFSCLENFRAYVDGFEPNGYIEGKLSQMSKEVNAKIEPFLKTTPRPIRLPDGPPCFQRS
KFLMDALKLSIEDPSHEGEGIPLYDAIKCMRTFFGWKEPYIVKPHEKGINPNYLLSWKQVLAELQDIEN
EEKIPRTKNMKKTSQKQWALGENMAPEKVDFDNCRDISDLKQYDSDEPELRSLSWQNEFNKACELTDS
IWIELDEIGEDVAPIEHIA SMRRNYFTA EVSHCRATEYIMKGVYINTALLNASCAAMDDFQLIPMISKCR
TKEGRRKTNLYGFIIKGRSHLRNDTDVVNFVSMEFSLTDPRLEPHKWEKYCVLEIGDMLLRSAIGQMSRP
15 MFLYVRTNGTSKIKMKWGMEMRPCLLQSLQQIESMVEAESSVKEKDMTKEFFENKSETWPIGESPKGVEE
GSIGKVCRTLLAKSVFNSLYASPQLEGFSAESRKLLLQVQALRDNLEPGTFDLGGGLYEAIEECLINDPWV
LLNASWFNSFLTHALR

SEQ ID NO:29

20 >gi|73921833|ref|YP_308877.1| PB1-F2 protein [Influenza A virus
(A/Korea/426/68 (H2N2))]
MGQEQDTPWTQSTEHINIQKRGSGQQTRKLERPNLTQLMDHYLRMTNQVDMHKQTASWKQWLSLRNHTQE
SLKIRVLKRWKLEFNKQEWTN

SEQ ID NO:30

25 >gi|73912683|ref|YP_308851.1| PB1 polymerase subunit [Influenza A virus
(A/Korea/426/68 (H2N2))]
MDVNPTLLFLKVPQA NAISTTFPYTGDPYPYSHGTGTGYTMDTVNRTHQYSEK GKWTTNTETGAPQLNPID
GPLPEDNEPSGYAQTD CVLEAMAFLEESHGPIFENSCL ETMEVIQQTRVDKLTQGRQTYDWT LN RNQPAA
30 TALANTIEVFRSNGLTANESGRLIDFLKDVI ESMDKEEMEITTHFQKR RVRDNMTKKMVTQRTIGKKKQ
RLNKR SYLIRALT LNTMTKDAERGKLRRAIATPGMQIRGFVHFVETLARNICEKLEQSGLPVGGNEKKA
KLANVVRKMMTNSQDTELSFTITGDNTKWNENQNPRVFLAMITYITRNQPEWFRNVLSIAPIMFSNKMAR
LGKGYMFESKSMKLRTQIPAEMLASIDLKYFNESTRKKIEKIRPLLIDGTVSLSPGMMMGFMNMLSTVLG
VSILNLGQKKYTKTTYWWDGLQSSDDFALIVNAPNHEGIQAGVNRFYRTCKLVGINMSKKKSYINRTGTF
35 EFTSFFYRYGFVANFSMELPSFGVSGINESADMSIGVTVIKNMINNDLGPATAQMALQLFIKD YRYTYR
CHRGDTQIQTRRSFELKKLWEQTRSKAGLLVSDGGSNLYNIRNLHIPEVCLKWELMDEDYQGRLCNPLNP
FVSHKEIESVNNAVVMPAHGPAKSMEYDAVATTHSWTPKRNRSILNTSQRGILEDEQMYQKCCNLFEKFF
PSSSYRRPVGISSMVEAMVSRARIDARIDFESGRIKKEEFAEIMKICSTIEELRRQK

SEQ ID NO:31

40 >gi|73921567|ref|YP_308869.1| non-structural protein NS2 [Influenza A virus
(A/Korea/426/68 (H2N2))]
MDSNTVSSFQDILLRMSKMQLGSSSED LN GMITQFESLKLYRDSLGEAVMRMGDLHSLQNRNGKWREQLG
QKFEEIRWLIEEVHRHLKITENSFEQITFMQALQLLFEVEQEIRTF SFQLI

SEQ ID NO:32

45 >gi|73921566|ref|YP_308870.1| non-structural protein NS1 [Influenza A virus
(A/Korea/426/68 (H2N2))]
MDSNTVSSFQVDCFLWHVRKQVVDQELGDAPFLDRLRRDQKSLRGRGSTLDLDIEAATR VGKQIVERILK
50 EESDEALKMTMASAPASRYLTDMTIEELS RDWFMLMPKQKVEGPLCIRIDQAIMDKNIM LKANFSVIFDR
LETLLLLRAFTEEGAIVGEISPLPSLPGHTIEDVKNAIGVLIGGLEWNDNTVRVSKTLQRF AWRSSNENG
RPPLTPKQKRKMARTIRSKVRRDKMAD

SEQ ID NO:33

55 >gi|73921307|ref|YP_308871.1| nucleoprotein [Influenza A virus
(A/Korea/426/68 (H2N2))]
MASQGTKRSYEQMETDGERQ NATEIRASVGKMIDGIGRFYIQMCTELKLSDYEGRLIQNSLTIERMVLSA
FDERRNKYLEEHP SAGKDPKKTGGPIYKRVDGKWMRELVL YDKEEIRRIWRQANNGDDATAGLTHMMIWH
SNLNDTTYQRTALVRTGMDPRMC SLMQGSTLPRRSGAAGAAVKGVGTMMELIRMIKRGINDRNFWRGE

- 5 NGRKTRSAYERMCNILKGKFQTAQAQRAMMDQVRESRNPNGNAEIEDLIIFLARSALILRGSVAHKSCLPACV
YGPAIASGYNFEKEGYSLVGIDPFKLLQNSQVYSLIRPNENPAHKSQLVWMACNSAAFEDLRVLSFIRGT
KVSPRGKLSTRGVQIASNENMDTMESSTLELRSRYWAIRTRSGGNTNQQRASAGQISVQPAFSVQRNLPF
DKPTIMAAFTGNTGRTSDMRAEIIRMMEGAKPEEMS FQGRGVFELSDEKATNP IVP SFDMSNEGSYFFG
DNAEEYDN
- SEQ ID NO:34
>gi|73921304|ref|YP_308872.1| neuraminidase [Influenza A virus
(A/Korea/426/68 (H2N2))]
10 MNPNQKIITIGSVSLTIATVCFLMQIAILVTTVTTLHFKQHECDSPASNQVMPCEPIIIERNITEIVYLN
TTIEKEICPEVVEYRNWSKPQCQITGFAPFSKDNSIRLSAGGDIWVTREPYVSCDPGKCYQFALGQGTTL
DNKHSNDTIHDRI PHRTLMLNELGVPFHLGTRQVCVAWSSSSCHDGKAWLHVCVTGDDKNATASFIYDGR
LMSIGSWSQNILRTQESCEVCINGTCTVVM TDGSASGRADTRILFIEEGKIVHISPLSGSAQHVEECSC
YPRYPDVRICIRDNWKGSNRPVIDINMEDYSIDSSYVCSGLVGDTPRNDDRSSNSNCRNPNNERGNPGVK
15 GWAFDNGDDVWMGRTISKDLRSGYETFKVIGGWSTPNSKSQINRQVIVDSNNWSGYSGIFSVEGKRCINR
CFYVELIRGRQQETRVWWTNSIVVFCGTSGTYGTGSWPDGANINFMPI
- SEQ ID NO:35
20 >gi|73919213|ref|YP_308844.1| nonstructural protein 2 [Influenza A virus
(A/New York/392/2004 (H3N2))]
MDSNTVSSFQDILLRMSKMQLGSSSEDINGMITQFESLKIYRDSLGEAVMRMGDLHLLQNRNGKWREQLG
QKFEEIRWLIEEVHRHLKTTENSFEQITFMQALQLLFEVEQEIRTF SFQLI
- SEQ ID NO:36
25 >gi|73919212|ref|YP_308845.1| nonstructural protein 1 [Influenza A virus
(A/New York/392/2004 (H3N2))]
MDSNTVSSFQVDCFLWHIRKQVVDQELSDAPFLDRLRRDQRS LRGRGNTLGLDIKAATHVGKQIVEKILK
EESDEALKMTMVSTPASRYITDMTIEELSRNWFMLMPKQKVEGPLCIRMDQAIM EKNI MLKANFSVIFDR
30 LETIVLLRAFTEEGAIVGEISPLPSFP GHTIEDVKNAIGVLIGGLEWNDNTVRVSKNLQRFWRSSNENG
GPPLTPKQKRKMARTARSKV
- SEQ ID NO:37
35 >gi|73919207|ref|YP_308839.1| hemagglutinin [Influenza A virus (A/New
York/392/2004 (H3N2))]
MKTIIALS YILCLVFAQKLPGNDNSTATLCLGHHAVPNGTIVKTITNDQIEVTNATELVQSSSTGGICDS
PHQILDGENCTLIDALLGDPQC DGFQNKWDLFVERSKAYSNCYPYDVPDYASLRSLVASSGTLEFNES
FNWTGVTQNGTSSACKRRSNN SFFSRLNWLTHLKFKYPALNVTMPNNEKF DKLYIWGVHHPGTDNDQISL
YAQASGRITVSTKRSQQTVI PSIGSRPRIRDVPSRISYWTIVKPGDILLINSTGNLIAPRGYFKIRSGK
40 SSIMRSDAPIGKCNSECITPNGSIPNDKPFQNVNRITYGACPRYVKQNTLKLATGMRNVPEKQTRGIFGA
IAGFIENGWEGMVDGWYGFRHQNSEGTGQAADLKSTQAAINQINGKLNRLIGKTNEKFHQIEKEFSEVEG
RIQDLEKYVEDTKIDLWSYNAELLVALENQHTIDLTDSEMKNKLFERTKKQLRENAEDMGNGCFKIYHKCD
NACIGSIRNGTYDHDVYRDEALNNRFQIKGVELKSGYKDWILWISFAISCFLLCVALLGFIMWACQKGN I
RCNICI
45
- SEQ ID NO:38
>gi|73919153|ref|YP_308840.1| matrix protein 2 [Influenza A virus (A/New
York/392/2004 (H3N2))]
MSLLTEVETPIRNEWGCRCNDSSDPLVVAASIIGILHLILWILDRLFFKCVYRLF KHGLKRG PSTEGVPE
50 SMREEYRKEQQNAVDADDSHFVSIELE
- SEQ ID NO:39
>gi|73919152|ref|YP_308841.1| matrix protein 1 [Influenza A virus (A/New
York/392/2004 (H3N2))]
55 MSLLTEVETYVLSIVPSGPLKAEIAQRLEDVFAGKNTDLEALMEWLKTRPILSPLTKGILGFVFTLT VPS
ERGLQRRRFVQNALNGNDPNNMDKAVKLYRK LKREITFHGAKEIALSYSAGALASCMGLIYNRMGAVTT
EVAFGLVCATCEQIADSQHRSHRQMVATTNPLIKHENRMV LASTTAKAMEQMAGSSEQAAEAMEIASQAR
QM VQAMRAVGTHPSSSTGLRDDLL ENLQTYQKRMGVQMQRFK

- SEQ ID NO:40
>gi|73919150|ref|YP_308848.1| PB1-F2 protein [Influenza A virus (A/New York/392/2004(H3N2))]
MEQEQDTPWTQSTEHTNIQRRGSGRQIQKLGHPNSTQLMDHYLRIMSQVDMHKQTVSWRLWPSLKNPTQV
5 SLRTHALKQWKSFNKQGWTN
- SEQ ID NO:41
>gi|73919149|ref|YP_308847.1| polymerase PB1 [Influenza A virus (A/New York/392/2004(H3N2))]
10 MDVNPTLLEFLKVPAQNAISTTFPYTGDPPYSHGTGTGYTMDTVNRTHQYSEK GKWTTNTETGAPQLNPID
GPLPEDNEPSGYAQTD CVLEAMAFLEESHPIFENS CLETMEVVQQTRVDKLTQGRQTYDWT LN RNQ PAA
TALANTIEVFRSNGLTANESGRLIDFLKDVME SMDKEEMEITTHFQRKRRVRDNMTKKMVTQRTIGKKKQ
RVNKRGYLIRALT LNTMTKDAERGKLRRAIATPGMQIRGFVYFVETLARSICEKLEQSGLPVGGNEKKA
15 KLANVVRKMMTNSQDTELSFTITGDNTKWNENQNPRMFLAMITYITKNQPEWFRNILSIAPIMFSNKMAR
LGKGYMFESKRMKLRTQIPAEMLASIDLKYFNESTRKKIEKIRPLLIDGTASLSPGMMMGMFNMLSTVLG
VSVLNLGQKKYTKTTYWWDGLQSSDDFALIVNAPNHEGIQAGVDRFYRTCKLVGINMSKKKSYINKTGT
EFTSFFYRYGFVANFSMELPSFGVSGINESADMSIGVTVIKNNMINNDLGPATAQMALQLFIKDYRYTYR
CHRGDTQIQTRRSFELKKLWDQTQSRAGLLVSDGGPNLYNIRNLHIPEVCLKWELMDENYRGRLCNPLNP
20 FVSHKEIESVNNAVVM PAHGPAKSMEYDAVATTHSWNPKRNR SILNTSQRGILEDEQMYQKCCNLFEKFF
PSSSYRRPIGISSMVEAMVSRARIDARIDFESGRIKKEEFSEIMKICSTIEELRRQK
- SEQ ID NO:42
>gi|73919147|ref|YP_308843.1| nucleocapsid protein [Influenza A virus (A/New York/392/2004(H3N2))]
25 MASQGTKRSYEQMETDGD RQ NATEIRASVGKMIDGIGRFYIQMCTELKLS DHEGR LIQNSLTIEKMVLSA
FDERRNKYLEEHPSAGKDPKKTGGPIYRRVDGKWMRELVL YDKEEI RRIWRQANNGEDATAGLTHIMIWH
SNLNDATYQRTALVRTGMDPRMCSLMQGSTLPRRSGAAGAAVKGIGTMVMELIRMV KRGINDRNFWRGE
NGRKTRSAYERM CNILKGKFQTAAQRAMVDQVRESRNP GNAEIEDLI FLARSALI LRGSVAHKSCLPACA
YGP AVSSGYDFEKEGYSLVGIDPFKLLQNSQIYSLIRPNENPAHKSQLVMMACHSAAFEDLRLLSFIRGT
30 KVS PRGKLSTRGVQIASNENMDNMGSSTLELRSGYWAIRTRSGGNTNQQRASAGQTSVQPTFSVQRNL PF
EKSTIMAAFTGNT EGRTSDMRAEIIRMMEGAKPEEV SFRRG RVFELSDEKATNP IVP SFDM SNEGSYFFG
DNAEEYDN
- SEQ ID NO:43
35 >gi|73919136|ref|YP_308842.1| neuraminidase [Influenza A virus (A/New York/392/2004(H3N2))]
MNP NQKIITIGSVSLTISTICFFMQIAILITTVTLHFKQYEFNSPPNNQV MLCEPTIIERNITEIVYLTN
TTIEKEMCPKLA EYRNWSKPQC DITGFAPFSKDNSIRLSAGGDIWVTREPYVSCDPDKCYQFALGQGTTL
NNVHSNDTVHDRTPYRTL LMNELGV PFHLG TKQVCI AWSSSSCHDGKAWLHVCVTGDDKNATASFIYNGR
40 LVDSIVSWSKILRTQESECV CINGTCTVVM TDGSASGKADTKILFIEEGKIIHTSTLSGSAQHVEECSC
YPRYPGVRCVCRDNWKGSNRPIVDINIKDYSIVSSYVCSGLVGDTPRKNDSSSSSHCLDPNNEEGGHGVK
GWAFDDGNDVVMGRTISEKLRSGYETFKVIEGWSKPN SKLQINRQVIVDRGNRSGYS GIFSVEGKSCINR
CFYVELIRGRKEETEVLWTSNSIVVFCGTS GTYGTGSWPDGADINLMPI
- SEQ ID NO:44
45 >gi|73919134|ref|YP_308846.1| polymerase PA [Influenza A virus (A/New York/392/2004(H3N2))]
MEDFVRQC FNP MIVELA EKAMKEYGEDLKIETNKFAAICTHLEVCFMYSDFHFINEQG GESIVVELDDPNA
LLKHRFEIIEGRDRTMAWTVVNSICNTTGAEKPKFLPDLYDYKENRFIEIGVTRREVHIYYLEKANKIKS
50 ENTHIHI FSFTGEEIATKADYTLDEESRARIKTRLFTIRQEMANRGLWDSFRQSERGEETIEEKFEISGT
MRRLADQSLPPKFSCLENFRAYVDGFEPNGCIEGKLSQMSKEVNAKIEPFLKTTPRPIKLPNGPPCYQRS
KFL LMDALKLSIEDPSHEGEGIPLYDAIKCIKTFFGWKEPYIVKPHEKGINSNYLLSWKQVLSELQDIEN
EEKIPRTKNMKKTSQ LK WALGENMAPEKVDFDNCRDISDLKQYDSDEPELRSLS SWIQNEFNKACELTDS
IWIELDEIGEDVAPIEYIASMRNYFTA EVSHCRATEYIMKGVYINTALLNASCAAMDDFQLIPMISKCR
55 TKEGRRKTNLYGFIIKGRSHLRNDTDVVNFVSMEFSLTDPRLEPHKWEKYCVLEIGDMLLRS AIGQISRP
MFLYVRTNGTSKVKMKWGMEMRRC LLQSLQQIESMIEAESSIKEKDMTKEFFENKSEAWPIGESPKGVEE
GSIGKVCRTLLAKSVFNSLYASPQLEGFSAESRKL L LVVQALRDNL EPGT FDLGGLYEAIEECLINDPWV
LLNASWFNSFLTHALK

- SEQ ID NO:45
>gi|73919060|ref|YP_308849.1| polymerase PB2 [Influenza A virus (A/New York/392/2004(H3N2))]
MERIKELRNLSQSRTREILTCTTVDHMAIIKKYTSGRQEKNP SLRMKWMAMKYPITADKRITEMVPER
5 NEQGQTLWSKMSDAGSDRVMVSPLAVTWNRNGPVASTVHYPKVYKTYFDKVERLKHGTFGPVHFRNQVK
IRRRVDINPGHADLSAKEAQDVIMEVVFNPNEVGARILTSESQLTITKEKKEELRDCKISPLMVAYMLERE
LVRKTRFLPVAGGTSSIIYIEVLHLTQGTCWEQMYTPGGEVRNDDVDQSLIIAARNIVRRAAVSADPLASL
LEMCHSTQIGGTRMVDILRQNPTEEQAVDICKAAMGLRISSSFSGGFTEFKRTSGSSVKKEEEVLTGNLQ
TLKIRVHEGYEEFTMVGKRATAILRKATRRLVQLIVSGRDEQSIAEAIIVAMVFSQEDCMIKAVRGDLNF
10 VNRRANQRLNPMHQLLRHFFQKDAKVLFFQNWGIEHIDSVGMVGVLDPDMPSTEMSMRGIRVSKMGVDEYSS
TERVVVSI DRFLRVRDQRGNVLLSPEEVSETQGTERLTITYSSSMWEINGPESVLVNTYQWII RNWEAV
KIQWSQNPAMLYNKMEFEPFQSLVPKAIRSQYSGFVRTLFQQMRDVLGTFDTTQIIKLLPFAAAPPKQSR
MQFSSLT VNVRGSGMRILVRGNSPVFNYNKTTKRLTILGKDAGTLIEDPDESTSGVESAVLRGFLIIGKE
15 DRRYGPALSINELSNLAKGEKANVLI GQGDVVLVMKRKRDS SILTDSQTATKRIRMAIN
- SEQ ID NO:46: CMV Protein IE122:

>gi|39841910|gb|AAR31478.1| UL122 [Human herpesvirus 5]
20 MESSAKRKMDPDNPDEGPSSKVPRPETPVT KATTF LQTMLRKEVNSQLSLGDPLFP ELAEESLKT FEQVT
EDCNENPEKDVLAELGDILAQAVNHAGIDSSSTGHTLTTHSCSVSSAPLNKPTPTSVAVTNTPLPGASAT
PELSPRKKPRKTTRPFKVI IKPPVPPAPIMLPLIKQEDIKPEPDFTIQYRNKI IDTAGCIVISDSEEEQG
EEVETRGATASSPSTGSGT PRVTSPTHPLSQMNHPPLPDPLARPD EDDSSSSSSSSSCSSASDSESESEEMK
CSSGGGASVTSSHHGRGGFGSAASSLLSCGHQSSGGASTGPRKKKSKRI SELDNEKVRNIMKDKNTPFCTPNVQTRRG
25 RVKIDEVSRMFRNTNRSLEYKNLPFTIPSMHQVLDEAIKACKTMQVNNKGIQIIYTRNHEVKSEVDAVRCRLGTMCNLA
LSTPFLMEHTMPVTHPPEVAQRTADACNEG VKAAWSL KELHTHQLCPRSSDYRNMI IHA**ATPVDLLGALNLC LPLMQKF**
PKQVMVRI FSTNQGGFMLPIYETA AKAYAVGQFEQPTETPPEDLDTL SLAIEAAIQDLRNKSQ
- 30 SEQ ID NO:126:

>gi|4927721|gb|AAD33253.1|AF125673_2 E7 [Human papillomavirus type 16]
MHGDTPTLHEYMLDLQPETTDLYCYEQ LNDSSSEEDEIDGPAGQAEPDRAHYNIVTFCKCDSTLRLCVQ
STH**VDIRTLEDLLMGTLGIVCPI**CSQKP
35

EXAMPLE 1

- 5 The peptides according to the invention used in the following examples were synthesized by Schafer-N as c-terminal amides using the Fmoc-strategy of Sheppard, (1978) J.Chem.Soc., Chem. Commun., 539.

Cell penetration assay

- 10 A set of peptides were biotinylated on N-terminal, and different combinations of aminoacids, with respect to length and type, were added to the sequence box X¹, X³ and X⁴ in the peptides as illustrated by the diagram below. The peptides were tested on cells grown from one individual blood donor.

- 15 Schematic diagram of amino acid sequence of the peptides according to the invention (Each X defines a sequence of amino acids):

X ¹	X ²	X ³	X ⁴	X ⁵
----------------	----------------	----------------	----------------	----------------

SP_2 to SP_17 are variants of 51n_Biotin, from a specific native domain on the HCV core protein.

20

Intracellular staining for biotinylated peptides

- 96-well U-bottom polystyrene plates (NUNC, cat no: 163320) were used for staining of human PBMCs. Briefly, 8ul of N- or C-terminally biotinylated peptides according to table 1 or table 2 (i.e. 5mM, 2.5mM & 1.25mM tested for each peptide) were incubated at 37°C for 2h
 25 with 40ul of PBMC (12.5 x 10⁶ cells/ml) from blood donors. Cells were then washed 3x with 150ul of Cellwash (BD, cat no: 349524), followed by resuspension of each cell pellet with 100ul of Trypsin-EDTA (Sigma, cat no: T4424), then incubated at 37°C for 5 min. Trypsinated cells were then washed 3x with 150ul of Cellwash (BD, cat no: 349524), followed by resuspension with BD Cytofix/Cytoperm™ plus (BD, cat no: 554715), then incubated at
 30 4°C for 20 min according to manufacturer. Cells were then washed 2x with 150ul PermWash (BD, cat no: 554715). Cells were then stained with Streptavidin-APC (BD, cat no: 554067) & Anti-hCD11c (eBioscience, cat no: 12-0116) according to manufacturer at 4°C for 30 min

aiming to visualize biotinylated peptides & dendritic cells, respectively. Cells were then washed 3x with 150ul PermWash, followed by resuspension in staining buffer (BD, cat no: 554656) before flow cytometry. Dendritic cells were gated as CD11c+ events outside lymphocyte region (i.e. higher FSC & SSC signals than lymphocytes). 200 000 total cells were acquired on a FACSCanto II flow cytometer with HTS loader, and histograms for both total cells & dendritic cells with respect to peptide-fluorescence (i.e. GeoMean) were prepared.

Extracellular staining for biotinylated peptides

96-well U-bottom polystyrene plates (NUNC, cat no: 163320) were used for staining of human PBMCs. Briefly, 8ul of N- or C-terminally biotinylated peptides according to table 1 or table 2 (i.e. 5mM, 2.5mM & 1.25mM tested for each peptide; all peptides manufactured by Schafer) were incubated at 37°C for 2h with 40ul of PBMC (12.5 x 10⁶ cells/ml) from blood donors. Cells were then washed 3x with 150ul of Cellwash (BD, cat no: 349524), then stained with Streptavidin-APC (BD, cat no: 554067) & Anti-hCD11c (eBioscience, cat no: 12-0116) according to manufacturer at 4°C for 30 min aiming to visualize biotinylated peptides & dendritic cells, respectively. Cells were then washed 3x with 150ul of Cellwash (BD, cat no: 349524), followed by resuspension in staining buffer (BD, cat no: 554656) before flow cytometry. Dendritic cells were gated as CD11c+ events outside lymphocyte region (i.e. higher FSC & SSC signals than lymphocytes). 200 000 total cells were acquired on a FACSCanto II flow cytometer with HTS loader, and histograms for both total cells & dendritic cells with respect to peptide-fluorescence (i.e. GeoMean) were prepared.

It was clearly seen that the arginine added on the beginning, middle and in the end improves the ability to enter the cell. It was observed that the variations of arginine decides the ability to enter the cell, best result are achieved with two in the beginning and two in the middle – but variations of different amount of arginines all show good result.

The data shown in the tables below are geomean-value of each testet peptide, as calculated by the FACS Duva software. The Geomean values by trypsinating/Cytofix/Cytoperm™;

Table 3

With Cytofix/Cytoperm and trypsination

BC20	Concentration			
	833µM	417 µM	208 µM	104 µM
51n_Biotin	ND	ND	ND	ND
SP 2	3136	5486		1211

SP 3	40456	25905	10788	4308
Sp 4	41491	28008	7540	2946
Sp 5	40534	46060	19065	6152
Sp 6	32577	22195	6514	1850
Sp 7	47922	30503	10481	2766
SP 8	31976	28606	9617	2917
SP 9	11344	6060	2945	1268
SP 10	43389	18607	7684	3709
SP 11	8262	3562	1277	601
SP 12	11089	10869	8679	4316
SP 13	568	394	332	
SP 14	1059	682	392	62
SP 15	1810	477	848	600
SP 16	4558	2240	1260	505
SP 17	30248	29967	12244	1500
SP 18	4869	1334	1971	1038
SP 19	17018	7201	3466	1625
SP 20	7319	2457	1268	445
- ve	0	0	0	0
+ ve	18694	26046	15340	7182
No peptide	224	106	78	168

As shown in table 4 the peptides according to the invention had the ability to allocate through the cell membrane.

5 Table 4: Results on DC-cells with low concentration of peptides (208 uM). The results are sorted in descending order (with respect to BC 16).

Peptide	BC 16	BC43
no pept	4473	2488
NO pept	4003	2359
NO pept	7799	
P-biotin	22058	33545
P-biotin	5568	26212
<i>P-biotin</i>	28169	
n-biotin	3745	1967
N-biotin	3591	3631

N-biotin	2271	
SP61_2	128330	150491
42	74920	65811
SP42_1	73973	
SP17	67428	
SP51_1*	66582	
BI10026	65196	43536
SP11b_1	64047	
42b	61968	19181
SP22	61597	18624
51 biotin	59789	33935
SP12	59274	31222
BI100260	47846	
SP61_3	47182	36803
V10	45430	
120b	44185	
100-22b	43537	28780
BI100260b	37702	
SP12_c	37100	43386
18b	29740	33597
SP13	28539	9129
BI10024b	26892	13002
SP21	26525	4294
61b biotin	25611	24416
SP61b_2	25489	
310-11-n-sc	25236	
100-12	24792	45622
SP5	21358	29146
SP4_c	20673	30786
SP7_c	20041	11832
SP20	19051	3383
190n	18616	23383
SP3_c	18505	26308
SP24	18181	63639
SP6_c	17964	38487
SP42b_1	17628	
SP23	17084	12801
SP5_c	16908	24403
SP7	16064	21414
SP8	15379	17836
BI-050sc5	13648	15549
190	12906	
SP61_4	12829	
190b	12391	20444

SP17_c	11830	
SP25	11552	
SP4	10751	17372
SP6	10233	14468
SP3	9480	10573
SP1_c	9234	
SP8_c	9159	23763
SP10	8627	9216
19	8417	6372
SP9	8356	10084
N13	7982	5953
51b	7000	6082
SP10_c	6904	
SP2	6677	6112
310-11	6537	
SP19	6522	6536
SP11	6490	13230
SP9_c	6180	7758
BI-050sc1	6056	6160
BI-050sc2	6019	11584
SP16	5730	3656
51n	5551	7721
SP18	5545	7751
310-11n	5436	
SP11_c	5371	7878
SP15	5076	2539
N10	4597	
SP14	3704	1629
SP2_c	323	13341
SP13_c		60307
42n		

Example 2

Positive CTL response may alternatively be assayed by ELISPOT assay.

Human IFN-gamma cytotoxic T-cell (CTL) response by ELISPOT assay

- Briefly, at day 1, PBMC samples from HCV patients were incubated in flasks (430 000 PBMCs/cm²) for 2h at 37°C, 5% CO₂ in covering amount of culture media (RPMI 1640 Fisher Scientific; Cat No. PAAE15-039 supplemented with L- Glutamine, (MedProbe Cat. No. 13E17-605E, 10% Foetal Bovine serum (FBS), Fisher Scientific Cat. No. A15-101) and

Penicillin/Streptomycin, (Fisher Acientific Cat. No. P11-010) in order to allow adherence of monocytes. Non-adherent cells were isolated, washed, and frozen in 10% V/V DMSO in FBS until further usage. Adherent cells were carefully washed with culture media, followed by incubation at 37°C until day 3 in culture media containing 2µg/ml final concentration of hrGM-CSF (Xiamen amoytop biotech co, cat no: 3004.9090.90) & 1µg/ml hrIL-4 (Invitrogen, Cat no: PHC0043), and this procedure was then repeated at day 6. At day 7, cultured dendritic cells (5 000-10 000 per well) were added to ELISPOT (Millipore multiscreen HTS) plates coated with 0.5µg/well anti-human γ Interferon™ together with thawed autologous non-adherent cells (200 000 per well), antigen samples (1-8ug/ml final concentration for peptide antigens; 5ug/ml final concentration for Concanavalin A (Sigma, Cat no: C7275) or PHA (Sigma, Cat no: L2769)) & anti-Anergy antibodies (0.03-0.05ug/ml final concentration for both anti-PD-1 (eBioscience, cat no: 16-9989-82) & anti-PD-L1 (eBioscience, cat no: 16-5983-82)). Plates were incubated overnight and spots were developed according to manufacturer. Spots were read on ELISPOT reader (CTL-ImmunoSpot® S5 UV Analyzer).

Table 5: ELISPOT of core proteins

		Patient A		Patient B		Patient C		Patient D	
	ConA	5090	5035	2605	2725	1355	700	529	512
	No Pep	20	25	10	25	10	0	4	1
	Irr.peptide	10	30	40	5	45	0	1	2
RRSRNLGKVIDTLTCBRLMGYIPLVGA	310-42b	90	115	60	80	70	25	13	11
RRIRNLGRVIETLTGBRLZGYIPLIGA	310-42	75	80	95	30	25	55	13	11
RRGYIPLVGAPLGBRVARALAHGVRV	310-51b	45	40	50	85	45	75	9	13
RRGYLPAVGAPIGBRVIRVIAHGLRL	310-51	40	75	50	60	100	35	9	9
RRNYATGNLPGRRGCSFSIFLLAL	310-61b	45	35	25	45	55	40	15	17
RRGGGQIIGGNYLIPRBGPBIGVRATB	310-11	ND	ND	ND	ND	ND	ND	11	14

Table 5 – Results from IFN-γ ELISPOT run on scaffold peptides with domains derived from HCV proteins

The peptides had in general a higher amount of T-cell spots compared to No Peptide and the irrelevant control peptide. This clearly indicates that the peptides have a positive immunological effect in HCV-patients.

Example 3

The REVEAL™ ProVE® Rapid Epitope Discovery System in Detail

Binding properties to HLA for the ninemers listed in table 6 were tested for the following HLA-classes:

HLA-A1, HLA-A2, HLA-A3, HLA-A11, HLA-A24, HLA-A29, HLA-B7, HLA-B8, HLA-B14, HLA-B15, HLA-B27, HLA-B35, HLA-B40.

- 5
- The peptides were synthesized as a Prospector PEPscreen®: Custom Peptide Library. Peptides 8-15 amino acids in length were synthesized in 0.5-2mg quantities with high average purity. Quality control by MALDI-TOF Mass Spectrometry was carried out on 100% of samples.
- 10
- The REVEAL™ binding assay determined the ability of each candidate peptide to bind to one or more MHC class I alleles and stabilizing the MHC-peptide complex. By comparing the binding to that of high and intermediate affinity T cell epitopes, the most likely immunogenic peptides in a protein sequence can be identified. Detection is based on the presence or absence of the native conformation of the MHC-peptide complex.
- 15
- Each peptide is given a score relative to the positive control peptide, which is a known T cell epitope. The score of the test peptide is reported quantitatively as a percentage of the signal generated by the positive control peptide, and the peptide is indicated as having a putative pass or fail result. Assay performance is confirmed by including an intermediate control peptide that is known to bind with weaker affinity to the allele under investigation.

20

Table 6:

HLA-type	A1	A2	A3	A11	A24	A29	B7	B8	B14	B15	B27	B35	B4
Positive Control	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100
Intermediate Control	8,41	8,79	10	14,95	10	30,52	12,00	56,60	24,21	56,88	10,49	93,56	76

AA placement in HCV-core protein	Linear Sequence													
163-171 " (mod)	NYATGNLPG NYVTGNIPG	11,16 N/D	3,94 N/D	8,65 N/D	4,69 N/D	1,13 N/D	34,26 N/D	0,51 N/D	9,96 N/D	0,79 N/D	6,06 N/D	2,33 N/D	6,44 N/D	75 I
164-172 " (mod)	YATGNLPGC YVTGNIPGS	8,10 13,33	22,46 6,88	3,67 5,34	2,16 5,18	1,20 2,71	18,82 10,29	0,76 2,70	9,51 12,40	0,41 0,76	5,09 6,16	2,44 1,98	4,52 4,42	77 88
165-173 " (mod)	ATGNLPGCS VTGNIPGST	7,13 16,18	4,62 8,44	3,64 2,52	2,54 4,85	0,64 1,60	22,84 15,05	1,12 2,24	10,73 13,12	0,87 0,75	3,81 5,15	2,56 1,58	7,00 4,00	90 103
166-174 " (mod)	TGNLPGCSF TGNIPGSTY	11,97 10,58	5,18 3,59	3,71 13,04	2,56 6,06	4,81 0,86	18,89 35,10	2,38 1,32	8,62 10,25	0,86 1,15	45,07 52,30	2,37 1,98	11,40 48,40	73 38
167-175 " (mod)	GNLPGCSFS GNIPGSTYS	6,44 7,54	6,11 5,07	3,54 5,04	2,41 1,54	1,13 0,00	13,48 24,38	1,84 0,94	8,71 6,32	0,81 0,34	3,96 4,80	13,44 1,58	4,80 7,86	80 55
168-176 " (mod)	NLPGCSFSI NIPGSTYSL	13,15 8,41	47,28 5,71	8,15 3,98	3,47 3,20	85,91 8,50	22,34 31,99	1,54 4,80	9,89 7,66	0,36 0,46	3,11 4,46	3,37 2,29	5,31 4,71	60 37

171-179 " (mod)	GCSFSIFLL GITFSIYLI	3,65	5,69	1,85	4,77	2,18	8,96	2,88	11,10	0,32	4,28	2,53	3,03	52
		9,61	7,27	1,42	4,88	3,05	10,42	2,07	8,94	0,63	8,15	3,29	9,27	68
172-180 " (mod)	CSFSIFLLA ITFSIYLIIV	9,69	23,38	5,68	9,00	1,37	6,45	2,22	9,85	0,93	11,66	2,99	4,96	33
		9,89	43,08	1,66	5,31	0,77	24,10	4,32	11,38	1,53	6,68	4,79	12,30	78
173-181 " (mod)	SFSIFLLAL TFSIYLIIVS	5,68	2,70	1,42	2,21	3,28	30,35	1,44	7,98	0,60	8,89	3,06	5,33	21
		9,93	2,02	0,91	2,89	0,60	8,58	1,28	9,52	0,37	8,63	1,49	2,48	23
174-182 " (mod)	SRNLGKVID IRNLGRVIE	15,30	4,94	3,46	4,45	0,71	9,94	1,28	16,71	0,67	11,25	1,28	4,59	33
		14,22	8,77	3,89	5,31	0,61	11,06	1,43	9,10	30,77	4,06	1,84	1,98	21
175-183 " (mod)	RNLGKVIDT RNLGRVIET	18,90	6,42	3,15	6,22	1,04	10,18	2,42	13,80	0,61	11,79	2,73	4,40	48
		15,93	5,96	5,52	13,68	0,51	15,13	0,71	15,51	0,26	8,84	0,91	2,34	27
176-184 " (mod)	NLGKVIDTL NLGRVIETL	18,14	7,70	6,41	12,81	1,16	10,67	2,64	16,86	0,57	12,59	2,64	4,39	53
		18,24	11,97	2,86	4,53	0,86	16,44	1,47	15,58	0,07	16,29	1,66	2,82	38
177-185 " (mod)	LGKVIDTLT LGRVIETLT	22,08	4,05	7,39	12,95	0,70	15,99	0,94	17,96	0,43	7,14	2,90	5,11	44
		16,33	5,90	7,40	20,27	0,97	12,31	2,66	14,74	0,28	16,38	2,54	3,43	62
178-186 " (mod)	GKVIDTLTC GRVIETLTS	14,77	4,89	3,69	3,96	0,41	14,26	0,25	13,63	0,16	7,31	0,84	3,11	19
		14,42	5,26	12,46	27,62	0,88	15,59	1,25	12,91	0,20	8,31	1,05	3,56	38
179-187 " (mod)	KVIDTLTCG RVIETLTSG	15,77	19,13	4,27	8,91	0,57	13,16	0,57	13,17	0,18	10,13	0,63	2,74	32
		11,61	9,82	2,23	3,25	0,96	15,52	0,50	11,83	0,11	37,99	2,32	1,90	32
180-188	VIDTLTCGF	25,85	5,14	2,93	6,00	0,95	13,90	0,45	10,00	0,11	9,49	1,11	3,26	27

WO 2012/072088

PCT/DK2011/050460

WO 2012/072088														PCT/DK2011/050460	
" (mod)	VIE ~~T~~ L ~~T~~ S ~~G~~ F	14,61	4,84	2,55	3,66	0,82	13,07	0,71	12,96	0,35	10,37	1,38	2,39	24	
131-139	ADLMGYIPL	31,05	8,06	1,93	3,35	0,61	14,51	0,73	11,40	0,13	8,30	1,01	2,08	26	
" (mod)	AEL[Nle]GYIPL	16,63	7,49	3,05	3,61	0,37	20,78	1,11	14,74	0,12	10,20	0,77	1,91	72	
132-140	DLMGYIPLV	16,67	103,34	4,66	5,01	1,14	25,64	4,84	22,65	0,23	7,67	1,05	2,49	39	
" (mod)	EL[Nle]GYIPLI	18,29	97,87	4,54	3,21	0,92	20,93	0,68	17,05	0,14	8,91	0,67	2,55	43	
133-141	LMGYIPLVG	14,17	31,94	2,53	2,13	0,51	24,26	0,41	20,75	0,08	9,43	1,44	0,81	33	
" (mod)	L[Nle]GYIPLIG	14,33	49,48	3,98	2,48	0,71	25,94	0,89	17,29	0,15	14,60	0,73	2,85	57	
134-142	MGYIPLVGA	11,34	10,71	3,40	4,79	0,35	19,54	0,80	14,88	0,73	8,85	0,83	0,85	40	
" (mod)	[Nle]GYIPLIGA	17,87	10,43	3,13	7,27	1,33	27,85	1,78	21,35	0,29	11,07	1,42	5,89	60	
135-143	GYIPLVGAP	15,62	7,28	3,46	3,84	0,80	29,74	1,37	15,94	0,47	11,02	1,76	10,35	72	
" (mod)	GYLPAVGAP	12,04	5,87	2,90	4,25	0,35	26,03	3,10	17,38	0,11	10,50	0,73	1,84	27	
136-144	YIPLVGAPL	21,80	35,00	5,29	3,54	4,51	31,09	41,87	33,10	0,13	54,57	0,69	99,81	30	
" (mod)	YLPavgAPI	15,76	75,80	2,58	3,07	4,65	28,40	7,70	17,92	0,12	49,63	0,51	2,03	64	
137-145	IPLVGAPLG	13,28	4,32	2,59	2,05	0,60	26,79	2,34	17,87	0,03	10,01	0,35	30,74	41	
" (mod)	LPAVGAPIG	20,57	5,60	3,02	4,95	0,61	25,32	8,12	15,87	0,12	10,10	0,49	1,06	35	
147-155	VARALAHGV	16,10	5,85	2,77	3,47	0,32	29,43	31,59	14,31	0,17	11,26	0,62	0,27	44	
" (mod)	VIRVIAHGL	18,66	8,60	4,29	4,54	0,28	27,12	29,93	32,81	0,55	9,64	1,16	2,51	52	
148-156	ARALAHGVR	20,62	6,72	2,35	3,83	1,15	31,96	4,70	18,91	1,10	13,58	3,55	2,45	50	
" (mod)	IRVIAHGLR	15,82	4,83	3,58	2,50	0,50	23,19	3,50	13,34	0,10	8,99	9,37	0,66	26	

149-157 " (mod)	RALAHGVRV RVIAHGLRL
27-35 " (mod)	GGQIVGGVY GGQIIIGGNY
28-36 " (mod)	GQIVGGVYL GQIIIGGNYL
29-37 " (mod)	QIVGGVYLL QIIIGGNLYLI
30-38 " (mod)	IVGGVYLLP IIGGNYLIP
31-39 " (mod)	VGGVYLLPR IGGNYLIPR
41-49 " (mod)	GPRLGVRAT GP[Cit]IGVRAT
42-50 " (mod)	PRLGVRATR P[Cit]IGVRAT[Cit]
43-51 " (mod)	RLGVRATRK [Cit]IGVRAT[Cit]R

20,77	8,69	2,51	4,45	0,84	29,73	2,55	18,08	0,46	11,84	6,37	6,07	70
11,60	18,68	13,32	17,75	0,79	23,22	102,96	12,60	0,42	106,77	15,57	1,43	24
13,97	6,88	3,47	2,71	1,78	20,46	3,99	14,66	0,32	110,34	0,66	0,54	30
24,04	7,06	2,23	2,40	0,39	14,38	1,09	10,52	0,25	90,62	0,24	2,22	15
15,07	55,85	4,80	4,46	0,50	13,73	6,56	14,53	0,37	182,45	1,01	3,24	24
20,65	12,96	3,01	2,77	0,47	14,17	2,39	11,90	0,22	173,66	0,64	2,60	31
20,25	57,04	2,36	1,96	1,01	13,36	28,66	13,15	0,37	15,07	0,60	4,01	24
21,63	73,54	6,30	4,06	12,60	13,60	52,04	10,61	0,40	29,59	0,91	8,99	83
22,11	6,78	1,88	3,10	0,30	12,38	24,16	11,48	1,65	15,00	0,68	2,68	15
21,14	8,27	3,88	2,81	10,95	13,88	15,47	12,18	1,17	19,93	0,97	5,35	65
17,88	8,73	3,41	4,16	0,34	14,20	1,15	11,57	0,44	13,22	0,79	4,02	55
17,90	7,04	2,88	3,57	0,89	14,85	20,49	14,23	0,32	15,02	0,82	2,54	25
13,85	6,23	2,45	1,60	0,05	12,66	124,37	13,18	0,94	10,01	0,68	1,83	15
18,60	9,16	3,48	2,72	0,16	11,88	5,38	9,99	0,13	7,29	0,59	2,25	15
17,77	6,85	3,63	2,73	0,97	13,35	42,64	11,46	0,14	10,04	0,86	1,65	14
18,83	7,36	2,52	1,72	0,21	14,14	1,91	9,47	0,19	12,64	0,94	2,45	45
20,39	6,71	112,23	102,85	0,28	12,83	1,67	10,21	0,07	11,74	33,13	2,83	25
19,68	8,17	25,16	8,27	0,50	16,67	14,59	12,09	0,17	14,45	0,86	8,51	18

Example 4

Intracellular staining:

Peptides according to the invention with X² and X⁴ derived from HCV, Influenza, or CMV were prepared and tested for intracellular staining in an experiment as described above in the “Cell

5 penetration assay”.

Peptide no.	Antigen	X ¹	X ²	X ³	X ⁴ (Different from X ² unless indicated)	X ⁵	median	n
1	Negative control		PVVHLTLRQAGDDFSR				1.00	
2	Neg control in scaffold	RRG	PVVHLTL	RRRG	QAGDDFS		4.12	
3	Positive control		RQIKIWFQNRMRMKWKK				2.73	
4	Positive control (tat)		YGRKKRRQRRR				4.43	
5	HCV (Native sequence)		11 amino acids	G	11 amino acids		1.84	
6	HCV	R	11 amino acids of native seq	RRR	11 amino acids of native seq	R	5.35	
7	HCV	R	11 amino acids (1 substitution)	RRR	11 amino acids (5 substitution)	R	11.59	
8	HCV	RR	11 amino acids of native seq	RR	11 amino acids of native seq		6.92	
9	HCV	RR	11 amino acids of native seq	RRR	11 amino acids of native seq		3.25	
10	HCV (Not scaffold)	EE	11 amino acids of native seq	EE	11 amino acids of native seq		1.16	
11	HCV (Not scaffold)	GG	11 amino acids of native seq	GG	11 amino acids of native seq		2.01	
12	HCV (Native sequence)		23 amino acids					3.01
13	HCV	RR	13 amino acids	RR	10 amino acids of native seq		21.43	
14	HCV	RR	13 amino acids	RR	10 amino acids		4.81	
15	HCV (Not scaffold)	EE	13 amino acids	EE	9 amino acids		1.16	
16	Influenza (Native sequence)		23 amino acids					3.70
17	Flu	BR	10 amino acids	BRGR	8 amino acids		20.77	
18	cmv	R	9 amino acids	BRR	=X ²	B	3.86	
19	HCV	W	11 amino acids of native seq	RR	11 amino acids of native seq		17.41	
20	HCV	RR	22 amino acids	RR	22 amino acids	R	13.25	

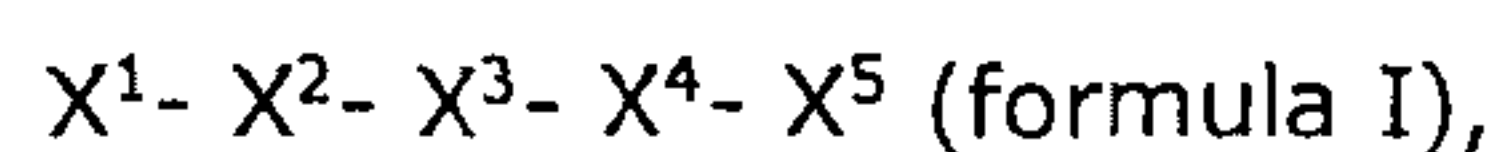
Extracellular staining for the same peptides 1-15:

Peptide no.	Antigen	median	n
1	Negative control	1.00	35
2	Neg control in scaffold	1.77	20
3	Positive control	9.15	19
5	HCV Native sequence	2.08	21
6	HCV	2.20	29
7	HCV	29.42	12
8	HCV	3.06	28
9	HCV	2.06	23
10	HCV (Not scaffold)	1.95	27
11	HCV (Not scaffold)	2.62	23
12	HCV (Native sequence)	1.76	14
13	HCV	16.80	7
14	HCV	17.00	16
15	HCV (Not scaffold)	2.38	12
19	HCV	1.37	19
20	HCV	126.5	7

Results: It was seen that peptides on scaffold form, compared to native peptides or peptides
5 with other amino acids in X1, X3 and X5 had improved uptake.

Claims

1. An isolated cell-penetrating peptide consisting of the following structure



wherein X^1 and X^3 independently define a linear sequence of any 2 or 3 arginines; X^2 and X^4 independently define a linear sequence of 8-15 amino acids of a single antigen of HCV, CMV, HPV, or Influenza virus or a variant sequence thereof comprising 1-10 substitutions, additions, or deletions relative to the antigen and sharing at least 50 % sequence identity to the corresponding native sequence of said antigen; said sequence X^4 being different from X^2 ; and wherein X^5 is an arginine or absent; and

wherein the sequence of amino acids defined by $X^1-X^2-X^3-X^4-X^5$ of formula I does not contain a continuous sequence of four or more arginine residues, which cell-penetrating peptide is of 18-60 amino acids, wherein the sequence of amino acids defined by $X^1-X^2-X^3-X^4-X^5$ of formula I is not found in the native sequence of said antigen, and

wherein the isolated peptide is a peptide monomer.

2. The isolated peptide according to claim 1, wherein X^2 and X^4 independently define a linear sequence sharing at least 60% or 70% sequence identity to the native sequence of said antigen.
3. The isolated peptide according to claim 1, wherein the peptide comprises one or more cysteines.
4. The isolated peptide according to claim 1, wherein the N- and/or C-terminal amino acid in X^2 is a hydrophilic or polar amino acid.
5. The isolated peptide according to claim 1, wherein the N-terminal amino acid in X^4 is a hydrophilic or polar amino acid.
6. The isolated peptide according to claim 1, wherein X^1 and/or X^3 consists of 2 or 3 arginines.
7. The isolated peptide according to claim 6, wherein X^1 consists of RR.
8. The isolated peptide according to claim 7, wherein X^3 consists of RR.

9. The isolated peptide according to claim 1, wherein X² and/or X⁴ is a linear sequence of 8-12 amino acids.
10. A dimer peptide comprising two peptides monomers, wherein each peptide monomer is as defined in claim 1.
11. A peptide combination comprising two or more peptides according to claim 1.
12. An immunogenic composition comprising at least one peptide according to claim 1 or a dimer peptide comprising two peptide monomers, wherein each peptide monomer comprises an isolated cell-penetrating peptide comprising the following structure

$$X^1-X^2-X^3-X^4-X^5 \text{ (formula I),}$$

wherein X¹ and X³ independently define a linear sequence of any 2 or 3 arginines; X² and X⁴ independently define a linear sequence of 8-15 amino acids from the same antigen selected from HCV, CMV, HPV, or Influenza virus or a variant sequence thereof comprising 1-10 substitutions, additions or deletions relative to the antigen and sharing at least 50 % sequence identity to the native sequence of said antigen; said sequence X⁴ being different from X²; and wherein X⁵ is an arginine or absent; which cell-penetrating peptide is of 18-60 amino acids, and wherein the sequence of amino acids defined by X¹-X²-X³-X⁴-X⁵ of formula I is not found in the native sequence of said antigen; and

wherein the sequence of amino acids defined by X¹-X²-X³-X⁴-X⁵ of formula I does not contain a continuous sequence of four or more arginine residues

in combination with a pharmaceutically acceptable diluent or vehicle.
13. The immunogenic composition according to claim 12 in the form of a vaccine composition.
14. The immunogenic composition of claim 12 further comprising an immunological adjuvant.
15. The peptide according to claim 1, wherein X⁵ is an arginine.
16. The immunogenic composition according to claim 12, wherein X⁵ is an arginine.

$$X^1 - X^2 - X^3 - X^4 - X^5 \text{ (formula I)}$$