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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 July 2012 (12.07.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/092934 A1

(51) International Patent Classification:
C07K 14/005 (2006.01) *C07K 14/18* (2006.01)
C07K 14/11 (2006.01) *A61K 39/12* (2006.01)
C07K 14/16 (2006.01)

(21) International Application Number:
PCT/DK2012/050010

(22) International Filing Date:
6 January 2012 (06.01.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11150323.1 6 January 2011 (06.01.2011) EP
61/475,988 15 April 2011 (15.04.2011) US

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

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **LANGE, Einer Tønnes** [NO/NO]; Strømdalbakken 6, NO-3744 Skien (NO). **GRØNVOLD, Maja Sommerfelt** [NO/NO]; Akland, NO-4950 Risør (NO). **SØRENSEN, Birger** [NO/NO]; Meierlia 3, NO-3744 Skien (NO). **LAWITZ, Karolina** [SE/SE]; Geijersgatan 16B, SE-21618 Limhamn (SE).

(74) Agents: **HANSEN, Carsten Borgund** et al.; Inspicos A/S, P.O. Box 45, Kogle Allé 2, 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))*

WO 2012/092934 A1

(54) Title: MONOMERIC AND MULTIMERIC IMMUNOGENIC PEPTIDES

X^1-X^2-X^3-X^4-X^5-X^6 (formula I)

(57) Abstract: The present invention relates to novel peptides and methods for inducing an immune response in a subject against an antigen and for treatment, diagnosis and prognosis of infections or autoimmune diseases including infections with HCV, HIV, CMV and Flu. The invention further relates to methods for identifying and providing peptides useful for the treatment and diagnosis.

Monomeric and multimeric immunogenic peptides

FIELD OF THE INVENTION

The present invention relates to novel peptides and methods for stimulating humoral or B-cell immunity, such as in the treatment, diagnosis and prognosis of infections with pathogens associated with infectious diseases including infections with HCV, HIV, CMV and Flu as well as antigens associated with autoimmune diseases. The invention further relates to methods for identifying and providing peptides useful in stimulating humoral or B-cell immunity, such as in the treatment and diagnosis of such diseases.

BACKGROUND OF THE INVENTION

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; the humoral immunity which involves the development of antibodies that can bind virus particles and certain antibodies that can neutralize infection, and the cell mediated immunity that leads to the development of cytotoxic T-cells that kill infected cells exposing viral epitopes in the context of HLA class I, in this way eliminating infected cells.

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 that they; are caused by viruses with stable virus antigens (i.e. very low mutation frequency, few subtypes), lack a reservoir in other animal species, 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, 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.

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 infection leads to chronic infection, which can result in long-term complications and even death.

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.

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 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 proteins enable the virus to escape immune surveillance by the host. As a consequence, most HCV-infected people proceed 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 5 % is at risk to develop liver carcinoma.

Influenza remains a significant cause of mortality and morbidity worldwide. The World Health Organisation (WHO) estimates that seasonal epidemics affect 3-5 million people 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, influenza A (H1N1) and A (H3N2) subtypes were circulating among humans in 2009. 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. Influenzaviruses 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 flu) 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, including multimeric, such as dimeric peptides, that may be used as immunogens to stimulate the humoral immunity in a subject.

In particular, it is an object of embodiments of the invention to provide peptides including multimeric, such as dimeric peptides comprising epitopes of an antigen that stimulates cells of the B lymphocyte lineage (B-cells) to secrete antibodies against this antigen.

The B-cell activation provided by the peptides according to the present invention may be both T cell-independent and T cell-dependent. Accordingly, the peptides according to the present invention or parts thereof may interact with B-cell receptors to activate the B-cells either

through a T helper cell dependent or independent manner leading to the production of specific antibodies. Furthermore, the peptides may be taken up by antigen presenting cells (macrophages and/or dendritic cells) such that epitopes within the peptides are correctly processed and presented to T-lymphocytes, such as a helper T cell, which in turn helps to activate the B cells in order to stimulate an effective immune response. The peptides may also be taken up by activated B-cells which can also act as antigen presenting cells. Peptides interact with the B-cells through the B-cell receptor and are then internalised into the cell. The epitopes within the peptides will be processed and presented to T-lymphocytes such as helper cells.

However, in some important aspects of the present invention, the peptides according to the present invention are designed to not effectively penetrate and be taken up by antigen presenting cells. Accordingly, in these aspects of the invention, the peptides according to the present invention may provide B-cell activation through interaction at the cell surface via the B-cell receptor. It is to be understood that in order to provide sustained B-cell stimulation, it is preferred that the peptides according to the present invention are designed to comprise a helper epitope that may be taken up by antigen presenting cells in order to stimulate CD4+ T-helper cells that can sustain effective humoral immunity in a subject.

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.

Further, it is an object of embodiments of the invention to provide peptides that may be used as antigens that can serve as targets in diagnostic assays.

SUMMARY OF THE INVENTION

The present invention pertains to a peptide design promoting efficient activation of a humoral immune response against antigens contained within this peptide design.

It has been found by the present inventor(s) that peptide constructs - amino acid sequences with a particular pattern or scaffold design, and in particular multimeric, such as dimeric peptides of this design - have the ability to effectively elicit a humoral immune response in a subject in response to the administration of these peptides.

The peptide constructs according to the present invention have been designed to be able to attach or bind to the cell surface. The peptide constructs or parts thereof may then be taken

up by the antigen presenting cells (such as macrophages and dendritic cells) and stimulate helper T-cells in order to elicit efficient and long lasting T-cell dependent B-cell activation. Alternatively the B-cells themselves may provide for the induction of help to activate the B-cells.

Accordingly the peptides according to the present invention may penetrate the cells and may be used to load cells with an immunogenically effective amount of a peptide or fragments of this peptide that can be presented by macrophages and dendritic cells. Accordingly these peptide constructs may elicit both a Cytotoxic T-lymphocyte immune (CTL) response and/or a humoral immune response.

So, in a first aspect the present invention relates to isolated monomeric peptides consisting of not more than 60 amino acids with the following structure

X^1-X^2-X^3-X^4-X^5-X^6 (formula I),

wherein X^1, X^3 and optional moiety X^5 each independently defines a linear sequence of any 1, 2, 3, 4, or 5 amino acid independently selected from glycine, arginine, norleucine, glutamine, serine, lysine, tryptophan, cysteine, or a derivative thereof; X^2, X^4, and optional moiety X^6 each independently defines a linear sequence of 5-17 amino acids, each having more than 50% sequence identity to a specific natural antigen.

In a second aspect the present invention relates to an isolated multimeric peptide comprising two or more monomeric peptides, each monomeric peptide independently consisting of not more than 60 amino acids with the following structure

X^1-X^2-X^3-X^4-X^5-X^6 (formula I),

wherein X^1, X^3 and optional moiety X^5 independently defines a linear sequence of any 1, 2, 3, 4, or 5 amino acid independently selected from glycine, arginine, norleucine, glutamine, serine, lysine, tryptophan, cysteine, or a derivative thereof; X^2, X^4, and optional moiety X^6 each independently defines a linear sequence of 5-17 amino acids, each having more than 50% sequence identity to a specific natural antigen, said monomeric peptides being covalently joined by one or more intermolecular bond.

In a third aspect the present invention relates to isolated monomeric peptides consisting of not more than 60 amino acids with the following structure

X^1-X^2-X^3-X^4-X^5-X^6 (formula I),

wherein X^1, X^3 and optional moiety X^5 each independently defines a linear sequence of any 1, 2, 3, 4, or 5 amino acid independently selected from glycine, arginine, norleucine, aspartic acid, glutamic acid, glutamine, serine, lysine, tryptophan, cysteine, ornithine, diaminopropionic acid or a derivative thereof; X^2, X^4, and optional moiety X^6 each independently defines a linear sequence of 5-17 amino acids, each having 50% or more sequence identity to a specific natural antigen.

In a further aspect the present invention relates to an isolated multimeric peptide comprising two or more monomeric peptides, each monomeric peptide independently consisting of not more than 60 amino acids with the following structure

X^1-X^2-X^3-X^4-X^5-X^6 (formula I),

wherein X^1, X^3 and optional moiety X^5 independently defines a linear sequence of any 1, 2, 3, 4, or 5 amino acid independently selected from glycine, arginine, norleucine, aspartic acid, glutamic acid, glutamine, serine, lysine, tryptophan, cysteine, ornithine, diaminopropionic acid or a derivative thereof; X^2, X^4, and optional moiety X^6 each independently defines a linear sequence of 5-17 amino acids, each having 50% or more sequence identity to a specific natural antigen, said monomeric peptides being covalently joined by one or more intermolecular bond.

In a further aspect the present invention relates to a composition comprising two or more compounds selected from a monomeric peptide according to the present invention, and an isolated multimeric peptide according to the present invention.

In a further aspect the present invention relates to the use of a peptide selected from a monomeric peptide according to the present invention, and an isolated multimeric peptide according to the present invention for inducing a humoral immune response in a subject.

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.

In a further aspect the present invention relates to an immunogenic composition comprising at least one monomeric peptide, an isolated multimeric peptide according to the invention, a peptide composition, the nucleic acid or polynucleotide, or the vector according the invention; in combination with a pharmaceutically acceptable diluent or vehicle and optionally an immunological adjuvant. In some embodiments this immunogenic composition is in the form of a vaccine composition.

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 monomeric peptide, an isolated multimeric peptide, a peptide composition, the nucleic acid or polynucleotide, or the vector, or the composition of the invention.

In a further aspect the present invention relates to a method for reducing and/or delaying the pathological effects of a disease antigen, such as an infectious agent in a subject infected with said agent or having said disease caused by said antigen, the method comprising administering an effective amount of at least one monomeric peptide, an isolated multimeric peptide, a peptide composition, the nucleic acid or polynucleotide, or the vector, or the composition according to the invention.

In a further aspect the present invention relates to a peptide according to the invention for use as a medicament, or for treating the pathological effects of a disease antigen, such as an infectious agent in a subject infected with said agent or having said disease caused by said antigen.

In a further aspect the present invention relates to a peptide according to the invention for use in a diagnostic assay. In a further aspect the present invention relates to a peptide according to the invention for use in an in vitro assay.

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.

As used herein a "multimeric peptide" or "oligomeric peptide" refers to an assembly of two or more different or identical linear peptide sequences or subunits, preferably interconnected or assembled by one or more chemical bond of a linker. Preferably the peptide sequences are interconnected by one or more, such as one covalent bond, such as an intermolecular disulfide (S-S) bond between two Cys residues, a methylated peptide bond between a N-ε-methylated Lys side-chain and the side-chain of an Asp or Glu residue, an oxime bond, or a thioether bond. The term includes a dimeric (or dimer) peptide suitably formed by a chemical linking of two linear peptide sequences. The term "multimeric peptide" further includes an assembly of 2, 3, 4, 5, 6, 7, 8, 9 or 10 different or identical peptide sequences. In some embodiments, the multimeric peptide is a dimeric peptide.

As used herein a "linker" refers to any compound suitable for assembly of the two or more different or identical linear peptide sequences or subunits into a multimeric peptide. The term includes any linker found useful in peptide chemistry. Since the multimeric peptide may be assembled or connected by standard peptide bonds in a linear way, the term linker also includes a "peptide spacer", also referred to as a "spacer".

In some embodiments, the linker is not a peptide sequence. In some embodiments, the linker is not a branched peptide sequence.

In some embodiments, the linker does not itself contain a peptide sequence derived from or identical to a natural antigen.

In some embodiments, the linker has a molecular weight of less than 10 kDa, such as less than 9 kDa, such as less than 8 kDa, such as less than 7 kDa, such as less than 6 kDa, such as less than 5 kDa, such as less than 4 kDa, such as less than 3 kDa, such as less than 2 kDa, such as less than 1.5 kDa, such as less than 1 kDa, such as less than 0.5 kDa, such as less than 0.2 kDa.In some embodiments, wherein the multimeric peptide is a dimeric peptide, the linker is not linking the two peptide sequences from one terminal cysteine in the first peptide to a second terminal cysteine in the second peptide.

In some embodiments, the linker is not linking the two or more peptide sequences through a terminal cysteine in any one of the peptides.

In some embodiments, the linker is not linking from a cysteine residue.

In some embodiments, in the peptide according to the present invention, X^1, X^3 and optional moiety X^5 in the same peptide is not identical in sequence.

In some embodiments, in the peptide according to the present invention, X^2, X^4 and optional moiety X^6 in the same peptide is not identical in sequence.

In some embodiments, in the multimeric peptide according to the present invention, X^1, X^3 and optional moiety X^5 in one peptide is not identical in sequence with X^1, X^3 and optional moiety X^5 in any other peptide.

In some embodiments, in the multimeric peptide according to the present invention, X^2, X^4 and optional moiety X^6 in one peptide is not identical in sequence with X^2, X^4 and optional moiety X^6 in any other peptide.

"HIV" generally denotes human immunodeficiency virus I.

"HIV disease" is composed of several stages including the acute HIV infection which often manifests itself as a flu-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 a 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, endoplasmatic reticulum, golgi apparatus, mitocondria and/or chloroplast, seemingly energy-independently. This capability to translocate across the plasma membrane of a "cell-penetrating peptide" 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 provided in the examples. The term "non-cell-penetrating peptide" as used herein refers to a peptide, which is not a cell-penetrating peptide.

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. Usually the peptide has been adapted or modified from the original 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, or 7 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, or 7 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, and/or X^6, that is defined as a sequence of 5-17 amino acids derived from an antigen, wherein the peptide sequence of the antigen comprises 1, 2, 3, 4, 5, 6, or 7 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 and/or X^6. 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. A capital D-letter subscript after the letter representing the amino acid residue designate herein amino acids specified to be in D-form, such as W_D referring to a tryptophan in D-form. A capital L-letter subscript after the letter representing the amino acid residue designate herein amino acids specified to be in L-form, such as W_L referring to a tryptophan in L-form.

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.

It is to be understood that for a multimeric peptide one or more, such as all peptide strands may have modified amino acids in the N-terminal of the amino acid sequences. The linker linking two or more peptide strands may be placed anywhere within the peptide strand, in particular if one or more of the peptide strands have modified amino acids in the N-terminal of the amino acid sequences. The linker may also serve to protect the peptide from degradation, which often is degradation from the N-terminal. Accordingly, the linker may be more freely placed if one or both peptide strands are protected from degradation.

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. 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 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 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 a 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 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, at least about 90, at least about 95, at least about 98, or at least about 99 percent sequence identity. In some embodiments, when measuring the sequence identity between two different peptide sequences, a gap of one or two amino acids is allowed when the two peptide sequences are aligned without having any influence on the value of sequence identity. In some embodiments, a residue position that is not identical differ by only a conservative amino acid substitution. 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 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 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 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); each herein incorporated by reference. "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 specific peptide sequence will exhibit 98% - 99% homogeneity for peptides 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. In some embodiments however, a second part of a peptide sequence may be bound to and continue from the side chain of a terminal amino acid in a

first part of an amino acid sequence. Also the term does not exclude that an amino acid within a sequence, such as within X^1, X^2, X^3, X^4, and/or X^5, 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^1, X^2, X^3, X^4, and/or X^5.

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.

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 an MHC molecule (in the case of T-cell receptor binding epitopes, i.e. "T-cell epitopes").

"B cell antigen" means any antigen that naturally is or could be engineered to be recognized by a B cell, and that triggers an immune response in a B cell (e.g., an antigen that is specifically recognized by a B cell receptor on a B cell).

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.

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.

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.

A "T helper lymphocyte epitope" (a T_H epitope), "T helper epitope" or "helper 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).

In the scaffold design according to the present invention, X^2, X^4, and X^6 defines a sequence of 5-17 amino acids derived from the antigen. This sequence of amino acids derived from an antigen may herein be referred to as an epitope.

The peptides according to the present invention may be a helper T lymphocyte (HTL) inducing peptide comprising HTL epitopes. A "HTL inducing peptide" is a HLA Class II binding peptide that is capable of inducing a HTL response. Also the peptides according to the present invention may in other embodiments be CTL inducing peptides comprising CTL epitopes in addition to or as an alternative to being a HTL inducing peptide. A "CTL inducing peptide" is a HLA Class I binding peptide that is capable of inducing a CTL response.

In other alternative embodiments, tryptophan or tryptophan derivatives are used in the sequence defined by X^1, X^3, or X^5. 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-Benzyloxytryptophan (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 "amino acid derivative", sometimes used in the context of a "derivative thereof" referring to a specific amino acid, means an amino acid compound, wherein one or more chemical groups has been modified, added or removed as compared to the amino acid to which the amino acid compound is a derivative of, while still having an amine group and a carboxylic acid group, as well as a side chain of an amino acid and still being able to form peptide bonds. In some embodiments an amino acid derivative is a standard amino acid that has only been modified in the side chain of the amino acid. In some embodiments an amino acid derivative is a non-natural amino acid such as Dpr. In some embodiments an amino acid is a modified moiety which is incorporated into the chemically synthesized peptide or polypeptide and that comprises an activatable group that is linkable, after activation, to another peptide, such as Dpr(Ser), Lys(Ser), or Ornithine(Ser).

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.

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.

The terms "patient" and "subject" refer to any human or animal 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."

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.

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 Chemsitry" (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-naturally basic amino acids are e.g. as described in US 6,858,396. Suitable positively charged amino acids includes un-natural alpha amino acids available from Bachem AG and includes alpha-amino-glycine, alpha,gamma-diaminobutyric acid, ornithine, alpha, beta-diaminoproprionic acid, alpha-difluoromethyl-ornithine, 4-amino-piperidine-4carboxylic 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.

The term "neutral amino acid" as used herein refers to an amino acid that has an isoelectric point 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 Chemsitry" (Macmillan, 1974).

Antigens

The specific natural antigen used in the peptide constructs according to the present invention may be a protein or peptide sequence derived from any B cell antigen, such as from any disease antigen, such as an infectious agent. Suitable antigens to be used according to the present invention include antigens derived from a bacteria, a mycobacterium, a virus, a parasite such as protozoa, a fungus, a cancer antigen, such as an oncogene, a prion, an atopic disease antigen, an addictive or abused substance or a toxin or an antigen of an autoimmune disease, such as rheumatoid arthritis, insulin dependent diabetes, multiple sclerosis and the like.

As used herein a "disease antigen" refers to any antigen confirmed or suspected to be involved in a specific disease.

In some embodiments, the antigen is an abused or addictive substance or a portion thereof, including, but are not limited to, nicotine, a narcotic, a cough suppressant, a tranquilizer, and a sedative. In some embodiments, the antigen is a toxin, such as a toxin from a chemical weapon or natural sources, or a pollutant.

Examples of bacteria for which antigens may be provided include, but are not limited to, M. tuberculosis, Mycobacterium, mycoplasma, neisseria and legionella. Examples of parasites include, but are not limited to, rickettsia and chlamydia.

Examples of an infectious disease antigen is TbH9 (also known as Mtb 39A), a tuberculosis antigen. Other tuberculosis antigens include, but are not limited to DPV (also known as Mtb8.4), 381, Mtb41, Mtb40, Mtb32A, MΛ9.9A, Mtb9.8, Mtblȯ, Mtb72f, Mtb59f, Mtb88f, Mtb71f, Mtb46f and Mtb31f ("f ' indicates that it is a fusion or two or more proteins).

Examples of cancer antigens may be a tumor associated antigen such as HER2, HER3 or HER4 receptor or one or more tumor-associated antigens or cell-surface receptors disclosed in US Publication No. 20080171040 or US Publication No. 20080305044 and are incorporated in their entirety by reference.

Other suitable cancer antigens that may be used by the present invention include CD proteins such as CD2, CD3, CD4, CD5, CD6, CD8, CD11, CD14, CD18, CD19, CD20, CD21, CD22, CD25, CD26, CD27, CD28, CD30, CD33, CD36, CD37, CD38, CD40, CD44, CD52, CD55, CD56, CD70, CD79, CD80, CD81, CD103, CD105, CD134, CD137, CD138, and CD152; members of the ErbB receptor family such as the EGF receptor, HER2, HER3 or HER4 receptor; cell adhesion molecules such as LFA-I, Mac1, pi 50.95, VLA-4, ICAM-I, VCAM, EpCAM, alpha4/beta7 integrin, and alpha v/beta3 integrin including either alpha or beta subunits thereof (e.g. anti-CD11a, anti-CD18 or anti-CD11b antibodies); growth factors such as VEGF; tissue factor (TF); TGF-β.; alpha interferon (alpha-IFN); an interleukin, such as IL-8; IgE; blood group antigens Apo2, death receptor; flk2/flt3 receptor; obesity (OB) receptor; mpl receptor; CTLA-4; protein C etc. In some embodiment the antigen is selected from IGF-IR, CanAg, EphA2, MUC1 , MUC16, VEGF, TF, CD19, CD20, CD22, CD27, CD33, CD37, CD38, CD40, CD44, CD56, CD138, CA6, Her2/neu, EpCAM, CRIPTO (a protein produced at elevated levels in a majority of human breast cancer cells), darpins, $alpha_v/beta_3$ integrin, $alpha_v/beta_5$ integrin, alpha y/beta integrin, TGF- β, CD11a, CD18, Apo2 and C242. In some embodiment the antigen is selected from a CD proteins such as CD3, CD4, CD8, CD19, CD20, CD27, CD34, CD37, CD38, CD46, CD56, CD70 and CD138; members of the ErbB receptor family such as the EGF receptor, HER2, HER3 or HER4 receptor; cell adhesion molecules such as LFA-I, Mac1, pl50.95, VLA-4, ICAM-I, VCAM, EpCAM, alpha4/beta7 integrin, and alpha v/beta3 integrin including either alpha or beta subunits thereof (e.g. anti-CD11a, anti-CD18 or anti-CD11b antibodies); growth factors such as VEGF; tissue factor (TF); TGF-β.; alpha interferon (alpha-IFN); an interleukin, such as IL-8; IgE; blood group antigens Apo2, death receptor; flk2/flt3 receptor; obesity (OB) receptor; mpl receptor; CTLA-4; protein C, etc. The most preferred targets herein are IGF-IR, CanAg, EGF-R, EGF-RvIII, EphA2, MUC1, MUC16, VEGF, TF, CD19, CD20, CD22, CD27, CD33, CD37, CD38, CD40, CD44, CD56, CD70, CD138,

CA6, Her2/neu, CRIPTO (a protein produced at elevated levels in a majority of human breast cancer cells), alpha$_v$/beta$_3$ integrin, alpha$_v$/beta$_5$ integrin, TGF- β, CD11a, CD18, Apo2, EpCAM and C242. In some embodiment the antigen is selected from a cellular oncogene, such as ras or myc.

Examples of viral antigens for use with the present invention include, but are not limited to, e.g., HIV, HCV, CMV, HPV, Flu, 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 El and E2 and other rubella virus components; rotaviral antigens such as VP7sc and other rotaviral components; 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, gpll, and other varicella zoster viral antigen components; Japanese encephalitis viral antigens such as proteins E, M-E, M-E-NSl, NSl, 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 viruses in virus families such as adenoviridae, retroviridae, picornaviridae, herpesviridae, rotaviruses (reoviridae), hantaviruses (Bunyaviridae), coronaviridiae, togaviridae, flavirviridae, rhabdoviridae, paramyxoviridae, orthomyxoviridae, bunyaviridae, arenaviridae, reoviridae, papilomaviridae, parvoviridae, poxviridae, hepadnaviridae, 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, flu, 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 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 (Immusystems 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 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, and more specifically from at least 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 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).

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.

Flu:

The epitopes to be incorporated into the scaffold design according to the present invention 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, 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 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 (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 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 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.

Human immunodeficiency virus (HIV):

For HIV, the epitopes to be incorporated into the scaffold design according to the present invention may be derived from viral proteins consisting of gp120, gp160, gp41, p24gag or p55gag the regulatory proteins (such as Tat, Rev, Nef) as well as the viral enzymes (such as polymerase, integrase or protease) derived from HIV, including members of the various genetic subtypes.

Human papillomavirus (HPV):

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 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 usefull for administration to both animals, such as mammals, and, in particular, to 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, such as recombinant virus particles, nanoparticles, such as nanogold, or cyclotides.

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(SO4)2.12H2O), or one of these in combination with 3-0-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). 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, 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).

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 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 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, polyamino acids, dendrimers, saponins, QS21, adsorption enhancing materials, fatty acids or, naked or particle absorbed cDNA.

The peptides may be delivered in oils such as Endocine™ and Montanide™ (Eurocine) – Montanide™ ISA 51 VG or Montanide™ ISA 720 VG (Seppic).

The adjuvant may be stimulators of the innate immune system that can be given separately from the peptide such as Leukotriene B4 (LTB4) and granulocyte macrophage colony stimulating factor (GM-CSF), such as Sargramostim/Leukine (glycosylated GM-CSF) and Molgramostim (nonglycosylated GM-CSF).

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. Preferable 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-CH2-CH2)$_n$O-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; 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

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 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 subject exposed to a disease antigen, such as an infectious agent, such as a pathogen, 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 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.

For example, a peptide of the invention may be used in a tetramer staining assay to assess peripheral blood mononuclear cells (PBMC) 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 reconstituted in the presence of the corresponding HLA heavy chain and beta2-microglobulin to generate a trimolecular 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 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 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.

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 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 molecule, i.e., antibodies that bind to a peptide-MHC complex.

In certain embodiments a first monomeric peptide and the at least one second monomeric peptide are associated via a linker; the linker may comprise any peptide linker, or peptide spacer, such as a glycine, a lysine or an arginine linker/spacer, a polyhistidinyl tag, Protein G, and Protein A but it is also possible to use a bis-maleimide linker/spacer, a disulfide linker, or a polyethylene glycol (PEG) linker. In practice, any linker found useful in peptide chemistry is also useful as a linker according to the present invention. Thus, the invention contemplates the use of "simple" linear peptides which are conjugated or fused to each other, but also peptide combinations where the individual peptides derived from a natural antigen are linked via non-peptide linkers. Use of multiple linker types are also within the scope of the present invention, and it is e.g. also a part of the invention to utilise linear peptides which include intrachain disulphide linkers.

Particularly interesting peptide combinations of the invention are set forth in the preamble to the examples.

In certain embodiments, at least one of the first and at least one second peptides in the peptide combination comprises an N- or C-terminal modification, such as an amidation, acylation, or acetylation.

Since the peptide combinations are contemplated as vaccine agents or diagnostic agents, they are in certain embodiments coupled to a carrier molecule, such as an immunogenic carrier. The peptides of the peptide combinations may thus be linked to other molecules either as recombinant fusions (e.g. via CLIP technology) or through chemical linkages in an oriented (e.g. using heterobifunctional cross-linkers) or nonoriented fashion. Linking to carrier molecules such as for example diphtheria toxin, latex beads (convenient in diagnostic and prognostic embodiments), and magnetic beads (also convenient in diagnostic and prognostic embodiments), polylysine constructs etc, are all possible according to the invention.

The immunogenic carrier is conveniently selected from carrier proteins such as those conventionally used in the art (e.g. diphtheria or tetanus toxoid, KLH etc.), but it is also possible to use shorter peptides (T-helper epitopes) which can induce T-cell immunity in larger proportions of a population. Details about such T-helper epitopes can e.g. be found in WO 00/20027, which is hereby incorporated by reference herein – all immunolgic carriers and "promiscuous" (i.e. universal) T-helper epitopes discussed therein are useful as immunogenic carriers in the present invention.

In certain embodiments, the carrier is a virus like particle, i.e. a particle sharing properties with virions without being infectious. Such virus-like particles may be provided chemically (e.g. Jennings and Bachmann Ann Rev Pharmacol. Toxicol. 2009. 49:303-26 Immunodrugs: Therapeutic VLP-based vaccines for chronic diseases) or using cloning techniques to generate fusion proteins (e.g. Peabody et al. J. Mol. Biol. 2008; 380: 252-63. Immunogenic display of diverse peptides on virus-like particles of RNA phage MS2). Another example is "Remune", an HIV vaccine originally made by Immune Response Corporation, which consists of formalin inactivated HIV that has been irradiated to destroy the viral genome.

In an embodiment, a nucleic acid is encoding one or more monomeric peptide of the multimeric, such as dimeric peptide according to the invention, where the encoded first peptide and the encoded at least one second peptide of a multimeric peptide are associated via a peptide linker, including a peptide spacer, and/or a disulphide bridge. The peptide linker/spacer is typically selected from the group consisting of a glycine, an arginine, a lysine

linker/spacer, or a glycine-lysine linker/spacer, but any peptide linker known in the art may be useful. The term peptide linker thus also is intended to denote coupling between the first and second peptide via a peptide bond. A peptide linker that links a first and second peptide by standard peptide bonds may also be referred to as a peptide spacer. Also, the first and second peptides may be linked via a peptide linker and a disulphide bond, as is the case when an intrachain disulphide bond is established.

In one embodiment, the nucleic acid according to the invention encodes the peptide combination, which is coupled (by fusion) to a carrier molecule, such as an immunogenic carrier; useful carriers are discussed above.

In some embodiments the linker is selected from the group consisting of a bis-maleimide linker, a disulfide linker, a polyethylene glycol (PEG) linker, a glycine linker/spacer, a lysine linker/spacer, and an arginine linker/spacer.

In some embodiments the multimeric peptide, such as a dimeric peptide contain a linker in the free amino group of the N-terminal of a monomeric peptide linking said monomeric peptide to another monomeric peptide.

In some embodiments the multimeric peptide, such as a dimeric peptide contain a linker in the free carboxyl group of the C-terminal of a monomeric peptide linking said monomeric peptide to another monomeric peptide.

At least two options for such linkers are described in A.R Jacobson et al, J. Med. Chem. 1989, 32, 1708-1717 and in D Giannotti et al, Journal of Medicinal Chemistry, 2000, Vol. 43, No. 22, the disclosures of which is hereby incorporated by reference.

Alternatively a link between the N-termini of peptides may be established by reacting with Br-$(CH_2)_n$-Br.

The length of the linker may be varied by the addition of glycine residues, for example Fmoc-NH-CH_2CH_2-NH-Gly-NH_2 may be used.

An example of such a synthesis, wherein a dimeric peptide is prepared by conjugation through succinic acid, may be as follows:

(H-Gly-Gly-Ala-Lys-Arg-Arg-Val-Val-Gln-Arg-Glu-Lys-Arg-Ala-Gly-Glu-Arg-Glu-Lys-Arg-Ala-NH_2)E(H-Gly-Gly-Ile-Glu-Glu-Glu-Gly-Gly-Arg-Asp-Arg-Asp-Arg-Gly-Gly-Glu-Gln-Asp-Arg-Asp-Arg-NH_2)F (Succinic acid linker between Gly^1E and Gly^1F)

This dimer was produced from the reaction of the following 2 monomers

Monomer E

H-Gly-Gly-Ala-Lys-Arg-Arg-Val-Val-Gln-Arg-Glu-Lys-Arg-Ala-Gly-Glu-Arg-Glu-Lys-Arg-Ala-NH_2 (SEQ ID NO:143).

Monomer F

H-Gly-Gly-Ile-Glu-Glu-Glu-Gly-Gly-Arg-Asp-Arg-Asp-Arg-Gly-Gly-Glu-Gln-Asp-Arg-Asp-Arg-NH_2 (SEQ ID NO:144).

The two monomers are reacted to give a heterodimer according to the reaction scheme outlined below; where the link is between N-terminal on Gly^1 of on chain E and the N-terminal on Gly^1 in chain F.

Monomers E and F are synthesized separately on a Sieber Amid resin. The Fmoc-groups on N-terminal Gly are removed while the peptides are still on resin. The peptides are cleaved from resin. The resulting protected peptide E is reacted with succinic acid anhydride and thereafter reacted with the protected peptide F. Protective groups are subsequently removed with 95% TFA. The formed heterodimer may be purified from un-reacted monomers by conventional purification methods known to the person skilled in the art.

An example of a synthesis, wherein a dimeric peptide is prepared by conjugation through di-amino propane, may be as follows:

(H-Gly-Gly-Ala-Lys-Arg-Arg-Val-Val-Gln-Arg-Glu-Lys-Arg-Ala-Gly-Glu-Arg-Glu-Lys-Arg-Ala-Gly-Gly)G(H-Gly-Gly-Ile-Glu-Glu-Glu-Gly-Gly-Arg-Asp-Arg-Asp-Arg-Gly-Gly-Glu-Gln-Asp-Arg-Asp-Arg-Gly-Gly)H trifluoroacetate salt (Diamino propane linker between Gly^{23} and Gly^{23})

This dimer was produced from the reaction of the following 2 protected monomers

Monomer G

H-Gly-Gly-Ala-Lys-Arg-Arg-Val-Val-Gln-Arg-Glu-Lys-Arg-Ala-Gly-Glu-Arg-Glu-Lys-Arg-Ala-Gly-Gly-COOH (SEQ ID NO:145)

Monomer H

H-Gly-Gly-Ile-Glu-Glu-Glu-Gly-Gly-Arg-Asp-Arg-Asp-Arg-Gly-Gly-Glu-Gln-Asp-Arg-Asp-Arg-Gly-Gly-COOH (SEQ ID NO:146)

The two monomers G and H are reacted to give a heterodimer according to the reaction scheme outlined below; where the link is between C-terminal on Gly^{23} of on chain G and the C-terminal on Gly^{23} in chain H.

Monomers G and H are synthesized separately on a 2-chlorotrityl resin. Boc-Gly-OH is coupled to the peptides on the resin before cleaving them of the resin. The resulting peptides are then Boc-protected, alternatively they may me acetylated before being cleaved of the resin. The resulting protected peptide G is reacted with Fmoc-diaminopropane, Fmoc is deprotected and G is coupled to the C-terminal of the protected peptide H via a peptide bond. Protective groups are subsequently removed with 95% TFA. The formed heterodimer may be purified from un-reacted monomers by conventional purification methods known to the person skilled in the art.

Method for synthesis of Cys-Lys bridge:

Exemplified with the preparation of BI400-B trifluoroacetate salt

(H-Gly-Ala-Lys-Arg-Arg-Val-Val-Gly-Gly-Cys(2-oxo-ethyl)-Gly-Gly-Ala-Lys-Arg-Arg-Val-Val-Gln-Arg-Glu-Lys-Arg-Ala-Gly-Glu-Arg-Glu-Lys-Arg-Ala-NH_2)A(H-Gly-Lys-Gly-Gly-Ile-Glu-Glu-Glu-Gly-Gly-Arg-Asp-Arg-Asp-Arg-Gly-Gly-Glu-Gln-Asp-Arg-Asp-Arg-NH_2)B trifluoroacetate salt (Thioether bond between $Cys(2\text{-oxo-ethyl})^9$A and Lys^2B)

This dimer was produced from the reaction of the following 2 protected monomers

Monomer A

H-Gly-Ala-Lys-Arg-Arg-Val-Val-Gly-Gly-Cys-Gly-Gly-Ala-Lys-Arg-Arg-Val-Val-Gln-Arg-Glu-Lys-Arg-Ala-Gly-Glu-Arg-Glu-Lys-Arg-Ala-NH_2 (SEQ ID NO:121)

Monomer B

H-Gly-Lys(bromoacetyl)-Gly-Gly-Ile-Glu-Glu-Glu-Gly-Gly-Arg-Asp-Arg-Asp-Arg-Gly-Gly-Glu-Gln-Asp-Arg-Asp-Arg-NH_2 ; (SEQ ID NO:122)

Or with the preparation of 400-Seq B trifluoroacetate salt

(H-Gly-Ala-Lys-Arg-Arg-Val-Val-Gly-Gly-Cys(2-oxo-ethyl)-Gly-Gly-Ala-Lys-Arg-Arg-Val-Val-Gln-Arg-Glu-Lys-Arg-Ala-Gly-Glu-Arg-Glu-Lys-Arg-Ala-NH_2)A(H-Gly-Lys-Gly-Gly-Ile-Glu-Glu-Glu-Gly-

Gly-Arg-Asp-Arg-Asp-Arg-Gly-Gly-Gln-Asp-Arg-Asp-Arg-NH_2)B trifluoroacetate salt (Thioether bond between Cys(2-oxo-ethyl)^{9}A and Lys^{2}B)

This dimer was produced from the reaction of the following 2 protected monomers

Monomer A

H-Gly-Ala-Lys-Arg-Arg-Val-Val-Gly-Gly-Cys-Gly-Gly-Ala-Lys-Arg-Arg-Val-Val-Gln-Arg-Glu-Lys-Arg-Ala-Gly-Glu-Arg-Glu-Lys-Arg-Ala-NH_2 (SEQ ID NO:121)

Monomer B

H-Gly-Lys(bromoacetyl)-Gly-Gly-Ile-Glu-Glu-Glu-Gly-Gly-Arg-Asp-Arg-Asp-Arg-Gly-Gly-Gln-Asp-Arg-Asp-Arg-NH_2 (SEQ ID NO:147)

The 2 monomers are reacted to give a heterodimer according to the reaction scheme outlined below; where the link is created between Lys 2 (bromoacetyl) side chain on chain B and Cys in chain A.

At neutral pH and room temperature, bromoacetyl moieties in buffered aqueous solutions are very reactive towards SH-containing moieties, such as the thiol group in cysteine. Thus, if a cysteine is present on the other peptide sequence, the SH will attack the bromoacetyl to form a intermolecular thioether bridge. When the reaction is buffered with a sodium-containing buffer, such as $NaHCO_3$, the only byproduct of the reaction is NaBr, an innocuous salt.

The formed heterodimer may be purified from un-reacted monomers by conventional purification methods known to the person skilled in the art.

Method for synthesis of oxime bond between two peptide sequences, an intermolecular bond:

Exemplified with the preparation of 400-Seq B* trifluoroacetate salt

(H-Gly-Ala-Lys-Arg-Arg-Val-Val-Gly-Gly-Dpr(COCHO)-Gly-Gly-Ala-Lys-Arg-Arg-Val-Val-Gln-Arg-Glu-Lys-Arg-Ala-Gly-Glu-Arg-Glu-Lys-Arg-Ala-NH_2)D(H-Gly-Lys(aminooxyacerylated)--Gly-Gly-Ile-Glu-Glu-Glu-Gly-Gly-Arg-Asp-Arg-Asp-Arg-Gly-Gly-Gln-Asp-Arg-Asp-Arg-NH_2)C trifluoroacetate salt (oxime is created between Dpr(COCHO)-)^{9}D and Lys(aminooxyacerylated)^{2}C)

This dimer is produced from the reaction of the following two monomers:

Monomer C

H-Gly-Lys(aminooxyacerylated)-Gly-Gly-Ile-Glu-Glu-Glu-Gly-Gly-Arg-Asp-Arg-Asp-Arg-Gly-Gly-Gln-Asp-Arg-Asp-Arg-NH_2 (SEQ ID NO:148)

Monomer D

H-Gly-Ala-Lys-Arg-Arg-Val-Val-Gly-Gly-Dpr(Ser)-Gly-Gly-Ala-Lys-Arg-Arg-Val-Val-Gln-Arg-Glu-Lys-Arg-Ala-Gly-Glu-Arg-Glu-Lys-Arg-Ala-NH_2. (SEQ ID NO:149)

The two monomers are reacted to give a heterodimer according to the reaction scheme outlined below; where the link is created between Lys^2 (aminooxyacetylated) side chain on chain C and oxidized Dpr(Ser) in chain D.

After removal of the Mtt group from Lys and while the peptide was still attached to the resin aminooxyacetylated (AoA) monomer C was synthesized by coupling aminooxyacetic acid to Lys. The peptide was then cleaved from the solid phase support and purified by conventional purification methods. The monomer D was, after cleavage from resin and purification, created by oxidation of the serinyl diaminopropionic acid residue (Dpr(Ser)) with periodate to the aldehyde function. Equimolar amounts of monomer A and B were dissolved in acetonitrile and acetate buffer (pH 4). After reaction for 16h at room temperature, the product C-oxime-D was isolated by conventional purification methods known to the person skilled in the art.

Dpr=diaminopropionic acid

Fmoc-Dpr (Boc-Ser(tBu))-OH Merck 04-12-1186

Method for synthesis of dimers with PEG-linker:

A multimeric, such as dimeric peptide, such as a heterodimeric peptide may be synthesized by, but are not restricted to the following protocol:

To the peptidyl resin containing deblocked Asp or Glu residue (monomer 1) is added HBTU, DIPEA and Trt-amino PEG amine in DMF. The mixture is allowed to couple over night. The resin is filtered from the solution and washed by standard protocol. The Trt group is removed from the Trt- PEGylated peptide. The monomer 2 containing deblocked Asp or Glu residue is then coupled to the exposed amino group using HBTU and DIPEA. After cleavage the desired product is purified using any suitable technique to give the desired multimeric peptide.

In some embodiments the isolated monomeric peptide contain intramolecular bonds, such as in the form of intramolecular Cys-Cys bonds. It is to be understood that the "intramolecular bond", used interchangeably with "intrachain bond", is a bond between two different amino acids within the same peptide chain, which however is not necessarily adjacent to each other

in the peptide sequence. Accordingly, in some embodiments, the isolated multimeric peptide according to the invention may contain both intramolecular bonds within one or more of the monomers, as well as an intermolecular bond between two chains of the multimeric peptide, such as a dimer. This intramolecular bond may be in the form of Cys-Cys bonds formed with cysteine residues within the same peptide sequence. In some embodiments the monomer contains an intramolecular bond derived from a Lys residue or other amino acid residue, such as a Ser, Cys, Asp or Glu that make the bond, such as a thioether bond or an oxime bond or through a PEG linker, to an amino acid residue on the other monomer peptide sequence.

Method for synthesis of multimeric peptides with PolyLys or MAPS:

PolyLys or MAPS (multiple antigen peptides) – has been extensively used over the last 20 years as a carrier protein to produce strong immunogenic response. The MAP system utilizes a peptidyl core of three or more radially branched lysine core to form a backbone for which the epitope sequences of interest can be built parallel using standard solid-phase chemistry.

The MAP system is a commercial product available from several companies such as AnaSpec, Bio-synthesis Inc. and others. The product, as offered in the catalogue only allows attachment of two (identical) peptide sequence to the polyLys core. It is however possible also to link two different peptide sequences by using different protecting groups for alfa- and epsylon-amino functional groups of lysine on the two different peptide sequences.

Use of the MAP system has been described in references including: Wang, C. Y et al. "Long-term high-titer neutralizing activity induced by octameric synthetic HIV antigen" Science 254, 285-288 (1991). Posnett, D. et al. "A novel method for producing anti-peptide antibodies" J. Biol. Chem. 263, 1719-1725 (1988), and in Tam, J. P. "Synthetic peptide vaccine design: synthesis and properties of a high-density multiple antigenic peptide system" PNAS USA 85, 5409-5413 (1988).

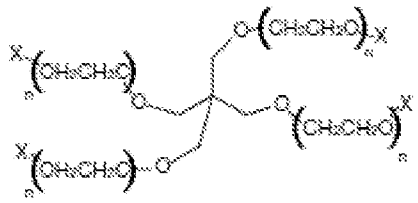
The MAP system could also be prepared by chemical (thioether, oxime, hydrazone) ligation of appropriately functionalized tetra- or octavalent polylysine constructs with the peptide antigen. By the use of this chemical ligation, the two peptide sequences being linked together would not have to be identical as they are synthesized separately.

Additionally a novel application of the MAP-based system is to synthesize on solid support a "probe" containing a poly(ethylene glycol) (PEG) chain in the dendritic arms of MAP.

Use of the MAP system will increase the size of a multimeric complex and may increase the immunogenic response.

Methods for the synthesis of multimeric peptides using PEG:

Suitable Multi-Arm Activated PEG to be used for a PEG linker are commercially available, e.g. a compound with the following structure:



Wherein X may be ethanethiol - CH2CH2SH (could be used to form S-S bridge with the epitope or a thioether link) or propylamine –CH2CH2CH2NH2, among others. These handles preferably allows for the linking of two identical peptide sequences and may be seen as a poly-monomeric epitope presenting construct. One could, however, anchor a dimer (two epitopes linked together) to the PEG above.

Method for synthesis of peptide- poly-L-Lys (PLL)-polyethylene glycol (PEG) construct:

Peptide- PLL-PEG constructs, may be synthesized by, but are not restricted to the following protocol:

Fmoc-Poly-L-Lys-resin (a commercial product) is de-protected with 20% piperifine-DMF. Fmoc-NH-PEG_4-COOH, in a mixed solvent of CH_2Cl_2-NMP is added followed by HBTU and DIPEA and the reaction is allowed to proceed for 24h. The resultant pegylated poly-L-Lys-resin is washed and the pegylation step is repeated. The reaction is monitored by Kaiser's ninhydrin test until a negative reading is obtained. After de-protection of Fmoc group, four identical peptide chains are synthesized directly on the branched poly-L-Lys-polyethylene glycol core by a stepwise solid-phase procedure. All residues activated with HBTU and DIPEA are allowed to couple for 2h. The coupling is monitored by Kaiser's ninhydrin test and is repeated if needed. After cleavage the desired product is purified using any suitable technique to give the desired peptide-construct.

Table 1 (Amino acids underlined refers to place of linker in dimeric molecules; Letter C in a large font refers to a cysteine residue optionally involved in an intramolecular bond with another cysteine residue in the same peptide sequence. Homoarginine is abbreviated Har, Norleucine is abbreviated as Nle or alternatively with the single letter "Z", N-ε-methylated Lys is abbreviated Lys(Me), Citrulline is abbreviated with the single letter "B", diaminopropionic acid is abbreviated with Dpr and serinyl diaminopropionic acid is abbreviated Dpr(Ser). Flu; abbreviation for Influenza)

Chain	Antigen	Reference ID	X1	X2	X3	X4	X5	X6	Position with reference to positions in SEQ ID NO:1, SEQ ID NO:3, and SEQ ID NO:4.			
									X2-SEQ	X4-SEQ	X6-SEQ	Protein
	Flu	BI100_CGnat	**RR**	SLLTEVETP	**GCG**	VETPIR	**G**	TPIRNEWG	2-10	7-12	9-16	M2
	Flu	BI100_CG	**RR**	SLZTDIETP	**GCG**	IDTPIR	**G**	TPIBQDWG	2-10	7-12	9-16	M2
	Flu	BI100-CGcyc	**WWGC**	TDIET	**CG**	IDTPIR	**G**	TPIBQDWG	5-9	7-12	9-16	M2
	Flu	BI100-Cyc_2	**RRG**	CSLLT	**C**	SLLTEVQTPIRN	**GRR**	SEWGSRSN	2-5	2-13	13-20	M2
A	Flu	BI150-Dimer	**RRZC**	SLLTEVQTPIRN	**GRR**	VETPIRN			2-13	7-13	-	M2
B	Flu	BI150-Dimer	**WWQC**	TPIRSEWGCRSN	**GRR**	SNDSS	**G**		9-20	19-23	-	M2
A	Flu	BI150-new	**WW**	SLZTDIETP	**GCG**	IDTPIR	**G**	TPIBQDWG	2-10	7-12	9-16	M2
B	Flu	BI150-new	**RR(Har)**	IDTPIR	**G**	TPIBQDWG	**KG**	SLZTDIETPG	7-12	9-16	2-11	M2
A	Flu	BI150-2mod	**R**	SLZTDIETP	**Dpr**	IDTPIR	G	TPIBQDWG	2-10	7-12	9-16	M2
B	Flu	BI150-2mod	**RR**	IDTPIR	**GG**	TPI(Har)QEW	**Dpr(Ser)**	SLZTDIETPG	7-12	9-15	2-11	M2

A	Flu	BI 150-dim_2	RR	SLZTDIETP	GCG	IDTPIR	G	TPIBQDWG	2-10	7-12	9-16	M2
B	Flu	BI 150-dim_2	Har	IDTPIR	G	TPIBQDWG	KG	SLZTDIETPG	7-12	9-16	2-11	M2
	HIV	BI450-AdjBT1	**W_DWGC**	AKRRV	**CGG**	AKRRVVQREKRA			501-505	501-512	-	gp120
	HIV	BI450-AdjBT2	**W_DWGC**	IEEEG	**CGG**	IEEEGGERDR			222-226	222-231	-	gp41
	HIV		**CGG**	AKRRVV	**GG**	AKRRVV	**G**	QREKRAV	501-506	501-506	507-513	
	HIV		**CGGG**	DQQLL	**GG**	AEEEIV	**GG**	IEEEGGERDRDR	257-261	266-271	221-232	
	HIV		**CGG**	AKRRVV	**GG**	AKRRVV	**GG**	QREKR	501-506	501-506	507-511	
	HIV		**CGGG**	DQQLL	**GG**	AEEEIV	**GG**	IEEEGG	257-261	266-271	222-227	
	HIV		**CGG**	AEEEVV	**GG**	DQQLL			266-271	257-261	-	
	HIV		**GCGG**	AKRRVV	**GG**	AKRRVV			501-506	501-506	-	
A	HIV	BI400-B (a-chain)	**G**	AKRRVV	**GGCGG**	AKRRVVQREKRA	**G**	EREKRA	501-506	501-512	507-512	gp120
B	HIV	BI400-B (b-chain)	**GKG**	GIEEE	**GG**	RDRDR	**GG**	EQDRDR	221-225	229-233	228-233	gp41

E	HIV		**GG**	AKRRVVQREKRA	**G**	EREKRA			501-512	507-512		gp120
F	HIV		**G**	GIEEE	**GG**	RDRDR	**GG**	EQDRDR	221-225	229-233	228-233	gp41
G	HIV		**GG**	AKRRVVQREKRA	**G**	EREKRA	**GG**		501-512	507-512		gp120
H	HIV		**G**	GIEEE	**GG**	RDRDR	**GG**	EQDRDRGG	221-225	229-233	228-235	gp41
A	HIV	400-Seq B (a-chain)	**G**	AKRRVV	**GGCGG**	AKRRVVQREKRA	**G**	EREKRA	501-506	501-512	507-512	gp120
B	HIV	400-Seq B (b-chain)	**GKG**	GIEEE	**GG**	RDRDR	**GG**	QDRDR	221-225	229-233	229-233	gp41
D	HIV	400-Seq B* (a-chain)	**G**	AKRRVV	**GG(Dpr (Ser))GG**	AKRRVVQREKRA	**G**	EREKRA	501-506	501-512	507-512	gp120
C	HIV	400-Seq B* (b-chain)	**GKG**	GIEEE	**GG**	RDRDR	**GG**	QDRDR	221-225	229-233	229-233	gp41
A	HIV	BI400-Bu1 (a-chain)	**G**	AKRRVV	**GGCGG**	AKRRVVQREKRA	**G**	EREKRA	501-506	501-512	507-512	gp120
B	HIV	BI400-Bu1 (b-chain)	**GKG**	GIEEE	**GG**	ERDRDR	**GG**	QDRDR	221-225	228-233	229-233	gp41
A	HIV	BI400-Bu2 (a-chain)	**G**	AKRRVV	**GGCGG**	AKRRVVEREKRA	**G**	QREKRA	501-506	501-512	507-512	gp120
B	HIV	BI400-Bu2 (b-chain)	**GKG**	GIEEE	**GG**	QDRDR	**GG**	RDRDR	221-225	229-233	229-233	gp41
A	HIV	BI400-Bu3 (a-chain)	**G**	AKRRVV	**GGCGG**	AKRRVVEREKRA	**G**	QREKRA	501-506	501-512	507-512	gp120

B	HIV	BI400-Bu3 (b-chain)	**GKG**	GIEEE	**GG**	EQDRDR	**GG**	ERDRD	221-225	228-233	228-232	gp41
A	HIV	SEQ400_B (Cyc)	GC	AKRRVV	CGGKG	AKRRVVQREKRA	**G**	EREKRA	501-506	501-512	507-512	gp120
B	HIV	SEQ400_B (Cyc)	**GKG**	GIEEE	**GG**	RDRDR	**GG**	EQDRDR	221-225	229-233	228-233	gp41
A	HIV	SEQ400_B (Cyc)	G**C**	AKRRVV	**C**GGKG	GAKRRVVQREKRA	G	EREKRA	501-506	501-512	506-512	gp120
B	HIV	SEQ400_B (Cyc)	GCGG	IEEEGGRDRDR	GG	QDRDR			222-233	229-233		gp41
A	HIV	BI400-bu1 (Cyc)	G	CAKRRVVC	GGKGG	AKRRVVQREKRA	G	EREKRA	501-506	501-512	507-512	gp120
B	HIV	BI400-bu1 (Cyc)	CGG	IEEEGGERDRDR	GG	QDRDR			222-233	229-233		gp41
A	HIV	BI400-bu2 (Cyc)	G	CAKRRVVC	GGKGG	AKRRVVEREKRA	G	QREKRA	501-506	501-512	507-512	gp120
B	HIV	BI400-bu2 (Cyc)	CGG	IEEEGGQDRDR	GG	RDRDR			222-233	229-233		gp41
A	HIV	BI400-bu3 (Cyc)	G	CAKRRVVC	GGKGG	AKRRVVEREKRA	G	QREKRA	501-506	501-512	507-512	gp120
B	HIV	BI400-bu3 (Cyc)	CGG	IEEEGGEQDRDR	GG	RDRDR			222-233	229-233		gp41
A	HIV	BI400-rev (Cyc)	G	CAKRRVVC	GGKGG	AKRRVVQREKRA	G	EREKRA	501-506	501-512	507-512	gp120
B	HIV	BI400-rev (Cyc)	CGG	EEEIGGRDRD	GG	RDRDQ			222-233	229-233		gp41

A	HIV	BI450-1 (a-chain)	GG	RLEPWKH	GC	GSQPKTA	G	HPGSQ	7-13	15-21	13-17	Tat
B	HIV	BI450-1 (b-chain)	GG	FHSQV	C	FITKGLGISYGRK			32-36	38-50	-	Tat
A	HIV	BI450-1_2 (a-chain)		RLEPWKH	GC	GSQPKTA	GWK	HPGSQ	7-13	15-21	13-17	Tat
B	HIV	BI450-1_2 (b-chain)	C	FITKGLGISY	G	FITKGLGISYGRK			38-47	38-50		Tat
A	HCV	BI 350-1 (a-chain)	RR	LLADARV**C**S	GG	LLADARVSA			342-350	342-350		E2
B	HCV	BI350-1 (b-chain)	R	GV(Nle)AGIAYFS	C	GVLAGIAYYS			163-172	163-172		E1
A	HCV	BI 350-1mod1	RR	GNWAKVL	K	NWAKVI			366-372	367-372	-	E1
B	HCV	BI350-1mod1	RRG	LLADARV	GCG	SGADRV	CS		342-348	342-348	-	E2
A	HCV	BI 350-1mod2	RR	GNWAKVL	Dpr	NWAKVI			366-372	367-372	-	E1
B	HCV	BI350-1mod2	RRG	LLADARV	G(Dpr(Ser))G	SGADRV	CS		342-348	342-348	-	E2
A	HCV		RR	GNWAKVL	Lys(Me)	NWAKVI			366-372	367-372	-	E1
B	HCV		RRG	LLADARV	GEG	SGADRV	CS		342-348	342-348	-	E2
A	HCV		RR	GNWAKVL	Lys(Me)	NWAKVI			366-372	367-372	-	E1
B	HCV		RRG	LLADARV	GDG	SGADRV	CS		342-348	342-348	-	E2

A	HCV		RR	GNWAKVL	E	NWAKVI			366-372	367-372	-	E1
B	HCV		RRG	LLADARV	G(Lys(Me))G	SGADRV	CS		342-348	342-348	-	E2
A	HCV		RR	GNWAKVL	D	NWAKVI			366-372	367-372	-	E1
B	HCV		RRG	LLADARV	G(Lys(Me))G	SGADRV	CS		342-348	342-348	-	E2

WO 2012/092934 PCT/DK2012/050010

Specific embodiments of the invention

In some embodiments, in the peptide according to the present invention, X^1, X^3 and optional moiety X^5 each independently defines a linear sequence of any 1, 2, 3, 4, or 5 amino acids.

In some embodiments, in the peptide according to the present invention, X^1, X^3 and optional moiety X^5 consist of one or more amino acids selected from glycine, arginine, norleucine, glutamine, serine, lysine, tryptophan, cysteine, or a derivative thereof. The derivate may be a derivative on any of the specified amino acids.

In some embodiments, in the peptide according to the present invention, X^1, X^3 and optional moiety X^5 consist of one or more amino acids selected from glycine, arginine, norleucine, aspartic acid, glutamic acid, glutamine, serine, lysine, tryptophan, cysteine, ornithine, diaminopropionic acid (Dpr) or a derivative thereof. The derivate may be a derivative on any of the specified amino acids.

In some embodiments, in the peptide according to the present invention, X^5, and/or moiety X^6 is not present.

In some embodiments, in the peptide according to the present invention, X^2 and/or X^4 and/or X^6 has more than 55%, such as more than 60%, such as more than 65%, such as more than 70%, such as more than 75%, such as more than 80%, such as more than 85%, such as more than 90%, such as more than 95%, such as more than 96%, such as more than 97%, such as more than 98%, such as more than 99%, such as 100% sequence identity to a specific natural antigen.

In some embodiments, in the peptide according to the present invention, the specific natural antigen is a protein or peptide sequence derived from a disease antigen, such as an infectious agent, such as bacteria, virus, parasite, fungus, or cancer antigens such as oncogene (lung, stomach, breast cancer) or an antigen causing an autoimmune disease such as diabetes, multiple sclerosis (MS), celiac disease, Myalgic Encephalomyelitis (ME), psoriasis, and/or Crohn's Disease.

Accordingly confirmed and suspected autoimmune diseases, where relevant antigens may be derived include Achlorhydra Autoimmune Active Chronic Hepatitis, Acute Disseminated Encephalomyelitis, Acute hemorrhagic leukoencephalitis, Addison's Disease, Agammaglobulinemia, Alopecia areata, Amyotrophic Lateral Sclerosis, Ankylosing Spondylitis, Anti-GBM/TBM Nephritis, Antiphospholipid syndrome, Antisynthetase syndrome, Arthritis,

Atopic allergy, Atopic Dermatitis, Autoimmune Aplastic Anemia, Autoimmune cardiomyopathy, Autoimmune hemolytic anemia, Autoimmune hepatitis, Autoimmune inner ear disease, Autoimmune lymphoproliferative syndrome, Autoimmune peripheral neuropathy, Autoimmune pancreatitis, Autoimmune polyendocrine syndrome Types I, II, & III, Autoimmune progesterone dermatitis, Autoimmune thrombocytopenic purpura, Autoimmune uveitis, Balo disease/Balo concentric sclerosis, Bechets Syndrome, Berger's disease, Bickerstaff's encephalitis, Blau syndrome, Bullous Pemphigoid, Castleman's disease, Chagas disease, Chronic Fatigue Immune Dysfunction Syndrome, Chronic inflammatory demyelinating polyneuropathy, Chronic recurrent multifocal ostomyelitis, Chronic lyme disease, Chronic obstructive pulmonary disease, Churg-Strauss syndrome, Cicatricial Pemphigoid, Coeliac Disease, Cogan syndrome, Cold agglutinin disease, Complement component 2 deficiency, Cranial arteritis, CREST syndrome, Crohns Disease (one of two types of idiopathic inflammatory bowel disease "IBD"), Cushing's Syndrome, Cutaneous leukocytoclastic angiitis, Dego's disease, Dercum's disease, Dermatitis herpetiformis, Dermatomyositis, Diabetes mellitus type 1, Diffuse cutaneous systemic sclerosis, Dressler's syndrome, Discoid lupus erythematosus, Eczema, Endometriosis, Enthesitis-related arthritis, Eosinophilic fasciitis, Epidermolysis bullosa acquisita, Erythema nodosum, Essential mixed cryoglobulinemia, Evan's syndrome, Fibrodysplasia ossificans progressiva, Fibromyalgia, Fibromyositis, Fibrosing aveolitis, Gastritis, Gastrointestinal pemphigoid, Giant cell arteritis, Glomerulonephritis, Goodpasture's syndrome, Graves' disease, Guillain-Barré syndrome (GBS), Hashimoto's encephalitis, Hashimoto's thyroiditis, Haemolytic anaemia, Henoch-Schonlein purpura, Herpes gestationis, Hidradenitis suppurativa, Hughes syndrome (See Antiphospholipid syndrome), Hypogammaglobulinemia, Idiopathic Inflammatory Demyelinating Diseases, Idiopathic pulmonary fibrosis, Idiopathic thrombocytopenic purpura (See Autoimmune thrombocytopenic purpura), IgA nephropathy (Also Berger's disease), Inclusion body myositis, Inflammatory demyelinating polyneuopathy, Interstitial cystitis, Irritable Bowel Syndrome (IBS), Juvenile idiopathic arthritis, Juvenile rheumatoid arthritis, Kawasaki's Disease, Lambert-Eaton myasthenic syndrome, Leukocytoclastic vasculitis, Lichen planus, Lichen sclerosus, Linear IgA disease (LAD), Lou Gehrig's Disease (Also Amyotrophic lateral sclerosis), Lupoid hepatitis, Lupus erythematosus, Majeed syndrome, Ménière's disease, Microscopic polyangiitis, Miller-Fisher syndrome, Mixed Connective Tissue Disease, Morphea, Mucha-Habermann disease, Muckle–Wells syndrome, Multiple Myeloma, Multiple Sclerosis, Myasthenia gravis, Myositis, Narcolepsy, Neuromyelitis optica (Also Devic's Disease), Neuromyotonia, Occular cicatricial pemphigoid, Opsoclonus myoclonus syndrome, Ord thyroiditis, Palindromic rheumatism, PANDAS (Pediatric Autoimmune Neuropsychiatric Disorders Associated with Streptococcus), Paraneoplastic cerebellar degeneration, Paroxysmal nocturnal hemoglobinuria (PNH), Parry Romberg syndrome, Parsonnage-Turner syndrome, Pars planitis, Pemphigus, Pemphigus vulgaris, Pernicious anaemia, Perivenous encephalomyelitis, POEMS syndrome, Polyarteritis nodosa, Polymyalgia rheumatica,

Polymyositis, Primary biliary cirrhosis, Primary sclerosing cholangitis, Progressive inflammatory neuropathy, Psoriasis, Psoriatic Arthritis, Pyoderma gangrenosum, Pure red cell aplasia, Rasmussen's encephalitis, Raynaud phenomenon, Relapsing polychondritis, Reiter's syndrome, Restless leg syndrome, Retroperitoneal fibrosis, Rheumatoid arthritis, Rheumatoid fever, Sarcoidosis, Schizophrenia, Schmidt syndrome, Schnitzler syndrome, Scleritis, Scleroderma, Sjögren's syndrome, Spondyloarthropathy, Sticky blood syndrome, Still's Disease, Stiff person syndrome, Subacute bacterial endocarditis (SBE), Susac's syndrome, Sweet syndrome, Sydenham Chorea, Sympathetic ophthalmia, Takayasu's arteritis, Temporal arteritis (also known as "giant cell arteritis"), Tolosa-Hunt syndrome, Transverse Myelitis, Ulcerative Colitis (one of two types of idiopathic inflammatory bowel disease "IBD"), Undifferentiated connective tissue disease, Undifferentiated spondyloarthropathy, Vasculitis, Vitiligo, Wegener's granulomatosis, Wilson's syndrome, and Wiskott-Aldrich syndrome.

In some embodiments, in the peptide according to the present invention, the specific natural antigen is a viral protein, such as a structural protein, such as a capsid protein, a regulatory protein, an enzymatic protein, and a proteolytic protein.

In some embodiments, in the peptide according to the present invention, the viral protein is a protein, such as a structural protein, such as a core or envelope protein, of a virus selected from the Hepatitis C virus; influenza virus such as an M2 protein, human immunodeficiency virus (HIV), cytomegalovirus (CMV), and Human papillomavirus (HPV).

In some embodiments, in the peptide according to the present invention, the viral protein is a viral protein of Hepatitis C virus selected from any one HCV consensus sequence of a specific genotype, such as 1, such as subtypes 1a and 1b, genotype 2, such as 2a and 2b, genotype 3, such as 3a, genotype 4, genotype 5, and genotype 6.

In some embodiments, in the peptide according to the present invention, the sequence of amino acids defined by X^1-X^2-X^3-X^4-X^5-X^6 is not found in the native sequence of said natural antigen.

In some embodiments, in the peptide according to the present invention, the sequence of amino acids defined by X^1-X^2-X^3-X^4 is not found in the native sequence of said natural antigen.

In some embodiments, in the peptide according to the present invention, the sequence of amino acids defined by X^1-X^2-X^3 is not found in the native sequence of said natural antigen.

In some embodiments, in the peptide according to the present invention, the monomeric peptide is of 18-60 amino acids, such as 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 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, such as of 36-60 amino acids, such as of 37-60 amino acids, such as of 38-60 amino acids, such as of 39-60 amino acids, such as of 40-60 amino acids, such as of 42-60 amino acids, such as of 44-60 amino acids, such as of 46-60 amino acids, such as of 48-60 amino acids, such as of 50-60 amino acids, such as of 52-60 amino acids, such as of 54-60 amino acids, such as of 56-60 amino acids, such as of 58-60 amino acids.

In some embodiments, in the peptide according to the present invention, the monomeric peptide is of 18-60 amino acids, such as 18-58 amino acids, such as 18-56 amino acids, such as 18-54 amino acids, such as 18-52 amino acids, such as 18-50 amino acids, such as 18-48 amino acids, such as 18-46 amino acids, such as 18-44 amino acids, such as 18-42 amino acids, such as 18-40 amino acids, such as 18-39 amino acids, such as 18-38 amino acids, such as 18-37 amino acids, such as 18-36 amino acids, such as 18-35 amino acids, such as 18-34 amino acids, such as 18-33 amino acids, such as 18-32 amino acids, such as 18-31 amino acids, such as 18-30 amino acids, such as 18-29 amino acids, such as 18-28 amino acids, such as 18-27 amino acids, such as 18-26 amino acids, such as 18-25 amino acids, such as 18-24 amino acids, such as 18-23 amino acids, such as 18-22 amino acids, such as 18-21 amino acids, such as 18-20 amino acids, such as 18-19 amino acids.

In some embodiments, in the peptide according to the present invention, the monomeric peptide consist of not more than about 55 amino acids, such as not more than about 50 amino acids, such as not more than about 45 amino acids, such as not more than about 40 amino acids, such as not more than about 38 amino acids, such as not more than about 36 amino acids, such as not more than about 34 amino acids, such as not more than about 32 amino acids, such as not more than about 30 amino acids, such as not more than about 28 amino acids, such as not more than about 26 amino acids, such as not more than about 24 amino acids, such as not more than about 22 amino acids, such as not more than about 20 amino acids, such as not more than about 18 amino acids, such as not more than about 16 amino acids, such as not more than about 14 amino acids, such as not more than about 12 amino acids, such as not more than about 10 amino acids.

In some embodiments, in the peptide according to the present invention, the monomeric peptide consist of at least about 10 amino acids, such as at least about 12 amino acids, such

as at least about 14 amino acids, such as at least about 16 amino acids, such as at least about 18 amino acids, such as at least about 20 amino acids, such as at least about 22 amino acids, such as at least about 24 amino acids, such as at least about 26 amino acids, such as at least about 28 amino acids, such as at least about 30 amino acids, such as at least about 32 amino acids, such as at least about 34 amino acids, such as at least about 36 amino acids, such as at least about 38 amino acids, such as at least about 40 amino acids, such as at least about 45 amino acids, such as at least about 50 amino acids, such as at least about 55 amino acids, such as at least about 60.

In some embodiments, in the peptide according to the present invention, X^2, X^4, and optional moiety X^6 each independently defines a linear sequence of 5-12 amino acids, such as 5-10 amino acids, such as 5-8 amino acids.

In some embodiments, in the peptide according to the present invention, X^1, X^3 and optional moiety X^5 each independently defines a linear sequence of any 1, 2, 3, 4, or 5 amino acid independently selected from glycine, arginine, tryptophan, and cysteine, and derivatives thereof. In some embodiments, in the peptide according to the present invention, the overall net charge of X^1-X^2-X^3-X^4-X^5-X^6 is equal to or above 0, such as above 1, 2, 3, 4, or 5.

In some embodiments, in the peptide according to the present invention, X^1, X^3 and optional moiety X^5 each independently defines a linear sequence of any 1, 2, 3, 4, or 5 amino acid independently selected from glycine, arginine, norleucine, aspartic acid, glutamic acid, glutamine, serine, lysine, tryptophan, cysteine, ornithine, diaminopropionic acid (Dpr) and derivatives thereof. In some embodiments, in the peptide according to the present invention, the overall net charge of X^1-X^2-X^3-X^4-X^5-X^6 is equal to or above 0, such as above 1, 2, 3, 4, or 5.

In some embodiments, in the peptide according to the present invention, X^1, X^3 and optional moiety X^5 each independently defines a linear sequence of any 1, 2, 3, 4, or 5 amino acid independently selected from glycine, arginine, tryptophan, and cysteine, and derivatives thereof. In some embodiments, in the peptide according to the present invention, the overall net charge of X^1-X^2-X^3-X^4-X^5-X^6 is equal to or above 0, such as above 1, 2, 3, 4, or 5.

In some embodiments, in the peptide according to the present invention, the overall net charge of any one sequence selected from X^1, X^2, X^3, X^4, X^5, and X^6 is equal to or above 0, such as above 1, 2, 3, 4, or 5.

It is to be understood that parts of the monomeric peptide, such as any one sequence selected from X^1, X^2, X^3, X^4, X^5, and X^6, such as X^3 or X^5, may have an overall net charge

below 0, if other parts of the monomeric peptide such as any one other sequence selected from X^1, X^2, X^3, X^4, X^5, and X^6 has an overall net charge above 0.

It is essential that parts of the monomeric peptide is able to attach to a cell surface, such as a helper cell, which binding is facilitated by a positive net charge of the sequence at the place of binding to the cell membrane.

In some embodiments, in the peptide according to the present invention, the monomeric peptide is capable of inducing a humoral immune response.

In some embodiments, in the peptide according to the present invention, the sequence X^1 and/or X^3 and/or X^5 is selected from K, Lys(Me), RRG, G(Dpr(Ser))G, Dpr, Dpr(ser), GG(Dpr(Ser))GG (SEQ ID NO:150), GEG, CS, GDG, E, G(Lys(Me))G, D, RR, WWGC (SEQ ID NO: 54), RRG, RRZC (SEQ ID NO: 55), WWQC (SEQ ID NO: 56), WW, RR-Har, Har, WDWGC (SEQ ID NO: 57), CGG, CGGG (SEQ ID NO: 58), GCGG (SEQ ID NO: 59), G, GKG, GC, GG, C, R, GCG, CG, GRR, GGCGG (SEQ ID NO: 60), CGGKG (SEQ ID NO: 61), CGGKGG (SEQ ID NO: 62), GGKGG (SEQ ID NO: 63), KG, and GWK.

In some embodiments, in the peptide according to the present invention, the sequence X^2 and/or X^4 and/or X^6 is selected from SLLTEVETP (SEQ ID NO: 64), SLZTDIETP (SEQ ID NO: 65), TDIET (SEQ ID NO: 66), CSLLT (SEQ ID NO: 67), SLLTEVQTPIRN (SEQ ID NO: 68), TPIRSEWGCRSN (SEQ ID NO: 69), IDTPIR (SEQ ID NO: 70), AKRRV (SEQ ID NO: 71), IEEEG (SEQ ID NO: 72), AKRRVV (SEQ ID NO: 73), DQQLL (SEQ ID NO: 74), AEEEVV (SEQ ID NO: 75), GIEEE (SEQ ID NO: 76), IEEEGGRDRDR (SEQ ID NO: 77), CAKRRVVC (SEQ ID NO: 78), IEEEGGERDRDR (SEQ ID NO: 79), IEEEGGQDRDR (SEQ ID NO: 80), IEEEGGEQDRDR (SEQ ID NO: 81), EEEIGGRDRD (SEQ ID NO: 82), RLEPWKH (SEQ ID NO: 83), FHSQV (SEQ ID NO: 84), FITKGLGISY (SEQ ID NO: 85), LLADARVCS (SEQ ID NO: 86), GV(Nle)AGIAYFS (SEQ ID NO: 87), VETPIR (SEQ ID NO: 88), VETPIRN (SEQ ID NO: 89), SNDSS (SEQ ID NO: 90), TPIBQDWG (SEQ ID NO: 91), AKRRVVQREKRA (SEQ ID NO: 92), IEEEGGERDR (SEQ ID NO: 93), AEEEIV (SEQ ID NO: 94), RDRDR (SEQ ID NO: 95), ERDRDR (SEQ ID NO: 96), AKRRVVEREKRA (SEQ ID NO: 97), QDRDR (SEQ ID NO: 98), EQDRDR (SEQ ID NO: 99), RDRDQ (SEQ ID NO: 100), GSQPKTA (SEQ ID NO: 101), FITKGLGISYGRK (SEQ ID NO: 102), LLADARVSA (SEQ ID NO: 103), GVLAGIAYYS (SEQ ID NO: 104), TPIRNEWG (SEQ ID NO: 105), SEWGSRSN (SEQ ID NO: 106), SLZTDIETPG (SEQ ID NO: 107), QREKRAV (SEQ ID NO: 108), QREKR (SEQ ID NO: 109), IEEEGG (SEQ ID NO: 110), EREKRA (SEQ ID NO: 111), QREKRA (SEQ ID NO: 112), ERDRD (SEQ ID NO: 113), HPGSQ (SEQ ID NO: 114), TPIXQEW (SEQ ID NO: 151), EQDRDRGG (SEQ ID NO: 152), GNWAKVL (SEQ ID NO: 153), LLADARV (SEQ ID NO: 154), NWAKVI (SEQ ID NO: 155), and SGADRV (SEQ ID NO: 156).

In some embodiments, in the peptide according to the present invention, the monomeric peptide comprises at least one, such as one, two, three, four, or five amino acids selected from a Cys, a Lys, a Ser, an Asp, and a Glu residue, or derivatives thereof.

In some embodiments, in the peptide according to the present invention, the monomeric peptide comprises at least one cysteine, such as one, two, three, four, or five cysteines.

Some cysteines may be involved in intramolecular Cys-Cys bonds, whereas others may be involved in the bonding to another peptide monomer, i.e. an intermolecular bond.

In some embodiments in the peptide according to the present invention, the sequence X^1 and/or X^3 and/or X^5 is as defined in table 1.

In some embodiments in the peptide according to the present invention, the sequence X^2 and/or X^4 and/or X^6 is as defined in table 1.

In some embodiments the peptide with the structure X^1-X^2-X^3-X^4-X^5-X^6 is as defined in table 1.

In some embodiments in the multimeric peptide according to the present invention is a dimeric peptide.

In some embodiments the multimeric, such as a dimeric peptide comprising two , three, four, five, six, seven, eight, nine or ten monomeric peptides with the structure X^1-X^2-X^3-X^4-X^5-X^6, is as defined in table 1.

In some embodiments in the peptide according to the present invention, the sequence X^2 and/or X^4 and/or X^6 defines a sequence of 4-17, such as 5-16, such as 5-15, such as 5-14, such as 5-13, such as 5-12, such as 5-10 amino acids.

In some embodiments in the peptide according to the present invention, the sequence X^2 and/or X^4 and/or X^6 defines a sequence of more than 5, such as more than 6, such as more than 7, such as more than 8, such as more than 9, such as more than 10, such as more than 11, such as more than 12, such as more than 13, such as more than 14, such as more than 15, such as more than 16 amino acids.

In some embodiments in the peptide according to the present invention, the sequence X^2 and/or X^4 and/or X^6 defines a sequence of less than 17, such as less than 16, such as less

than 15, such as less than 14, such as less than 13, such as less than 12, such as less than 11, such as less than 10, such as less than 9, such as less than 8, such as less than 7, such as less than 6 amino acids.

In some embodiments in the peptide according to the present invention, the sequence X^1 and/or X^3 and/or X^5 defines a sequence that contain one or more amino acid selected from glycine (G), arginine (R), norleucine (Nle), aspartic acid (D), glutamic acid (E), glutamine (Q), serine (S), lysine (K), tryptophan (W), cysteine (C), Ornithine, diaminopropionic acid (Dpr) and a derivative thereof.

In some embodiments in the peptide according to the present invention, the monomeric peptide contain one or more intramolecular bond, such as one or more Cys-Cys bond.

In some embodiments in the peptide according to the present invention, the monomeric peptide has delayed proteolytic degradation in the N-terminal, such as by incorporation of the first 1, 2, or 3 amino acids in the N-terminal in the D-form, or by incorporation of the first 1, 2, or 3 amino acids in the N-terminal in beta or gamma form.

In some embodiments, in the multimeric, such as a dimeric peptide according to the present invention, the two or more monomeric peptides are identical in sequence.

In some embodiments, in the multimeric, such as dimeric peptide according to the present invention, the two or more monomeric peptides are different in sequence.

In some embodiments, in the multimeric, such as dimeric peptide according to the present invention, one, two or more of the peptide strands of the multimeric, such as dimeric peptide has delayed proteolytic degradation in the N-terminal, such as by incorporation of the first 1, 2, or 3 amino acids in the N-terminal in the D-form, or by incorporation of the first 1, 2, or 3 amino acids in the N-terminal in beta or gamma form.

In some embodiments, in the multimeric, such as dimeric peptide according to the present invention, the linker is placed within any sequence selected from X^1, X^2, X^3, X^4, X^5, and X^6, such as in X^1, X^2 or X^3 of the first monomeric peptide to anywhere on the at least one second monomeric peptide, such as within the sequence of X^1, X^2 or X^3.

In some embodiments, in the multimeric, such as dimeric peptide according to the present invention, the linker is placed at an amino acid position selected from 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57,

58, 59, 60 of the first monomeric peptide to a position selected from 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60 of the at least one second monomeric peptide.

In some embodiments, in the multimeric, such as dimeric peptide according to the present invention, the linker is placed in X^1, X^2 or X^3 of the first monomeric peptide to anywhere on the at least one second monomeric peptide, such as within the sequence of X^1, X^2 or X^3.

In some embodiments, in the multimeric, such as dimeric peptide according to the present invention, the multimeric, such as dimeric peptide contain a helper epitope of at least 12 amino acids, such as at least 13, 14, 15 or 17 amino acids, which helper epitope consist of a combined sequence of amino acids, which is a sequence of amino acids from the first monomeric peptide, and a sequence of amino acids from the at least one second monomeric peptide, such as between 2-12 amino acids from the first monomeric peptide and 2-12 amino acids from the at least one second monomeric peptide.

It is to be understood that an epitope may not only be present within the sequence of the monomeric peptide. An epitope may also be present with a combination of amino acids of the first and the at least one second monomeric peptide in a multimeric, such as dimeric peptide sequence, wherein this combination of amino acids forms a sequence that span from the first to the at least one second monomeric peptide sequence. This epitope may be a continuous sequence of amino acids or it may be a three-dimensional epitope with amino acids found in both monomeric peptides.

In some embodiments, in the multimeric, such as dimeric peptide according to the present invention, the intermolecular bond is a disulfide (S-S) bond between two Cys residues.

In some embodiments, in the multimeric, such as dimeric peptide according to the present invention, the intermolecular bond is a methylated peptide bond between a N-ε-methylated Lys side-chain and the side-chain of an Asp or Glu residue.

In some embodiments, in the multimeric, such as dimeric peptide according to the present invention, the intermolecular bond is a thioether bond between a Cys residue in the first monomeric peptide and a modified Lys residue in the at least one second monomeric peptide.

In some embodiments, in the multimeric, such as dimeric peptide according to the present invention, the intermolecular bond is an oxime bond.

In some embodiments, in the multimeric, such as dimeric peptide according to the present invention, the intermolecular bond is an oxime bond between a derivatized Lys residue in the first monomeric peptide and a derivatized Ser residue in the at least one second monomeric peptide.

In some embodiments, in the multimeric, such as dimeric peptide according to the present invention, the intermolecular bond is an oxime bond between a derivatized lysine, ornitine or diaminopropionic acid residue in the first monomeric peptide and a derivatized serine moiety, such as a serine residue , such as in a serinyl diaminopropionic acid residue, such as in a serinyl lysin residue or such as in a serinyl ornitine residue, in the at least one second monomeric peptide.

In some embodiments, in the multimeric, such as dimeric peptide according to the present invention, the monomeric peptides are linked by a polyethylene glycol (PEG) linker, such as through an Asp or a Glu residue in the first monomeric peptide and an Asp or a Glu residue in the at least one second monomeric peptide.

In some embodiments, in the multimeric, such as dimeric peptide according to the present invention, any one of the monomeric peptides is independently as defined herein.

In some embodiments, the peptide according to the present invention is essentially a non-cell-penetrating peptide. In other embodiments, the peptide according to the present invention is a cell-penetrating peptide. In some embodiments, the peptide according to the present invention is able to attach to the cell membrane of an antigen presenting cell.

It is to be understood that when referring to the peptides ability to attach to and enter a cell, such as an antigen presenting cell, it may be with reference to the complete sequence of the peptide as well as a fragment thereof, such as a fragment representing an epitope. Accordingly, it may be the case that the entire sequence is essentially a non-cell-penetrating peptide, whereas a fragment of the peptide is able to efficiently enter a cell, such as an antigen presenting cell.

In some embodiments, the peptide according to the present invention is not a peptide or a dimeric peptide as specifically disclosed in International Patent Application No: PCT/DK2011/050460.

In some embodiments, the peptide according to the present invention is not a peptide or a dimeric peptide as specifically disclosed in International Patent Application No: PCT/EP2010/059513, such as one selected from:

CGGAKRRVVGGAKRRVVGQREKRAV (SEQ ID NO:115)

CGGGDQQLLGGAEEEIVGGIEEEGGERDRDR (SEQ ID NO:116)

CGGAKRRVVGGAKRRVVGGQREKR (SEQ ID NO:117)

CGGGDQQLLGGAEEEIVGGIEEEGG (SEQ ID NO:118)

CGGAEEEVVGGDQQLL (SEQ ID NO:119)

GCGGAKRRVVGGAKRRVV (SEQ ID NO:120)

GAKRRVVGGCGGAKRRVVQREKRAGEREKRA (SEQ ID NO:121)

GKGGIEEEGGRDRDRGGEQDRDR (SEQ ID NO:122)

GAKRRVVGGCGGAKRRVVQREKRAGEREKRA (SEQ ID NO:123)

GKGGIEEEGGERDRDRGGQDRDR (SEQ ID NO:124)

GAKRRVVGGCGGAKRRVVEREKRAGQREKRA (SEQ ID NO:125)

GKGGIEEEGGQDRDRGGRDRDR (SEQ ID NO:126)

GAKRRVVGGCGGAKRRVVEREKRAGQREKRA (SEQ ID NO:127)

GKGGIEEEGGEQDRDRGGERDRD (SEQ ID NO:128)

In some embodiments, the peptide according to the present invention is not a dimeric peptide selected from (The peptides are linked via the underlined amino acid):

CGGAKRRVVGGAKRRVVGQREKRAV
|
CGGGDQQLLGGAEEEIVGGIEEEGGERDRDR;

CGGAKRRVVGGAKRRVVGGQREKR
|
CGGGDQQLLGGAEEEIVGGIEEEGG;

CGGAEEEVVGGDQQLL
|
GCGGAKRRVVGGAKRRVV;

GAKRRVVGGCGGAKRRVVQREKRAGEREKRA
|
GKGGIEEEGGRDRDRGGEQDRDR;

GAKRRVVGGCGGAKRRVVQREKRAGEREKRA
|
GKGGIEEEGGERDRDRGGQDRDR;

GAKRRVVGGCGGAKRRVVEREKRAGQREKRA
|
GKGGIEEEGGQDRDRGGRDRDR;

GAKRRVVGGCGGAKRRVVEREKRAGQREKRA
|
GKGGIEEEGGEQDRDRGGERDRD;

CGGAKRRVVGGAKRRVVGQREKRAV
|
CGGGDQQLLGGAEEEIVGGIEEEGG;

CGGAKRRVVGGAKRRVVGQREKRAV
|
GCGGAKRRVVGGAKRRVV;

CGGAKRRVVGGAKRRVVGGQREKR
|
CGGGDQQLLGGAEEEIVGGIEEEGGERDRDR;

CGGAKRRVVGGAKRRVVGGQREKR
|
GCGGAKRRVVGGAKRRVV;

CGGAEEEVVGGDQQLL
|
CGGGDQQLLGGAEEEIVGGIEEEGGERDRDR;

CGGAEEEVVGGDQQLL
|
CGGGDQQLLGGAEEEIVGGIEEEGG;

GAKRRVVGGCGGAKRRVVQREKRAGEREKRA
|
GKGGIEEEGGQDRDRGGRDRDR;

GAKRRVVGGCGGAKRRVVQREKRAGEREKRA
|
GKGGIEEEGGEQDRDRGGERDRD;

GAKRRVVGGCGGAKRRVVEREKRAGQREKRA
|
GKGGIEEEGGRDRDRGGEQDRDR; or

GAKRRVVGGCGGAKRRVVEREKRAGQREKRA
|
GKGGIEEEGGERDRDRGGQDRDR.

Numbered embodiments according to the present invention:

1. An isolated monomeric peptides consisting of not more than 60 amino acids with the following structure

X^1-X^2-X^3-X^4-X^5-X^6 (formula I),

wherein X^1, X^3 and optional moiety X^5 each independently defines a linear sequence of any 1, 2, 3, 4, or 5 amino acid independently selected from glycine, arginine, norleucine, glutamine, serine, lysine, tryptophan, cysteine, or a derivative thereof; X^2, X^4, and optional moiety X^6 each independently defines a linear sequence of 5-17 amino acids, each having more than 50% sequence identity to a specific natural antigen.

2. The isolated monomeric peptide according to embodiment 1, wherein X^2 and/or X^4 and/or X^6 has more than 55%, such as more than 60%, such as more than 65%, such as more than 70%, such as more than 75%, such as more than 80%, such as more than 85%, such as more than 90%, such as more than 95%, such as more than 96%, such as more than 97%, such as more than 98%, such as more than 99%, such as 100% sequence identity to a specific natural antigen.

3. The isolated monomeric peptide according to embodiments 1 or 2, wherein said specific natural antigen is a protein or peptide sequence derived from a disease antigen, such as an infectious agent, such as bacteria, virus, parasite, fungus, or cancer antigens such as oncogene (lung, stomach, breast cancer) or an antigen causing an autoimmune disease such as diabetes, multiple sclerosis (MS), celiac disease, Myalgic Encephalomyelitis (ME), psoriasis, and/or Crohn's Disease.

4. The isolated monomeric peptide according to any one of embodiments 1-3, wherein said specific natural antigen is a viral protein, such as a structural protein, such as a capsid protein, a regulatory protein, an enzymatic protein, and a proteolytic protein.

5. The isolated monomeric peptide according to any one of embodiments 1-4, wherein said viral protein is a protein, such as a core protein, of a virus selected from the Hepatitis C virus; influenza virus, such as an M2 protein, human immunodeficiency virus (HIV), cytomegalovirus (CMV), and Human papillomavirus (HPV).

6. The isolated monomeric peptide according to embodiment 5, wherein said viral protein is a viral protein of Hepatitis C virus selected from any one HCV consensus sequence of a specific genotype, such as 1, such as subtypes 1a and 1b, genotype 2, such as 2a and 2b, genotype 3, such as 3a, genotype 4, genotype 5, and genotype 6.

7. The isolated monomeric peptide according to any one of embodiments 1-3, wherein the sequence of amino acids defined by X^1-X^2-X^3-X^4-X^5-X^6 is not found in the native sequence of said natural antigen.

8. The isolated monomeric peptide according to any one of embodiments 1-7, which monomeric peptide is of 18-60 amino acids, such as 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 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, such as of 36-60 amino acids, such as of 37-60 amino acids, such as of 38-60 amino acids, such as of 39-60 amino acids, such as of 40-60 amino acids, such as of 42-60 amino acids, such as of 44-60 amino acids, such as of 46-60 amino acids, such as of 48-60 amino acids, such as of 50-60 amino acids, such as of 52-60 amino acids, such as of 54-60 amino acids, such as of 56-60 amino acids, such as of 58-60 amino acids.

9. The isolated monomeric peptide according to any one of embodiments 1-8, which monomeric peptide is of 18-60 amino acids, such as 18-58 amino acids, such as 18-56 amino acids, such as 18-54 amino acids, such as 18-52 amino acids, such as 18-50 amino acids, such as 18-48 amino acids, such as 18-46 amino acids, such as 18-44 amino acids, such as 18-42 amino acids, such as 18-40 amino acids, such as 18-39 amino acids, such as 18-38 amino acids, such as 18-37 amino acids, such as 18-36 amino acids, such as 18-35 amino acids, such as 18-34 amino acids, such as 18-33 amino acids, such as 18-32 amino acids, such as 18-31 amino acids, such as 18-30 amino acids, such as 18-29 amino acids, such as

18-28 amino acids, such as 18-27 amino acids, such as 18-26 amino acids, such as 18-25 amino acids, such as 18-24 amino acids, such as 18-23 amino acids, such as 18-22 amino acids, such as 18-21 amino acids, such as 18-20 amino acids, such as 18-19 amino acids.

10. The isolated monomeric peptide according to any one of embodiments 1-9, which monomeric peptide consist of not more than about 55 amino acids, such as not more than about 50 amino acids, such as not more than about 45 amino acids, such as not more than about 40 amino acids, such as not more than about 38 amino acids, such as not more than about 36 amino acids, such as not more than about 34 amino acids, such as not more than about 32 amino acids, such as not more than about 30 amino acids, such as not more than about 28 amino acids, such as not more than about 26 amino acids, such as not more than about 24 amino acids, such as not more than about 22 amino acids, such as not more than about 20 amino acids, such as not more than about 18 amino acids, such as not more than about 16 amino acids, such as not more than about 14 amino acids, such as not more than about 12 amino acids, such as not more than about 10 amino acids.

11. The isolated monomeric peptide according to any one of embodiments 1-10, which monomeric peptide consist of at least about 10 amino acids, such as at least about 12 amino acids, such as at least about 14 amino acids, such as at least about 16 amino acids, such as at least about 18 amino acids, such as at least about 20 amino acids, such as at least about 22 amino acids, such as at least about 24 amino acids, such as at least about 26 amino acids, such as at least about 28 amino acids, such as at least about 30 amino acids, such as at least about 32 amino acids, such as at least about 34 amino acids, such as at least about 36 amino acids, such as at least about 38 amino acids, such as at least about 40 amino acids, such as at least about 45 amino acids, such as at least about 50 amino acids, such as at least about 55 amino acids, such as at least about 60.

12. The isolated monomeric peptide according to any one of embodiments 1-11, wherein the overall net charge of X^1-X^2-X^3-X^4-X^5-X^6 is equal to or above 0, such as above 1, 2, 3, 4, or 5.

13. The isolated monomeric peptide according to any one of embodiments 1-12, wherein said monomeric peptide is capable of inducing a humoral immune response.

14. The isolated monomeric peptide according to any one of embodiments 1-13, wherein said sequence X^1 and/or X^3 and/or X^5 is selected from RR, WWGC (SEQ ID NO: 54), RRG, RRZC (SEQ ID NO: 55), WWQC (SEQ ID NO: 56), WW, RR-Har, Har, WDWGC (SEQ ID NO: 57), CGG, CGGG (SEQ ID NO: 58), GCGG (SEQ ID NO: 59), G, GKG, GC, GG, C, R, GCG, CG,

GRR, GGCGG (SEQ ID NO: 60), CGGKG (SEQ ID NO: 61), CGGKGG (SEQ ID NO: 62), GGKGG (SEQ ID NO: 63), KG, and GWK.

15. The isolated monomeric peptide according to any one of embodiments 1-14, wherein said sequence X^2 and/or X^4 and/or X^6 is selected from SLLTEVETP (SEQ ID NO: 64), SLZTDIETP (SEQ ID NO: 65), TDIET (SEQ ID NO: 66), CSLLT (SEQ ID NO: 67), SLLTEVQTPIRN (SEQ ID NO: 68), TPIRSEWGCRSN (SEQ ID NO: 69), IDTPIR (SEQ ID NO: 70), AKRRV (SEQ ID NO: 71), IEEEG (SEQ ID NO: 72), AKRRVV (SEQ ID NO: 73), DQQLL (SEQ ID NO: 74), AEEEVV (SEQ ID NO: 75), GIEEE (SEQ ID NO: 76), IEEEGGRDRDR (SEQ ID NO: 77), CAKRRVVC (SEQ ID NO: 78), IEEEGGERDRDR (SEQ ID NO: 79), IEEEGGQDRDR (SEQ ID NO: 80), IEEEGGEQDRDR (SEQ ID NO: 81), EEEIGGRDRD (SEQ ID NO: 82), RLEPWKH (SEQ ID NO: 83), FHSQV (SEQ ID NO: 84), FITKGLGISY (SEQ ID NO: 85), LLADARVCS (SEQ ID NO: 86), GV(Nle)AGIAYFS (SEQ ID NO: 87), VETPIR (SEQ ID NO: 88), VETPIRN (SEQ ID NO: 89), SNDSS (SEQ ID NO: 90), TPIBQDWG (SEQ ID NO: 91), AKRRVVQREKRA (SEQ ID NO: 92), IEEEGGERDR (SEQ ID NO: 93), AEEEIV (SEQ ID NO: 94), RDRDR (SEQ ID NO: 95), ERDRDR (SEQ ID NO: 96), AKRRVVEREKRA (SEQ ID NO: 97), QDRDR (SEQ ID NO: 98), EQDRDR (SEQ ID NO: 99), RDRDQ (SEQ ID NO: 100), GSQPKTA (SEQ ID NO: 101), FITKGLGISYGRK (SEQ ID NO: 102), LLADARVSA (SEQ ID NO: 103), GVLAGIAYYS (SEQ ID NO: 104), TPIRNEWG (SEQ ID NO: 105), SEWGSRSN (SEQ ID NO: 106), SLZTDIETPG (SEQ ID NO: 107), QREKRAV (SEQ ID NO: 108), QREKR (SEQ ID NO: 109), IEEEGG (SEQ ID NO: 110), EREKRA (SEQ ID NO: 111), QREKRA (SEQ ID NO: 112), ERDRD (SEQ ID NO: 113), and HPGSQ (SEQ ID NO: 114).

16. The isolated monomeric peptide according to any one of embodiments 1-15, wherein said monomeric peptide comprises at least one amino acid selected from a Cys, a Lys, a Ser, an Asp, and a Glu residue, or derivatives thereof.

17. The isolated monomeric peptide according to any one of embodiments 1-16, wherein said sequence X^1 and/or X^3 and/or X^5 is as defined in table 1.

18. The isolated monomeric peptide according to any one of embodiments 1-17, wherein said sequence X^2 and/or X^4 and/or X^6 is as defined in table 1.

19. The isolated monomeric peptide according to any one of embodiments 1-18, wherein said sequence X^2 and/or X^4 and/or X^6 defines a sequence of 4-17, such as 5-16, such as 5-15, such as 5-14, such as 5-13, such as 5-12, such as 5-10 amino acids.

20. The isolated monomeric peptide according to any one of embodiments 1-19, which monomeric peptide contain one or more intramolecular bond, such as one or more Cys-Cys bond.

21. The isolated monomeric peptide according to any one of embodiments 1-20, which monomeric peptide has delayed proteolytic degradation in the N-terminal, such as by incorporation of the first 1, 2, or 3 amino acids in the N-terminal in the D-form, or by incorporation of the first 1, 2, or 3 amino acids in the N-terminal in beta or gamma form.

22. An isolated multimeric, such as dimeric peptide comprising two or more monomeric peptides, each monomeric peptide independently consisting of not more than 60 amino acids with the following structure

X^1-X^2-X^3-X^4-X^5-X^6 (formula I),

wherein X^1, X^3 and optional moiety X^5 independently defines a linear sequence of any 1, 2, 3, 4, or 5 amino acid independently selected from glycine, arginine, norleucine, glutamine, serine, lysine, tryptophan, cysteine, or a derivative thereof; X^2, X^4, and optional moiety X^6 each independently defines a linear sequence of 5-17 amino acids, each having more than 50% sequence identity to a specific natural antigen, said monomeric peptides being covalently joined by one or more intermolecular bond.

23. The isolated dimeric peptide according to embodiment 22, wherein two or more monomeric peptides are identical in sequence.

24. The isolated dimeric peptide according to embodiment 22, wherein two or more monomeric peptides are different in sequence.

25. The isolated multimeric, such as dimeric peptide according to any one of embodiments 22-24, wherein one or more peptide strands of the multimeric, such as dimeric peptide has delayed proteolytic degradation in the N-terminal, such as by incorporation of the first 1, 2, or 3 amino acids in the N-terminal in the D-form, or by incorporation of the first 1, 2, or 3 amino acids in the N-terminal in beta or gamma form.

26. The isolated multimeric, such as dimeric peptide according to any one of embodiments 22-25, wherein the linker is placed within any sequence selected from X^1, X^2, X^3, X^4, X^5, and X^6, such as in X^1, X^2 or X^3 of the first monomeric peptide to anywhere on the at least one second monomeric peptide, such as within the sequence of X^1, X^2 or X^3.

27. The isolated multimeric, such as dimeric peptide according to any one of embodiments 22-26, which multimeric, such as dimeric peptide contain a helper epitope of at least 12 amino acids, such as at least 13, 14, 15 or 17 amino acids, which helper epitope consist of a combined sequence of amino acids, which is a sequence of amino acids from the first monomeric peptide, and a sequence of amino acids from the at least one second monomeric peptide, such as between 2-12 amino acids from the first monomeric peptide and 2-12 amino acids from the at least one second monomeric peptide.

28. The isolated multimeric, such as dimeric peptide according to any one of embodiments 22-27, wherein said intermolecular bond is selected from a disulfide (S-S) bond between two Cys residues.

29. The isolated multimeric, such as dimeric peptide according to any one of embodiments 22-28, wherein said intermolecular bond is a thioether bond between a Cys residue in the first monomeric peptide and a modified Lys residue in the at least one second monomeric peptide.

30. The isolated multimeric, such as dimeric peptide according to any one of embodiments 22-29, wherein said intermolecular bond is an oxime bond between a derivatized Lys residue in the first monomeric peptide and a derivatized Ser residue in the at least one second monomeric peptide.

31. The isolated multimeric, such as dimeric peptide according to any one of embodiments 22-30, wherein said intermolecular bond is a methylated peptide bond between a N-ε-methylated Lys side-chain in the first monomeric peptide and the side-chain of an Asp or Glu residue in the at least one second monomeric peptide.

32. The isolated multimeric, such as dimeric peptide according to any one of embodiments 22-31, wherein said monomeric peptides are linked by a polyethylene glycol (PEG) linker, such as through an Asp or a Glu residue in the first monomeric peptide and an Asp or a Glu residue in the at least one second monomeric peptide, or by a polyLys core.

33. The isolated multimeric, such as dimeric peptide according to any one of embodiments 22-31, wherein any one of said monomeric peptide is independently as defined in any one of embodiments 1-21.

34. Composition comprising two or more compounds selected from a monomeric peptide as defined in any one of embodiments 1-21, and an isolated multimeric, such as dimeric peptide as defined in any one of embodiments 22-33.

35. Use of a peptide selected from a monomeric peptide as defined in any one of embodiments 1-21, and an isolated multimeric, such as dimeric peptide as defined in any one of embodiments 22-33 for inducing a humoral immune response in a subject.

36. An isolated nucleic acid or polynucleotide encoding a peptide according to any one of embodiments 1-33.

37. A vector comprising the nucleic acid or polynucleotide according to embodiment 36.

38. A host cell comprising the vector according to embodiment 37.

39. An immunogenic composition comprising at least one monomeric peptide according to any one of embodiments 1-21, an isolated multimeric, such as dimeric peptide according to any one of embodiments 22-33, a peptide composition according to embodiment 34, the nucleic acid or polynucleotide according to embodiment 36, or the vector according to embodiment 37; in combination with a pharmaceutically acceptable diluent or vehicle and optionally an immunological adjuvant.

40. The immunogenic composition according to embodiment 39 in the form of a vaccine composition.

41. A method for inducing an immune response in a subject against an antigen which comprises administration of at least one monomeric peptide according to any one of embodiments 1-21, an isolated multimeric, such as dimeric peptide according to any one of embodiments 22-33, a peptide composition according to embodiment 34, the nucleic acid or polynucleotide according to embodiment 36, or the vector according to embodiment 37; or the composition according to any one of embodiments 39-40.

42. A method for reducing and/or delaying the pathological effects of a disease antigen, such as an infectious agent in a subject infected with said agent or having said disease caused by said antigen, the method comprising administering an effective amount of at least one monomeric peptide according to any one of embodiments 1-21, an isolated multimeric, such as dimeric peptide according to any one of embodiments 22-33, a peptide composition according to embodiment 34, the nucleic acid or polynucleotide according to embodiment 36, or the vector according to embodiment 37; or the composition according to any one of embodiments 39-40.

43. A peptide according to any one of embodiments 1-33 for use as a medicament.

44. A peptide according to any one of embodiments 1-33 for treating the pathological effects of a disease antigen, such as an infectious agent in a subject infected with said agent or having said disease caused by said antigen.

45. A peptide according to any one of embodiments 1-33 for use in an in vitro assay, such as an ELISA assay, such as for diagnostic purposes.

46. Use of a peptide according to any one of embodiments 1-33 for in vitro assay, such as an ELISA assay, such as for diagnostic purposes.

Sequence list (amino acids in bold represents suitable antigenic sequences that may be used as any of X^2 and/or X^4 and/or X^6 as defined in formula I of the present invention)

SEQ ID NO:1: Flu M2
>gi|21693176|gb|AAM75162| /Human/M2/H1N1/Puerto Rico/1934/// matrix protein M2 [Influenza A virus (A/Puerto Rico/8/34/Mount Sinai(H1N1))]
MSLLTEVETPIRNEWGCRCNGSSDPLAIAANIIGILHLTLWILDRLFFKCIYRRFKYGLKGGPSTEGVPK
SMREEYRKEQQSAVDADDGHFVSIELE

SEQ ID NO:2: >gi|1906383|gb|AAB50256.1| tat protein [Human immunodeficiency virus
1]MEPVDPRLEPWKHPGSQPKTACTNCYCKKCCFHCQVCFITKALGISYGRKKRRQRRRAHQNSQTHQASLS
KQPTSQPRGDPTGPKE

SEQ ID NO:3: >B.FR.1983.HXB2-LAI-IIIB-BRU (gp120)
MRVKEKYQHLWRWGWRWGTMLLGMLMICSATEKLWVTVYYGVPVWKEATTTLFCASDAKAYDTEVHNVWATHACV
PTDPNPQEVVLVNVTENFNMWKNDMVEQMHEDIISLWDQSLKPCVKLTPLCVSLKCTDLKNDTNTNSSSGRMIME
KGEIKNCSFNISTSIRGKVQKEYAFFYKLDIIPIDNDTTSYKLTSCNTSVITQACPKVSFEPIPIHYCAPAGFAI
LKCNNKTFNGTGPCTNVSTVQCTHGIRPVVST**QLLLNGSLAEEEVVIR**SVNFTDNAKTIIVQLNTSVEINCTRPN
NNTRKRIRIQRGPGRAFVTIGKIGNMRQAHCNISRAKWNNTLKQIASKLREQFGNNKTIIFKQSSGGDPEIVTHS
FNCGGEFFYCNSTQLFNSTWFNSTWSTEGSNNTEGSDTITLPCRIKQIINMWQKVGKAMYAPPISGQIRCSSNIT
GLLLTRDGGNSNNESEIFRPGGGDMRDNWRSELYKYKVVKIEPLGV**APTKAKRRVVQREKR**

SEQ ID NO:4: HIV gp41
>B.FR.1983.HXB2-LAI-IIIB-BRU (ACC No. K03455)
AVGIGALFLGFLGAAGSTMGAASMTLTVQARQLLSGIVQQQNNLLRAIEAQQHLLQLTVWGIKQLQARILAVERY
LKDQQLLGIWGCSGKLICTTAVPWNASWSNKSLEQIWNHTTWMEWDREINNYTSLIHSLIEESQNQQEKNEQELL
ELDKWASLWNWFNITNWLWYIKLFIMIVGGLVGLRIVFAVLSIVNRVRQGYSPLSFQTHLPTPRGP**DRPEGIEEE**
GGERDRDRSIRLVNGSLALIWDDLRSLCLFSYHRLRDLLLIVTRIVELLGRRGWEALKYWWNLLQYWSQELKNSA
VSLLNATAIAVAEGTDRVIEVVQGACRAIRHIPRRIRQGLERILL

SEQ ID NO:5: >1b._._.AB016785._ (HCV-E1)
YEVRNVSGVYHVTNDCSNSSIVYGAADMIMHTPGCVPCVRENNSSRCWVALTPTLAARNRSIPTTTIRRHVDLLV
GAAAFCSAMYVGDLCGSVFLVSQLFTFSPRRYETVQDCNCSLYPGHVSGHRMAWDMMMNWSPTAALVVSQLLRIP
QAVVDMVTGAHW**GVLAGLAYYS**MVGNWAKVLIVMLLFAGVDG

SEQ ID NO:6: >1b._._.AB016785.AB016785
TTHVTGGQTGRTTLGITAMFAFGPHQKLQLINTNGSWHINRTALNCNDSLNTGFLAALFYARKFNSSGCPERMAS
CRPIDKFVQGWGPITHAVPDNLDQRPYCWHYAPQPCGIIPASQVCGPVYCFTPSPVVVGTTDRFGAPTYTWGENE
TDVLLLNNTRPPQGNWFGCTWMNGTGFAKTCGGPPCNIGGVGNNTLTCPTDCFRKHPEATYTKCGSGPWLTPRCM

VDYPYRLWHYPCTVNFTIFKVRMYVGGVEHRLTAACNWTRGERCDLEDRDRSELSPLLLSTTEWQVLPCSFTTLP
ALSTGLIHLHQNIVDVQYLYGVGSAVVSIVIKWEYILLLFL**LLADARVCA**CLWMMLLIAQAEA

SEQ ID NO:7: Accession no AF009606; Hepatitis C virus subtype 1a polyprotein gene, complete cds.

MSTNPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRL
GVRATRKTSERSQPRGRRQPIPKARRPEGRTWAQPGYPWPLYGNEGCGWAGWLLSPRG
SRPSWGPTDPRRRSRNLGKVIDTLTCGFADLM**GYIPLVGAPLG**GA**ARALAHGVRVL**ED
GVNYATGNLPGCSFSIFLLALLSCLTVPASAYQVRNSSGLYHVTNDCPNSSIVYEAAD
AILHTPGCVPCVREGNASRCWVAVTPTVATRDGKLPTTQLRRHIDLLVGSATLCSALY
VGDLCGSVFLVGQLFTFSPRRHWTTQDCNCSIYPGHITGHRMAWDMMMNWSPTAALVV
AQLLRIPQAIMDMIAGAHWGVLAGIAYFSMVGNWAKVLVVLLLFAGVDAETHVTGGSA
GRTTAGLVGLLTPGAKQNIQLINTNGSWHINSTALNCNESLNTGWLAGLFYQHKFNSS
GCPERLASCRRLTDFAQGWGPISYANGSGLDERPYCWHYPPRPCGIVPAKSVCGPVYC
FTPSPVVVGTTDRSGAPTYSWGANDTDVFVLNNTRPPLGNWFGCTWMNSTGFTKVCGA
PPCVIGGVGNNTLLCPTDCFRKHPEATYSRCGSGPWITPRCMVDYPYRLWHYPCTINY
TIFKVRMYVGGVEHRLEAACNWTRGERCDLEDRDRSELSPLLLSTTQWQVLPCSFTTL
PALSTGLIHLHQNIVDVQYLYGVGSSIASWAIKWEYVVLLFLLLADARVCSCLWMMLL
ISQAEAALENLVILNAASLAGTHGLVSFLVFFCFAWYLKGRWVPGAVYAFYGMWPLLL
LLLALPQRAYALDTEVAASCGGVVLVGLMALTLSPYYKRYISWCMWWLQYFLTRVEAQ
LHVWVPPLNVRGGRDAVILLMCVVHPTLVFDITKLLLAIFGPLWILQASLLKVPYFVR
VQGLLRICALARKIAGGHYVQMAIIKLGALTGTYVYNHLTPLRDWAHNGLRDLAVAVE
PVVFSRMETKLITWGADTAACGDIINGLPVSARRGQEILLGPADGMVSKGWRLLAPIT
AYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTR
TIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRG
DSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGLFRAAVCTRGVAKAVDFIPVENLETT
MRSPVFTDNSSPPAVPQSFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATL
GFGAYMSKAHGVDPNIRTGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHS
TDATSILGIGTVLDQAETAGARLVVLATATPPGSVTVSHPNIEEVALSTTGEIPFYGK
AIPLEVIKGGRHLIFCHSKKKCDELAAKLVALGINAVAYYRGLDVSVIPTSGDVVVVS
TDALMTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTTLPQDAVSRTQRRGRTGRG
KPGIYRFVAPGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYMNTPGLPVCQ
DHLEFWEGVFTGLTHIDAHFLSQTKQSGENFPYLVAYQATVCARAQAPPPSWDQMWKC
LIRLKPTLHGPTPLLYRLGAVQNEVTLTHPITKYIMTCMSADLEVVTSTWVLVGGVLA
ALAAYCLSTGCVVIVGRIVLSGKPAIIPDREVLYQEFDEMEECSQHLPYIEQGMMLAE
QFKQKALGLLQTASRQAEVITPAVQTNWQKLEVFWAKHMWNFISGIQYLAGLSTLPGN
PAIASLMAFTAAVTSPLTTGQTLLFNILGGWVAAQLAAPGAATAFVGAGLAGAAIGSV
GLGKVLVDILAGYGAGVAGALVAFKIMSGEVPSTEDLVNLLPAILSPGALVVGVVCAA
ILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPESDAAARVTAILSSLTVTQLLR
RLHQWISSECTTPCSGSWLRDIWDWICEVLSDFKTWLKAKLMPQLPGIPFVSCQRGYR
GVWRGDGIMHTRCHCGAEITGHVKNGTMRIVGPRTCRNMWSGTFPINAYTTGPCTPLP
APNYKFALWRVSAEEYVEIRRVGDFHYVSGMTTDNLKCPCQIPSPEFFTELDGVRLHR
FAPPCKPLLREEVSFRVGLHEYPVGSQLPCEPEPDVAVLTSMLTDPSHITAEAAGRRL
ARGSPPSMASSSASQLSAPSLKATCTANHDSPDAELIEANLLWRQEMGGNITRVESEN
KVVILDSFDPLVAEEDEREVSVPAEILRKSRRFARALPVWARPDYNPPLVETWKKPDY
EPPVVHGCPLPPPRSPPVPPPRKKRTVVLTESTLSTALAELATKSFGSSSTSGITGDN
TTTSSEPAPSGCPPDSDVESYSSMPPLEGEPGDPDLSDGSWSTVSSGADTEDVVCCSM
SYSWTGALVTPCAAEEQKLPINALSNSLLRHHNLVYSTTSRSACQRQKKVTFDRLQVL
DSHYQDVLKEVKAAASKVKANLLSVEEACSLTPPHSAKSKFGYGAKDVRCHARKAVAH
INSVWKDLLEDSVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALY
DVVSKLPLAVMGSSYGFQYSPGQRVEFLVQAWKSKKTPMGFSYDTRCFDSTVTESDIR
TEEAIYQCCDLDPQARVAIKSLTERLYVGGPLTNSRGENCGYRRCRASGVLTTSCGNT
LTCYIKARAACRAAGLQDCTMLVCGDDLVVICESAGVQEDAASLRAFTEAMTRYSAPP
GDPPQPEYDLELITSCSSNVSVAHDGAGKRVYYLTRDPTTPLARAAWETARHTPVNSW
LGNIIMFAPTLWARMILMTHFFSVLIARDQLEQALNCEIYGACYSIEPLDLPPIIQRL
HGLSAFSLHSYSPGEINRVAACLRKLGVPPLRAWRHRARSVRARLLSRGGRAAICGKY
LFNWAVRTKLKLTPIAAAGRLDLSGWFTAGYSGGDIYHSVSHARPRWFWFCLLLLAAG
VGIYLLPNR

SEQ ID NO:8: HCV core protein, H77, Accession AF009606
Genbank number: 2316097
>gi|2316098|gb|AAB66324.1| polyprotein [Hepatitis C virus subtype 1a]
MSTNPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRGRRQPIPKARR
PEGRTWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSWGPTDPRRRSRNLGKVIDTLTCGFADLMGYIPLVGAPLGGAAR
ALAHGVRVLEDGVNYATGNLPGCSFSIFLLALLSCLTVPASA

SEQ ID NO:9:
Hepatitis C virus mRNA, complete cds; ACCESSION M96362 M72423; Hepatitis C virus subtype 1b
MSTNPKPQRKTKRNTNRRPQDIKFPGGGQIVGGVYLLPRRGPRL
GVRATRKTSERSQPRGRRQPIPKARRPEGRAWAQPGYPWPLYGNEGLGWAGWLLSPRG
SRPSWGPTDPRRK**SRNLGKVIDTLT**CGFAD**LMGYIPLVGAPLG**G**VARALAHGVRV**LED
GV**NYATGNLPGCSFSIFLLAL**LSCLTTPVSAYEVRNASGMYHVTNDCSNSSIVYEAAD
MIMHTPGCVPCVREDNSSRCWVALTPTLAARNASVPTTTLRRHVDLLVGVAAFCSAMY
VGDLCGSVFLVSQLFTFSPRRHETVQDCNCSIYPGRVSGHRMAWDMMMNWSPTTALVV
SQLLRIPQAVVDMVTGSHWGILAGLAYYSMVGNWAKVLIAMLLFAGVDGTTHVTGGAQ
GRAASSLTSLFSPGPVQHLQLINTNGSWHINRTALSCNDSLNTGFVAALFYKYRFNAS
GCPERLATCRPIDTFAQGWGPITYTEPHDLDQRPYCWHYAPQPCGIVPTLQVCGPVYC
FTPSPVAVGTTDRFGAPTYRWGANETDVLLLNNAGPPQGNWFGCTWMNGTGFTKTCGG
PPCNIGGVGNNTLTCPTDCFRKHPGATYTKCGSGPWLTPRCLVDYPYRLWHYPCTVNF
TIFKVRMYVGGAEHRLDAACNWTRGERCDLEDRDRSELSPLLLSTTEWQVLPCSFTTL
PALSTGLIHLHQNIVDIQYLYGIGSAVVSFAIKWEYIVLLFLLLADARVCACLWMMLL
VAQAEAALENLVVLNAASVAGAHGILSFIVFFCAAWYIKGRLVPGAAYALYGVWPLLL
LLLALPPRAYAMDREMAASCGGAVFVGLVLLTLSPHYKVFLARFIWWLQYLITRTEAH
LQVWVPPLNVRGGRDAIILLTCVVHPELIFDITKYLLAIFGPLMVLQAGITRVPYFVR
AQGLIRACMLARKVVGGHYVQMVFMKLAALAGTYVYDHLTPLRDWAHTGLRDLAVAVE
PVVFSDMETKVITWGADTAACGDIILALPASARRGKEILLGPADSLEGQGWRLLAPIT
AYSQQTRGLLGCIITSLTGRDKNQVEGEVQVVSTATQSFLATCINGVCWTVFHGAGSK
TLAGPKGPITQMYTNVDQDLVGWPAPPGARSLTPCTCGSSDLYLVTRHADVIPVRRRG
DGRGSLLPPRPVSYLKGSSGGPLLCPSGHAVGILPAAVCTRGVAMAVEFIPVESMETT
MRSPVFTDNPSPPAVPQTFQVAHLHAPTGSGKSTRVPAAYAAQGYKVLVLNPSVAATL
GFGAYMSKAHGIDPNLRTGVRTITTGAPITYSTYGKFLADGGGSGGAYDIIMCDECHS
TDSTTIYGIGTVLDQAETAGARLVVLSTATPPGSVTVPHLNIEEVALSNTGEIPFYGK
AIPIEAIKGGRHLIFCHSKKKCDELAAKLSGLGLNAVAYYRGLDVSVIPTSGDVVVVA
TDALMTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVSRSQRRGRTGRG
RAGIYRFVTPGERPSGMFDSSVLCECYDAGCAWYELTPAETSVRLRAYLNTPGLPVCQ
DHLEFSEGVFTGLTHIDAHFLSQTKQAGENFPYLVAYQATVCARAQAPPPSWDEMWRC
LIRLKPTLHGPTPLLYRLGAVQNEVTLTHPITKFIMTCMSADLEVVTSTWVLVGGVLA
ALAAYCLTTGSVVIVGRIILSGKPAIIPDREVLYQEFDEMEECASHLPYFEQGMQLAE
QFKQKALGLLQTATKQAEAAAPVVESKWRALETFWAKHMWNFISGIQYLAGLSTLPGN
PAIRSPMAFTASITSPLTTQHTLLFNILGGWVAAQLAPPSAASAFVGAGIAGAAVGTI
GLGKVLVDILAGYGAGVAGALVAFKIMSGEMPSAEDMVNLLPAILSPGALVVGIVCAA
ILRRHVGPGEGAVQWMNRLIAFASRGNHVSPRHYVPESEPAARVTQILSSLTITQLLK
RLHQWINEDCSTPCSSSWLREIWDWICTVLTDFKTWLQSKLLPRLPGVPFFSCQRGYK
GVWRGDGIMHTTCPCGAQITGHVKNGSMRIVGPKTCSNTWYGTFPINAYTTGPCTPSP
APNYSKALWRVAAEEYVEVTRVGDFHYVTGMTTDNVKCPCQVPAPEFFTEVDGVRLHR
YAPACRPLLREEVVFQVGLHQYLVGSQLPCEPEPDVAVLTSMLTDPSHITAETAKRRL
ARGSPPSLASSSASQLSAPSLKATCTTHHDSPDADLIEANLLWRQEMGGNITRVESEN
KVVILDSFDPLRAEDDEGEISVPAEILRKSRKFPPALPIWAPPDYNPPLLESWKDPDY
VPPVVHGCPLPPTKAPPIPPPRRKRTVVLTESTVSSALAELATKTFGSSGSSAIDSGT
ATAPPDQASGDGDRESDVESFSSMPPLEGEPGDPDLSDGSWSTVSEEASEDVVCCSMS
YTWTGALITPCAAEESKLPINPLSNSLLRHHNMVYATTSRSAGLRQKKVTFDRLQVLD
DHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSAKSKFGYGAKDVRSLSSRAVTHI
RSVWKDLLEDTETPISTTIMAKNEVFCVQPEKGGRKPARLIVFPDLGVRVCEKMALYD
VVSTLPQAVMGSSYGFQYSPKQRVEFLVNTWKSKKCPMGFSYDTRCFDSTVTENDIRV
EESIYQCCDLAPEAKLAIKSLTERLYIGGPLTNSKGQNCGYRRCRASGVLTTSCGNTL
TCYLKATAACRAAKLRDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPG
DPPQPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWL
GNIIMYAPTLWARMILMTHFFSILLAQEQLEKTLDCQIYGACYSIEPLDLPQIIERLH
GLSAFSLHSYSPGEINRVASCLRKLGVPPLRAWRHRARSVRAKLLSQGGRAATCGKYL

FNWAVRTKLKLTPIPAASRLDLSGWFVAGYSGGDIYHSLSRARPRWFMLCLLLLSVGV
GIYLLPNR

SEQ ID NO:10, nucleocapsid protein of influenza A virus

```
  1 MASQGTKRSY EQMETSGERQ NATEIRASVG RMVGGIGRFY IQMCTELKLS DHEGRLIQNS
 61 ITIERMVLSA FDERRNKYLE EHPSAGKDPK KTGGPIYRRR DGKWMRELIL YDKEEIRRIW
121 RQANNGEDAT AGLTHMMIWH SNLNDATYQR TRALVRTGMD PRMCSLMQGS TLPRRSGAAG
181 AAVKGVGTMV MELIRMIKRG INDRNFWRGE NGRRTRIAYE RMCNILKGKF QTAAQRAMMD
241 QVRESRNPGN AEIEDLIFLA RSALILRGSV AHKSCLPACV YGLAVASGYD FEREGYSLVG
301 IDPFRLLQNS QVFSLIRPNE NPAHKSQLVW MACHSAAFED LRVSSFIRGT RVVPRGQLST
361 RGVQIASNEN METMDSSTLE LRSRYWAIRT RSGGNTNQQR ASAGQISVQP TFSVQRNLPF
421 ERATIMAAFT GNTEGRTSDM RTEIIRMMEN ARPEDVSFQG RGVFELSDEK ATNPIVPSFD
481 MSNEGS
```

SEQ ID NO:11

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

MSLLTEVETPIRNEWGCRCNDS**SDPLVVAASIIGILHLIL**WILDR**LFFKCVYRLFKHGL**KR**GPSTEGVPE** 70
SMREEYRKEQQNAVDADDSHFVSIELE

SEQ ID NO:12

>gi|73919147|ref|YP_308843.1| nucleocapsid protein [Influenza A virus (A/New York/392/2004(**H3N2**)**)**]

```
MASQGTKRSYEQMETDGDRQNATEIRASVGKMIDGIGRFYIQMCTELKLSDHEGRLIQNSLTIEKMVLSA 70
FDERRNKYLEEHPSAGKDPKKTGGPIYRRVDGKWMRELVLYDKEEIRRIWRQANNGEDATAGLTHIMIWH 140
SNLNDATYQRTRALVRTGMDPRMCSLMQGSTLPRRSGAAGAAVKGIGTMVMELIRMVKRGINDRNFWRGE 210
NGRKTR
```
SAYERMCNILKGK**FQTAAQRAMV**DQVRESRNPGNAEIED**LIFLARSALILRGSVAHK**SCLPACA 280
```
YGPAVSSGYDFEKEGYSLVGIDPFKLLQNSQIYSLIRPNENPAHKSQLVWMACHSAAFEDLRLLSFIRGT 350
KVSPRGKLSTRGVQIASNENMDNMGSSTLELRSGYWAIRTRSGGNTNQQRASAGQTSVQPTFSVQRNLPF 420
EKSTIMAAFTGNTEGRTSDMRAEIIRMMEGAKPEEVSFRGRGVFELSDEKATNPIVPSFDMSNEGSYFFG 490
DNAEEYDN
```
--

SEQ ID NO:13

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

MSLLTEVETPIRNEWGCRCNGS**SDPLAIAANIIGILHLIL**WILDR**LFFKCIYRRFKYGL**KG**GPSTEGVPK**
SMREEYRKEQQSAVDADDGHFVSIELE

SEQ ID NO:14

>gi|8486130|ref|NP_040982.1| nucleocapsid protein [Influenza A virus (A/Puerto Rico/8/34**(H1N1)**)]

```
MASQGTKRSYEQMETDGERQNATEIRASVGKMIGGIGRFYIQMCTELKLSDYEGRLIQNSLTIERMVLSA
FDERRNKYLEEHPSAGKDPKKTGGPIYRRVNGKWMRELILYDKEEIRRIWRQANNGDDATAGLTHMMIWH
SNLNDATYQRTRALVRTGMDPRMCSLMQGSTLPRRSGAAGAAVKGVGTMVMELVRMIKRGINDRNFWRGE
NGRKTR
```
IAYERMCNILKGK**FQTAAQKAMM**DQVRESRDPGNAEFED**LTFLARSALILRGSVAHKS**CLPACV
```
YGPAVASGYDFEREGYSLVGIDPFRLLQNSQVYSLIRPNENPAHKSQLVWMACHSAAFEDLRVLSFIKGT
KVVPRGKLSTRGVQIASNENMETMESSTLELRSRYWAIRTRSGGNTNQQRASAGQISIQPTFSVQRNLPF
DRTTVMAAFTGNTEGRTSDMRTEIIRMMESARPEDVSFQGRGVFELSDEKAASPIVPSFDMSNEGSYFFG
DNAEEYDN
```
--

SEQ ID NO:15
>gi|73912687|ref|YP_308853.1| membrane protein M2 [Influenza A virus (A/Korea/426/68**(H2N2)**)]
MSLLTEVETPIRNEWGCRCNDS**SDPLVVAASIIGILHFIL**WILDR**LFFKCIYRFFKHGL**KR**GPSTEGVPE**
SMREEYRKEQQSAVDADDSHFVSIELE

SEQ ID NO:16
>gi|73921307|ref|YP_308871.1| nucleoprotein [Influenza A virus (A/Korea/426/68**(H2N2)**)]
MASQGTKRSYEQMETDGERQNATEIRASVGKMIDGIGRFYIQMCTELKLSDYEGRLIQNSLTIERMVLSA
FDERRNKYLEEHPSAGKDPKKTGGPIYKRVDGKWMRELVLYDKEEIRRIWRQANNGDDATAGLTHMMIWH
SNLNDTTYQRTRALVRTGMDPRMCSLMQGSTLPRRSGAAGAAVKGVGTMVMELIRMIKRGINDRNFWRGE
NGRKTR**SAYERMCNIL**KGK**FQTAAQRAMM**DQVRESRNPGNAEIED**LIFLARSALILRGSVAHKS**CLPACV
YGPAIASGYNFEKEGYSLVGIDPFKLLQNSQVYSLIRPNENPAHKSQLVWMACNSAAFEDLRVLSFIRGT
KVSPRGKLSTRGVQIASNENMDTMESSTLELRSRYWAIRTRSGGNTNQQRASAGQISVQPAFSVQRNLPF
DKPTIMAAFTGNTEGRTSDMRAEIIRMMEGAKPEEMSFQGRGVFELSDEKATNPIVPSFDMSNEGSYFFG
DNAEEYDN

SEQ ID NO:17
>gi|330647|gb|AAA45994.1| pp65 [Human herpesvirus 5]
MASVLGPISGHVLKAVFSRGDTPVLPHETRLLQTGIHVRVSQPSLILVSQYTPDSTPCHRGDNQLQVQHT 70
YFTGSEVENVSVNVHNPTGRSICPSQEPMSIYVYALPLKMLNIPSINVHHYPSAAERKHRHLPVADAVIH 140
ASGKQMWQARLTVSGLAWTRQQNQWKEPDVYYTSAFVFPTKDVALRHVVCAHELVCSMENTRATKMQVIG 210
DQYVKVYLESFCEDVPSGKLFMHVTLGSDVEEDLTMTRNPQPFMRPHERNGFTVLCPKNMIIKPGKISHI 280
MLDVAFTSHEHFGLLCPKSIPGLSISGNLLMNGQQIFLEVQAIRETVELRQYDPVAALFFFDIDLLLQRG 350
PQYSEHPTFTSQYRIQGKLEYRHTWDRHDEGAAQGDDDVWTSGSDSDEELVTTERKTPRVTGGGAMAGAS 420
TSAGRKRKSASSATACTAGVMTRGRLKAESTVAPEEDTDEDSDNEIHNPAVFTWPPWQAGILAR**NLVPMV** 490
ATVQGQNLKYQEFFWDANDIYRIFAELEGVWQPAAQPKRRRHRQDALPGPCIASTPKKHRG 541

SEQ ID NO:18
>gi|33330937|gb|AAQ10712.1| putative transforming protein E6 [Human papillomavirus type 16]
MHQKRTAMFQDPQERPGKLPQLCTELQTTIHDIILECVYCKQQLLRRE**VYDFAFRDLCIVY**RDGNPYAVC 70
DKCLKFYSKISEYRHYCYSVYGTTLEQQYNKPLCDLLIRCINCQKPLCPEEKQRHLDKKQRFHNIRGRWT 140
GRCMSCCRSSRTRRETQL

SEQ ID NO:19
>gi|56583270|ref|NP_040979.2| matrix protein 2 [Influenza A virus (A/Puerto Rico/8/34(H1N1))]
MSLLTEVETPIRNEWGCRCNGSSDPLAIAANIIGILHLILWILDRLFFKCIYRRFKYGLKGGPSTEGVPK
SMREEYRKEQQSAVDADDGHFVSIELE

SEQ ID NO:20
>gi|8486139|ref|NP_040987.1| PB2 protein [Influenza A virus (A/Puerto Rico/8/34(H1N1))]
MERIKELRNLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPALRMKWMMAMKYPITADKRITEMIPER
NEQGQTLWSKMNDAGSDRVMVSPLAVTWWNRNGPMTNTVHYPKIYKTYFERVERLKHGTFGPVHFRNQVK
IRRRVDINPGHADLSAKEAQDVIMEVVFPNEVGARILTSESQLTITKEKKEELQDCKISPLMVAYMLERE
LVRKTRFLPVAGGTSSVYIEVLHLTQGTCWEQMYTPGGEVKNDDVDQSLIIAARNIVRRAAVSADPLASL
LEMCHSTQIGGIRMVDILKQNPTEEQAVGICKAAMGLRISSSFSFGGFTFKRTSGSSVKREEEVLTGNLQ
TLKIRVHEGYEEFTMVGRRATAILRKATRRLIQLIVSGRDEQSIAEAIIVAMVFSQEDCMIKAVRGDLNF
VNRANQRLNPMHQLLRHFQKDAKVLFQNWGVEPIDNVMGMIGILPDMTPSIEMSMRGVRISKMGVDEYSS
TERVVVSIDRFLRVRDQRGNVLLSPEEVSETQGTEKLTITYSSSMMWEINGPESVLVNTYQWIIRNWETV
KIQWSQNPTMLYNKMEFEPFQSLVPKAIRGQYSGFVRTLFQQMRDVLGTFDTAQIIKLLPFAAAPPKQSR
MQFSSFTVNVRGSGMRILVRGNSPVFNYNKATKRLTVLGKDAGTLTEDPDEGTAGVESAVLRGFLILGKE
DRRYGPALSINELSNLAKGEKANVLIGQGDVVLVMKRKRDSSILTDSQTATKRIRMAIN

SEQ ID NO:21
>gi|8486137|ref|NP_040986.1| polymerase PA [Influenza A virus (A/Puerto Rico/8/34(H1N1))]
MEDFVRQCFNPMIVELAEKTMKEYGEDLKIETNKFAAICTHLEVCFMYSDFHFINEQGESIIVELGDPNA
LLKHRFEIIEGRDRTMAWTVVNSICNTTGAEKPKFLPDLYDYKENRFIEIGVTRREVHIYYLEKANKIKS
EKTHIHIFSFTGEEMATKADYTLDEESRARIKTRLFTIRQEMASRGLWDSFRQSERGEETIEERFEITGT
MRKLADQSLPPNFSSLENFRAYVDGFEPNGYIEGKLSQMSKEVNARIEPFLKTTPRPLRLPNGPPCSQRS
KFLLMDALKLSIEDPSHEGEGIPLYDAIKCMRTFFGWKEPNVVKPHEKGINPNYLLSWKQVLAELQDIEN
EEKIPKTKNMKKTSQLKWALGENMAPEKVDFDDCKDVGDLKQYDSDEPELRSLASWIQNEFNKACELTDS
SWIELDEIGEDVAPIEHIASMRRNYFTSEVSHCRATEYIMKGVYINTALLNASCAAMDDFQLIPMISKCR
TKEGRRKTNLYGFIIKGRSHLRNDTDVVNFVSMEFSLTDPRLEPHKWEKYCVLEIGDMLLRSAIGQVSRP
MFLYVRTNGTSKIKMKWGMEMRRCLLQSLQQIESMIEAESSVKEKDMTKEFFENKSETWPIGESPKGVEE
SSIGKVCRTLLAKSVFNSLYASPQLEGFSAESRKLLLIVQALRDNLEPGTFDLGGLYEAIEECLINDPWV
LLNASWFNSFLTHALS

SEQ ID NO:22
>gi|8486133|ref|NP_040984.1| nonstructural protein NS1 [Influenza A virus (A/Puerto Rico/8/34(H1N1))]
MDPNTVSSFQVDCFLWHVRKRVADQELGDAPFLDRLRRDQKSLRGRGSTLGLDIETATRAGKQIVERILK
EESDEALKMTMASVPASRYLTDMTLEEMSREWSMLIPKQKVAGPLCIRMDQAIMDKNIILKANFSVIFDR
LETLILLRAFTEEGAIVGEISPLPSLPGHTAEDVKNAVGVLIGGLEWNDNTVRVSETLQRFAWRSSNENG
RPPLTPKQKREMAGTIRSEV

SEQ ID NO:23
>gi|8486132|ref|NP_040983.1| nonstructural protein NS2 [Influenza A virus (A/Puerto Rico/8/34(H1N1))]
MDPNTVSSFQDILLRMSKMQLESSSEDLNGMITQFESLKLYRDSLGEAVMRMGDLHSLQNRNEKWREQLG
QKFEEIRWLIEEVRHKLKVTENSFEQITFMQALHLLLEVEQEIRTFSFQLI

SEQ ID NO:24
>gi|8486128|ref|NP_040981.1| neuraminidase [Influenza A virus (A/Puerto Rico/8/34(H1N1))]
MNPNQKIITIGSICLVVGLISLILQIGNIISIWISHSIQTGSQNHTGICNQNIITYKNSTWVKDTTSVIL
TGNSSLCPIRGWAIYSKDNSIRIGSKGDVFVIREPFISCSHLECRTFFLTQGALLNDRHSNGTVKDRSPY
RALMSCPVGEAPSPYNSRFESVAWSASACHDGMGWLTIGISGPDNGAVAVLKYNGIITETIKSWRKKILR
TQESECACVNGSCFTIMTDGPSDGLASYKIFKIEKGKVTKSIELNAPNSHYEECSCYPDTGKVMCVCRDN
WHGSNRPWVSFDQNLDYQIGYICSGVFGDNPRPKDGTGSCGPVYVDGANGVKGFSYRYGNGVWIGRTKSH
SSRHGFEMIWDPNGWTETDSKFSVRQDVVAMTDWSGYSGSFVQHPELTGLDCIRPCFWVELIRGRPKEKT
IWTSASSISFCGVNSDTVDWSWPDGAELPFTIDK

SEQ ID NO:25
>gi|8486126|ref|NP_040980.1| haemagglutinin [Influenza A virus (A/Puerto Rico/8/34(H1N1))]
MKANLLVLLCALAAADADTICIGYHANNSTDTVDTVLEKNVTVTHSVNLLEDSHNGKLCRLKGIAPLQLG
KCNIAGWLLGNPECDPLLPVRSWSYIVETPNSENGICYPGDFIDYEELREQLSSVSSFERFEIFPKESSW
PNHNTTKGVTAACSHAGKSSFYRNLLWLTEKEGSYPKLKNSYVNKKGKEVLVLWGIHHPSNSKDQQNIYQ
NENAYVSVVTSNYNRRFTPEIAERPKVRDQAGRMNYYWTLLKPGDTIIFEANGNLIAPRYAFALSRGFGS
GIITSNASMHECNTKCQTPLGAINSSLPFQNIHPVTIGECPKYVRSAKLRMVTGLRNIPSIQSRGLFGAI
AGFIEGGWTGMIDGWYGYHHQNEQGSGYAADQKSTQNAINGITNKVNSVIEKMNIQFTAVGKEFNKLEKR
MENLNKKVDDGFLDIWTYNAELLVLLENERTLDFHDSNVKNLYEKVKSQLKNNAKEIGNGCFEFYHKCDN
ECMESVRNGTYDYPKYSEESKLNREKVDGVKLESMGIYQILAIYSTVASSLVLLVSLGAISFWMCSNGSL
QCRICI

SEQ ID NO:26
>gi|8486123|ref|NP_040978.1| matrix protein 1 [Influenza A virus (A/Puerto Rico/8/34(H1N1))]
MSLLTEVETYVLSIIPSGPLKAEIAQRLEDVFAGKNTDLEVLMEWLKTRPILSPLTKGILGFVFTLTVPS
ERGLQRRRFVQNALNGNGDPNNMDKAVKLYRKLKREITFHGAKEISLSYSAGALASCMGLIYNRMGAVTT
EVAFGLVCATCEQIADSQHRSHRQMVTTTNPLIRHENRMVLASTTAKAMEQMAGSSEQAAEAMEVASQAR

QMVQAMRTIGTHPSSSAGLKNDLLENLQAYQKRMGVQMQRFK

SEQ ID NO:27
>gi|83031685|ref|YP_418248.1| PB1-F2 protein [Influenza A virus (A/Puerto Rico/8/34(H1N1))]
MGQEQDTPWILSTGHISTQKRQDGQQTPKLEHRNSTRLMGHCQKTMNQVVMPKQIVYWKQWLSLRNPILV FLKTRVLKRWRLFSKHE

SEQ ID NO:28
>gi|8486135|ref|NP_040985.1| polymerase 1 PB1 [Influenza A virus (A/Puerto Rico/8/34(H1N1))]
MDVNPTLLFLKVPAQNAISTTFPYTGDPPYSHGTGTGYTMDTVNRTHQYSEKARWTTNTETGAPQLNPID GPLPEDNEPSGYAQTDCVLEAMAFLEESHPGIFENSCIETMEVVQQTRVDKLTQGRQTYDWTLNRNQPAA TALANTIEVFRSNGLTANESGRLIDFLKDVMESMKKEEMGITTHFQRKRRVRDNMTKKMITQRTIGKRKQ RLNKRSYLIRALTLNTMTKDAERGKLKRRAIATPGMQIRGFVYFVETLARSICEKLEQSGLPVGGNEKKA KLANVVRKMMTNSQDTELSLTITGDNTKWNENQNPRMFLAMITYMTRNQPEWFRNVLSIAPIMFSNKMAR LGKGYMFESKSMKLRTQIPAEMLASIDLKYFNDSTRKKIEKIRPLLIEGTASLSPGMMMGMFNMLSTVLG VSILNLGQKRYTKTTYWWDGLQSSDDFALIVNAPNHEGIQAGVDRFYRTCKLHGINMSKKKSYINRTGTF EFTSFFYRYGFVANFSMELPSFGVSGSNESADMSIGVTVIKNNMINNDLGPATAQMALQLFIKDYRYTYR CHRGDTQIQTRRSFEIKKLWEQTRSKAGLLVSDGGPNLYNIRNLHIPEVCLKWELMDEDYQGRLCNPLNP FVSHKEIESMNNAVMMPAHGPAKNMEYDAVATTHSWIPKRNRSILNTSQRGVLEDEQMYQRCCNLFEKFF PSSSYRRPVGISSMVEAMVSRARIDARIDFESGRIKKEEFTEIMKICSTIEELRRQK

SEQ ID NO:29
>gi|8486130|ref|NP_040982.1| nucleocapsid protein [Influenza A virus (A/Puerto Rico/8/34(H1N1))]
MASQGTKRSYEQMETDGERQNATEIRASVGKMIGGIGRFYIQMCTELKLSDYEGRLIQNSLTIERMVLSA FDERRNKYLEEHPSAGKDPKKTGGPIYRRVNGKWMRELILYDKEEIRRIWRQANNGDDATAGLTHMMIWH SNLNDATYQRTRALVRTGMDPRMCSLMQGSTLPRRSGAAGAAVKGVGTMVMELVRMIKRGINDRNFWRGE NGRKTRIAYERMCNILKGKFQTAAQKAMMDQVRESRDPGNAEFEDLTFLARSALILRGSVAHKSCLPACV YGPAVASGYDFEREGYSLVGIDPFRLLQNSQVYSLIRPNENPAHKSQLVWMACHSAAFEDLRVLSFIKGT KVVPRGKLSTRGVQIASNENMETMESSTLELRSRYWAIRTRSGGNTNQQRASAGQISIQPTFSVQRNLPF DRTTVMAAFTGNTEGRTSDMRTEIIRMMESARPEDVSFQGRGVFELSDEKAASPIVPSFDMSNEGSYFFG DNAEEYDN

SEQ ID NO:30
>gi|73918826|ref|YP_308855.1| polymerase 2 [Influenza A virus (A/Korea/426/1968(H2N2))]
MERIKELRNLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPSLRMKWMMAMKYPITADKRITEMVPER NEQGQTLWSKMSDAGSDRVMVSPLAVTWWNRNGPMTSTVHYPKIYKTYFEKVERLKHGTFGPVHFRNQVK IRRRVDINPGHADLSAKEAQDVIMEVVFPNEVGARILTSESQLTITKEKKEELQDCKISPLMVAYMLERE LVRKTRFLPVAGGTSSVYIEVLHLTQGTCWEQMYTPGGEVRNDDVDQSLIIAARNIVRRAAVSADPLASL LEMCHSTQIGGTRMVDILRQNPTEEQAVDICKAAMGLRISSSFSFGGFTFKRTSGSSIKREEEVLTGNLQ TLKIRVHEGYEEFTMVGKRATAILRKATRRLVQLIVSGRDEQSIAEAIIVAMVFSQEDCMIKAVRGDLNF VNRANQRLNPMHQLLRHFQKDAKVLFQNWGIEHIDNVMGMIGVLPDMTPSTEMSMRGIRVSKMGVDEYSS TERVVVSIDRFLRVRDQRGNVLLSPEEVSETQGTEKLTITYSSSMMWEINGPESVLVNTYQWIIRNWETV KIQWSQNPTMLYNKMEFEPFQSLVPKAIRGQYSGFVRTLFQQMRDVLGTFDTTQIIKLLPFAAAPPKQSR MQFSSLTVNVRGSGMRILVRGNSPVFNYNKTTKRLTILGKDAGTLTEDPDEGTSGVESAVLRGFLILGKE DRRYGPALSINELSTLAKGEKANVLIGQGDVVLVMKRKRDSSILTDSQTATKRIRMAIN

SEQ ID NO:31
>gi|73919145|ref|YP_308850.1| hemagglutinin [Influenza A virus (A/Korea/426/68(H2N2))]
MAIIYLILLFTAVRGDQICIGYHANNSTEKVDTILERNVTVTHAKDILEKTHNGKLCKLNGIPPLELGDC SIAGWLLGNPECDRLLSVPEWSYIMEKENPRYSLCYPGSFNDYEELKHLLSSVKHFEKVKILPKDRWTQH TTTGGSWACAVSGKPSFFRNMVWLTRKGSNYPVAKGSYNNTSGEQMLIIWGVHHPNDEAEQRALYQNVGT

YVSVATSTLYKRSIPEIAARPKVNGLGRRMEFSWTLLDMWDTINFESTGNLVAPEYGFKISKRGSSGIMK
TEGTLENCETKCQTPLGAINTTLPFHNVHPLTIGECPKYVKSEKLVLATGLRNVPQIESRGLFGAIAGFI
EGGWQGMVDGWYGYHHSNDQGSGYAADKESTQKAFNGITNKVNSVIEKMNTQFEAVGKEFSNLEKRLENL
NKKMEDGFLDVWTYNAELLVLMENERTLDFHDSNVKNLYDKVRMQLRDNVKELGNGCFEFYHKCDNECMD
SVKNGTYDYPKYEEESKLNRNEIKGVKLSSMGVYQILAIYATVAGSLSLAIMMAGISFWMCSNGSLQCRI
CI

SEQ ID NO:32
>gi|73912688|ref|YP_308854.1| membrane protein M1 [Influenza A virus
(A/Korea/426/68(H2N2))]
MSLLTEVETYVLSIVPSGPLKAEIAQRLEDVFAGKNTDLEALMEWLKTRPILSPLTKGILGFVFTLTVPS
ERGLQRRRFVQNALNGNGDPNNMDRAVKLYRKLKREITFHGAKEVALSYSAGALASCMGLIYNRMGAVTT
EVAFAVVCATCEQIADSQHRSHRQMVTTTNPLIRHENRMVLASTTAKAMEQMAGSSEQAAEAMEVASQAR
QMVQAMRAIGTPPSSSAGLKDDLLENLQAYQKRMGVQMQRFK

SEQ ID NO:33
>gi|73912687|ref|YP_308853.1| membrane protein M2 [Influenza A virus
(A/Korea/426/68(H2N2))]
MSLLTEVETPIRNEWGCRCNDSSDPLVVAASIIGILHFILWILDRLFFKCIYRFFKHGLKRGPSTEGVPE
SMREEYRKEQQSAVDADDSHFVSIELE

SEQ ID NO:34
>gi|73912685|ref|YP_308852.1| polymerase PA [Influenza A virus
(A/Korea/426/68(H2N2))]
MEDFVRQCFNPMIVELAEKAMKEYGEDLKIETNKFAAICTHLEVCFMYSDFHFINEQGESIMVELDDPNA
LLKHRFEIIEGRDRTMAWTVVNSICNTTGAEKPKFLPDLYDYKENRFIEIGVTRREVHIYYLEKANKIKS
ENTHIHIFSFTGEEMATKADYTLDEESRARIKTRLFTIRQEMANRGLWDSFRQSERGEETIEERFEITGT
MRRLADQSLPPNFSCLENFRAYVDGFEPNGYIEGKLSQMSKEVNAKIEPFLKTTPRPIRLPDGPPCFQRS
KFLLMDALKLSIEDPSHEGEGIPLYDAIKCMRTFFGWKEPYIVKPHEKGINPNYLLSWKQVLAELQDIEN
EEKIPRTKNMKKTSQLKWALGENMAPEKVDFDNCRDISDLKQYDSDEPELRSLSSWIQNEFNKACELTDS
IWIELDEIGEDVAPIEHIASMRRNYFTAEVSHCRATEYIMKGVYINTALLNASCAAMDDFQLIPMISKCR
TKEGRRKTNLYGFIIKGRSHLRNDTDVVNFVSMEFSLTDPRLEPHKWEKYCVLEIGDMLLRSAIGQMSRP
MFLYVRTNGTSKIKMKWGMEMRPCLLQSLQQIESMVEAESSVKEKDMTKEFFENKSETWPIGESPKGVEE
GSIGKVCRTLLAKSVFNSLYASPQLEGFSAESRKLLLVVQALRDNLEPGTFDLGGLYEAIEECLINDPWV
LLNASWFNSFLTHALR

SEQ ID NO:35
>gi|73921833|ref|YP_308877.1| PB1-F2 protein [Influenza A virus
(A/Korea/426/68(H2N2))]
MGQEQDTPWTQSTEHINIQKRGSGQQTRKLERPNLTQLMDHYLRTMNQVDMHKQTASWKQWLSLRNHTQE
SLKIRVLKRWKLFNKQEWTN

SEQ ID NO:36
>gi|73912683|ref|YP_308851.1| PB1 polymerase subunit [Influenza A virus
(A/Korea/426/68(H2N2))]
MDVNPTLLFLKVPAQNAISTTFPYTGDPPYSHGTGTGYTMDTVNRTHQYSEKGKWTTNTETGAPQLNPID
GPLPEDNEPSGYAQTDCVLEAMAFLEESHPGIFENSCLETMEVIQQTRVDKLTQGRQTYDWTLNRNQPAA
TALANTIEVFRSNGLTANESGRLIDFLKDVIESMDKEEMEITTHFQRKRRVRDNMTKKMVTQRTIGKKKQ
RLNKRSYLIRALTLNTMTKDAERGKLKRRAIATPGMQIRGFVHFVETLARNICEKLEQSGLPVGGNEKKA
KLANVVRKMMTNSQDTELSFTITGDNTKWNENQNPRVFLAMITYITRNQPEWFRNVLSIAPIMFSNKMAR
LGKGYMFESKSMKLRTQIPAEMLASIDLKYFNESTRKKIEKIRPLLIDGTVSLSPGMMMGMFNMLSTVLG
VSILNLGQKKYTKTTYWWDGLQSSDDFALIVNAPNHEGIQAGVNRFYRTCKLVGINMSKKKSYINRTGTF
EFTSFFYRYGFVANFSMELPSFGVSGINESADMSIGVTVIKNNMINNDLGPATAQMALQLFIKDYRYTYR
CHRGDTQIQTRRSFELKKLWEQTRSKAGLLVSDGGSNLYNIRNLHIPEVCLKWELMDEDYQGRLCNPLNP
FVSHKEIESVNNAVVMPAHGPAKSMEYDAVATTHSWTPKRNRSILNTSQRGILEDEQMYQKCCNLFEKFF
PSSSYRRPVGISSMVEAMVSRARIDARIDFESGRIKKEEFAEIMKICSTIEELRRQK

SEQ ID NO:37

>gi|73921567|ref|YP_308869.1| non-structural protein NS2 [Influenza A virus (A/Korea/426/68(H2N2))]
MDSNTVSSFQDILLRMSKMQLGSSSEDLNGMITQFESLKLYRDSLGEAVMRMGDLHSLQNRNGKWREQLG
QKFEEIRWLIEEVRHRLKITENSFEQITFMQALQLLFEVEQEIRTFSFQLI

SEQ ID NO:38
>gi|73921566|ref|YP_308870.1| non-structural protein NS1 [Influenza A virus (A/Korea/426/68(H2N2))]
MDSNTVSSFQVDCFLWHVRKQVVDQELGDAPFLDRLRRDQKSLRGRGSTLDLDIEAATRVGKQIVERILK
EESDEALKMTMASAPASRYLTDMTIEELSRDWFMLMPKQKVEGPLCIRIDQAIMDKNIMLKANFSVIFDR
LETLILLRAFTEEGAIVGEISPLPSLPGHTIEDVKNAIGVLIGGLEWNDNTVRVSKTLQRFAWRSSNENG
RPPLTPKQKRKMARTIRSKVRRDKMAD

SEQ ID NO:39
>gi|73921307|ref|YP_308871.1| nucleoprotein [Influenza A virus (A/Korea/426/68(H2N2))]
MASQGTKRSYEQMETDGERQNATEIRASVGKMIDGIGRFYIQMCTELKLSDYEGRLIQNSLTIERMVLSA
FDERRNKYLEEHPSAGKDPKKTGGPIYKRVDGKWMRELVLYDKEEIRRIWRQANNGDDATAGLTHMMIWH
SNLNDTTYQRTRALVRTGMDPRMCSLMQGSTLPRRSGAAGAAVKGVGTMVMELIRMIKRGINDRNFWRGE
NGRKTRSAYERMCNILKGKFQTAAQRAMMDQVRESRNPGNAEIEDLIFLARSALILRGSVAHKSCLPACV
YGPAIASGYNFEKEGYSLVGIDPFKLLQNSQVYSLIRPNENPAHKSQLVWMACNSAAFEDLRVLSFIRGT
KVSPRGKLSTRGVQIASNENMDTMESSTLELRSRYWAIRTRSGGNTNQQRASAGQISVQPAFSVQRNLPF
DKPTIMAAFTGNTEGRTSDMRAEIIRMMEGAKPEEMSFQGRGVFELSDEKATNPIVPSFDMSNEGSYFFG
DNAEEYDN

SEQ ID NO:40
>gi|73921304|ref|YP_308872.1| neuraminidase [Influenza A virus (A/Korea/426/68(H2N2))]
MNPNQKIITIGSVSLTIATVCFLMQIAILVTTVTLHFKQHECDSPASNQVMPCEPIIIERNITEIVYLNN
TTIEKEICPEVVEYRNWSKPQCQITGFAPFSKDNSIRLSAGGDIWVTREPYVSCDPGKCYQFALGQGTTL
DNKHSNDTIHDRIPHRTLLMNELGVPFHLGTRQVCVAWSSSSCHDGKAWLHVCVTGDDKNATASFIYDGR
LMDSIGSWSQNILRTQESECVCINGTCTVVMTDGSASGRADTRILFIEEGKIVHISPLSGSAQHVEECSC
YPRYPDVRCICRDNWKGSNRPVIDINMEDYSIDSSYVCSGLVGDTPRNDDRSSNSNCRNPNNERGNPGVK
GWAFDNGDDVWMGRTISKDLRSGYETFKVIGGWSTPNSKSQINRQVIVDSNNWSGYSGIFSVEGKRCINR
CFYVELIRGRQQETRVWWTSNSIVVFCGTSGTYGTGSWPDGANINFMPI

SEQ ID NO:41
>gi|73919213|ref|YP_308844.1| nonstructural protein 2 [Influenza A virus (A/New York/392/2004(H3N2))]
MDSNTVSSFQDILLRMSKMQLGSSSEDLNGMITQFESLKIYRDSLGEAVMRMGDLHLLQNRNGKWREQLG
QKFEEIRWLIEEVRHRLKTTENSFEQITFMQALQLLFEVEQEIRTFSFQLI

SEQ ID NO:42
>gi|73919212|ref|YP_308845.1| nonstructural protein 1 [Influenza A virus (A/New York/392/2004(H3N2))]
MDSNTVSSFQVDCFLWHIRKQVVDQELSDAPFLDRLRRDQRSLRGRGNTLGLDIKAATHVGKQIVEKILK
EESDEALKMTMVSTPASRYITDMTIEELSRNWFMLMPKQKVEGPLCIRMDQAIMEKNIMLKANFSVIFDR
LETIVLLRAFTEEGAIVGEISPLPSFPGHTIEDVKNAIGVLIGGLEWNDNTVRVSKNLQRFAWRSSNENG
GPPLTPKQKRKMARTARSKV

SEQ ID NO:43
>gi|73919207|ref|YP_308839.1| hemagglutinin [Influenza A virus (A/New York/392/2004(H3N2))]
MKTIIALSYILCLVFAQKLPGNDNSTATLCLGHHAVPNGTIVKTITNDQIEVTNATELVQSSSTGGICDS
PHQILDGENCTLIDALLGDPQCDGFQNKKWDLFVERSKAYSNCYPYDVPDYASLRSLVASSGTLEFNNES
FNWTGVTQNGTSSACKRRSNNSFFSRLNWLTHLKFKYPALNVTMPNNEKFDKLYIWGVHHPGTDNDQISL
YAQASGRITVSTKRSQQTVIPSIGSRPRIRDVPSRISIYWTIVKPGDILLINSTGNLIAPRGYFKIRSGK
SSIMRSDAPIGKCNSECITPNGSIPNDKPFQNVNRITYGACPRYVKQNTLKLATGMRNVPEKQTRGIFGA

IAGFIENGWEGMVDGWYGFRHQNSEGTGQAADLKSTQAAINQINGKLNRLIGKTNEKFHQIEKEFSEVEG
RIQDLEKYVEDTKIDLWSYNAELLVALENQHTIDLTDSEMNKLFERTKKQLRENAEDMGNGCFKIYHKCD
NACIGSIRNGTYDHDVYRDEALNNRFQIKGVELKSGYKDWILWISFAISCFLLCVALLGFIMWACQKGNI
RCNICI

SEQ ID NO:44
>gi|73919153|ref|YP_308840.1| matrix protein 2 [Influenza A virus (A/New York/392/2004(H3N2))]
MSLLTEVETPIRNEWGCRCNDSSDPLVVAASIIGILHLILWILDRLFFKCVYRLFKHGLKRGPSTEGVPE
SMREEYRKEQQNAVDADDSHFVSIELE

SEQ ID NO:45
>gi|73919152|ref|YP_308841.1| matrix protein 1 [Influenza A virus (A/New York/392/2004(H3N2))]
MSLLTEVETYVLSIVPSGPLKAEIAQRLEDVFAGKNTDLEALMEWLKTRPILSPLTKGILGFVFTLTVPS
ERGLQRRRFVQNALNGNGDPNNMDKAVKLYRKLKREITFHGAKEIALSYSAGALASCMGLIYNRMGAVTT
EVAFGLVCATCEQIADSQHRSHRQMVATTNPLIKHENRMVLASTTAKAMEQMAGSSEQAAEAMEIASQAR
QMVQAMRAVGTHPSSSTGLRDDLLENLQTYQKRMGVQMQRFK

SEQ ID NO:46
>gi|73919150|ref|YP_308848.1| PB1-F2 protein [Influenza A virus (A/New York/392/2004(H3N2))]
MEQEQDTPWTQSTEHTNIQRRGSGRQIQKLGHPNSTQLMDHYLRIMSQVDMHKQTVSWRLWPSLKNPTQV
SLRTHALKQWKSFNKQGWTN

SEQ ID NO:47
>gi|73919149|ref|YP_308847.1| polymerase PB1 [Influenza A virus (A/New York/392/2004(H3N2))]
MDVNPTLLFLKVPAQNAISTTFPYTGDPPYSHGTGTGYTMDTVNRTHQYSEKGKWTTNTETGAPQLNPID
GPLPEDNEPSGYAQTDCVLEAMAFLEESHPGIFENSCLETMEVVQQTRVDKLTQGRQTYDWTLNRNQPAA
TALANTIEVFRSNGLTANESGRLIDFLKDVMESMDKEEMEITTHFQRKRRVRDNMTKKMVTQRTIGKKKQ
RVNKRGYLIRALTLNTMTKDAERGKLKRRAIATPGMQIRGFVYFVETLARSICEKLEQSGLPVGGNEKKA
KLANVVRKMMTNSQDTELSFTITGDNTKWNENQNPRMFLAMITYITKNQPEWFRNILSIAPIMFSNKMAR
LGKGYMFESKRMKLRTQIPAEMLASIDLKYFNESTRKKIEKIRPLLIDGTASLSPGMMMGMFNMLSTVLG
VSVLNLGQKKYTKTTYWWDGLQSSDDFALIVNAPNHEGIQAGVDRFYRTCKLVGINMSKKKSYINKTGTF
EFTSFFYRYGFVANFSMELPSFGVSGINESADMSIGVTVIKNNMINNDLGPATAQMALQLFIKDYRYTYR
CHRGDTQIQTRRSFELKKLWDQTQSRAGLLVSDGGPNLYNIRNLHIPEVCLKWELMDENYRGRLCNPLNP
FVSHKEIESVNNAVVMPAHGPAKSMEYDAVATTHSWNPKRNRSILNTSQRGILEDEQMYQKCCNLFEKFF
PSSSYRRPIGISSMVEAMVSRARIDARIDFESGRIKKEEFSEIMKICSTIEELRRQK

SEQ ID NO:48
>gi|73919147|ref|YP_308843.1| nucleocapsid protein [Influenza A virus (A/New York/392/2004(H3N2))]
MASQGTKRSYEQMETDGDRQNATEIRASVGKMIDGIGRFYIQMCTELKLSDHEGRLIQNSLTIEKMVLSA
FDERRNKYLEEHPSAGKDPKKTGGPIYRRVDGKWMRELVLYDKEEIRRIWRQANNGEDATAGLTHIMIWH
SNLNDATYQRTRALVRTGMDPRMCSLMQGSTLPRRSGAAGAAVKGIGTMVMELIRMVKRGINDRNFWRGE
NGRKTRSAYERMCNILKGKFQTAAQRAMVDQVRESRNPGNAEIEDLIFLARSALILRGSVAHKSCLPACA
YGPAVSSGYDFEKEGYSLVGIDPFKLLQNSQIYSLIRPNENPAHKSQLVWMACHSAAFEDLRLLSFIRGT
KVSPRGKLSTRGVQIASNENMDNMGSSTLELRSGYWAIRTRSGGNTNQQRASAGQTSVQPTFSVQRNLPF
EKSTIMAAFTGNTEGRTSDMRAEIIRMMEGAKPEEVSFRGRGVFELSDEKATNPIVPSFDMSNEGSYFFG
DNAEEYDN

SEQ ID NO:49
>gi|73919136|ref|YP_308842.1| neuraminidase [Influenza A virus (A/New York/392/2004(H3N2))]
MNPNQKIITIGSVSLTISTICFFMQIAILITTVTLHFKQYEFNSPPNNQVMLCEPTIIERNITEIVYLTN
TTIEKEMCPKLAEYRNWSKPQCDITGFAPFSKDNSIRLSAGGDIWVTREPYVSCDPDKCYQFALGQGTTL
NNVHSNDTVHDRTPYRTLLMNELGVPFHLGTKQVCIAWSSSSCHDGKAWLHVCVTGDDKNATASFIYNGR
LVDSIVSWSKKILRTQESECVCINGTCTVVMTDGSASGKADTKILFIEEGKIIHTSTLSGSAQHVEECSC

YPRYPGVRCVCRDNWKGSNRPIVDINIKDYSIVSSYVCSGLVGDTPRKNDSSSSSHCLDPNNEEGGHGVK
GWAFDDGNDVWMGRTISEKLRSGYETFKVIEGWSKPNSKLQINRQVIVDRGNRSGYSGIFSVEGKSCINR
CFYVELIRGRKEETEVLWTSNSIVVFCGTSGTYGTGSWPDGADINLMPI

SEQ ID NO:50
>gi|73919134|ref|YP_308846.1| polymerase PA [Influenza A virus (A/New York/392/2004(H3N2))]
MEDFVRQCFNPMIVELAEKAMKEYGEDLKIETNKFAAICTHLEVCFMYSDFHFINEQGESIVVELDDPNA
LLKHRFEIIEGRDRTMAWTVVNSICNTTGAEKPKFLPDLYDYKENRFIEIGVTRREVHIYYLEKANKIKS
ENTHIHIFSFTGEEIATKADYTLDEESRARIKTRLFTIRQEMANRGLWDSFRQSERGEETIEEKFEISGT
MRRLADQSLPPKFSCLENFRAYVDGFEPNGCIEGKLSQMSKEVNAKIEPFLKTTPRPIKLPNGPPCYQRS
KFLLMDALKLSIEDPSHEGEGIPLYDAIKCIKTFFGWKEPYIVKPHEKGINSNYLLSWKQVLSELQDIEN
EEKIPRTKNMKKTSQLKWALGENMAPEKVDFDNCRDISDLKQYDSDEPELRSLSSWIQNEFNKACELTDS
IWIELDEIGEDVAPIEYIASMRRNYFTAEVSHCRATEYIMKGVYINTALLNASCAAMDDFQLIPMISKCR
TKEGRRKTNLYGFIIKGRSHLRNDTDVVNFVSMEFSLTDPRLEPHKWEKYCVLEIGDMLLRSAIGQISRP
MFLYVRTNGTSKVKMKWGMEMRRCLLQSLQQIESMIEAESSIKEKDMTKEFFENKSEAWPIGESPKGVEE
GSIGKVCRTLLAKSVFNSLYASPQLEGFSAESRKLLLVVQALRDNLEPGTFDLGGLYEAIEECLINDPWV
LLNASWFNSFLTHALK

SEQ ID NO:51
>gi|73919060|ref|YP_308849.1| polymerase PB2 [Influenza A virus (A/New York/392/2004(H3N2))]
MERIKELRNLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPSLRMKWMMAMKYPITADKRITEMVPER
NEQGQTLWSKMSDAGSDRVMVSPLAVTWWNRNGPVASTVHYPKVYKTYFDKVERLKHGTFGPVHFRNQVK
IRRRVDINPGHADLSAKEAQDVIMEVVFPNEVGARILTSESQLTITKEKKEELRDCKISPLMVAYMLERE
LVRKTRFLPVAGGTSSIYIEVLHLTQGTCWEQMYTPGGEVRNDDVDQSLIIAARNIVRRAAVSADPLASL
LEMCHSTQIGGTRMVDILRQNPTEEQAVDICKAAMGLRISSSFSFGGFTFKRTSGSSVKKEEEVLTGNLQ
TLKIRVHEGYEEFTMVGKRATAILRKATRRLVQLIVSGRDEQSIAEAIIVAMVFSQEDCMIKAVRGDLNF
VNRANQRLNPMHQLLRHFQKDAKVLFQNWGIEHIDSVMGMVGVLPDMTPSTEMSMRGIRVSKMGVDEYSS
TERVVVSIDRFLRVRDQRGNVLLSPEEVSETQGTERLTITYSSSMMWEINGPESVLVNTYQWIIRNWEAV
KIQWSQNPAMLYNKMEFEPFQSLVPKAIRSQYSGFVRTLFQQMRDVLGTFDTTQIIKLLPFAAAPPKQSR
MQFSSLTVNVRGSGMRILVRGNSPVFNYNKTTKRLTILGKDAGTLIEDPDESTSGVESAVLRGFLIIGKE
DRRYGPALSINELSNLAKGEKANVLIGQGDVVLVMKRKRDSSILTDSQTATKRIRMAIN

SEQ ID NO:52: CMV Protein IE122:
>gi|39841910|gb|AAR31478.1| UL122 [Human herpesvirus 5]
MESSAKRKMDPDNPDEGPSSKVPRPETPVTKATTFLQTMLRKEVNSQLSLGDPLFPELAEESLKTFEQVT
EDCNENPEKDVLAELGDILAQAVNHAGIDSSSTGHTLTTHSCSVSSAPLNKPTPTSVAVTNTPLPGASAT
PELSPRKKPRKTTRPFKVIIKPPVPPAPIMLPLIKQEDIKPEPDFTIQYRNKIIDTAGCIVISDSEEEQG
EEVETRGATASSPSTGSGTPRVTSPTHPLSQMNHPPLPDPLARPDEDSSSSSSSCSSASDSESESEEMK
CSSGGGASVTSSHHGRGGFGSAASSSLLSCGHQSSGGASTGPRKKKSKRISELDNEKVRNIMKDKNTPFCTPNVQTRRG
RVKIDEVSRMFRNTNRSLEYKNLPFTIPSMHQVLDEAIKACKTMQVNNKGIQIIYTRNHEVKSEVDAVRCRLGTMCNLA
LSTPFLMEHTMPVTHPPEVAQRTADACNEGVKAAWSLKELHTHQLCPRSSDYRNMIIHA**ATPVDLLGALNLCLPLM**QKF
PKQVMVRIFSTNQGGFMLPIYETAAKAYAVGQFEQPTETPPEDLDTLSLAIEAAIQDLRNKSQ

SEQ ID NO:53:
>gi|4927721|gb|AAD33253.1|AF125673_2 E7 [Human papillomavirus type 16]
MHGDTPTLHEYMLDLQPETTDLYCYEQLNDSSEEEDEIDGPAGQAEPDRAHYNIVTFCCKCDSTLRLCVQ
STH**VDIRTLEDLL**M**GTLGIVCPI**CSQKP

EXAMPLE 1

The peptides according to the present invention may be synthesized by Schafer-N as c-terminal amides using the Fmoc-strategy of Sheppard, (1978) J.Chem.Soc., Chem. Commun., 539.

Cell penetration assay

A set of peptides is biotinylated on N-terminal, and different combinations of amino acids, with respect to length and type, are added to the sequence box X^1, X^2, X^3, X^4, X^5 and X^6 in the peptides as illustrated by the diagram below. The peptides are tested on cells grown from one individual blood donor.

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	X^6

Intracellular staining for biotinylated peptides

96-well U-bottom polystyrene plates (NUNC, cat no: 163320) are used for staining of human PBMCs. Briefly, 8ul of N- or C-terminally biotinylated peptides according to table 1 (i.e. 5mM, 2.5mM & 1.25mM tested for each peptide) are incubated at 37°C for 2h with 40ul of PBMC (12.5 x 106 cells/ml) from blood donors. Cells are 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 are 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 4°C for 20 min according to manufacturer. Cells are then washed 2x with 150ul PermWash (BD, cat no: 554715). Cells are 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 are then washed 3x with 150ul PermWash, followed by resuspension in staining buffer (BD, cat no: 554656) before flow cytometry. Dendritic cells are gated as CD11c+ events outside lymphocyte region (i.e. higher FSC & SSC signals than lymphocytes). 200 000 total cells are 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) are prepared.

Extracellular staining for biotinylated peptides

96-well U-bottom polystyrene plates (NUNC, cat no: 163320) are used for staining of human PBMCs. Briefly, 8ul of N- or C-terminally biotinylated peptides according to table 1 (i.e. 5mM, 2.5mM & 1.25mM tested for each peptide; all peptides manufactured by Schafer) are incubated at 37°C for 2h with 40ul of PBMC (12.5 x 106 cells/ml) from blood donors. Cells are 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 are 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 are gated as CD11c+ events outside lymphocyte region (i.e. higher FSC & SSC signals than lymphocytes). 200 000 total cells are 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) are prepared.

It may then be seen if the peptide has the ability to enter the cell.

Example 2

Positive CTL response may 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 are 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 are isolated, washed, and frozen in 10% V/V DMSO in FBS until further usage. Adherent cells are 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 is then repeated at day 6. At day 7, cultured dendritic cells (5 000-10 000 per well) are 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

are incubated overnight and spots are developed according to manufacturer. Spots are read on ELISPOT reader (CTL-ImmunoSpot® S5 UV Analyzer).

Example 3

The REVEAL & ProVE® Rapid Epitope Discovery System in Detail

Binding properties to HLA for the peptides according to the present invention may be tested for the following HLA-types in class I: 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, and the following HLA-types in class II: HLA-DR1, HLA-DR3, HLA-DR4, HLA-DR7, HLA-DR11, HLA-DR13, HLA-DR15. The peptides are synthesized as a Prospector PEPscreen®: Custom Peptide Library. Peptides 8-15 amino acids in length are synthesized in 0.5-2mg quantities with high average purity. Quality control by MALDI-TOF Mass Spectrometry is carried out on 100% of samples..

The REVEAL™ binding assay determines 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 may be identified. Detection is based on the presence or absence of the native conformation of the MHC-peptide complex.
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.

Example 4: Preparation of dimeric peptides
Amino acids that link two monomeric peptide sequences are underlined.

Influenza (M2e):

Constructs derived from the extracellular domain on influenza protein M2 (M2e-domain)

Native domain:
MSLLTEVETPIRNEWGCRCNDSSD

The following sequences was prepared or under preparation. The different parts, X^1-X^6, are divided by brackets.

BI100_CGnat [RR][SLLTEVETP][GCG][VETPIR][G][TPIRNEWG]
BI100_CG [RR][SL(Nle)TDIETP][GCG][IDTPIR][G][TPIBQDWG]

BI100-CGcyc [WWGC][TDIET][CG][IDTPIR][G][TPIBQDWG]
BI100-Cyc_2 [RRGC][SLLT][C][SLLTEVQTPIRN][GRR] [SEWGSRSN]

Bi150-Dimer [RR(Nle)C][SLLTEVQTPIRN][GRR][VETPIRN]
|
[WWQC][TPIRSEWGCRSN]GRR[SNDSSG]

BI150-new [WW][SL(Nle)TDIETP][GCG][IDTPIR][G][TPIBQDWG]
|
[RR(Har)][IDTPIR][G][TPIBQDWG][KG][SL(Nle)TDIETPG]

BI150-2mod [R][SLZTDIETP][Dpr(Aoa)][IDTPIR][G][TPIBQDWG]
|
[RR][IDTPIR][GG][TPI(Har)QEW][Dpr(Ser)][SLZTDIETPG]

This construct links the monomeric peptides via a Dpr(Aoa) in the first peptide to an oxidized by $NaIO_4$ Dpr(Ser) residue in the second.
Dpr(Aoa) = N-α-Fmoc-N-β-(N-t.-Boc-amino-oxyacetyl)-L-diaminopropionic acid

Explanation:
The brackets used in the sequences are meant to indicate the different parts/boxes. For BI100_CGnat / BI100_CG, the boxes will have the following amino-acid sequences:

Part X^1	RR	
Part X^2	**SLLTEVETP/SL(Nle)TDIETP**	(aa 2-10 on M2e native domain)
Part X^3	GCG	
Part X^4	**VETPIR/IDTPIR**	(aa 7-12 on M2e native domain)
Part X^5	G	
Part X^6	**TPIRNEWG/TPIBQDWG**	(aa 9-16 on M2e native domain)

The boxes on part of the other sequences can be found in a similar manner

Example C5-sequences

BI450-AdjBT1 W_DW_LGCAKRRVCGGAKRRVVQREKRA

BI450-AdjBT2 W_DW_LGCIEEEGCGGIEEEGGERDR

BI400-B GAKRRVVGG**C**GGAKRRVVQREKRAGEREKRA

|

G**K**GGIEEEGGRDRDRGGEQDRDR

Examples of disulfide linked constructs can be, but are not restricted to, the following linked peptide sequences:

CGGAKRRVVGGAKRRVVGQREKRAV (SEQ ID NO:115)
|
CGGGDQQLLGGAEEEIVGGIEEEGGERDRDR (SEQ ID NO:116)

CGGAKRRVVGGAKRRVVGGQREKR (SEQ ID NO:117)
|
CGGGDQQLLGGAEEEIVGGIEEEGG (SEQ ID NO:118)

CGGAEEEVVGGDQQLL (SEQ ID NO:119)
|
GCGGAKRRVVGGAKRRVV (SEQ ID NO:120)

The above disulfide linked constructs may e.g. be synthesised by titration of 2-pyridinesulfenyl (SPyr)-protected cysteine-containing peptides with thiol-unprotected peptides. This has proven to be a superior procedure to selectively generate disulfide-linked peptide heterodimers preventing the formation of homodimers (Schutz A *et al.*, Tetrahedron, Volume 56, Issue 24, 9 June 2000, Pages 3889-3891). Similar constructs may be made where SEQ ID NO: 115 is disulphide linked to SEQ ID NOs 118 or 120, or where SEQ ID NO: 117 is disulphide linked to SEQ ID NOs: 116 or 120, or where SEQ ID NO: 119 is disulphide linked to SEQ ID NOs: 116 or 118.

Examples of thio-esther linked constructs can be, but are not restricted to, the following linked peptide sequences, which have all been obtained from Bachem (UK) Ltd:

GAKRRVVGGCGGAKRRVVQREKRAGEREKRA (SEQ ID NO: 121)
|
GKGGIEEEGGRDRDRGGEQDRDR (SEQ ID NO: 122)

(the peptides are linked via the underlined Cys and Lys residues; the entire construct is termed BI400-B herein).

GAKRRVVGGCGGAKRRVVQREKRAGEREKRA (SEQ ID NO: 121)
|
GKGGIEEEGGERDRDRGGQDRDR (SEQ ID NO: 124)

(the peptides are linked via the underlined Cys and Lys residues; the entire construct is termed BI400-Bu1 herein).

GAKRRVVGGCGGAKRRVVEREKRAGQREKRA (SEQ ID NO: 125)
|
GKGGIEEEGGQDRDRGGRDRDR (SEQ ID NO: 126)

(the peptides are linked via the underlined Cys and Lys residues; the entire construct is termed BI400-Bu2 herein).

GAKRRVVGGCGGAKRRVVEREKRAGQREKRA (SEQ ID NO: 125)
|
GKGGIEEEGGEQDRDRGGERDRD (SEQ ID NO: 128)

(the peptides are linked via the underlined Cys and Lys residues; the entire construct is termed BI400-Bu3 herein).

The Cys-Lys linker is typically established in the form of a thioether bond between a cysteine in one peptide and a bromoacetyl derivatized lysine in the other peptide.

Similar constructs may be made where SEQ ID NO: 121 is Cys-Lys linked to SEQ ID NOs 126 or 128, or where SEQ ID NO: 125 is Cys-Lys linked to SEQ ID NOs: 122 or 124.

Examples of other linked constructs can be, but are not restricted to, the following linked peptide sequences,

GAKRRVVGGSGGAKRRVVQREKRAGEREKRA (SEQ ID NO: 129)
|
GKGGIEEEGGRDRDRGGEQDRDR (SEQ ID NO: 122)

(the peptides are linked via the underlined Ser and Lys residues).

The Ser-Lys linker is typically established in the form of an oxime bond between oxidized (aldehyde) serine in one peptide and (aminooxyacerylated)derivatized lysine in the other peptide.

Example 5

The construction of HCV dimeric peptides:

BI350-1mod1:

RRGNWAKVLKNWAKVI (SEQ ID NO: 130)
|
RRGLLADARVGCGSGADRVCS (SEQ ID NO: 131)

This construct links the monomeric peptides via a lysine residue in the first peptide to a cysteine residue in the second peptide by using a sulpho-SMCC linker.

BI350-1mod2:

RRGNWAKVL(Dpr)NWAKVI (SEQ ID NO: 132)
|
RRGLLADARVG(Dpr(Ser))GSGADRVCS (SEQ ID NO: 133)

This construct links the monomeric peptides via a Dpr(Aoa) in the first peptide to an oxidized by $NaIO_4$ Dpr(Ser) residue in the second peptide.
Dpr(Aoa) = N-α-Fmoc-N-β-(N-t.-Boc-amino-oxyacetyl)-L-diaminopropionic acid

Alternatively, the K or C may be substituted with an N-ε-methylated Lys, which is linked to Asp or Glu.

Accordingly, an N-ε-methylated Lys may be linked to Asp or Glu by a side-chain to side-chain peptide bond, wherein the N methylation makes the bond more stable.

The sequences would then be (Lys(Me) refers to an N-ε-methylated Lys residue):

RRGNWAKVL-Lys(Me)-NWAKVI (SEQ ID NO: 134)
|
RRGLLADARVGEGSGADRVCS (SEQ ID NO: 135)

or

RRGNWAKVL-Lys(Me)-NWAKVI (SEQ ID NO: 134)
|
RRGLLADARVGDGSGADRVCS (SEQ ID NO: 136)

or alternatively, if the bond is reversed:

RRGNWAKVLENWAKVI (SEQ ID NO: 137)
|

RRGLLADARVG-Lys(Me)-GSGADRVCS (SEQ ID NO: 138)

or

RRGNWAKVLDNWAKVI (SEQ ID NO: 139)
|
RRGLLADARVG-Lys(Me)-GSGADRVCS (SEQ ID NO: 138)

The construction of an Influenza dimeric peptides:

BI150-2mod

R-SLZTDIETP-(Dpr)-IDTPIRGTPIBQDWG (SEQ ID NO: 140)
|
RR-IDTPIR-GG-TPI(Har)QEW-Dpr(Ser)-SLZTDIETPG (SEQ ID NO: 141)

Dpr is diaminopropionic acid, Dpr(Ser) is serinyl diaminopropionic acid.
(to form a Dpr-Dpr(Ser) oxime bond, as described elsewhere in this application)

Native sequence peptide from influenza M2e protein, which the above construct targets, used for testing:

BI100-cg2 MSLLTEVETPIRNEWGCRC (SEQ ID NO: 142)

EXAMPLE 6

Immunological studies

Rabbit immunizations

New Zealand White female rabbits (n=3) was immunized intradermally at weeks 0, 2 & 6 with 1 ml of BI400-B vaccine consisting of 500 µg BI400-B in 50% V/V Freund´s adjuvant (i.e. Complete Freund´s adjuvant used for priming, followed by boostings with Incomplete Freund´s adjuvant). Individual blood serum was isolated for ELISA.

Direct ELISA for human or rabbit sera

50-100 µl of BI400-B (pre-incubated in Coating buffer - 0.05M Na_2CO_3 pH9.6; denoted CB - in cold at 16 µg/ml for each peptide 1-3 days prior to coating) or just CB (background control) is used for coating wells in microtiter plates at 4°C overnight. The microtiter plates are then washed 3x with washing buffer (PBS + 1% v/v Triton-X100; denoted WB), followed by 2h blocking at room temperature (RT) with 200 µl/well of blocking buffer (PBS + 1% w/v BSA). Plates are then washed 3x with WB, followed by 1h incubation at 37°C with 50-70 ul/well of added human (or rabbit) sera (serial dilutions ranging from 1:1 – 1:250 in dilution buffer (PBS + 1% v/v Triton-X100 + 1% w/v BSA; denoted DB)). Plates are then washed 6x with WB, followed by 1h incubation at RT with 70 µl/well of Alkaline Phosphatase-conjugated Protein G (3µg/ml in DB; Calbiochem 539305). Plates are then washed 6x with WB, followed by 10-60 min incubation at room temperature with 100 µl/well of 0.3% w/v of Phenophtalein monophosphate (Sigma P-5758). Plates are finally quenched by adding 100 µl/well of Quench solution (0.1M TRIS + 0.1M EDTA + 0.5M NaOH + 0.01% w/v NaN_3; pH14), followed by ELISA reader (ASYS UVM 340) at 550 nm.

Results

The results from the immunization studies with BI400-B demonstrate that it is possible to generate peptides that elicit an efficient antibody response. The specificity of these antibody responses may be confirmed in competitive Elisa. Antibodies generated to BI400-B in animal models are comparable with antibodies elicited in natural HIV infection and associated with longterm nonprogression. These results show that these peptides are suitable for diagnostics as well as the development of a vaccine targeting HIV-induced immune activation as well as other pathogens.

EXAMPLE 7

Influenza specific M2e response by ELISPOT assay

At day one, PBMC samples from blood donors were thawed, washed with warm medium and incubated in flasks (250000PBMCs/cm2) for 24 hours at 37°C, 5% CO2 in covering amount of culture media (RPMI 1640 with ultra-glutamine, Lonza, BE12-702F701; 10% Foetal Bovine serum (FBS), Fisher Scientific Cat. No. A15-101; Penicillin/Streptomycin, Fisher Scientific Cat. No. P11-010) to allow the cells to recover after thawing. At day two, the cells were added to a Falcon Microtest Tissue Culture plate, 96well flat bottom, at 500 000 cells per well in a volume of 200µl total medium. Parallel wells were added the indicated stimuli in duplicate or left with medium as a control for 6 days at 37°C, 5% CO_2. After the six day of incubation, 100µl of the cell suspension were transferred to an ELISPOT (Millipore

multiscreen HTS) plate coated with 1µg/ml native influenza M2e protein. After a 24 hour incubation, the plate was washed four times with PBS + 0,05% Tween20, and a fifth time with PBS, 200µl/well. A mouse Anti-human IgG or IgM biotin (Southern Biotech 9040-08 and 9020-08) was diluted in PBS with 0.5% FBS and incubated for 90 minutes at 37°C. The washing was repeated as described, before 80µl Streptavidin-Alkaline-Phosphatase (Sigma Aldrich, S289) was added each well and incubated at 60 minutes in the dark, at room temperature. The wells were then washed 2 times with PBS + 0.05% Tween20 and 4 times with PBS, 200µl/well, before the substrate, Vector Blue Alkaline Phosphatase Substrate kit III (Vector Blue, SK-5300) was added and let to develop for 7 minutes at room temperature. The reaction was stopped with running water, the plates let dry and the sport enumerated by an ELISPOT reader (CTL-ImmunoSpot® S5 UV Analyzer).

Influenza specific M2e response by ELISA

100µl of antigen as indicated (pre-incubated in Coating buffer - 0.05M Na_2CO_3 pH9.6; denoted CB - in cold at 8µg/ml 1-3 days) or just CB (background control) was used for coating wells in microtiter plates at 4°C. The microtiter plates are then washed 3x with washing buffer (PBS + 1% v/v Triton-X100; denoted WB), followed by 2h blocking at room temperature (RT) with 200 µl/well of blocking buffer (PBS + 1% w/v BSA). Plates are then washed 3x with WB, followed by 1h incubation at 37°C with 50-70 ul/well of added human (or rabbit or sheep) sera (serial dilutions ranging from 1:5 – 1:250 in dilution buffer (PBS + 1% v/v Triton-X100 + 1% w/v BSA; denoted DB)). Plates are then washed 6x with WB, followed by 1h incubation at RT with 70 µl/well of Alkaline Phosphatase-conjugated Protein G (3µg/ml in DB; Calbiochem 539305) or goat anti-mouse IgG biotin (1µg/ml, Southern Biotech, 1030-08. In case of the goat anti-mouse IgG biotin, the plates were washed one extra step as described, before addition of 100µl Streptavidin-Alkaline-Phosphatase (1µg/ml, Sigma Aldrich, S289) and incubated 1 hour at RT. Plates are then washed 6x with WB, followed by 10-60 min incubation at room temperature with 100 µl/well of 0.3% w/v of Phenophtalein monophosphate (Sigma P-5758). Plates are finally quenched by adding 100 µl/well of Quench solution (0.1M TRIS + 0.1M EDTA + 0.5M NaOH + 0.01% w/v NaN_3; pH14), followed by a measurement with a ELISA reader (ASYS UVM 340) at 550 nm. The strength of the sera, i.e. the magnitude of the humoral immune response, is then reported as the dilution of sera that result in the described Optical Density (OD) value, or the OD value at the indicated dilution of sera.

RESULTS

In table 2 show the relative amount of number of B-cells that proliferated compared to baseline proliferation of unstimulated PBMC. The table clearly shows that the BI150-2mod was recognised by human memory cells as the response was IgG specific, meaning that the BI150-2mod can stimulate class-switched B cells in a way that allow the antibodies produced to bind to the native sequence that BI150-2mod is derived from. A large variation between individual donors is to be expected as there will be a large variation in the human population with regard to their pre-existing humoral memory towards the influenza M2 protein. From table 2 it can also be noted that the BI150-2mod was also able to induce proliferation of IgM B cells.

Table 2. B cell response in human PBMC stimulated with BI150-2mod against native sequence

		PBMC from blood donors							
		BC-39	BC-42	BC-28	BC-35	BC-31	BC-34	Mean	Mec
IgG	Survival stimuli*	2200 %	120 %	111 %	6625 %	300 %	100 %	1576 %	210
	Survival stimuli+ BI150-2mod	2260 %	440 %	133 %	1638 %	300 %	950 %	953 %	695
IgM	Survival stimuli	3400 %	3067 %	22900 %	300 %	400 %	3100 %	5528 %	308
	Survival stimuli+ BI150-2mod	3400 %	6783 %	7600 %	1700 %	467 %	1300 %	3542 %	354

*Surivival stimuli: rh-sCD40 Ligand, Immuno Tools, 11343345, ODN 2006 Type B CpG oligonucleotide-Human TLR9 ligand, Invivogen Sigma, Tlrl-2006, rh IL-21, Immunotools, 11340213

The vaccination of BALB/c mice was done s.c. with 100µg peptide at weeks 1, 3, 7, serum from week 9 with 1mg Aluminium in the form of 1.3% Alhydrogel. In table 3, the data show that the BI350-1mod1 induced a specific immune response in BALB/c mice as the vaccine antigen was recognised. It was also evident that the resulting immune response comes as a consequence of a class switching event as the immune response was IgG specific.

Table 3. Serum IgG response (OD) in mice vaccinated with BI350-1mod1 +Alhydrogel directed against vaccine antigen

	Mean (n=6)	SEM
BI350-1mod1	0.26	0.09

The vaccination of sheep was done with 500µg of peptide in PBS with the following adjuvants: ISA51 (vaccination week 1, 2, 3, 4, serum from week 6) and Freunds Adjuvant (vaccination weeks 1, 12, 26, serum from week 28). In table 4, the data show the dilution of serum that correspond to an OD value three times the assay background. The table clearly shows that the BI100-CG and BI100-CGcyc induced an immune response by itself or with two different standard adjuvants. This indicate that the constructs can be combined with different adjuvants and induce an IgG response, a signal that an immunological memory has been elicited in sheep.

Table 4. Dilution of serum three times background IgG response in sheep vaccinated with BI100 constructs directed against BI100-cg2 antigen

	BI100-CG	BI100CGcyc
Freunds	11000	125
ISA51	125	25
None	25	25

CLAIMS

1. An isolated monomeric peptides consisting of not more than 60 amino acids with the following structure

X^1-X^2-X^3-X^4-X^5-X^6 (formula I),

wherein X^1, X^3 and optional moiety X^5 each independently defines a linear sequence of any 1, 2, 3, 4, or 5 amino acid independently selected from glycine (G), arginine (R), norleucine, aspartic acid (D), glutamic acid (E), glutamine (Q), serine (S), lysine (K), tryptophan (W), cysteine (C), ornithine, diaminopropionic acid (Dpr) or a derivative thereof; X^2, X^4, and optional moiety X^6 each independently defines a linear sequence of 5-17 amino acids, each having 50% or more sequence identity to a specific natural antigen.

2. The isolated monomeric peptide according to claim 1, wherein X^2 and/or X^4 and/or X^6 has more than 55%, such as more than 60%, such as more than 65%, such as more than 70%, such as more than 75%, such as more than 80%, such as more than 85%, such as more than 90%, such as more than 95%, such as more than 96%, such as more than 97%, such as more than 98%, such as more than 99%, such as 100% sequence identity to a specific natural antigen.

3. The isolated monomeric peptide according to claims 1 or 2, wherein said specific natural antigen is a protein or peptide sequence derived from a disease antigen, such as an infectious agent, such as bacteria, virus, parasite, fungus, or cancer antigens such as oncogene (lung, stomach, breast cancer) or an antigen causing an autoimmune disease such as diabetes, multiple sclerosis (MS), celiac disease, Myalgic Encephalomyelitis (ME), psoriasis, and/or Crohn's Disease.

4. The isolated monomeric peptide according to any one of claims 1-3, wherein said specific natural antigen is a viral protein, such as a structural protein, such as a capsid protein, a regulatory protein, an enzymatic protein, and a proteolytic protein.

5. The isolated monomeric peptide according to any one of claims 1-4, wherein said viral protein is a protein, such as a core protein or an envelope protein, of a virus selected from the Hepatitis C virus, influenza virus, such as an M2 protein, human immunodeficiency virus (HIV), cytomegalovirus (CMV), and Human papillomavirus (HPV).

6. The isolated monomeric peptide according to claim 5, wherein said viral protein is a viral protein of Hepatitis C virus selected from any one HCV consensus sequence of a specific

genotype, such as 1, such as subtypes 1a and 1b, genotype 2, such as 2a and 2b, genotype 3, such as 3a, genotype 4, genotype 5, and genotype 6.

7. The isolated monomeric peptide according to any one of claims 1-3, wherein the sequence of amino acids defined by X^1-X^2-X^3-X^4-X^5-X^6 is not found in the native sequence of said natural antigen.

8. The isolated monomeric peptide according to any one of claims 1-7, which monomeric peptide is of 18-60 amino acids, such as 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 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, such as of 36-60 amino acids, such as of 37-60 amino acids, such as of 38-60 amino acids, such as of 39-60 amino acids, such as of 40-60 amino acids, such as of 42-60 amino acids, such as of 44-60 amino acids, such as of 46-60 amino acids, such as of 48-60 amino acids, such as of 50-60 amino acids, such as of 52-60 amino acids, such as of 54-60 amino acids, such as of 56-60 amino acids, such as of 58-60 amino acids.

9. The isolated monomeric peptide according to any one of claims 1-8, which monomeric peptide is of 18-60 amino acids, such as 18-58 amino acids, such as 18-56 amino acids, such as 18-54 amino acids, such as 18-52 amino acids, such as 18-50 amino acids, such as 18-48 amino acids, such as 18-46 amino acids, such as 18-44 amino acids, such as 18-42 amino acids, such as 18-40 amino acids, such as 18-39 amino acids, such as 18-38 amino acids, such as 18-37 amino acids, such as 18-36 amino acids, such as 18-35 amino acids, such as 18-34 amino acids, such as 18-33 amino acids, such as 18-32 amino acids, such as 18-31 amino acids, such as 18-30 amino acids, such as 18-29 amino acids, such as 18-28 amino acids, such as 18-27 amino acids, such as 18-26 amino acids, such as 18-25 amino acids, such as 18-24 amino acids, such as 18-23 amino acids, such as 18-22 amino acids, such as 18-21 amino acids, such as 18-20 amino acids, such as 18-19 amino acids.

10. The isolated monomeric peptide according to any one of claims 1-9, which monomeric peptide consist of not more than about 55 amino acids, such as not more than about 50 amino acids, such as not more than about 45 amino acids, such as not more than about 40 amino acids, such as not more than about 38 amino acids, such as not more than about 36 amino acids, such as not more than about 34 amino acids, such as not more than about 32 amino acids, such as not more than about 30 amino acids, such as not more than about 28 amino acids, such as not more than about 26 amino acids, such as not more than about 24

amino acids, such as not more than about 22 amino acids, such as not more than about 20 amino acids, such as not more than about 18 amino acids, such as not more than about 16 amino acids, such as not more than about 14 amino acids, such as not more than about 12 amino acids, such as not more than about 10 amino acids.

11. The isolated monomeric peptide according to any one of claims 1-10, which monomeric peptide consist of at least about 10 amino acids, such as at least about 12 amino acids, such as at least about 14 amino acids, such as at least about 16 amino acids, such as at least about 18 amino acids, such as at least about 20 amino acids, such as at least about 22 amino acids, such as at least about 24 amino acids, such as at least about 26 amino acids, such as at least about 28 amino acids, such as at least about 30 amino acids, such as at least about 32 amino acids, such as at least about 34 amino acids, such as at least about 36 amino acids, such as at least about 38 amino acids, such as at least about 40 amino acids, such as at least about 45 amino acids, such as at least about 50 amino acids, such as at least about 55 amino acids, such as at least about 60.

12. The isolated monomeric peptide according to any one of claims 1-11, wherein the overall net charge of X^1-X^2-X^3-X^4-X^5-X^6 is equal to or above 0, such as above 1, 2, 3, 4, or 5.

13. The isolated monomeric peptide according to any one of claims 1-12, wherein said monomeric peptide is capable of inducing a humoral immune response.

14. The isolated monomeric peptide according to any one of claims 1-13, wherein said sequence X^1 and/or X^3 and/or X^5 is selected from K, Lys(Me), RRG, G(Dpr(Ser))G, Dpr, Dpr(ser), GG(Dpr(Ser))GG (SEQ ID NO:150), GEG, CS, GDG, E, G(Lys(Me))G, D, RR, WWGC (SEQ ID NO: 54), RRG, RRZC (SEQ ID NO: 55), WWQC (SEQ ID NO: 56), WW, RR-Har, Har, WDWGC (SEQ ID NO: 57), CGG, CGGG (SEQ ID NO: 58), GCGG (SEQ ID NO: 59), G, GKG, GC, GG, C, R, GCG, CG, GRR, GGCGG (SEQ ID NO: 60), CGGKG (SEQ ID NO: 61), CGGKGG (SEQ ID NO: 62), GGKGG (SEQ ID NO: 63), KG, and GWK.

15. The isolated monomeric peptide according to any one of claims 1-14, wherein said sequence X^2 and/or X^4 and/or X^6 is selected from SLLTEVETP (SEQ ID NO: 64), SLZTDIETP (SEQ ID NO: 65), TDIET (SEQ ID NO: 66), CSLLT (SEQ ID NO: 67), SLLTEVQTPIRN (SEQ ID NO: 68), TPIRSEWGCRSN (SEQ ID NO: 69), IDTPIR (SEQ ID NO: 70), AKRRV (SEQ ID NO: 71), IEEEG (SEQ ID NO: 72), AKRRVV (SEQ ID NO: 73), DQQLL (SEQ ID NO: 74), AEEEVV (SEQ ID NO: 75), GIEEE (SEQ ID NO: 76), IEEEGGRDRDR (SEQ ID NO: 77), CAKRRVVC (SEQ ID NO: 78), IEEEGGERDRDR (SEQ ID NO: 79), IEEEGGQDRDR (SEQ ID NO: 80), IEEEGGEQDRDR (SEQ ID NO: 81), EEEIGGRDRD (SEQ ID NO: 82), RLEPWKH (SEQ ID NO: 83), FHSQV (SEQ ID NO: 84), FITKGLGISY (SEQ ID NO: 85), LLADARVCS (SEQ ID NO: 86),

GV(Nle)AGIAYFS (SEQ ID NO: 87), VETPIR (SEQ ID NO: 88), VETPIRN (SEQ ID NO: 89), SNDSS (SEQ ID NO: 90), TPIBQDWG (SEQ ID NO: 91), AKRRVVQREKRA (SEQ ID NO: 92), IEEEGGERDR (SEQ ID NO: 93), AEEEIV (SEQ ID NO: 94), RDRDR (SEQ ID NO: 95), ERDRDR (SEQ ID NO: 96), AKRRVVEREKRA (SEQ ID NO: 97), QDRDR (SEQ ID NO: 98), EQDRDR (SEQ ID NO: 99), RDRDQ (SEQ ID NO: 100), GSQPKTA (SEQ ID NO: 101), FITKGLGISYGRK (SEQ ID NO: 102), LLADARVSA (SEQ ID NO: 103), GVLAGIAYYS (SEQ ID NO: 104), TPIRNEWG (SEQ ID NO: 105), SEWGSRSN (SEQ ID NO: 106), SLZTDIETPG (SEQ ID NO: 107), QREKRAV (SEQ ID NO: 108), QREKR (SEQ ID NO: 109), IEEEGG (SEQ ID NO: 110), EREKRA (SEQ ID NO: 111), QREKRA (SEQ ID NO: 112), ERDRD (SEQ ID NO: 113), HPGSQ (SEQ ID NO: 114), TPIXQEW (SEQ ID NO: 151), EQDRDRGG (SEQ ID NO: 152), GNWAKVL (SEQ ID NO: 153), LLADARV (SEQ ID NO: 154), NWAKVI (SEQ ID NO: 155), and SGADRV (SEQ ID NO: 156).

16. The isolated monomeric peptide according to any one of claims 1-15, wherein said monomeric peptide comprises at least one amino acid selected from a Cys, a Lys, a Ser, an Asp, and a Glu residue, or derivatives thereof.

17. The isolated monomeric peptide according to any one of claims 1-16, wherein said sequence X^1 and/or X^3 and/or X^5 is as defined in table 1.

18. The isolated monomeric peptide according to any one of claims 1-17, wherein said sequence X^2 and/or X^4 and/or X^6 is as defined in table 1.

19. The isolated monomeric peptide according to any one of claims 1-18, wherein said sequence X^2 and/or X^4 and/or X^6 defines a sequence of 4-17, such as 5-16, such as 5-15, such as 5-14, such as 5-13, such as 5-12, such as 5-10 amino acids.

20. The isolated monomeric peptide according to any one of claims 1-19, which monomeric peptide contain one or more intramolecular bond, such as one or more Cys-Cys bond.

21. The isolated monomeric peptide according to any one of claims 1-20, which monomeric peptide has delayed proteolytic degradation in the N-terminal, such as by incorporation of the first 1, 2, or 3 amino acids in the N-terminal in the D-form, or by incorporation of the first 1, 2, or 3 amino acids in the N-terminal in beta or gamma form.

22. An isolated multimeric, such as dimeric peptide comprising two or more monomeric peptides, each monomeric peptide independently consisting of not more than 60 amino acids with the following structure

X^1-X^2-X^3-X^4-X^5-X^6 (formula I),

wherein X^1, X^3 and optional moiety X^5 independently defines a linear sequence of any 1, 2, 3, 4, or 5 amino acid independently selected from glycine, arginine, norleucine, aspartic acid, glutamic acid, glutamine, serine, lysine, tryptophan, cysteine, ornithine, diaminopropionic acid or a derivative thereof; X^2, X^4, and optional moiety X^6 each independently defines a linear sequence of 5-17 amino acids, each having 50% or more sequence identity to a specific natural antigen, said monomeric peptides being covalently joined by one or more intermolecular bond.

23. The isolated dimeric peptide according to claim 22, wherein two or more monomeric peptides are identical in sequence.

24. The isolated dimeric peptide according to claim 22, wherein two or more monomeric peptides are different in sequence.

25. The isolated multimeric, such as dimeric peptide according to any one of claims 22-24, wherein one or more peptide strands of the multimeric, such as dimeric peptide has delayed proteolytic degradation in the N-terminal, such as by incorporation of the first 1, 2, or 3 amino acids in the N-terminal in the D-form, or by incorporation of the first 1, 2, or 3 amino acids in the N-terminal in beta or gamma form.

26. The isolated multimeric, such as dimeric peptide according to any one of claims 22-25, wherein the linker is placed within any sequence selected from X^1, X^2, X^3, X^4, X^5, and X^6, such as in X^1, X^2 or X^3 of the first monomeric peptide to anywhere on the at least one second monomeric peptide, such as within the sequence of X^1, X^2 or X^3.

27. The isolated multimeric, such as dimeric peptide according to any one of claims 22-26, which multimeric, such as dimeric peptide contain a helper epitope of at least 12 amino acids, such as at least 13, 14, 15 or 17 amino acids, which helper epitope consist of a combined sequence of amino acids, which is a sequence of amino acids from the first monomeric peptide, and a sequence of amino acids from the at least one second monomeric peptide, such as between 2-12 amino acids from the first monomeric peptide and 2-12 amino acids from the at least one second monomeric peptide.

28. The isolated multimeric, such as dimeric peptide according to any one of claims 22-27, wherein said intermolecular bond is selected from a disulfide (S-S) bond between two Cys residues.

29. The isolated multimeric, such as dimeric peptide according to any one of claims 22-28, wherein said intermolecular bond is a thioether bond between a Cys residue in the first monomeric peptide and a modified Lys residue in the at least one second monomeric peptide.

30. The isolated multimeric, such as dimeric peptide according to any one of claims 22-29, wherein said intermolecular bond is an oxime bond between a derivatized Lys residue in the first monomeric peptide and a derivatized Ser residue in the at least one second monomeric peptide.

31. The isolated multimeric, such as dimeric peptide according to any one of claims 22-30, wherein said intermolecular bond is a peptide bond between a N-methylated Lys side-chain in the first monomeric peptide and the side-chain of an Asp or Glu residue in the at least one second monomeric peptide.

32. The isolated multimeric, such as dimeric peptide according to any one of claims 22-31, wherein said intermolecular bond is an oxime bond between an aldehyde moiety, produced by oxidation of a serine residue in the first monomeric peptide and a free aminooxy group of a modified amino acid (aminooxy acid), such as derivataized diaminopropionic acid, Lysine or Ornithine in in the second monomeric peptide

33. The isolated multimeric, such as dimeric peptide according to any one of claims 22-32, wherein said monomeric peptides are linked by a polyethylene glycol (PEG) linker, such as through an Asp or a Glu residue in the first monomeric peptide and an Asp or a Glu residue in the at least one second monomeric peptide, or by a polyLys core.

34. The isolated multimeric, such as dimeric peptide according to any one of claims 22-33, wherein any one of said monomeric peptide is independently as defined in any one of claims 1-21.

35. Composition comprising two or more compounds selected from a monomeric peptide is as defined in any one of claims 1-21, and an isolated multimeric, such as dimeric peptide as defined in any one of claims 22-34.

36. Use of a peptide selected from a monomeric peptide is as defined in any one of claims 1-21, and an isolated multimeric, such as dimeric peptide as defined in any one of claims 22-34 for inducing a humoral immune response in a subject.

37. An isolated nucleic acid or polynucleotide encoding a peptide according to any one of claims 1-34.

38. A vector comprising the nucleic acid or polynucleotide according to claim 37.

39. A host cell comprising the vector according to claim 38.

40. An immunogenic composition comprising at least one monomeric peptide according to any one of claims 1-21, an isolated multimeric, such as dimeric peptide according to any one of claims 22-34, a peptide composition according to claim 35, the nucleic acid or polynucleotide according to claim 37, or the vector according to claim 38; in combination with a pharmaceutically acceptable diluent or vehicle and optionally an immunological adjuvant.

41. The immunogenic composition according to claim 40 in the form of a vaccine composition.

42. A method for inducing an immune response in a subject against an antigen which comprises administration of at least one monomeric peptide according to any one of claims 1-21, an isolated multimeric, such as dimeric peptide according to any one of claims 22-34, a peptide composition according to claim 35, the nucleic acid or polynucleotide according to claim 37, or the vector according to claim 38; or the composition according to any one of claims 40-41.

43. A method for reducing and/or delaying the pathological effects of a disease antigen, such as an infectious agent in a subject infected with said agent or having said disease caused by said antigen, the method comprising administering an effective amount of at least one monomeric peptide according to any one of claims 1-21, an isolated multimeric, such as dimeric peptide according to any one of claims 22-34, a peptide composition according to claim 35, the nucleic acid or polynucleotide according to claim 37, or the vector according to claim 38; or the composition according to any one of claims 40-41.

44. A peptide according to any one of claims 1-34 for use as a medicament.

45. A peptide according to any one of claims 1-34 for treating the pathological effects of a disease antigen, such as an infectious agent in a subject infected with said agent or having said disease caused by said antigen.

46. A peptide according to any one of claims 1-34 for use in an in vitro assay, such as an ELISA assay, such as for diagnostic purposes.

47. Use of a peptide according to any one of claims 1-34 for in vitro assay, such as an ELISA assay, such as for diagnostic purposes.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/DK2012/050010

A. CLASSIFICATION OF SUBJECT MATTER
IPC (2006.01): C07K 14/005, C07K 14/11, C07K 14/16, C07K 14/18, A61K 39/12

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: C07K, A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
DK, NO, SE, FI; Classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPODOC, WPI, FULL TEXT PATENT DOCUMENTS IN ENGLISH, GERMAN AND FRENCH LANGUAGE, DGENE, EBI-SECURELINE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
E,X	WO 2011/000962 A2 (BIONOR IMMUNO AS [NO/NO]) 2011.01.06 See SEQ ID NO: 23, 25, 27, 29, 31, 32, 33, 38, 39, 40, 42, 43, 44 and45	1-47
X	WO 2010/056796 A2 (THERACLONE SCIENCES, INC. [US/US]) 20 10.05.20 See [0006]-[0017]; SEQ ID NO:1549 on page 86	1-47
X	UniProtKB/TrEMBL Accession no. C9E989, entered into UniProtKB 2009.11.03 [Retrieved from http://www.uniprot.org/uniprot/C9E989 on 2012.01.26]	1-5, 8-11
A	EP 1878742 A2 (BIONOR IMMUNO AS) 2008.01.16 See [0015]-[0021]; example 14	
A	WO 00/29008 A2 (BOARD OF REGENTS, THE UNIVERSITY OF TEXAS SYSTEM [US/US]) 2000.05.25 See abstract and page 4, lines 10-17.	

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
26/01/2012	10/02/2012
Name and mailing address of the ISA/ Nordic Patent Institute, Helgeshøj Allé 81, DK-2630 Taastrup, Denmark Facsimile No. +45 43 50 80 08	Authorized officer Tine Haarmark Nielsen Telephone No. +45 43 50 80 57

Form PCT/ISA/210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/DK2012/050010

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.: 39, 42-43 (partly)
because they relate to subject matter not required to be searched by this Authority, namely:
Claims 39, 42-43 (partly) relate to a subject matter considered by this Authority to be covered by the provision of Rule 39.1(iv)/67.1(iv) PCT. The claims are directed to a method of treatment of the human or animal body. The search has been carried out based on the alleged effect of the compound/composition. Likewise, an opinion on novelty, inventive step and industrial applicability will be given taking these alleged effects into account.

2. ☒ Claims Nos.: 1-47 (partly)
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
The present claim 1 relates to an extremely large number of possible compounds. Support and disclosure in the sense of Rule 6 and 5 PCT is to be found however for only a small proportion of the compounds claimed, see Table 1 and Example 6-7.
Continued on Extra Sheet.

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

Continuation of Box II.2:
The non-compliance with the substantive is to such an extent, that the search was performed taking into consideration the non-compliance in determining the extent of the search of claim 1 (PCT Guidelines 9.19 and 9.23).

The search of claim 1 was restricted to those claimed compounds which appear to be supported by the description and closely related compounds to these compounds, in particularly the peptides with the sequences disclosed in Table 1. The same applies to claims 2-47.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/DK2012/050010

Patent document Cited in search report	Publication date	Patent family member(s)	Publication date
WO 2011000962 A2	2011.01.06	none	
WO 2010056796 A1	2010.05.20	CA 2743306 A1	2010.05.20
		AU 2009314091 A1	2010.05.20
		US 2010135955 A1	2010.06.03
		TW 201029663 A	2010.08.16
		EP 2346529 A1	2011.07.27
EP 1878742 A2	2008.01.16	NO 20004413 A	2002.03.05
		CA 2421193 A1	2002.03.14
		WO 0220554 A2	2002.03.14
		AU 8270601 A	2002.03.22
		EP 1322665 A2	2003.07.02
		BR 0114036 A	2003.07.22
		SK 4022003 A3	2003.09.11
		CN 1460111 A	2003.12.03
		AR 033998 A1	2004.01.21
		JP 2004508381 A	2004.03.18
		US 2004115615 A1	2004.06.17
		ZA 200301740 A	2004.06.21
		NZ 524556 A	2005.01.28
		HU 0303134 A2	2005.03.29
		EA 006308 B1	2005.10.27
		AU 2001282706B B2	2007.03.22
		HK 1059940 A1	2007.06.01
		NZ 535148 A	2007.06.29
		AU 2007202739 A1	2007.07.05
		US 2007237783 A1	2007.10.11
		CN 101104636 A	2008.01.16
		AT 406380T T	2008.09.15
		ES 2312465T T3	2009.03.01
		AR 071262 A2	2010.06.09
		JP 2011219481 A	2011.11.04
WO 0029008 A2	2000.05.25	CA 2350911 A1	2000.05.25
		AU 2150900 A	2000.06.05
		EP 1131633 A2	2001.09.12
		CN 1333873 A	2002.01.30
		JP 2002529112 A	2002.09.10
		US 6656471 B1	2003.12.02
		US 2005032039 A1	2005.02.10