REPLIKIN PEPTIDES AND USES THEREOF

[0001] Replikins are a newly discovered class of peptides that share structural characteristics. Replikins have been found in viruses, bacteria, fungus, cancer associated proteins, plants and unicellular parasites and their use as targets in the development of methods of treating or preventing diseases. Replikins are useful in the detection of these diseases. Also this invention relates to the use of Replikins to stimulate growth of plants used for food.

[0002] Rapid replication is characteristic of virulence in certain bacteria, viruses and malignancies, but no chemistry common to rapid replication in different organisms has been described previously. This application describes a new class of protein structures related to rapid replication which the applicants have discovered. This new family of conserved small proteins related to rapid replication, named Replikins, can be used to predict and control rapid replication in multiple organisms and diseases and to induce rapid replication in plant and animal life.

[0003] We constructed an algorithm to search for Replikins. In applying the algorithm not only was the function of the epitope revealed - rapid replication, but an entire family of homologues whose function is related to rapid replication was discovered, which we named Replikins.

[0004] The algorithm is based on the following: 1) Evidence that the immune system looks to parts rather than a whole protein in recognition. Protein chains are first hydrolyzed by the immune system into smaller pieces, frequently six (6) to ten (10) amino acids long, as part of the immune systems' process of recognition of foreign structures against which it may mount an immune defense. By way of example, the immune system recognizes the presence of disease by chopping up proteins of the disease agent into smaller peptide sequences and reading them. This principle is used as a basis for the algorithm with which to search for homologues of the malignin cancer epitope, once the structure of the epitope was known; 2) The specific structure of the malignin epitope, in which two of the three lysines (K's) are eight residues apart is in accordance with the apparent 'rules' used by the immune

system for recognition referred to above (6-10 amino acids long); 3) The fact that the malignin cancer epitope was shown to be a very strong antigen, that is – a generator of a strong immune response; that there are three lysines (K's) in the 10-mer peptide glioma Replikin and that K's are known to bind frequently to DNA and RNA, potential targets for the entry of viruses; and 4) One histidine (H) is included in the sequence of the malignin epitope, between the two K's which are eight (8) residues apart, suggesting a connection to the metals of redox systems which are required to provide the energy for replication.

[0005] Engineered enzymes and catalytic antibodies, possessing tailored binding pockets with appropriately positioned functional groups have been successful in catalyzing a number of chemical transformations, sometimes with impressive efficiencies. Just as two or more separate proteins with specific and quite different functions are now often recognized to be synthesized together by organisms, and then separately cleaved to 'go about their separate functions', so the Replikin structure is a unique protein with a unique function that appears to be recognized separately by the immune system and may be now rationally engineered-e.g. synthesized to produce a functional unit.

[0006] From a proteomic point of view, this template based on the newly determined glioma peptide sequence has led to the discovery of a wide class of proteins with related conserved structures and a particular function, in this case replication. Examples of the increase in Replikin concentration with virulence of a disease appear in diseases including, influenza, HIV, cancer and tomato leaf curl virus. This class of structures is related to the phenomenon of rapid replication in organisms as diverse as yeast, algae, plants, the gemini curl leaf tomato virus, HIV and cancer.

[0007] In addition to detecting the presence of Replikins in rapidly replicating organisms, we found that 1) Replikin concentration (number of Replikins per 100 amino acids) and 2) Replikin compositions in specific functional states dependant on rapid replication, provide the basis for the finding that Replikins are related quantitatively as well as qualitatively to the rate of replication of the organism in

which they reside. Examples of these functional proofs include the relationship found between rapid replication and virulence in glioblastoma cells, between Replikins in influenza virus and the prediction of influenza pandemics and epidemics, and the relationship between Replikin concentration and rapid replication in HIV.

[0008] The first functional basis for Replikins' role in rapid replication was found in the properties of the glioma Replikin, a 10KD peptide called Malignin in brain glioblastoma multiforme (glioma) - a 250KD cell protein. Antimalignin antibody increased in concentration in serum (AMAS), measured by an early stage diagnostic test for cancer now used for most or all cell types. Malignin was so named because in tissue culture the expression of this peptide and its concentration per milligram membrane protein extractable increased with increased rate of cell division per unit time. Not only is there an increase in the amount of malignin in proportion to the cell number increase but the amount of malignin is enriched, that is - increased ten fold whereas the cell number increased only five fold.

[0009] The structure of malignin protein was determined through hydrolysis and mass spectrometry which revealed what proved to be a novel 16 mer peptide sequence. We searched for the 16 mer peptide sequence which we have named a Glioma Replikin protein in databases for the healthy human genome and found that it was not present in these databases.

[0010] As such, the fixed requirement algorithm was used to search in other organisms for the Glioma Replikin protein or homologues thereof. Over 4,000 protein sequences in the "Pub Med" database were searched and homologues were found in viruses and plant forms specifically associated with rapid replication. Homologues of such Replikin proteins occurred frequently in proteins called 'replicating proteins' by their investigators.

[0011] Homologues of the Replikin sequence were found in all tumor viruses (that is viruses that cause cancer), and in 'replicating proteins' of algae, plants, fungi, viruses and bacteria.

[0012] That malignin is enriched ten-fold compared to the five-fold increase in cell number and membrane protein concentration in rapid replication of glioma cells suggests an integral relationship of the Replikins to replication. When the glioma replikin was synthesized in vitro and administered as a synthetic vaccine to rabbits, abundant antimalignin antibody was produced - establishing rigourously the antigenic basis of the antimalignin antibody in serum (AMAS) test, and providing the first potential synthetic cancer vaccine and the prototype for Replikin vaccines in other organisms.

[0013] The demonstration of the relationship of the Replikins to replication and the natural immune response to cancer Replikins (overriding cell type) based upon the shared specificity of cancer Replikins, permits passive augmentation of immunity with antimalignin antibody and active augmentation with synthetic Replikin vaccines.

[0014] A study of 8,090 serum specimens from cancer patients and controls has demonstrated that the concentration of antimalignin antibody increases with age in healthy individuals, as the incidence of cancer in the population increases, and increases further two to three-fold in early malignancy, regardless of cell type. In vitro this antibody is cytotoxic to cancer cells at picograms (femtomoles) per cancer cell, and in vivo the concentration of antimalignin antibody relates quantitatively to the survival of cancer patients. As shown in glioma cells, the stage in cancer at which cells only have been transformed to the immortal malignant state but remain quiescent or dormant, now can be distinguished from the more active life-threatening replicating state which is characterized by the increased concentration of Replikins. In addition, clues to the viral pathogenesis of cancer may be found in the fact that glioma glycoprotein 10B has a 50% reduction in carbohydrate residues when compared to the normal 10B. This reduction is associated with virus entry in other instances and so may be evidence of the attachment of virus for the delivery of virus Replikins to the 10B of glial cells as a step in the transformation to the malignant state.

[0015] The sharing of immunological specificity by diverse members of the class, as demonstrated with antimalignin antibody for the glioma and related cancer Replikins, suggests that B cells and their product antibodies may recognize Replikins by means of a similar recognition 'language'. With the discovery of the Replikins, this shared immunological specificity may explain what was previously difficult to understand: why the antimalignin antibody is elevated in all cancers, and is cytotoxic to cancer cells and related to survival of cancer patients in most or all cell types. Thus antimalignin antibody is produced against cancer Replikins, which share immunological specificity and which are related to the phenomenon of rapid replication, not to cell type.

[0016] The recognition of the cancer Replikins, whether those in viruses known to cause cancer, or those in transforming proteins, or those isolated in cancer cell proteins (see Table 2, sections on cancer Replikins) is sufficiently general that the antimalignin antibody in serum test (AMAS Test) is an effective general cancer test. Yet there is sufficient individuality and difference in the fine structure (primary amino acid sequence) of each of the cancer replikins that they can be assayed specifically in tissues and in fluids by diagnostic methods common in the art, such as mass spectrometry. Once the particular Replikin is identified in the sequences of the cancer proteins by the methods discovered and described here, the Replikin is synthesized and acts as a standard for assays of tissues and fluids for the same structure. For example, the definitive and highly specific mass spectrometry analysis for one such cancer protein, the first defined cancer cell Replikin, the Glioma Replikin 'kagvaflhkk', is shown in Table 1. This specific measurement of the cancer Replikins permits the diagnostic specification of the tissue or organ type affected by the cancer and its specifc treatment. Thus, for example, the Glioma Replikin occurs only in glioblastoma multiforme, a malignant brain tumor, one of the most malignant of all tumors with a mortality of over 90%, for which no effective treatment is available. If its Glioma Replikin is measured in serum, the presence of brain malignancy is detected. It is also now possible to target the Glioma Replikin specifically with chemical, radiological and other treatments. The same novel diagnostic and therapeutic methods are now available for ovarian cancer Replikin and the other cancer Replikins as listed, by example, only in Table 2.

[0017] A second functional basis for the Replikins' role in rapid replication is the study of data from the past 100 years on influenza virus hemagglutinin protein sequences and epidemiology of influenza epidemics and pandemics. To date, only serological hemagglutinin and antibody classification, but no strain-specific conserved peptide sequences have previously been described in influenza, and no changes in concentration and composition of any strain-specific peptide sequences have been described previously which correlate with epidemiologically documented epidemics or rapid replication.

[0018] A four to ten-fold increase in the concentration of strain-specific influenza Replikins in one of each of the four major strains, influenza B, (A)H1N1, (A)H2N2 and (A)H3N2 was found, and that such increase of Replikin concentration was related to influenza epidemics caused specifically by each strain from 1902 to 2001. These increases in concentration were then shown to be due to the reappearance of at least one specific Replikin composition from 1 to up to 64 years after its disappearance, plus the emergence of new strain-specific Replikin compositions. Previously, no strain-specific chemical structures were known with which to predict which strains would predominate in coming influenza seasons, nor to devise annual mixtures of whole-virus strains for vaccines. The recent sharp increase in H3N2 Replikin concentration (1997 to 2000), the largest in H3N2's history, and the reappearance of specific Replikin compositions which were last seen in the high mortality H3N2 pandemic of 1968 and in the two high mortality epidemics of 1975 and 1977, but were absent for 20-25 years, together may be a warning of coming epidemics.

[0019] Synthetic Replikins are new vaccines. This high degree of conservation of Replikin structures observed whereby the identical structure can persist for 100 years, or reappear after an absence of from one to 64 years reappears indicates that what was previously thought to be change in virulence due to random substitution of amino acids in influenza proteins is more likely to be change due to an organized process of conservation of Replikins. In fact, if random substitutions of each amino acid occurred, the chance against an average length influenza

Replikin sequence being conserved for one year (let alone 84) is calculated to be in the order of 2 to the 27th power to 1.

[0020] The significant conservation of Replikins is not unique to influenza virus, for example, it is also present in foot and mouth disease virus type O and in HIV, as well as in wheat. More recently, significant conservation of Replikins is present in coronavirus nucleocapsid proteins.

[0021] A third functional basis for Replikins' role in rapid replication is the increase in Replikin concentration shown to be related to rapid replication in HIV. The Replikin concentration in the slow-growing low-titre strain of HIV (NS1, "Bru"), prevalent in early stage infection, was found to be one-sixth of the Replikin concentration in the rapidly-growing high-titre strain of HIV (SI, "Lai"), prevalent in late stage HIV infection.

[0022] Other examples are given of the relation of Replikins to rapid replication. For example, in tomato curl leaf gemini virus, which devastates tomato crops, the first 161 amino acids of the 'replicating protein', which have been shown to bind to DNA, contain five Replikins.

[0023] In malaria, legendary for rapid replication, trypanosomes are released from the liver in tens of thousands from one trypanosome. Multiple, novel, almost 'flamboyant' Replikin structures with concentrations of up to 111 overlapping Replikins per 100 amino acids are found therein.

[0024] The increase in Replikin concentration in influenza epidemics is functionally comparable to the glioma Replikin's increase in concentration during rapid replication of malignant glioma cells and comparable to rapid replication in HIV and in a diverse range of other organisms. Replikins thus are associated with and appear to be part of the structural bases of rapid replication in different organisms.

[0025] Replikin concentration and composition therefore provide new methods to detect and to control the process of replication, which is central to the survival and dominance of each biological population. The discovery of these new proteins related to rapid replication provides new opportunities 1) for detection of pathogens by qualitative and quantitative determinations of Replikins, 2) for the control of a broad range of diseases in which rapid replication is a key factor by targeting native Replikins and by using synthetic Replikins as vaccines, and 3) for the use of Replikins to foster growth of algal and plant foods.

[0026] There is a significant number of diseases and pathogens which have proved difficult to detect and treat and for which there is no effective vaccine. Thus, for each disorder there is a need for developing a target that will provide effective methods of detecting, treating or preventing these diseases and pathogens.

SUMMARY OF THE INVENTION

[0027] The present invention provides a method for identifying nucleotide or amino acid sequences that include a Replikin sequence. The method is referred to herein as a 3-point-recognition method. By use of the "3-point recognition" method, namely, peptides comprising from 7 to about 50 amino acids including (1) at least one lysine residue located six to ten amino acid residues from a second lysine residue; (2) at least one histidine residue; and (3) at least 6% lysine residues (Replikin) – constituting a new class of peptides was revealed in algae, yeast, fungi, amoebae, bacteria, plant and virus proteins having replication, transformation, or redox functions.

[0028] In one aspect of the invention there are provided isolated or synthesized peptides containing a Replikin sequence. The peptides comprise from 7 to about 50 amino acids including (1) at least one lysine residue located six to ten amino acid residues from a second lysine residues; (2) at least one histidine residue; and (3) at least 6% lysine residues.

[0029] The present invention also provides methods for detecting the presence of a contaminating organism in a body sample or environmental sample comprising:

(1) isolating nucleic acids from the body sample or environmental sample;

(2) screening the nucleic acids for the presence of a Replikin structure; and

(3) correlating the presence of a Replikin structure with the presence of the contaminating organism.

[0030] In another aspect of the invention there is provided a process for stimulating the immune system of a subject to produce antibodies that bind specifically to a Replikin sequence, said process comprising administering to the subject an effective amount of a dosage of a composition comprising at least one Replikin peptide. One embodiment comprises at least one peptide that is present in an emerging strain of the organism if such new strain emerges.

[0031] The present invention also provides antibodies that bind specifically to a Replikin, as defined herein, as well as antibody cocktails containing a plurality of antibodies that specifically bind to Replikins. In one embodiment of the invention, there are provided compositions comprising an antibody or antibodies that specifically bind to a Replikin and a pharmaceutically acceptable carrier.

[0032] In one aspect of the invention there are provided isolated, or separated from other proteins, recombinant, or synthesized peptides or other methods containing a viral Replikin sequence. The viral Replikin peptides comprise from 7 to about 50 amino acids including (1) at least one lysine residue located six to ten amino acid residues from a second lysine residue; (2) at least one histidine residue; and (3) at least 6% lysine residues. (viral Replikin).

[0033] The present application also provides isolated, or separated from nucleocapsid proteins, amongst others, recombinant, or synthesized peptides or other methods containing a viral Replikin sequence. The viral nucleocapsid Replikin peptides comprise from 7 to about 50 amino acids including (1) at least one lysine residue located six to ten amino acid residues from a second lysine residue; (2) at least one histidine residue; and (3) at least 6% lysine residues.

[0034] The present invention also provides methods for detecting the presence of a contaminating virus in a body sample or environmental sample comprising:

(1) isolating nucleic acids from the body sample or environmental sample;

(2) screening the nucleic acids for the presence of a viral Replikin structure; and

(3) correlating the presence of viral Replikin structures, their concentration and composition, with the presence of the contaminating virus.

[0035] In another aspect of the invention there is provided a process for stimulating the immune system of a subject to produce antibodies that bind specifically to a viral Replikin sequence, said process comprising administering to the subject an effective amount of a dosage of a composition comprising at least one Replikin peptide. One embodiment comprises at least one peptide that is present in an emerging strain of the virus if such new strain emerges.

[0036] The present invention also provides antibodies that bind specifically to a viral Replikin, as defined herein, as well as antibody cocktails containing a plurality of antibodies that specifically bind to viral Replikins. In one embodiment of the invention, there are provided compositions comprising an antibody or antibodies that specifically bind to a viral Replikin and a pharmaceutically acceptable carrier.

[0037] The present invention also provides therapeutic compositions comprising one or more of isolated virus peptides having from 7 to about 50 amino acids comprising: (1) at least one lysine residue located six to ten residues from a second lysine residue; (2) at least one histidine residue; and (3) at least 6% lysine residues, and a pharmaceutically acceptable carrier.

[0038] In another aspect of the invention there is provided an antisense nucleic acid molecule complementary to a virus Replikin mRNA sequence, said Replikin mRNA sequence denoting from 7 to about 50 amino acids comprising:

(1) at least one lysine residue located six to ten residues from a second lysine residue;

(2) at least one histidine residue; and

(3) at least 6% lysine residues.

[0039] In yet another aspect of the invention there is provided a method of simulating the immune system of a subject to produce antibodies to viruses, said method comprising: administering an effective amount of at least one virus Replikin peptide having from 7 to about 50 amino acids comprising (1) at least one lysine residue located six to ten amino acid residues from a second lysine residue; (2) at least one histidine residue; (3) and at least 6% lysine residues.

[0040] In another aspect, there is provided a method of selecting a virus peptide for inclusion in a preventive or therapeutic virus vaccine comprising:

(1) obtaining at least one isolate of each strain of a plurality of strains of said virus;

(2) analyzing the amino acid sequence of the at least one isolate of each strain of the plurality of strains of the virus for the presence and concentration of Replikin sequences;

(3) comparing the concentration of Replikin sequences in the amino acid sequence of the at least one isolate of each strain of the plurality of strains of the virus to the concentration of Replikin sequences observed in the amino acid sequence of each of the strains at least one earlier time period to provide the concentration of Replikins for at least two time periods, said at least one earlier time period being within about six months to about three years prior to step (1);

(4) indentifying the strain of the virus having the highest increase in concentration of Replikin sequences during the at least two time periods; and

(5) selecting at least one Replikin sequence present in the strain of the virus peptide identified in step (4) as a peptide for inclusion in the virus vaccine.

[0041] The present invention also provides a method of making a preventive or therapeutic virus vaccine comprising:

(1) identifying a strain of a virus as an emerging strain,

(2) selecting at least one Replikin sequence present in the emerging strain as a peptide template for the virus vaccine manufacture,

(3) synthesizing peptides having the amino acid sequence of the at least one Replikin sequence selected in step (2), and

(4) combining a therapeutically effective amount of the peptides of step (3) with a pharmaceutically acceptable carrier and/or adjuvant.

[0042] In another aspect, the invention is directed to a method of identifying an emerging strain of a virus for diagnostic, preventive or therapeutic purposes comprising:

(1) obtaining at least one isolate of each strain of a plurality of strains of the virus;

(2) analyzing the amino acid sequence of the at least one isolate of each strain of the plurality of strains of the virus for the presence and concentration of Replikin sequences;

(3) comparing the concentration of Replikin sequences in the amino acid sequence of the at least one isolate of each strain of the plurality of strains of the virus to the concentration of Replikin sequences observed in the amino acid sequence of each of the strains at least one earlier time period to provide the concentration of Replikins for at least two time periods, said at least one earlier time period being within about six months to about three years prior to step (1); and

(4) identifying the strain of the virus having the highest increase in concentration of Replikin sequences during the at least two time periods.

[0043] In yet another aspect of the invention, there is provided a preventive or therapeutic virus vaccine comprising at least one isolated Replikin present in a protein of an emerging strain of the virus and a pharmaceutically acceptable carrier and/or adjuvant.

[0044] Also provided by the present invention is a method of preventing or treating a virus infection comprising administering to a patient in need thereof a preventive or therapeutic virus vaccine comprising at least one isolated Replikin present in a protein of an emerging strain of the virus and a pharmaceutically acceptable carrier and/or adjuvant.

INFLUENZA

[0045] Influenza is an acute respiratory illness of global importance. Despite international attempts to control influenza virus outbreaks through vaccination, influenza infections remain an important cause of morbidity and mortality. Worldwide influenza epidemics and pandemics have occurred at irregular and previously unpredictable intervals throughout history and it is expected that they will continue to occur in the future. The impact of both pandemic and epidemic influenza is substantial in terms of morbidity, mortality and economic cost.

[0046] Influenza vaccines remain the most effective defense against influenza virus, but because of the ability of the virus to mutate and the availability of non-human host reservoirs, it is expected that influenza will remain an emergent or re-emergent infection. Global influenza surveillance indicates that influenza viruses may vary within a country and between countries and continents during an influenza season. Virological surveillance is of importance in monitoring antigenic shift and drift. Disease surveillance is also important in assessing the impact of epidemics. Both types of information have provided the basis of the vaccine composition and the correct use of antivirals. However, to date there has been only annual post hoc hematological classification of the increasing number of emerging influenza virus strains, and no specific chemical structure of the viruses has been identified as an indicator of approaching influenza epidemics or pandemics. Currently, the only basis for annual classification of influenza virus as active, inactive or prevalent in a given year is the activities of the virus hemagglutinin and neuraminidase proteins. No influenza viral chemical structure has been identified prior to this application that can be used for quantitative warning of epidemics or pandemics or to design more effective and safer vaccines.

[0047] Because of the annual administration of influenza vaccines and the short period of time when a vaccine can be administered, strategies directed at improving vaccine coverage are of critical importance.

[0048] In one aspect of the invention there are provided isolated or synthesized influenza virus peptides containing a Replikin sequence. The influenza Replikin virus peptides comprise from 7 to about 50 amino acids including (1) at least one lysine residue located six to ten amino acid residues from a second lysine residue; (2) at least one histidine residue; and (3) at least 6% lysine residues. (Influenza Replikin).

[0049] In another aspect of the invention, there is provided a process for stimulating the immune system of a subject to produce antibodies that bind specifically to an influenza virus Replikin sequence, said process comprising administering to the subject an effective amount of dosage of a composition comprising at least one influenza virus Replikin peptide. In a preferred embodiment the composition comprises at least on peptide that is present in an emerging strain of influenza virus.

[0050] The present invention also provides antibodies that bind specifically to an influenza virus Replikin, as defined herein, as well as antibody cocktails containing a plurality of antibodies that specifically bind to influenza virus Replikins. In one embodiment of the invention, there are provided compositions comprising an antibody or antibodies that specifically bind to an influenza Replikin and a pharmaceutically acceptable carrier.

[0051] The present invention also provides therapeutic compositions comprising one or more of isolated influenza virus peptides having from 7 to about 50 amino acids comprising:

(1) at least one lysine residue located six to ten residues form a second lysine residue;

(2) at least one histidine residue; and

(3) at least 6% lysine residues, and a pharmaceutical acceptable carrier.

[0052] In another aspect of the invention there is provided an antisense nucleic acid molecule complementary to an influenza virus hemagglutinin Replikin mRNA sequence, said Replikin mRNA sequence denoting from 7 to about 50 amino acids comprising:

(1) at least one lysine residue located six to ten residues from a second lysine residue;

(2) at least one histidine residue; and

(3) at least 6% lysine residues.

[0053] In yet another aspect of the invention there is provided a method of simulating the immune system of a subject to produce antibodies to influenza virus comprising administering an effective amount of at least one influenza virus Replikin peptide having from 7 to about 50 amino acids comprising:

(1) at least one lysine residue located six to ten amino acid residues from a second lysine residue;

(2) at least one histidine residue; and

(3) at least 6% lysine residues.

[0054] In another aspect, there is provided a method of selecting an influenza virus peptide for inclusion in a preventive or therapeutic influenza virus vaccine comprising:

(1) obtaining at least one isolate of each strain of a plurality of strains of influenza virus;

(2) analyzing the hemagglutinin amino acid sequence of the at least one isolate of each strain of the plurality of strains of influenza virus for the presence and concentration of Replikin sequences;

(3) comparing the concentration of Replikin sequences in the hemagglutinin amino acid sequence of the at least one isolate of each strain of the plurality of strains of influenza virus to the concentration of Replikin sequences observed in the hemagglutinin amino acid sequence of each of the strains at least one earlier time period to

provide the concentration of Replikins for at least two time periods, said at least one earlier time period being within about six months to about three years prior to step (1);

(4) identifying the strain of influenza virus having the highest increase in concentration of Replikin sequences during the at least two time periods;

(5) selecting at least one Replikin sequence present in the strain of influenza virus peptide identified in step (4) as a peptide for inclusion in an influenza virus vaccine.

[0055] The present invention also provides a method of making a preventive or therapeutic influenza virus vaccine comprising:

(1) identifying a strain of influenza virus as an emerging strain;

(2) selecting at least one Replikin sequence present in the emerging strain as a peptide template for influenza virus vaccine manufacture,

(3) synthesizing peptides having the amino acid sequence of the at least one Replikin sequence selected in step (2), and

(4) combining a therapeutically effective amount of the peptides of step (3) with a pharmaceutically acceptable carrier and/or adjuvant.

[0056] In another aspect, the invention is directed to a method of identifying an emerging strain of influenza virus for diagnostic, preventive or therapeutic purposes comprising:

(1) obtaining at least one isolate of each strain of a plurality of strains of influenza virus;

(2) analyzing the hemagglutinin amino acid sequence of the at least one isolate of each strain of the plurality of strains of influenza virus for the presence and concentration of Replikin sequences;

(3) comparing the concentration of Replikin sequences in the hemagglutinin amino acid sequence of the at least one isolate of each strain of the plurality of strains of influenza virus to the concentration of Replikin sequences observed in the hemagglutinin amino acid sequence of each of the strains at least one earlier time period to

provide the concentration of Replikins for at least two time periods, said at least one earlier time period being within about six months to about three years prior to step (1); and

(4) identifying the strain of influenza virus having the highest increase in concentration of Replikin sequences during the at least two time periods.

[0057] In yet another aspect of the invention, there is provided a preventive or therapeutic influenza virus vaccine comprising at least one isolated Replikin present in the hemagglutinin protein of an emerging strain of influenza virus and a pharmaceutically acceptable carrier and/or adjuvant.

[0058] Also provided by the present invention is a method of preventing or treating influenza virus infection comprising administering to a patient in need thereof a preventive or therapeutic vaccine comprising at least one isolated Replikin present in the hemagglutinin protein of an emerging strain of influenza virus and a pharmaceutically acceptable carrier and/or adjuvant.

TRYPANOSOMES

[0059] In one aspect of the invention there are provided isolated or synthesized trypanosome peptides containing a Replikin sequence. The trypanosome Replikin peptides comprise from 7 to about 50 amino acids including (1) at least one lysine residue located six to ten amino acid residues from a second lysine residue; (2) at least one histidine residue; and (3) at least 6% lysine residues. (Trypanosome Replikins).

MALARIA

[0060] One trypanosome disorder which has proved difficult to treat and for which there is no effective vaccine is malaria. Malaria causes much death, and physical and economic hardship in tropical regions. Malaria is caused mainly by Plasmodium falciparum, which has proved to be extremely resistant to treatment and to date, a vaccine for malaria has remained elusive. Thus there is a need for effective malaria vaccines and methods of treating or preventing the disease. This

application provides the basis for such vaccines and methods of treatment and prevention. All of the methods described above for production of and treatment with Replikin virus vaccines and Replikin influenza virus vaccines are applicable to the production of and treatment with Replikin malaria vaccines.

[0061] In the present invention, there are provided vaccines and methods for preventing or treating malaria. The malaria vaccines comprise at least one isolated Plasmodium falciparum Replikin. The present invention also provides methods for treating or preventing malaria comprising administering to a patient an effective amount of preventive or therapeutic vaccine comprising at least one isolated Plasmodium falciparum Replikin.

[0062] Also provided by the present invention are antibodies, antibody cocktails and compositions that comprise antibodies that specifically bind to a Replikin or Replikins present in a malaria antigen of Plasmodium falciparum.

[0063] Another example of a trypanosome which may be treated under the present invention as is the case for malaria, the Replikins of Treponema Pallidum (syphilis), can be used for detection, prevention, treatment of syphilis.

BACTERIA

[0064] In one aspect of the invention there are provided isolated or synthesized bacterial peptides containing a Replikin sequence (bacterial Replikins). The bacterial peptides comprise from 7 to about 50 amino acids including (1) at least one lysine residue located six to ten amino acid residues from a second lysine residue; (2) at least one histidine residue; and (3) at least 6% lysine residues. (bacterial Replikins). US Application Serial No. 10/105,232 filed March 26, 2002 is incorporated by reference in its entirety, including but not limited to the bacterial sequence listing and information.

[0065] The present invention also provides methods for detecting the presence of a contaminating bacterial organism in a body sample or environmental sample comprising:

(1) isolating nucleic acids from the body sample or environmental sample;

(2) screening the nucleic acids for the presence of a Replikin structure; and

(3) correlating the presence of a Replikin structure with the presence of the contaminating organism.

[0066] In another aspect of the invention there is provided a process for stimulating the immune system of a subject to produce antibodies that bind specifically to a bacterial Replikin sequence, said process comprising administering to the subject an effective amount of a dosage of a composition comprising at least one bacterial Replikin peptide. One embodiment comprises at least one bacterial peptide that is present in an emerging strain of the bacterial organism if such new strain emerges.

[0067] The present invention also provides antibodies that bind specifically to a bacterial Replikin, as defined herein, as well as antibody cocktails containing a plurality of antibodies that specifically bind to bacterial Replikins. In one embodiment of the invention, there are provided compositions comprising an antibody or antibodies that specifically bind to a bacterial Replikin and a pharmaceutically acceptable carrier.

[0068] The present invention also provides therapeutic compositions comprising one or more of isolated bacterial peptides having from 7 to about 50 amino acids comprising:

(1) at least one lysine residue located six to ten residues from a second lysine residue;

(2) at least one histidine residue;

(3) at least 6% lysine residues; and

(4) a pharmaceutically acceptable carrier.

[0069] In another aspect of the invention there is provided an antisense nucleic acid molecule complementary to a bacterial Replikin mRNA sequence, said Replikin mRNA sequence denoting from 7 to about 50 amino acids comprising:

(1) at least one lysine residue located six to ten residues from a second lysine residue;

(2) at least one histidine residue; and

(3) at least 6% lysine residues.

[0070] In yet another aspect of the invention there is provided a method of simulating the immune system of a subject to produce antibodies to bacteria comprising administering an effective amount of at least one bacterial Replikin peptide having from 7 to about 50 amino acids comprising:

(1) at least one lysine residue located six to ten amino acid residues from a second lysine residue;

(2) at least one histidine residue; and

(3) at least 6% lysine residues.

[0071] In another aspect, there is provided a method of selecting a bacterial Replikin peptide for inclusion in a preventive or therapeutic bacterial vaccine comprising:

(1) obtaining at least one isolate of each strain of a plurality of strains of the bacteria;

(2) analyzing the amino acid sequence of the at least one isolate of each strain of the plurality of strains of the bacteria for the presence and concentration of bacterial Replikin sequences;

(3) comparing the concentration of bacterial Replikin sequences in the amino acid sequence of the at least one isolate of each strain of the plurality of strains of the bacteria to the concentration of bacterial Replikin sequences observed in the amino acid sequence of each of the strains at least one earlier time period to provide the concentration of bacterial Replikins for at least two time periods, said at least one earlier time period being within about six months to about three years prior to step (1), or earlier in rapidly mutating bacteria;

(4) indentifying the strain of the bacteria having the highest increase in concentration of bacterial Replikin sequences during the at least two time periods; and

(5) selecting at least one bacterial Replikin sequence present in the strain of the bacterial peptide identified in step (4) as a peptide for inclusion in the bacterial vaccine.

[0072] The present invention also provides a method of making a preventive or therapeutic bacterial vaccine comprising:

(1) identifying a strain of a bacteria as an emerging strain;

(2) selecting at least one bacterial Replikin sequence present in the emerging strain as a peptide template for the bacterial vaccine manufacture;

(3) synthesizing peptides having the amino acid sequence of the at least one bacterial Replikin sequence selected in step (2); and

(4) combining a therapeutically effective amount of the peptides of step (3) with a pharmaceutically acceptable carrier and/or adjuvant.

[0073] In another aspect, the invention is directed to a method of identifying an emerging strain of bacteria for diagnostic, preventive or therapeutic purposes comprising:

(1) obtaining at least one isolate of each strain of a plurality of strains of the bacteria;

(2) analyzing the amino acid sequence of the at least one isolate of each strain of the plurality of strains of the bacteria for the presence and concentration of bacterial Replikin sequences;

(3) comparing the concentration of bacterial Replikin sequences in the amino acid sequence of the at least one isolate of each strain of the plurality of strains of the bacteria to the concentration of bacterial Replikin sequences observed in the amino acid sequence of each of the strains at least one earlier time period to provide the concentration of bacterial Replikins for at least two time periods, said at least one

earlier time period being within about six months to about three years prior to step (1); and

(4) identifying the strain of the bacteria having the highest increase in concentration of bacterial Replikin sequences during the at least two time periods.

[0074] In yet another aspect of the invention, there is provided a preventive or therapeutic bacterial vaccine comprising at least one isolated bacterial Replikin present in a protein of an emerging strain of the bacteria and a pharmaceutically acceptable carrier and/or adjuvant.

[0075] Two important sub-species of bacteria, classified under mycobacteria, are Mycobacterium leprae (leprosy) whose 30-s ribosomal protein has a C-terminal Replikin and Mycobacterium tuberculosis (tuberculosis) whose ATPase has three Replikins.

Replikin in 30s ribosomal protein s6 of Mycobacterium leprae (leprosy) is:
kvmrtdkh

Replikins in the ATPase of Mycobacterium tuberculosis are:
hprpkvaaalkdsyrlk
hprpkvaaalk
ksaqkwpdkflagaaqvah

Replikins in the B-D-galactosidase of E. coli:
hawqhqgktlfisrk
hqgktlfisrk

Replikins in Agrobacterium tumefaciens:
hsdqqlavmiaakrlddyk
hlldhpasvgqldlramlaveevkidnpvymek
hpasvgqldlramlaveevkidnpvymek
kevmakncnikepaglttnqeafngdpralaqylmniah

kncnikcpaglttnqcafngdpralaqylmniah

hhdtysiedlaqlihdakaarvrvivk

hdtysiedlaqlihdakaarvrvivk

hdakaarvrvivk

kigqgakpgcggqlpspkvtvciaaarggtpgvelvsppphh

kigqgakpgcggqlpspkvtvciaaarggtpgvelvsppph

kaseitktlasgamshgalvaaaheavahgtnmvggmsnsgcggch

kaseitktlasgamshgalvaaaheavah

kaseitktlasgamshgalvaaah

kaseitktlasgamsh

kryfpnvktpvggvtfaviaqavadwh

hhiaaglgfgasavyplgvqfraeekfgadadkafkrfakaackslmk

hhiaaglgfgasavyplgvqfraeekfgadadkafkrfakaaekslmk

hhiaaglgfgasavyplgvqfraeekfgadadkafkrfakaaek

hhiaaglgfgasavyplgvqfraeekfgadadkafkrfak

hhiaaglgfgasavyplgvqfraeekfgadadk

hiaaglgfgasavyplgvqfraeekfgadadkafkrfakaaekslmk

hiaaglgfgasavyplgvqfraeekfgadadkafkrfakaaek

hiaaglgfgasavyplgvqfraeekfgadadkafkrfak

hiaaglgfgasavyplgvqfraeekfgadadk

kfglydaafeksscgvgfitrkdgvqth

[0076] Also provided by the present invention is a method of preventing or treating a bacterial infection comprising administering to a patient in need thereof a preventive or therapeutic vaccine comprising at least one isolated bacterial Replikin present in a protein of an emerging strain of the bacteria and a pharmaceutically acceptable carrier and/or adjuvant.

FUNGUS

[0077] In one aspect of the invention there are provided isolated or synthesized fungal peptides containing a Replikin sequence. The fungal Replikin peptides comprise from 7 to about 50 amino acids including (1) at least one lysine residue

located six to ten amino acid residues from a second lysine residue; (2) at least one histidine residue; and (3) at least 6% lysine residues (fungal Replikins).

[0078] All of the methods described above for production of and treatment with bacterial Replikin vaccines are applicable to the production of and treatment with fungal Replikin vaccines.

[0079] In another aspect of the invention there is provided a process for stimulating the immune system of a subject to produce antibodies that bindspecifically to a *fungal Replikin sequence, said process comprising* administering to the subject an effective amount of a dosage of a composition comprising at least one fungal Replikin peptide.

[0080] The present invention also provides antibodies that bind specifically to a fungal Replikin, as defined herein, as well as antibody cocktails containing a plurality of antibodies that specifically bind to viral Replikins. In one embodiment of the invention, there are provided compositions comprising an antibody or antibodies that specifically bind to a fungal Replikin and a pharmaceutically acceptable carrier.

[0081] The present invention also provides therapeutic compositions comprising one or more of isolated fungal peptides having from 7 to about 50 amino acids comprising:

(1) at least one lysine residue located six to ten residues from a second lysine residue;

(2) at least one histidine residue;

(3) at least 6% lysine residues; and

(4) a pharmaceutically acceptable carrier.

[0082] In another aspect of the invention there is provided an antisense nucleic acid molecule complementary to an fungal Replikin mRNA sequence, said Replikin mRNA sequence having from 7 to about 50 amino acids comprising:

(1) at least one lysine residue located six to ten residues from a second lysine residue;

(2) at least one histidine residue; and

(3) at least 6% lysine residues.

[0083] In another aspect of the invention there is provided a process for stimulating the immune system of a subject to produce antibodies that bind specifically to a fungal Replikin sequence, said process comprising administering to the subject an effective amount of a dosage of a composition comprising at least one Replikin peptide.

INCREASING REPLICATION

[0084] In yet another aspect of the invention there is provided a method for increasing the replication rate of an organism comprising transforming a gene encoding an enzyme or other protein having a replication function in the organism with at least one Replikin structure.

DEFINITIONS

[0085] As used herein, the term "peptide" or "protein" refers to a compound of two or more amino acids in which the carboxyl group of one is united with an amino group of another, forming a peptide bond. The term peptide is also used to denote the amino acid sequence encoding such a compound. As used herein, "isolated" or "synthesized" peptide or biologically active portion thereof refers to a peptide that is substantially free of cellular material or other contaminating peptides from the cell or tissue source from which the peptide is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized by any method, or substantially free from contaminating peptides when synthesized by recombinant gene techniques.

[0086] As used herein, a Replikin peptide or Replikin protein is an amino acid sequence having 7 to about 50 amino acids comprising:

(1) at least one lysine residue located six to ten amino acid residues from a second lysine residue;

(2) at least one histidine residue;

(3) at least 6% lysine residues.

Similarly, a Replikin sequence is the amino acid sequence encoding such a peptide or protein.

[0087] As used herein, "emerging strain" as used herein refers to a strain of a virus, bacterium, fungus, or other organisms identified as having an increased increasing concentration of Replikin sequences in one or more of its protein sequences relative to the concentration of Replikins in other strains of such organism. The increase or increasing concentration of Replikins occurs over a period of at least about six months, and preferably over a period of at least about one year, most preferably over a period of at least about three years or more, for example, in influenza virus, but may be a much shorter period of time for bacteria and other organisms.

[0088] As used herein, "mutation" refers to change in this structure and properties of an organism caused by substitution of amino acids. In contrast, the term "conservation" as used herein, refers to conservation of particular amino acids due to lack of substitution.

BRIEF DESCRIPTION OF THE DRAWINGS

[0089] Figure 1 is a bar graph depicting the frequency of occurrence of Replikins in various organisms.

[0090] Figure 2 is a graph depicting the percentage of malignin per milligram total membrane protein during anaerobic replication of glioblastoma cells.

[0091] Figure 3 is a bar graph showing amount of antimalignin antibody produced in response to exposure to the recognin 16-mer.

[0092] Figure 4A is a photograph of a blood smear taken with ordinary and fluorescent light. Figure 4B is a photograph of a blood smear taken with ordinary and fluorescent light illustrating the presence of two leukemia cells. Figure 4C is a

photograph of a dense layer of glioma cells in the presence of antimalignin antibody. Figure 4D and Figure 4E are photographs of the layer of cells in Figure 4C taken at 30 and 45 minutes following addition of antimalignin antibody.

[0093] Figure 4F is a bar graph showing the inhibition of growth of small cell lung carcinoma cells in vitro by antimalignin antibody.

[0094] Figure 5 is a plot of the amount of antimalignin antibody present in the serum of patients with benign or malignant breast disease pre-and post surgery.

[0095] Figure 6 is a box diagram depicting an embodiment of the invention wherein a computer is used to carry out the 3-point-recognition method of identifying Replikin sequences.

[0096] Figure 7 is a graph showing the concentration of Replikins observed in hemagglutinin of influenza B and influenza A strain, H1N1, on a year by year basis from 1940 through 2001.

[0097] Figure 8 is a graph of the Replikin concentration observed in hemagglutinin of influenza A strains, H2N2 and H3N2, as well as an emerging strain defined by its constituent Replikins, designated H3N2(R), on a year by year basis from 1950 to 2001.

[0098] Figure 9 is a graph depicting the Replikin count per year for specific Replikin strains.

[0099] Figure 10 is a chart depicting the mean Replikin count per year for nucleocapsid coronavirus isolates.

[0100] Figure 11 is a chart depicting the Replikin count per year for H5N1 Hemagglutinins.

DETAILED DESCRIPTION OF THE INVENTION

[0101] The identification of a new family of small peptides related to the phenomenon of rapid replication, referred to herein as Replikins, provides targets for detection of pathogens in a sample and developing therapies, including vaccine development. In general, knowledge of and identification of this family of peptides enables development of effective therapies and vaccines for any organism that harbors Replikins. Identification of this family of peptides also provides for the detection of viruses and virus vaccine development.

[0102] For example, identification of this family of peptides provides for the detection of influenza virus and provides new targets for influenza treatment. Identification of this family of peptides also provides for example, for the detection of malaria and provides new targets for malaria vaccine development. Further examples provided by the identification of this family of peptides include the detection of infectious disease Replikins, cancer immune Replikins and structural protein Replikins.

[0103] Rapid replication is characteristic of virulence in certain bacteria, viruses and malignancies, but no chemistry common to rapid replication in different organisms has been described. We have found a family of conserved small protein sequences related to rapid replication, which we have named Replikins. Such Replikins offer new targets for developing effective detection methods and therapies. The first Replikin found was the glioma Replikin, which was identified in brain glioblastoma multiforme (glioma) cell protein called malignin.

[0104] Hydrolysis and mass spectrometry of malignin revealed the novel 16 mer peptide sequence which contains the glioma Replikin. This Replikin was not found in databases for the normal healthy human genome and therefore appeared to be derived from some source outside the body.

[0105] We have devised an algorithm to search for the glioma Replikin or homologue thereof. Homologues were not common in over 4,000 protein

sequences, but were found, surprisingly, in all tumor viruses, and in the replicating proteins of algae, plants, fungi, viruses and bacteria.

[0106] We have identified that both 1) Replikin concentration (number of Replikins per 100 amino acids) and 2) Replikin composition correlate with the functional phenomenon of rapid replication. These relationships provide functional basis for the determination that Replikins are related quantitatively as well as qualitatively to the rate of replication.

[0107] The first functional basis for Replikins role to rapid replication is seen in glioma replication. The fact that glioma malignin is enriched ten-fold compared to the five-fold increase in cell number and membrane protein concentration in rapid replication of glioma cells suggests an integral relationship of the Replikins to replication. When the glioma Replikin was synthesized in vitro and administered as a synthetic vaccine to rabbits, abundant antimalignin antibody was produced. This establishes the antigenic basis of the antimalignin antibody in serum (AMAS) test, and provides the first potential synthetic cancer vaccine and the prototype for Replikin vaccines in other organisms. With the demonstration of this natural immune relationship of the Replikins to replication and this natural immune response to cancer Replikins, which overrides cell type, based upon the shared specificity of cancer Replikins and rapid replication, both passive augmentation of this immunity with antimalignin antibody and active augmentation with synthetic Replikin vaccines now is possible.

[0108] The relationship between the presence of antimalignin antibody and survival in patients was shown in a study of 8,090 serum specimens from cancer patients. The study showed that the concentration of antimalignin antibody increases with age, as the incidence of cancer in the population increases, and increases further two to three-fold in early malignancy, regardless of cell type. In vitro, the antimalignin antibody is cytotoxic to cancer cells at picograms (femtomoles) per cancer cell, and in vivo the concentration of antimalignin antibody relates quantitatively to the survival of cancer patients. As shown in glioma cells, the stage in cancer at which cells have only been transformed to the immortal

malignant state but remain quiescent or dormant, now can be distinguished from the more active life-threatening replicating state, which is characterized by the increased concentration of Replikins. In addition, clues to the viral pathogenesis of cancer may be found in the fact that glioma glycoprotein 10B has a 50% reduction in carbohydrate residues when compared to the normal 10B. This reduction is associated with virus entry in other instances, and so may be evidence of the attachment of virus for the delivery of virus Replikins to the 10B of glial cells as a step in the transformation to the malignant state.

[0109] Our study concerning influenza virus hemagglutinin protein sequences and influenza epidemiology over the past 100 years, has provided a second functional basis for the relations of Replikins to rapid replication. Only serological hemagglutinin and antibody classification, but no strain-specific conserved peptide sequences have previously been described in influenza. Further, no changes in concentration and composition of any strain-specific peptide sequences have been described previously that correlate with epidemiologically documented epidemics or rapid replication. In this study, a four to ten-fold increase in the concentration of strain-specific influenza Replikins in one of each of the four major strains, influenza B, (A)HIN1, (A)H2N2 and, (A)H3N2 is shown to relate to influenza epidemics caused by each strain from 1902 to 2001.

[0110] We then showed that these increases in concentration are due to the reappearance of at least one specific Replikin composition from 1 to up to 64 years after its disappearance, plus the emergence of new strain-specific Replikin compositions. Previously, no strain-specific chemical structures were known with which to predict the strains that would predominate in coming influenza seasons, nor to devise annual mixtures of whole-virus strains for vaccines. The recent sharp increase in H3N2 Replikin concentration (1997 to 2000), the largest in H3N2's history, and the reappearance of specific Replikin compositions that were last seen in the high mortality H3N2 pandemic of 1968, and in the two high mortality epidemics of 1975 and 1977, but were absent for 20-25 years, together may be a warning of coming epidemics. This high degree of conservation of Replikin structures observed, whereby the identical structure can persist for 100 years, or

reappear after an absence of from one to 64 years, indicate that what was previously thought to be change due to random substitution of amino acids in influenza proteins is more likely to be change due to an organized process of conservation of Replikins.

[0111] The conservation of Replikins is not unique to influenza virus but was also observed in other sources, for example in foot and mouth disease virus, type 0, HIV tat, and wheat.

[0112] A third functional basis for Replikins' role in rapid replication is seen in the increase in rapid replication in HIV. Replikin concentration was shown to be related to rapid replication in HIV. We found the Replikin concentration in the slow growing low-titre strain of HIV (NS1, "Bru"), which is prevalent in early stage infection, to be one-sixth of the Replikin concentration in the rapidly-growing high-titre strain of HIV (SI, "Lai")(prevalent in late stage HIV infection).

[0113] Further examples demonstrate the relationship of Replikins to rapid replication. In the "replicating protein," of tomato curl leaf gemini virus, which devastates tomato crops, the first 161 amino acids, the sequence that has been shown to bind to DNA, was shown to contain five Replikins. In malaria, legendary for rapid replication when trypanosomes are released from the liver in the tens of thousands from one trypanosome, multiple, novel, almost 'flamboyant' Replikin structures have been found with concentrations of up to 36 overlapping Replikins per 100 amino acids.

[0114] The conservation of any structure is critical to whether that structure provides a stable invariant target to attack and destroy or to stimulate. When a structure is tied in some way to a basic survival mechanism of the organism, the structures tend to be conserved. A varying structure provides an inconstant target, which is a good strategy for avoiding attackers, such as antibodies that have been generated specifically against the prior structure and thus are ineffective against the modified form. This strategy is used by influenza virus, for example, so that a previous vaccine may be quite ineffective against the current virulent virus.

REPLIKINS AS STABLE TARGETS FOR TREATMENT

[0115] Both bacteria and HIV have both Replikin and non-Replikin amino acids. In HIV, for example, there has been a recent increase in drug-resistance from 9% to 13% due to mutation, that is substitution of non-Replikin amino acids. (See detailed analysis of TAT protein of HIV discussed herein). In bacteria, the development of 'resistant strains' is due to a similar mechanism. However, we have found that Replikin structures do not mutate or change to the same degree as non Replikin amino acids (see also discussion of foot and mouth disease virus conservation of Replikins discussed herein; further see discussion of conservation of coronavirus Replikins discussed herein). The Replikin structures, as opposed to the non-Replikin structures are conserved and thus provide new constant targets for treatment.

[0116] Certain structures too closely related to survival functions apparently cannot change constantly. Because an essential component of the Replikin structure is histidine (h), which is know for its frequent binding to metal groups in redox enzymes and probable source of energy needed for replication, and since this histidine structure remains constant, this structure remains all the more attractive a target for destruction or stimulation.

[0117] From a proteomic point of view, inventors construction of a template based on the newly determined glioma peptide sequence led them to the discovery of a wide class of proteins with related conserved structures and a particular function, in this case replication. Examples of the increase in Replikin concentration with virulence of a disease include, influenza, HIV, cancer and tomato leaf curl virus. This newly recognized class of structures is related to the phenomenon of rapid replication in organisms as diverse as yeast, algae, plants, the gemini curl leaf tomato virus, HIV and cancer.

[0118] Replikin concentration and composition provide new quantitative methods to detect and control the process of replication, which is central to the survival and dominance of each biological population. The sharing of immunological specificity by diverse members of the class, as demonstrated with

antimalignin antibody for the glioma and related cancer Replikins, suggests that B cells and their product antibodies may recognize Replikins by means of a similar recognition language.

[0119] Examples of peptide sequences of cancer Replikins or as containing a Replikin, i.e., a homologue of the glioma peptide, kagvaflhkk, may be found in such cancers of, but not limited to, the lung, brain, liver, soft-tissue, salivary gland, nasopharynx, esophagus, stomach, colon, rectum, gallbladder, breast, prostate, uterus, cervix, bladder, eye, forms of melanoma, lymphoma, leukemia, and kidney.

[0120] Replikins provide for: 1) detection of pathogens by qualitative and quantitative determinations of Replikins; 2) treatment and control of a broad range of diseases in which rapid replication is a key factor by targeting native Replikins and by using synthetic Replikins as vaccines; and 3) fostering increased growth rates of algal and plant foods.

[0121] The first Replikin sequence to be identified was the cancer cell Replikin found in a brain cancer protein, malignin, which was demonstrated to be enriched ten-fold during rapid anaerobic replication of glioblastoma multiforme (glioma) cells. (Figure 2) Malignin is a 10KDa portion of the 250 KDa glycoprotein 10B, which was isolated in vivo and in vitro from membranes of glioblastoma multiforme (glioma) cells. Hydrolysis and mass spectroscopy of malignin revealed a 16-mer peptide sequence, ykagvaflhkkndide (SEQ ID NO.:4), which is referred to herein as the glioma Replikin and which includes the shorter peptide, kagvaflhkk (SEQ ID NO.: 1), both of which apparently are absent in the normal human genome.

Table 1

16-mer peptide sequence ykagvaflhkkndide obtained from malignin by hydrolysis and mass spectrometry

Seq ID NO.	Fragment Identified	MH+ (mass)	Sequence	Method By Which Fragment Obtained			
				Auto-hydrolysis of malignin free in solution	Auto-hydrolysis of malignin immobilized on bromoacetyl cellulose	Micro-waved 5 seconds	Micro-waved 30 seconds
19	1-3	381.21	()yka(g)				+
20	1-5	537.30	()ykagv(a)		+		
21	2-6	445.28	(y)kagva(f)		+		
22	2-7	592.35	(y)kagvaf(1)			+	
23	4-11	899.55	(a)gvaflhkk(n)				+
24	5-7	336.19	(g)vaf(1)				+
25	6-7	237.12	(v)af(1)	+			
26	6-10	615.36	(v)aflhk(k)				+
27	6-10	615.36	(v)aflhk(k)	+			
28	6-12	857.50	(v)aflhkkn(d)		+		
29	6-12	857.50	(v)afhkkn(d)	+			
30	7-8	279.17	(a)fl(h)			+	
31	10-16	861.43	(h)kkndide()		+		
32	11-14	489.27	(k)kndi(d)		+		
33	12-15	476.2-	(k)ndid(e)	+			

[0122] When the 16-mer glioma Replikin was synthesized and injected as a synthetic vaccine into rabbits, abundant antimalignin antibody was produced. (Bogoch et al., Cancer Detection and Prevention, 26 (Suppl. 1): 402 (2002)). The concentration of antimalignin antibody in serum in vivo has been shown to relate quantitatively to the survival of cancer patients. (Bogoch et al., Protides of Biological Fluids, 31:739-747 (1984). In vitro antimalignin antibodies have been shown to be cytotoxic to cancer cells at a concentration of picograms (femtomolar) per cancer cell. (Bogoch et al., Cancer Detection and Prevention, 26 (Suppl. 1): 402 (2002).

[0123] Studies carried out by the inventors showed that the glioma Replikin is not represented in the normal healthy human genome. Consequently, a search for the origin and possible homologues of the Replikin sequence was undertaken by analysis of published sequences of various organisms.

[0124] By using the 16-mer glioma Replikin sequence as a template and constructing a recognition proteomic system to visually scan the amino acid sequences of proteins of several different organisms, a new class of peptides, the Replikins, was identified. The present invention provides a method for identifying nucleotide or amino acid sequences that include a Replikin sequence. The method is referred to herein as a 3-point-recognition method. The three point recognition method comprises: a peptide from 7 to about 50 amino acids including (1) at least one lysine residue located six to ten amino acid residues from a second lysine residue; (2) at least one histidine residue; and (3) at least 6% lysine residues. (Replikin). These peptides or proteins constitute a new class of peptides in species including algae, yeast, fungi, amoebae, bacteria, plant, virus and cancer proteins having replication, transformation, or redox functions. Replikin peptides have been found to be concentrated in larger 'replicating' and 'transforming' proteins (so designated by their investigators, See Table 2) and cancer cell proteins. No sequences were found to be identical to the malignin 16-mer peptide.

Table 2

Examples of Replikins in various organisms - prototype: Glioma Replikin* kagvaflhkk (SEQ ID No.:1)

	SEQ ID NO.		
Algae:	34	Caldophera prolifera	kaskftkh
	35	Isolepisprolifera	kaqaetgeikgh
Yeast:	36	Schizosaccharomyces pombe	ksflkypkkhk
	37	Oryza sativa	kkaygnelhk
	2	Sacch. cerevisiae replication binding protein	hsikrelgiifdk
Fungi:	38	Isocitrate lyase ICL1,Penicillium marneffei	kvdivthqk
	39	DNA-dependent RNA polymerase II, Discula destructiva	kleedaayhrkk
	40	Ophiostoma novo-ulm 1,RNA in Dutch elm disease fungus	kvilplrgnikgiffkh
Amoeba:	41	Entamoeba invadens, histone H2B	klilkgdlnkh
Bacteria:	42	Pribosomal protein replication factor, Helicobacter pylori Replication-associated protein Staph. aureus	ksvhaflk
	10	Mycoplasma pulmonic, chromosome replication	kkektthnk
	43	Macrophage infectivity potentiator, L. legionella	kvhffqlkk
	90	Bacillus anthracis	kihlisvkk
	91	Bacillus anthracis	hvkkekeknk
	92	Bacillus anthracis	khivkievk
	93	Bacillus anthracis	kkkkikdiygkdallh
	94	Bacillus anthracis	kwekikqh
	95	Bacillus anthracis	kklqipppiepkkddiih
	96	Bacillus anthracis	hnryasnivesayllilnew-knniqsdlikk
	97	Bacillus anthracis	havddyagyllidknqsdlv-tnskk
	98	Bacillus anthracis	haerlkvqknapk
Plants:	44	Arabidopsis thaliana, prolifera	kdhdfdgdk
	45	Arabidopsis thaliana, cytoplasmic ribosomal	kmkglkqkkah
	46	Arabidopsis thaliana, DNA binding protein	kelssttqeksh
Viruses:	9	Replication associated protein A [Maize streak virus]	kekkpskdeimrdiish
	11	Bovine herpes virus 4, DNA replication protein	hkinitngqk
	12	Meleagrid herpesvirus 1, replication binding protein	hkdlyrllmk
	47	Feline immunodeficiency	hlkdyklvk
	3	Foot and Mouth Disease (O)	hkqkivapvk
	5	HIV Type 1	kefnegkegh
	7	HIV Type 2	kewnegkegh
	99	Small Pox Virus (Variola)	khynnitwyk
	100	Small Pox Virus (Variola)	kysqtgkeliih
	101	Small Pox Virus (Variola)	hyddvrikndivvsrck
	102	Small Pox Virus (Variola)	hrfklildski
	103	Small Pox Virus (Variola)	kerghnyyfek

Tumor Viruses:	48	Rous sarcoma virus tyrosine-protein kinase	kklrhck
	49	v-yes, avian sarcoma	kklrhdk
	50	c-yes, colon cancer, malignant melanoma	kklrhdk
	51	v-srcC, avian sarcoma	kklrhck
	52	c-src, colon, mammary, panrcreatic cancer	kklrhck
	53	Neuroblastoma RAS viral (v-ras) oncogene	kqahclak
	54	VP1 (major capsid protein) [Polyamavirus sp.]	kthrfskh
	55	Sindbis	knlhckik
	56	E1 [Human papilloamavirus type 71]	khrpllqlk
	57	v-erbB from AEV and c-erb	kspnhvk
	58	v-fms (feline sarcoma)	knihlckk
	59	c-fms (acute and chronic myclomonocytic tumors)	knihlckk
	60	large t-antigen I [Polyomavirus sp.l	kphlaqslck
	61	middle t-antigen [Polyomavirus sp,]-	kqhrelkdk
	62	small t-antigen [Polyomavirus spJ,	kqhrelkdk
	63	v-abl, murine acute leukemia	kvpvtisptlkh
	64	Human T-cell lymphotropic virus typo 2	ksllievdkdish
	65	c-kit, GI tumors, small cell lung carcinoma	kagitimvkreyh
	18	*Hepatitis C*	hyppkpgcivpak
Trans-Forming Proteins:	66	Transforming protein myb	ksgkhlgk
	67	Transforming protein myc, Burkitt lymphoma	krreqlkhk
	68	Ras-related GTP-binding protein	ksfevikvih
	69	Transforming protein ras (teratocarcinoma)	kkkhtvkk
	70	TRAF-associated NF-kB activator TANK	kaqkdhlsk
	71	RFP transforming protein	hlkrvkdlkk
	72	Transforming protein D (S.C.)	kygspkhrlik
	73	Papilloma virus type 11, transforming protein	klkhilgkarfik
	74	Protein tryosine kinase (EC 2.7.1.l12slk	kgdhvkhykirk
	75	Transforming protein (axl(-))	keklrdvmvdrhk
	76	Transforming protein (N-myc)	klqarqqqllkkieh
	77	Fibroblast growth factor 4 (Kaposi sarcoma)	kkgnrvsptmkvth
Cancer Cell Proteins:	78	Matrix metaloproteinase 7 (uterine)	keiplhfrk
	79	Transcription factor 7-like	kkkphikk
	80	Breast cancer antigen NY-BR-87	ktrhdplak
	81	BRCA-1-Associated Ring Domain Protein (breast)	khhpkdnlik
	82	'Autoantigen from a breast tumor'	khkrkkfrqk
	83	Glioma Replikin (this study)	kagvaflhkk
	84	Ovarian cancer antigen	khkrkkfrqk
	85	EE L leukemia	kkksklhkdk
	86	Proto-oncogene tyrosine-protein kinase C-ABLE	hkselcpalprk
	87	Adenomatosis polyposis coli	kkkkpsrlkgdnck
	88	Gastric cancer transforming protein	ktkkgnrvsptmkvth
	89	Transforming protein (K-RAS 2B),lung	khkekmskdgkkkkkksk

[0125] Identification of an amino acid sequence as a Replikin or as containing a Replikin, i.e., a homologue of the glioma peptide, kagvaflhkk, requires that the three following requirements be met. According to the three point recognition system the sequences have three elements: (1) at least one lysine residue located six to ten residues from another lysine residue; (2) at least one histidine residue; and (3) a composition of at least 6% lysine within an amino acid sequence of 7 to about 50 residues.

[0126] Databases were searched using the National Library of Medicine keyword "PubMed" descriptor for protein sequences containing Replikin

sequences. Over 4,000 protein sequences were visually examined for homologues. Sequences of all individual proteins within each group of PubMed-classified proteins were visually scanned for peptides meeting the three above-listed requirements. An infrequent occurrence of homologues was observed in "virus peptides" as a whole (1.5%) (N=953), and in other peptides not designated as associated with malignant transformation or replication such as "brain peptides" and "neuropeptides" (together 8.5%) (N=845). However, surprisingly, homologues were significantly more frequently identified in large "replicating proteins," which were identified as having an established function in replication in bacteria, algae, and viruses. Even more surprising was the finding that Replikin homologues occurred in 100% of "tumor viruses" (N=250), in 97% of "cancer proteins" (N=401), and in 85% of "transforming viruses" (N=248). These results suggest that there are shared properties of cancer pathogenesis regardless of cell type and suggest a role of viruses in carcinogenesis, i.e., conversion of cells from a transformed albeit dormant state to a more virulent actively replicating state.

[0127] Homologues of the following amino acid sequence, kagvaflhkk, as defined by the three point recognition method, were found in such viruses, or viral peptides, as, but not limited to, adenovirus, lentivirus, a-virus, retrovirus, andeno-associated virus, human immunodeficiency virus, hepatitis virus, influenza virus, maize streak virus, herpes virus, bovine herpes virus, feline immunodeficiency virus, foot and mouth disease virus, small pox virus, rous sarcoma virus, neuroblastoma RAS viral oncogene, polyamavirus, sindbis, human papilloma virus, myelomonocytic tumor virus, murine acute leukemia, T-cell lymphotropic virus, and tomato leaf curl virus.

[0128] Furthermore, homologues of the amino acid sequence kagvaflhkk are present in known classes of coronavirus, which are members of a family of enveloped viruses that replicate in the cytoplasm of host cells. Additionally, the homologue of the amino acid sequence kagvatlhkk are present in the recently identified class of coronavirus responsible for severe acute respiratory syndrome, or SARS. The replikin is located in the nucleocapsid whole protein sequence of the SARS coronovirus. In addition, the location of the replikins is present in other

members of the coronavirus class and, more specifically, are also present in the nucleocapsid protein sequences from these coronaviruses.

[0129] Replikins are present in such bacteria as, but not limited to, Acetobacter, Achromobacter, Actinomyces, Aerobacter, Alcaligenes, Arthrobacter, Azotobacter, Bacillus, Brevibacterium, Chainia, Clostridium, Corynebacterium, Erwinia, Escheria, Lebsiella, Lactobacillus, Haemophilus, Flavobacterium, Methylomonas, Micrococcus, Mycobacterium, Micronomspora, Mycoplasma, Neisseria, Nocardia, Proteus, Pseudomonas, Rhizobium, Salmonella, Serratia, Staphylococcus, Streptocossus, Streptomyces, Streptosporangium, Streptovirticillium, Vibrio peptide, and Xanthomas.

[0130] Replikins are present in such fungi as, but not limited to, Penicillium, Discula, Ophiostoma novo-ulim, Mycophycophta, Phytophthora infestans, Absidia, Aspergillus, Candida, Cephalosporium, Fusarium, Hansenula, Mucor, Paecilomyces, Pichia, Rhizopus, Torulopsis, Trichoderma, and Erysiphe.

[0131] Replikins are present in such yeast as, but not limited to, Saccharomyces, Cryptococcus, including Cryptococcusneoformas, Schizosaccharomyces, and Oryza.

[0132] Replikins are present in algae such as, but not limited to, Caldophera, Isolepisprolifera, Chondrus, Gracilaria, Gelidium, Caulerpa, Laurencia, Cladophexa, Sargassum, Penicillos, Halimeda, Laminaria, Fucus, Ascophyllum, Undari, Rhodymenia, Macrocystis, Eucheuma, Ahnfeltia, and Pteroclasia.

[0133] Replikins are present in amoeba such as, but not limited to, Entamoeba (including Entamoeba invadens), Amoebidae, Acanthamoeba and Naegleria.

[0134] Replikins are present in plants such as, but not limited to, Arabidopsis, wheat, rice, and maize.

AUXILIARY SPECIFICATIONS

[0135] To permit classification of subtypes of Replikins, additional or "auxiliary specifications" to the basic "3-point-recognition" requirements may be added: (a) on a structural basis, such as the common occurrence of adjacent di- and polylysines in cancer cell proteins (e.g., transforming protein P21B(K-RAS 2B), lung, Table 2, SEQ ID NO.: 89), and other adjacent di-amino acids in TOLL-like receptors, or b) on a functional basis, such as exhibiting ATPase, tyrosine kinase or redox activity as seen in Table 2.

FUNCTIONAL DERIVATIVES

[0136] "Functional derivatives" of the Replikins as described herein are fragments, variants, analogs, or chemical derivatives of the Replikins, which retain at least a portion of the immunological cross reactivity with an antibody specific for the Replikin. A fragment of the Replikin peptide refers to any subset of the molecule. Variant peptides may be made by direct chemical synthesis, for example, using methods well known in the art. An analog of a Replikin to a non-natural protein substantially similar to either the entire protein or a fragment thereof. Chemical derivatives of a Replikin contain additional chemical moieties not normally a part of the peptide or peptide fragment.

[0137] As seen in Figure 2, during anaerobic respiration when the rate of cell replication is increased, malignin is enriched. That is, malignin is found to increase not simply in proportion to the increase in cell number and total membrane proteins, but is enriched as much as ten-fold in concentration, starting with 3% at rest and reaching 30% of total membrane protein. This clear demonstration of a marked increase in Replikin concentration with glioma cell replication points to, and is consistent with, the presence of Replikins identified with the 3-point recognition method in various organisms. For example, Replikins were identified in such proteins as "Saccharomyces cerevisiae replication binding protein" (SEQ ID NO.: 2) (hsikrelgiifdk); the "replication associated protein A of maize streak virus" (SEQ ID NO.: 8) (kyivcarcahk) and (SEQ ID NO.: 9) (kekkpskdeimrdiish); the "replication-associated protein of Staphylococcus aureus" (SEQ ID NO.: 10) (kkektthnk); the "DNA replication protein of bovine herpes virus 4" (SEQ ID NO.:

11) (hkinitngqk); and the "Mealigrid herpes virus 1 replication binding protein" (SEQ ID NO.: 12) (hkdlyrllmk). Previous studies of tomato leaf curl gemini virus show that the regulation of virus accumulation appears to involve binding of amino acids 1-160 of the "replicating protein" of that virus to leaf DNA and to other replication protein molecules during virus replication. Analysis of this sequence showed that amino acids 1-135 of this "replicating protein" contain a replikin count (concentration) as high as 20.7 (see section on tomato leaf curl Gemini virus.)

[0138] Table 2 shows that Replikin-containing proteins also are associated frequently with redox functions, and protein synthesis or elongation, as well as with cell replication. The association with metal-based redox functions, the enrichment of the Replikin-containing glioma malignin concentration during anaerobic replication, and the cytotoxicity of antimalignin at low concentrations (picograms/cell) (Figure 4c-f), all suggest that the Replikins are related to central respiratory survival functions, have been found less often subjected to the mutations characteristic of non-Replikin amino acids.

[0139] Of particular interest, it was observed that at least one Replikin per 100 amino acids was found to be present in the hemagglutinin proteins of almost all of the individual strains of influenza viruses examined. The Replikin sequences that were observed to occur in the hemagglutinin proteins of isolates of each of the four prevalent strains of influenza virus, influenza B, H1N1, H2N2, and H3N2, for each year that amino acid sequence data are availabe (1902-2001), are shown in Tables 3, 4, 5 and 6.

Table 3

Replikin Sequences present in hemagglutinins of Influenza B viruses in each year for which amino acid sequences were available (1940-2001).

Influenza B Replikins	Year Detected in Influenza B strain
Peak in Figure 7:	E
kshfanlk (SEQ ID NO. 104)	1940,43,51,59,75,76,77,89,90,93,97,98,99,00,01
kshfanlkgtk (SEQ ID NO. 105)	1940,43,51,59,75,76,77,89,90,93,97,98,99,00,01
kshfanlkgtktrgklcpk (SEQ ID NO. 106)	1940,43,51,59,75,76,77,89,90,93,97,98,99,00,01
hekygglnk (SEQ ID NO. 107)	1940,43,51,59,75,76,77,89,90,93,97,98,99,00,01
hekygglnksk (SEQ ID NO. 108)	1940,43,51,59,75,76,77,89,90,93,97,98,99,00,01
hekygglnkskpyytgehak (SEQ ID NO. 109)	1940,43,51,59,75,76,77,89,90,93,97,98,99,00,01
hakaigncpiwvk (SEQ ID NO. 110)	1940,43,51,59,75,76,77,89,90,93,97,98,99,00,01
hakaigncpiwvktplklangtk (SEQ ID NO. 111)	1940,43,51,59,75,76,77,89,90,93,97,98,99,00,01
hakaigncpiwvktplklangtkyrppak (SEQ ID NO. 112)	1940,43,51,59,75,76,77,89,90,93,97,98,99,00,01
hakaigncpiwvktplklangtkyrppakllk (SEQ ID NO. 113)	1940,43,51,59,75,76,77,89,90,93,97,98,99,00,01
k(a/v)silhevk (SEQ ID NO. 119)	1940, 59, 90,93
kvwcasgrskvikgslpligeadclh (SEQ ID NO. 123)	1940,43, 59,75,76,77,89,90, 98,99,00
kpyytgehak (SEQ ID NO. 124)	1940, 59, 89,90,93,97,98, 01
hgvavaadlkstqeaink (SEQ ID NO. 128)	1940, 59, 00
hgvavaadlkstqeainkdtistqeaink (SEQ ID NO. 129)	1940

hsdneiqmvklygdsk (SEQ ID NO. 116)	
hsdneiqdkmvklygdskpqk (SEQ ID NO. 117)	
kyggltkskpyytgeh (SEQ ID NO. 122)	
kcmgtipsakasilhevk (SEQ ID NO. 125)	1943, 75,76,77, 93
klygdskpqkftssangvtth (SEQ ID NO. 130)	1943, 75,76,77, 93,97, 00
hsdnetqmaklygdskpqk (SEQ ID NO. 131)	1943, 75,76,77, 93
hfanlkgtqtrgk (SEQ ID NO. 132)	1959
hfanlkgtktrgk (SEQ ID NO. 114)	1976, 89,90, 99,00,01
hfanlkgtktrgklcpk (SEQ ID NO. 115)	1976, 90 00,01
kprsalkckgfh (SEQ ID NO. 133)	1988
kctgtipsakasilhevk (SEQ ID NO. 121)	1993
hnvinaekapggpyk (SEQ ID NO. 126)	1993,97, 00
hsdnetqmaklygdsk (SEQ ID NO. 127)	1993,97, 00
hsdneiqmvklygdskpqk (SEQ ID NO. 118)	1997,98, 00
kctgtipsakasilh (SEQ ID NO. 120)	2000
kskpyytgehakai(g/a)ncpiwvk (SEQ ID NO. 134)	2000

1. Influenza B has not been responsible for any human pandemic.
2. Abbreviation for years: e.g., "43" = 1943, "01" = 2001.
3. The first year that a given Replikin appears is indicated at the beginning of the series of years in which that Replikin has been found.
4. Overlapping Replikin sequences are listed separately.
5. Return of replikins, absent for several years, in the two years before the epidemic of 1977, underlined, correlates with increased total Replikin concentration (Replikin Count = number of Replikins per 100 amino acid residues). See Figure 7.

Table 4

H1N1 Replikin Sequences present in H1N1 hemagglutinins of influenza viruses in each year for which amino acid sequences were available (1918-2000)

H1N1 Replikin *Year Detected in Influenza*

H1N1 Strain

Peak in Figure7 :	P1	E1	E1.1, 1.2, 1.3	E1.4	)
hp(v/t)tigecpkyv(r/k)(s/t)(t/a)k (SEQ ID NO. 135)	1918,25,28,30,31,35,47,48,51,52,55,56,57,59,63,77,79,80,81,85,87,88,89,91,92,95,96,97,98,99,00				
hdsnvknly(e/g)kv(k/r)(n/s)qk(k/r)nnak (SEQ ID NO. 136)	1918,	28,30,31,	77,79,80,	88, 91, 95,	98
hdsnvknly(e/g)kv(k/r)(n/s)qlk (SEQ ID NO. 137)	1918,	28,30,31,	77,79,80,	88, 91, 95,	98
hkc(nn/dd)(a/t/e)cmesv(r/k)ngtydypkyseesklnre(e/k)idgvk (SEQ ID NO. 138)	1918,	30, 35,	77, 80,		98
hkc(nn/dd)(a/t/e)cmesv(r/k)ngtydypkyseesk (SEQ ID NO. 139)	1918,	30, 35,	77, 80,		98
hqn(e/g)qgsgyaadqksiqnai(d/n)gitnkvnsviekmntqftavgkefnklek (SEQ ID NO. 140)	1918,	28,30,31,35,	59, 79,	95	
hqn(e/g)qgsgyaadqksiqnai(d/n)gitnkvnsviek (SEQ ID NO. 141)	1918,	28,30,31,35,	59, 79,	95	
hqn(e/g)qgsgyaadqksiqnai(d/n)gitnk (SEQ ID NO. 142)	1918,	28,30,31,35,	59, 79,	95	
kfeifpktsswpnh (SEQ ID NO. 143)	1918,		77		
kg(n/s/t)sypkl(n/s)ksy(v/t)nnkgkevlviwgvh (SEQ ID NO. 144)	1918,	35,	77,	96	
ksy(v/t)nnkgkevlviwgvh (SEQ ID NO. 145)	1918,	35,	77,	96	
hkcnnecmesvkngtydypkyseesklnrekidgvk (SEQ ID NO. 146)		1928, 31,		95	
hkcnnecmesvkngtydypkyseesk (SEQ ID NO. 147)		1928, 31,		95	
hkcnnecmesvkngtydypk (SEQ ID NO. 148)		1928, 31,		95	
hkcnnecmesvk (SEQ ID NO. 149)		1928, 31,		95	
hngkssfy(k/r)nilwlt(e/g)knglypnlsksyvnnkek (SEQ ID NO. 150)		1928,		95,	00
hngkssfy(k/r)nilwlt(e/g)knglypnlsksyvnnk (SEQ ID NO. 151)		1928, 31,		95,	00

hngkssfy(k/r)nllwlt(e/g)knglypnlsk (SEQ ID NO. 152) 1928, 31, 95, 00

hngkssfy(k/r)nllwlt(e/g)k (SEQ ID NO. 153) 1928, 31, 95, 00

kssfyknllwltekuglypnlsksyvnnkekevlvlwgvh (SEQ ID NO. 154) 1928, 31, 95

knllwltekuglypnlsksyvnnkekevlvlwgvh (SEQ ID NO. 155) 1928, 31, 95

knglypnlsksyvnnkekevlvlwgvh (SEQ ID NO. 156) 1928, 31, 95,96, 00

ksy(v/a)nnkekev(l/-)(v/-)lwgvh (SEQ ID NO. 157) 1928, 31, 51, 95,96, 98, 00

kesswpnhevtk (SEQ ID NO. 158) 1928, 31, 95

het(t/n)kgvtaacpyagassfyrnllwlvkkensypklsksyvnnk (SEQ ID NO. 159) 1930, 35

het(t/n)kgvtaacpyagassfyrnllwlvkkensypklsk (SEQ ID NO. 160) 1930, 35

kfeifpktsswpnevlvlwgvh (SEQ ID NO. 161) 1930

kerswpkh (SEQ ID NO. 162) 1947, 51,52,55,56, 79, 82

klsksyvnnkekevlvlwqvh (SEQ ID NO. 163) 1947, 51

knnkekevlvlwqvh (SEQ ID NO. 164) 1947

h(k/n)(g/q)kssfy(r/k)nllwltekng(i/s)yp(n/t)lsksyannkek (SEQ ID NO. 165) 1948 79, 89, 96

h(k/n)(g/q)kssfy(r/k)nllwltek (SEQ ID NO. 166) 1948 79, 89, 96

hakkssfyk (SEQ ID NO. 167) 1951, 57,59

hngklcrlkgk (SEQ ID NO. 168) 1951,52,55,56,57,59, 79,

hyklnn(q/g)kk (SEQ ID NO. 169) 1956, 00

hdiyrdeainnrfqiqgvkdtqgyk (SEQ ID NO. 170) 1956

kgngcfeifhk (SEQ ID NO. 171) 1956

klnrliektndkyhqiek (SEQ ID NO. 172) 1956

klnrliektndkyh (SEQ ID NO. 173) 1956

kchtdkgsisttk (SEQ ID NO. 174) 1956

kinngdyaklyiwgvh (SEQ ID NO. 175) 1956

hngklcrkgiaplqlgk (SEQ ID NO. 176) 1959, 82

hetnrqvtaacpyagansffrnliwlvkkessypklsk (SEQ ID NO. 177) 1963, 81

hetnrqvtaacpyagansffrnliwlvkkessypk (SEQ ID NO. 178) 1963, 81

hpptstdqqslyqnadayifvgsskynrkfk (SEQ ID NO. 179) 1963, 81

hpptstdqqslyqnadayifvgsskynrkfkpeia (SEQ ID NO. 180) 1963, 81

hdiyrdeainnrfqiqgvkitqgyk (SEQ ID NO. 181) 1977,79, 91

hqneqgsgyaadqkstqnaidgitnkvnsviekmntqftavgk (SEQ ID NO. 182) 1977

hqneqgsgyaadqkstqnaidgitnkvnsviek (SEQ ID NO. 183) 1977

hqneqgsgyaadqkstqnaingitnkvnsviekmntqftavgkefnklek (SEQ ID NO. 184) 1979, 91

hngklcrlkgiaplqlgk (SEQ ID NO. 185) 1979

hkcnnecmesvk (SEQ ID NO. 186) 1979

kfeifpkasswpnh (SEQ ID NO. 187) 1981

hdsnvknlyekvrsqlmnak (SEQ ID NO. 188) 1981

kvnsvikkmntqfaavgkefnh (SEQ ID NO. 189) 1981

khngklck (SEQ ID NO. 190) 1981

kkgtsypklsksythnkgkevlvlwgvh (SEQ ID NO. 191) 1981

kgtsypklsksythnkgkevlvlwgvh (SEQ ID NO. 192) 1981

klsksythnkgkevlvlwgvh (SEQ ID NO. 193) 1981

ksythnkgkevlvlwgvh (SEQ ID NO. 194) 1981

kgvtascshk (SEQ ID NO. 195) 1985,87

kgvtascshkgrssfyrnllwlteknglypnlsk (SEQ ID NO. 196) 1985,87

kgnsypklsksyvnnkekevlvlwgih (SEQ ID NO. 197) 1988

kefnhlek (SEQ ID NO. 198) 1988

hpptstdqqslyqnadayvfvgsskynkkfkpeiatrpk (SEQ ID NO. 199) 1988

hpptstdcqslyqnadayvfvgsskynkkfk (SEQ ID NO. 200)	1988
hegkssfymllwltekegsypkiknsyvnk (SEQ ID NO. 201)	1991
hegkssfymllwltekegsypk (SEQ ID NO. 202)	1991
hkcdnecmesvmgtydypkyseesk (SEQ ID NO. 203)	1991
kesswpnhtvtk (SEQ ID NO. 204)	1991,92
knliwlteknglypnlsksyvnnkekeilvlwgvh (SEQ ID NO. 205)	1991,92, 96
hngkssfy(k/m)(n/-)llwlt(e/g)(-/k)knglypnlsk (SEQ ID NO. 206)	1991,92, 96, 00
hngkssfyknilwltek (SEQ ID NO. 207)	1991,92, 96
htvtkgvtascshngkssfyknilwlteknglypnlsksyvnnkekevlvlwgvh (SEQ ID NO. 208)	1995
htvt(k/g)gv(t/s)ascshngkssfy(k/m)(n/-)llwlt(e/g)k(-n/k)glypnlsk (SEQ ID NO. 209)	1995, 00
htvtkgvtascshngkssfyknilwltek (SEQ ID NO. 210)	1995
kyvrstkdrnvtglrnipsiqsrglfgaiagfieggwtgmidgwygyh (SEQ ID NO. 211)	1995
hqneqgsgyaadqkstqnaingitnkvnsiiekmntqftavgk (SEQ ID NO. 212)	1995
hqneqgsgyaadqkstqnaingitnkvnsiiek (SEQ ID NO. 213)	1995
hqneqgsgyaadqkstqnaingitnk (SEQ ID NO. 214)	1995
hsgarsfymllwivkkgnsypk (SEQ ID NO. 215)	1996
hsgarsfymllwivkkgnsypklnk (SEQ ID NO. 216)	1996
hsgarsfymllwivkkgnsypklnksytndk (SEQ ID NO. 217)	1996
hsgarsfymllwivkkgnsypklnksytndkgk (SEQ ID NO. 218)	1996
htvskgvttscshngk (SEQ ID NO. 219)	1996
katswpnhetik (SEQ ID NO. 220)	1996
kqvttscshnqk (SEQ ID NO. 221)	1996
kgnsypklnksytndkgkevlviwgvh (SEQ ID NO. 222)	1996
klnksytndkgkevlviwgvh (SEQ ID NO. 223)	1996

ksytndkgkevlviwgvh (SEQ ID NO. 224)	<u>1996</u>
hnqkssfymllwlt(e/q)knglypnlsksy(v/a)annkek (SEQ ID NO. 225)	1997,98,99
hpitigecpkyvrsak (SEQ ID NO. 226)	1997
hqneqgsgyaadqkstqnaingitnkvnsviekmntqfeavgk (SEQ ID NO. 227)	1998
hqneqgsgyaadqkstqnaingitnkvnsviek (SEQ ID NO. 228)	1998
hngkssfymllwlteknglypnlsksyvnnkek (SEQ ID NO. 229)	1998

1. Influenza H1N1 was responsible for the human pandemic (global distribution) of 1918.
2. Abbreviation for years: eg. "96" = 1996.
3. The first year that a given Replikin appears is indicated at the beginning of the series of years in which that Replikin has been found in this work.
4. Overlapping Replikin sequences are listed separately.
5. *Increase in number of new Replikin structures occurs in years of epidemics (underlined): eg. 1918 and 1977 and correlates with increased total Replikin concentration (number of Replikins per 100 amino acid residues). See Figure 7.*

Table 5

Replikin Sequences present in hemagglutinins of Influenza H2N2 viruses in years 1957-2000

Influenza H2N2 Replikins | ***Year Detected in Influenza H2N2 strain***

(Peak in Figure 8: P2 E2)

khfekvkilpk (SEQ ID NO. 230)	1957,58,59,60,61,64,65,68, 78,83,84,91
khllssvkhfekvk (SEQ ID NO. 231)	1957,58,59,60,61, 83,84,91
ha(k/q/m)(d/n)ilekthagk (SEQ ID NO. 232)	1957,58,59,60,61,64,65,68, 78,83,84,91, 95
ha(k/q/m)(d/n)ilekthagklc(k/r) (SEQ ID NO. 233)	1957,58,59,60,61,64,65,68, 78,83,84,91, 95
hnvhpltigecpkyvksek (SEQ ID NO. 234)	1957,58,59, 65,68
hpltigecpkyvksek (SEQ ID NO. 235)	1957,58,59, 65,68,64,65,68,78,83,84,91
khlissvkhfekvkilpk (SEQ ID NO. 236)	1957,58,59,60,61,64,65,68, 78
krqssgimktegtlencetkcqtplgainttlpfhnvh (SEQ ID NO. 237)	1957, 59, 83
kgsnyp(v/i)ak(g/r)synntsgeqmliiwq(v/i)h (SEQ ID NO. 238)	1957,58,59, 61, 83, 91, 95
httlgqsracavsgnpsffrnmvwltekgsnypvak (SEQ ID NO. 239)	1957
khfekvk (SEQ ID NO. 240)	1957, 59, 65
kiskrgssgimktegtlencetkcqtplgainttlpfh (SEQ ID NO. 241)	1957, 59, 65, 91
krgssgimktegtlencetkcqtplgainttlpfh (SEQ ID NO. 242)	1957, 59, 65, 91
ktegtlencetkcqtplgainttlpfh (SEQ ID NO. 243)	1957, 59, 65, 91
kiskrgssgimktegtlencetkcqtplgainttlpfh (SEQ ID NO. 244)	1957, 59, 65, 91
ktegtlencetkcqtplgainttlpfhn(v/i)h (SEQ ID NO. 245)	1957, 59, 65, 91
kiskrgssgimktegtlencetkcqtplgainttlpfh (SEQ ID NO. 246)	1957, 59, 65, 91
k(e/g)snypvakgsynntsgeqmliiwgvh (SEQ ID NO. 247)	1957, 60, 65

hpltigecpkyvksek (SEQ ID NO. 248)	1957,	60,	65	
kcqtplgaikttlpfh (SEQ ID NO. 249)	1957,		65	
hhsndqgsgyaadkestqka(f/i)dgitnkvnsviek-				
-mntqfeavgkff(n/s)nlekleninkk (SEQ ID NO. 250)		1961,	65,68,	83,84
hsndqgsgyaadkestqka(f/i)dgitnkvnsviek-				
-mntqfeavgkff(n/s)nlekleninkk (SEQ ID NO. 251)		1961,	65, 68,	83,84
hsndqgsgyaadkestqka(f/i)dgitnk (SEQ ID NO. 252)		1961,	65, 68,	83,84
hdsnvrnlydkvrmqlrdnak (SEQ ID NO. 253)		1964,	68,76,	84,91
hkcddecmnsvkngtydypkinrneikgvk (SEQ ID NO. 254)		1964,	65,68,76,	83,84,91
hkcddecmnsvkngtydypklnrneik (SEQ ID NO. 255)		1964,	65,68,76,	83,84,91
hkcddecmnsvkngtydypk (SEQ ID NO. 256)		1964,	65,68,76,	83,84,91
hkcddecmnsvk (SEQ ID NO. 257)		1964,	65,68,76,	83,84,91
kgsnypvakgsynntngeqiiiiwgvh (SEQ ID NO. 258)			1976,78	
hsndqgsgyaadkestqkavdgitnkvnsviekmntqfeavgk (SEQ ID NO. 259)			1976,	91
krgssginktegtlencetkcqtplgainttlpfh (SEQ ID NO. 260)			1976,78,	83,84
hpltigecpkyvksek (SEQ ID NO. 261)			1976	
hakdilekthngkick (SEQ ID NO. 262)			1976	

1. Influenza H2N2 was responsible for the human pandemic (global distribution) of 1957.

2. Abbreviation for years: eg. "58" =1958.

3. The first year that a given Replikin appears is indicated at the beginning of the series of years in which that Replikin has been found in this work.

4. Overlapping Replikin sequences are listed separately.

5. Increase in number of new Replikin structures occurs in years of epidemics (underlined): eg. 1957 and 1965 and correlates with increased total Replikin concen tration (number of Replikins per 100 amino acid residues). See Figure 8.

Table 6

H3N2 Replikin Sequences present in H3N2 hemagglutinins of Influenza viruses in each year for which amino acid sequences were available (1968-2000)

Influenza H3N2 Replikins	Year Detected in Influenza H3N2 strain Influenza Replikins (Peak in Figure 8: P3 E3 E4)
hdvyrdealnnrfqikgvelksgyk (SEQ ID NO. 263)	1968,72, 75 96,97,98
htidltdsemnkdfertrk (SEQ ID NO. 264)	1968
kfhqiek (SEQ ID NO. 265)	1968,72, 75, 77 96,97,98
ktnekfh(g/q)iek (SEQ ID NO. 266)	1968 86 98
klnr(v/l)iektnekfh (SEQ ID NO. 267)	1968,72, 75, 77 97,98
hqiekefsevegriqdlekyvedtk (SEQ ID NO. 268)	1968,72, 98
kicnnphk (SEQ ID NO. 269)	1975
klnrvikknekfh (SEQ ID NO. 270)	1975
hd(i,v)yrdealnnrfqik(g/q)ve(r/k)s(q/g)yk (SEQ ID NO. 271)	1975,76,77, 86
hqiekefsevegriqdlekyvedtk (SEQ ID NO. 272)	1975
kyvedtkidlwsynaellvalenqh (SEQ ID NO. 273)	1975
kyvkqnslkltgmrnvpekqtrglfgaiagfiengwegmidgwygfrh (SEQ ID NO. 274)	1975
kefsevegriqdlekyvedtkidlwsynaellvalenqh (SEQ ID NO. 275)	1975 2000
hqn(s/e)(e/q)g(t/s)g(q/y)aad(l/q)k--stq(a/n)a(i/l)d(q/g)l(n/t)(g/n)k(l/v)n(r/s)vi(e/c)k (SEQ ID NO. 276)	1975 2000
hcd(g/q)f(q,r)nekwdlf(v,/i)er(s/t)k (SEQ ID NO. 277)	1975,76,77,78,80,81,82,83,84,85,86,88,89,90,91,92,93,94,95,96,97,98

htidltdsennkklfertrk (SEQ ID NO. 278)	1977,
ksgstypvlkvtmpnndnfdklyiwgvh (SEQ ID NO. 279)	1977
klnwltksgntypvlnvtmpnndnfdklviwgvh (SEQ ID NO. 280)	1982
htidltdsennklfektrk (SEQ ID NO. 281)	1986
klnrliektnekfhqtek (SEQ ID NO. 282)	1987
htgkssvmrsdapidfcnsecitpnqsipndkpfqnvnkitygacpk (SEQ ID NO. 283)	1994
htgkssvmrsdapidfcnsecitpnqsipndkpfqnvnk (SEQ ID NO. 284)	1994
hpstdsdqtslyvrasgrvtvstkrsqqtvipk (SEQ ID NO. 285)	1994
kyvedtkidlwsynaellvalenqh (SEQ ID NO. 286)	1997,98
klfertrkqlrenaedmgngcfkiyh (SEQ ID NO. 287)	1998
krrsiksffsrlnwlh (SEQ ID NO. 288)	1998
hpvtigecpky(v/r)kstk (SEQ ID NO. 289)	2000
kgnsypklsklsksyiinkkkevlviwgih (SEQ ID NO. 290)	2000
klsklsks(v/y)iinkkkevlviwgih (SEQ ID NO. 291)	2000
klsks(v/y)iinkkkevlviwgih (SEQ ID NO. 292)	2000

1. Influenza H3N2 was responsible for the human pandemic (global distribution) of 1968.
2. Abbreviation for years: eg. "77" =1977.
3. The first year that a given Replikin appears is indicated at the beginning of the series of years in which that Replikin has been found.
4. Overlapping Replikin sequences are listed separately.
5. Increase in number of new Replikin structures occurs in years of epidemics (underlined) : eg. 1975 and correlates with increased total Replikin concentration (number of Replikins per 100 amino acid residues). See **Figure 8**.

[0140] Both the concentration and type, *i.e.*, composition of Replikins observed, were found to relate to the occurrence of influenza pandemics and epidemics. The concentration of Replikins in influenza viruses was examined by visually scanning the hemagglutinin amino acid sequences published in the National Library of Medicine "PubMed" data base for influenza strains isolated world wide from human and animal reservoirs year by year over the past century, i.e., 1900 to 2001. These Replikin concentrations (number of Replikins per 100 amino acids, mean +/- SD) were then plotted for each strain.

[0141] The concentration of Replikins was found to directly relate to the occurrence of influenza pandemics and epidemics. The concentration of Replikins found in influenza B hemagglutinin and influenza A strain, H1N1, is shown in Figure 7, and the concentration of Replikins found in the two other common influenza virus A strains, H2N2 and H3N2 is shown in Figure 8 (H2N2, H3N2). The data in Figure 8 also demonstrate an emerging new strain of influenza virus as defined by its constituent Replikins (H3N2(R)).

[0142] Each influenza A strain has been responsible for one pandemic: in 1918, 1957, and 1968, respectively. The data in Figures 7 and 8 show that at least one Replikin per 100 amino acids is present in each of the influenza hemagglutinin proteins of all isolates of the four common influenza viruses examined, suggesting a function for Replikins in the maintenance of survival levels of replication. In the 1990s, during the decline of the H3N2 strain, there were no Replikins in many isolates of H3N2, but a high concentration of new Replikins appeared in H3N2 isolates, which define the emergence of the H3N2(R) strain.

[0143] Several properties of Replikin concentration are seen in Figure 7 and Figure 8 to be common to all four influenza virus strains. First, the concentration is cyclic over the years, with a single cycle of rise and fall occurring over a period of two to thirty years. This rise and fall is consistent with the known waxing and waning of individual influenza virus strain predominance by hemagglutinin and neuraminidase classification. Second, peak Replikin concentrations of each influenza virus strain previously shown to be responsible for a pandemic were

observed to relate specifically and individually to each of the three years of the pandemics. For example, for the pandemic of 1918, where the influenza virus strain, H1N1, was shown to be responsible, a peak concentration of the Replikins in H1N1 independently occurred (P1); for the pandemic of 1957, where H2N2 emerged and was shown to be responsible, a peak concentration of the Replikins in H2N2 occurred (P2); and for the pandemic of 1968, where H3N2 emerged and was shown to be the cause of the pandemic, a peak concentration of the Replikins in H3N2 occurred (P3). Third, in the years immediately following each of the above three pandemics, the specific Replikin concentration decreased markedly, perhaps reflecting the broadly distributed immunity generated in each case. Thus, this post-pandemic decline is specific for H1N1 immediately following the pandemic (P1) for which it was responsible, and is not a general property of all strains at the time. An increase of Replikin concentration in influenza B repeatedly occurred simultaneously with the decrease in Replikin concentration in H1N1, e.g., EB1 in 1951 and EB2 in 1976, both associated with influenza B epidemics having the highest mortality. (Stuart-Harris, et al., Edward Arnold Ltd. (1985). Fourth, a secondary peak concentration, which exceeded the primary peak increase in concentration, occurred 15 years after each of the three pandemics, and this secondary peak was accompanied by an epidemic: 15 years after the 1918 pandemic in an H1N1 'epidemic' year (E1); eight years after the 1957 pandemic in an H2N2 'epidemic' year (E2); and occurred seven years after the 1968 pandemic in an H3N2 'epidemic' year (E3). These secondary peak concentrations of specific Replikins may reflect recovery of the strain. Fifth, peaks of each strain's specific Replikin concentration frequently appear to be associated with declines in Replikin concentration of one or both other strains, suggesting competition between strains for host sites. Sixth, there is an apparent overall tendency for the Replikin concentration of each strain to decline over a period of 35 years (H2N2) to 60 years (influenza B). This decline cannot be ascribed to the influence of vaccines because it was evident in the case of influenza B from 1940 to 1964, prior to common use of influenza vaccines. In the case of influenza B, Replikin recovery from the decline is seen to occur after 1965, but Replikin concentration declined again between 1997 and 2000 (Figure 7). This correlates with the low occurrence of influenza B in recent case isolates. H1N1 Replikin concentration peaked in 1978-1979 (Figure 7)

together with the reappearance and prevalence of the H1N1 strain, and then peaked in 1996 coincident with an H1N1 epidemic. (Figure 7). H1N1 Replikin concentration also declined between 1997 and 2000, and the presence of H1N1 strains decreased in isolates obtained during these years. For H2N2 Replikins, recovery from a 35 year decline has not occurred (Figure 8), and this correlates with the absence of H2N2 from recent isolates. For H3N2, the Replikin concentration of many isolates fell to zero during the period from 1996 to 2000, but other H3N2 isolates showed a significant, sharp increase in Replikin concentration. This indicates the emergence of a substrain of H3N2, which is designated herein as H3N2(R).

[0144] Figures 7 and 8 demonstrate that frequently, a one to three year stepwise increase is observed before Replikin concentration reaches a peak. This stepwise increase proceeds the occurrence of an epidemic, which occurs concurrently with the Replikin peak. Thus, the stepwise increase in concentration of a particular strain is a signal that particular strain is the most likely candidate to cause an epidemic or pandemic.

[0145] Currently, Replikin concentration in the H3N2(R) strain of influenza virus is increasing (Figure 8, 1997 to 2000). Three similar previous peak increases in H3N2 Replikin concentration are seen to have occurred in the H3N2-based pandemic of 1968 (Figure 8), when the strain first emerged, and in the H3N2-based epidemics of 1972 and 1975 (Figure 8). Each of these pandemic and epidemics was associated with excess mortality. (Ailing, et al., Am J. Epidemiol.,113(1):30-43 (1981). The rapid ascent in concentration of the H3N2(R) subspecies of the H3N2 Replikins in 1997-2000, therefore, statistically represents an early warning of an approaching severe epidemic or pandemic. An H3N2 epidemic occurred in Russia in 2000 (Figure 8, E4); and the CDC report of December 2001 states that currently, H3N2 is the most frequently isolated strain of influenza virus worldwide. (Morbidity and Mortality Weekly Reports (MMWR), Center for Disease Control; 50(48):1084-68 (Dec.7, 2001).

[0146] In each case of influenza virus pandemic or epidemic new Replikins emerge. There has been no observation of two of the same Replikins in a given hemagglutinin in a given isolate. To what degree the emergence of a new Replikin represents mutations versus transfer from another animal or avian pool is unknown. In some cases, each year one or more of the original Replikin structures is conserved, while at the same time, new Replikins emerge. For example, in influenza virus B hemagglutinin, five Replikins were constantly conserved between 1919 and 2001, whereas 26 Replikins came and went during the same period (some recurred after several years absence). The disappearance and re-emergence years later of a particular Replikin structure suggests that the Replikins return from another virus host pool rather than through de novo mutation.

[0147] In the case of H1N1 Replikins, the two Replikins present in the P1 peak associated with the 1918 pandemic were not present in the recovery E1 peak of 1933, which contains 12 new Replikins. Constantly conserved Replikins, therefore, are the best choice for vaccines, either alone or in combination. However, even recently appearing Replikins accompanying one year's increase in concentration frequently persist and increase further for an additional one or more years, culminating in a concentration peak and an epidemic, thus providing both an early warning and time to vaccinate with synthetic Replikins (see for example, H1N1 in the early 1990's, Figure 7; see also, for example, H5N1 1995-2002, FIGURE 11, "Replikin Count" (number of Replikins per 100 amino acids) refers to Replikin concentration).

[0148] The data in FIGURES 7, 8 and 11 demonstrate a direct relationship between the presence and concentration of a particular Replikin in influenza protein sequences and the occurrence of pandemics and epidemics of influenza. Thus, analysis of the influenza virus hemagglutinin protein sequence for the presence and concentration of Replikins provides a predictor of influenza pandemics and/or epidemics, as well as a target for influenza vaccine formulation. It is worth nothing again (see paragraph **[0109]**) with reference to this data, previously, no strain-specific chemical structures were known with which to predict the strains that would

predominate in coming influenza seasons, nor to devise annual mixtures of whole-virus strains for vaccines.

[0149] Similar to the findings of strain-specific Replikin Count increases in the influenza group one to three years prior to the occurrence of a strain-specific epidemics, the increase in Replikin Count of the coronavirus nucleocapsid protein has also been identified. Replikin Counts of the coronavirus nucleocapsid protein has increased as follows: 3.1 (±1.8) in 1999; 3.9(±1.2) in 2000; 3.9 (±1.3) in 2001; and 5.1 (±3.6) in 2002. This pre-pandemic increase supports the finding that a coronavirus is responsible for the current (2003) SARS pandemic. (See Table 7)

[0150] Thus, monitoring Replikin structure and Replikin Count provides a means for developing synthetic strain-specific preventive vaccination and antibody therapies against the 1917-1918 Goose Replikin and its modified and accompanying Replikins as observed in both influenza and coronavirus strains.

[0151] Figure 10 depicts the automated Replikin analysis of nucleocapsid coronavirus proteins for which the protein sequence is available on isolates collected from 1962 to 2003. Each individual protein is represented by an accession number and is analyzed for the presence of Replikins. The Replikin Count (number of Replikins per 100 amino acid) is automatically calculated as part of the automated Replikin analysis. For each year, the mean (± Standard deviation (S.D.)) Replikin Count per year is automatically calculated for all Replikin Counts that year. This example of early warning of increasing replication, before an epidemic, of a particular protein (the nucleocapsid protein) in a particular virus strain (the coronavirus) is comparable to the increase seen in strains of influenza virus preceding influenza epidemics and pandemics (FIGURES 7, 8 and 11). It may be seen that the Replikin Count rose from 1999 to 2002, consistent with the SARS coronavirus pandemic, which emerged at the end of 2002 and has persisted into 2003. Figure 9 provides a graph of the Replikin Counts for several virus strains, including the coronavirus nucleocapsid Replikin, from 1917 to 2002.

Table 7

Replikin Sequence																													'Multi-K' Replikins: Length	% Untreated Mortality	ORGANISM
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			Amino Acid position
A. INFLUENZA, SARS AND OTHER CORONAVIRUSES																															
k	k	g	t	s	y	p	k	l	s	k	s	y	t	n	n	k	g	k	e	v	l	v	i	w	g	v	h	h	29		1917-18 Goose Replikin (SEQ ID NO. 743)
k	k	g	t	s	y	p	k	l	s	k	s	y	t	n	n	k	g	k	e	v	l	v	i	w	g	v	h	h	29	2.5	1918 Human Influenza (SEQ ID NO. 744)
l	k	e	d	l	y	p	k	l	r	k	s	v	v	h	n	k	k	k	e	v	l	v	i	w	g	i	h	h	29		1919-2001 H1N1, H1N2 (SEQ ID NO. 745)
l	k	e	n	s	y	p	k	l	r	k	s	i	i	i	n	k	k	k	e	v	l	v	i	w	g	i	h	h			H3N2 Influenza (SEQ ID NO. 746)
k	k	g	t	s	y	p	k	l	s	k	s	y	t	n	n	k	k	k	e	v	l	v	i	w	g	v	h	h	29		2001 H1N2 Influenza (SEQ ID NO. 747)
k	k	n	s	a	y	p	t	l	k	r	s	y	n	n	t	n	q	e	d	i	l	v	i	w	g	i	h	h	>37		1996-2001 H5N1 Influenza (SEQ ID NO. 748)
k	k	s	a	k	t	g	t	p	k	p	s	r	n	q	s	p	a	s	s	q	t	s	a	k	s	l	a	h	>37		2000 Human coronavirus 229E (SEQ ID NO. 794)[1]
k	k	l	g	v	d	t	e	k	q	q	q	r	s	k	s	k	e	r	s	n	s	k	t	r	d	t	t	p	>37		2003 Canine coronavirus (SEQ ID NO. 795)[2]
k	n	g	l	y	p	n	l	s	k	s	y	a	n	n	k	e	k	e	v	l	i	l	w	g	v	h	h		28		2002 H1N2 (SEQ ID NO. 749)
k	k	i	n	s	p	q	p	k	f	e	g	s	g	v	p	d	n	e	n	l	k	t	s	q	q	h			27		Avian bronchitis coronavirus (SEQ ID NO. 715)
k	t	g	n	a	k	l	q	r	k	k	e	k	k	n	k	r	e	t	t	l	q	q	h						24		Porcine epidemic diarrhea coronavirus (SEQ ID NO. 716)
k	h	l	d	a	y	k	t	f	p	p	t	e	p	k	k	d	k	k	k	k									21		2003 Human SARS nucleocapsid (SEQ ID NO. 712)

[1] Human coronavirus 229E 2000, SEQ ID NO. 794: kksaktgtpkpsrnqspassqtsakslarsqssetkeqkh

[2] Canine coronavirus 2003, SEQ ID NO. 795: kklgvdtekqqqrsrskskersnskitrdttpknenkh

k h r e f v f k n k d g f l y v y k	19		2003 Human SARS spike protein (SEQ ID NO. 717)
k e e l d k y f k n h	11		2003 Human SARS spike protein (SEQ ID NO. 718)
k y r y l r h g k	9		2003 Human SARS spike protein (SEQ ID NO. 719)
k k g a k l l h k	9	55	2003 SARS envelope protein (SEQ ID NO. 720)
k h l d a y k	7	55	**2003 Human SARS nucleocapsid protein (SEQ ID NO. 743)**

B. OTHER VIRUSES, BACTERIA, MALARIA AND CANCER REPLIKINS

h l v c g k k g i g l s g r k k	19		HIV-TAT (SEQ ID NO. 613)
k k i t n i t t k f e g l e k c c k h	19		Monkeypox virus (SEQ ID NO. 721)
k k l k k s l k l l s f y h p k k	17		African swine fever virus (SEQ ID NO. 722)
k n r i e r l k k e y s s t w h	16		West Nile Virus (SEQ ID NO. 723)
k s r g i p i k k g h	11		Nipah virus, v-protein (SEQ ID NO. 724)
k s r i m p i k k g h	11		Hendra virus, V-protein (SEQ ID NO. 725)
k k f l n g f k h h	10		Sindbis virus (SEQ ID NO. 726)
k k k s k k h k d k	10		EEL Leukemia (SEQ ID NO. 85)
k h h p k d n l i k	10		BRCA-1 Breast cancer (SEQ ID NO. 81)
k h k r k k f r q k	10		Ovarian cancer (SEQ ID NO. 84)
k a g v a f i h k k	10	>90%	Glioma Replikin (SEQ ID NO. 83)
k i h l i s v k k	9		Smallpox virus (SEQ ID NO. 727)

k l i s i h e k	8		Smallpox virus (SEQ ID NO. 728)
k l r e e h e k	8		B. anthracis, HATPase (SEQ ID NO. 729)
k h k k q i v k	8		Plasm. Falciparum ATPase (SEQ ID NO. 750)
k k h a t v l k	8	>90%	Ebola virus polymerase (SEQ ID NO. 730
k k e d d e k h	8		P. falciparum blood trophozoites (SEQ ID NO.)
k h k e k m s k	8	>90%	(K-RAS 2B) lung cancer (SEQ ID NO. 731)
k k l r h e k	7		Rous sarcoma virus (SEQ ID NO 48)
k k l r h e k	7		c-src, colon, breast cancer (SEQ ID NO. 52)
k k l r h d k	7		c-yes, melanoma, colon cancer (SEQ ID NO. 50)

SARS and H3N2-Fujian influenza virus replikins traced back to a 1918 pandemic replikin

[0152] The origin of the SARS virus is as yet unknown. We report evidence that certain SARS virus peptides can be traced back through homologous peptides in several strains of influenza virus isolates from 2002 to a sequence in the strain of the 1918 influenza pandemic responsible for the deaths of over 20 million people.

[0153] By quantitative analysis of primary protein sequences of influenza virus and other microorganisms recorded through the last century we have found a new class of peptide structures rich in lysines and histidine, related to the phenomenon of rapid replication itself and to epidemics, rather than to the type of organism (eg. Table 1) and named them Replikins. We have found a new class of peptide structures with the following obligatory algorithm: at least two lysines 6 to 10 residues apart, lysine concentration 6% or greater, one histidine, in 7 to 50 amino acids. Because these peptides relate to the phenomenon of rapid replication itself and to epidemics, we named them Replikins. We have found a quantitative correlation of strain-specific replikin concentration (replikin count = number of replikins per 100 amino acids) in the hemagglutinin protein with influenza epidemics and pandemics (Figure 7). No previous correlation of influenza epidemics with strain-specific viral protein chemistry have been reported. Conservation, condensation and concentration of replikin structure also has been found in influenza (eg. in Table 7a), HIV and malaria. The detection of replikins in SARS coronavirus, in addition to tracing its possible evolution, has permitted the synthesis of small SARS antigens for vaccines.

[0154] We have found a quantitative correlation of strain-specific replikin concentration (count) in the influenza hemagglutinin proteins with influenza epidemics and with each of the three pandemics of the last century, in 1918, 1957, and 1968. A similar course was observed for each of these three pandemics: after a strain-specific high replikin count, an immediate decline followed, then a 'rebound' increase with an accompanying epidemic occurred. Also, a 1 to 3 year warning increase in count preceded most epidemics.

[0155] We found that the replikin in the hemagglutinin of an influenza virus isolated from a goose in 1917 (which we named the Goose Replikin) appeared in the next year in the H1N1 strain of influenza responsible for the 1918 pandemic, with only two substitutions as follows: **kkg**(t/s)syp**k**l**s**k**s**y(t/v)nn**kgk**evlvlwgv**hh.** Table 7a shows that the influenza 1917 Goose Replikin (GR) then was essentially conserved for 85 years, despite multiple minor substitutions and apparent translocations to other influenza strains. We have found that the 1917 influenza GR demonstrated apparent mobility between several influenza strains, appearing in H1N1 (the pandemic of 1918), in H2N2 (pandemic of 1957-58), in H3N2 (pandemic of 1968, epidemic in China and Russia 2000, Fujian strain epidemic 2003) and in H5N1 (epidemic in China 1997). In 1997 its structure was restored in H1N2 exactly to its 1918 structure **kkg**ssyp**k**l**s**k**s**yvnn**kgk**evlvlwgv**hh**.

[0156] The SARS coronavirus first appeared in the 2002-2003 influenza season. The dual origin in 2002 of SARS replikins, from influenza GR and coronavirus replikins (or from some unknown shared precursor) is suggested by the following events, all of which occurred in 2002: 1) a condensation for the first time in 85 years is seen in the GR-H1N2 Replikin sequence from 29 to 28 amino acids (Table 7a)(A similar condensation was found in H3N2 Fujian from 29 to 27 amino acids in the current epidemic (Table 7a)); 2) the replikin count of GR-H1N2 showed a marked decline consistent with GR moving out of H1N2; 3) the replikin count of coronavirus nucleocapsid proteins showed a marked increase; and 4) SARS coronavirus appeared in 2002-2003 with replikins containing the following motifs: **'kkg'** and **'k-k'**, previously seen in GR 1918 and GR-H1N2 2001; **'k-kk'**, **'kk'** and **'kl'** seen in influenza GR-H1N2 2001; **'kk'** seen in the avian bronchitis coronavirus replikin; and **'kk-kk-k'**, **'k-k'**, **'kk'**, **'kl'** and **'kt'** seen in the replikin of porcine epidemic diarrhea coronavirus (Table 7a) (SARS is believed to have made its first appearance in humans as the epidemic pneumonia which erupted in a crowded apartment house where there was a severe back-up of fecal sewage, which was then airborne by ventilating fans).

Table 7a.

Goose Replikin (GR) sequences in different influenza strains from 1917 to 2003: SARS and H3N2-Fujian appearance 2002-2003.

Replikins related to the Goose Replikin:
Continuous amino acid sequences
Shared motif and/or position–
Amino acid substitutions– clear background
'Condensed' indicates condensation of acids)
sequence length in H1N2 and H3N2-Fujian

Replikin Length (Number of amino acids)

Virus or other organism containing replikin (Complete replikins except for Fujian strain)

	SEQ.ID NO.		
kkgtsypklsksytnnkgkevlvlwgvhh		29	**1917 H1N1 Influenza Goose Replikin (GR)**
[illegible]	743	29	**1918 GR in H1N1Human Influenza**
[illegible]	744	29	1958 GR H2N2 Influenza
[illegible]	745	29	1964,1965,1968 GR in H2N2 Influenza
kkgtsypklsksytnnkgkevlvlwgvhh		29	1976,'77,'80,'81,'85 GR in H1N1 Influenza
[illegible]		29	1996-2001 GR in H5N1 Influenza
[illegible]	746	29	1996 GR in H1N1 Influenza
[illegible]	747	29	1997,1998 GR in H1N1 Influenza
[illegible]	748	29	1999 GR in H1N2 Influenza
[illegible]	749	29	2000 GR in H1N2 Influenza
[illegible]	750	29	2001 GR in H1N2 Influenza
[illegible]		29	**2001 GR in H1N2 Influenza**
[illegible]		28	**2002 GR in H1N2 Influenza (condensed)**
khldaykttfppteppkdkkks		21	**2002-3 Human SARS nucleocapsid protein**
[illegible]	751	29	**1968-2001 GR in H3N2 Influenza (complete)**
kleykypalnvtmpnndkfdklyiwgvhh	752	29	**1996 H3N2 Fujian Influenza (incomplete)**
kykypaldnvtmpnnekfdklyiwgvhh	753	27	**2003 H3N2 Fujian(condensed,incomplete)**
[illegible]		24	Porcine epidemic diarrhea coronavirus
[illegible]		27	Avian bronchitis coronavirus
kknvksakqlphlkvlldvrgakqlph	754	27	2000 shrimp white spot syndrome virus
[illegible]		27	Avian bronchitis coronavirus
khlrefvfnkdgflyvyk	755	20	2002-3 Human SARS spike protein
[illegible]	756	14	2002-3 Human SARS nucleocapsid protein
khlrefvfnkdgflyvyk		20	2002-3 Human SARS spike protein
[illegible]		14	2002-3 Human SARS nucleocapsid protein
keeldkyfkn		11	2002-3 Human SARS spike protein
[illegible]		9	2002-3 Human SARS envelope protein
kyrylrhgk		9	2002-3 Human SARS spike protein
khlday		7	2002-3 Human SARS nucleocapsid protein
ksrgipikkgh		11	Nipah virus, v-protein
ksrimpikkgh		11	Hendra virus, v-protein
kkflnqfkhh		10	Sindbis virus
kkkskkhkdk		10	EEL leukemia
khhpkdnlik		10	BRCA-1 breast cancer
khkrkkfrqk		10	Ovarian cancer
[illegible]		10	Glioma Replikin
kihlisvk		9	Smallpox virus

[illegible]	9	HIV TAT protein
[illegible]	8	Smallpox virus
[illegible]	8	B.anthracis, HATPase
[illegible]	8	Ebola virus polymerase

[0157] The recent increasingly high replikin count peaks of the 1917 Goose Replikin (Figure 7), now in H1N2 (Table 7a), approaching the 1917 replikin count, could be a warning of a coming pandemic which may already have begun since the SARS virus and the H3N2-Fujian virus are the current carriers of the short replikin derivatives of the Goose Replikin seen in Table 7 and 7a to be associated with high mortality.

[0158] Since the Goose Replikin has at least an 85 year history involving most or all of the A-strains of influenza and SARS, it and its components are conserved vaccine candidates for pan-strain protection. Condensed short SARS replikins, 7 to 21 amino acids long, enriched in % lysine and histidine compared to the Goose Replikin, occurred in association with the higher mortality rate of SARS (10-55%) when compared to that (2.5%) of the Goose Replikin, 29 amino acids long. Short replikins here mixed with long replikins in SARS may be responsible for high mortality. This is also the case for replikins of other organisms such as the ebola and smallpox viruses and anthrax bacteria (Table 7a). These short SARS replikins showed surprising homology with short replikins of other organisms such as smallpox, anthrax, and ebola which are associated with even higher untreated mortality rates (Table 7a).

[0159] Short synthetic vaccines, besides being much more rapidly produced (days rather than months), and far less expensive, should avoid the side effects attendant on the contamination and the immunological interference engendered by multiple epitopes of thousands of undesired proteins in current whole virus vaccines in general. In any case for influenza, current whole virus vaccines are ineffective in more than half of the elderly. But would short replikins be sufficiently immunogenic? The short glioma replikin 'kagvaflhkk' proved to be a successful basis for a synthetic anti-glioblastoma multiforme and anti-bronchogenic carcinoma vaccine. It produced anti-malignin antibody, which is cytotoxic to cancer cells at

picograms/cell and relates quantitatively to the survival of cancer patients. In order to prepare for a recurrent SARS attack, which appears likely because of the surge we found in the coronavirus nucleocapsid replikin count in 2002, We synthesized four SARS short replikins, found in nucleocapsid, spike, and envelope proteins. We found that these synthetic short SARS replikins when injected into rabbits also produced abundant specific antibody. For example, the 21 amino acid SARS nucleocapsid replikin antibody binds at dilutions greater than 1 in 204,800. Because of previous unsuccessful attempts by others to achieve with various small peptides a strong immune response without the unwanted side effects obtained with a whole protein or the thousands of proteins or nucleic acids as in smallpox vaccine, the ability of small synthetic replikin antigens to achieve strong immune responses is significant for the efficacy of these SARS vaccines.

[0160] We examined the relationship of Replikin structure in influenza and SARS viruses to increased mortality, with results as shown in Table 7 The relation of high mortality to short or condensed Replikin sequences is seen in the high mortality organisms shown in Section B of Table 7, in viruses other than influenza and SARS, and in bacteria, malaria and cancer. In support of the unifying concept of Replikin structure and of the relation of Replikins to rapid replication rather than any cell type or infectious organism, in addition to the prevalence of the basic Replikin structure in a broad range of viral, bacterial, malarial and cancer organisms in which replication is crucial to propagation and virulence, the following homologous sequences have been observed: note the "k"s in positions 1 and 2, note the alignment of "k"s as they would present to DNA, RNA or other receptor or ligand for incorporation or to stimulate rapid replication, note the frequency of "double k"s and "multiple k"s , note the frequency of "g" in position 3 and the occurrence of the triplets "kkg", "hek", "hdk" and "hkk" in the most condensed shortened Replikins associated with the highest mortality organisms, cancer cells and genes as diverse as the smallpox virus, the anthrax virus, Rous sarcoma virus and glioblastome multiforme (glioma), c-src in colon and breast cancer, and c-yes in melanoma and colon cancer. Note also the almost identical Replikin structure for two recently emerging high mortality viruses in Australia and Southeast Asia, Nipah and Hendrah viruses. These two viruses are reported to have similar or identical

antibodies formed against them but no structural basis has been known for this up till now, with our finding of their two almost identical Replikins, for this similar antibody. Table 7 also shows the relationship of five SARS Replikins of 2003 which we have found both to the influenza Goose Replikin of 1917 and to two coronaviruses, the avian bronchitis coronavirus and the porcine epidemic diarrhea virus. The first 2003 human SARS Replikin in Table 7 shows certain sequence homologies to the influenza virus goose 1917 and human 1918 Replikins through an intermediary structure of influenza H1N2 in 2002 (*e.g.*, see Replikin "k" in positions 1, 18 and 19). The 1917 Goose Replikin sequence is seen in Table 7 to have been largely conserved despite many substitutions in amino acids which are not crucial to the definition of Replikins through 1999 (substitutions are show in italics). The original 29 amino acid 1917 Replikin sequence was then found to have been almost exactly restored to its structure of 1917-1918 in the 2001 H1N2 Replikin. However, the 2002 H1N2 influenza Replikin has been shortened from 29 to 28 amino acids and the "shift to the left" of amino acids kevl(i/v)wg (v/i)hh is clearly evident. In 2003, one Replikin was further shortened (or compacted) to the 21 amino acid Replikin of the first listed 2003 human SARS virus. The % k of the 2003 SARS Replikin is now 38.1 % (8/21) in comparison to 20.7 % of the Goose Replikin and the 1918 Human Pandemic Replikin. Compared to the influenza 29 amino acid Replikin, three SARS Replikins were found to be further shortened (or compacted) to 19, 11 and 9 amino acid long sequences, respectively. In the SARS 9 amino acid sequences shown, the % k is 44.4% (4/9). With the shortening of the SARS Replikin, the SARS mortality rate in humans rose to 10% in the young and 55.5 % in the elderly compared to the 2.5 % mortality in the 1918 influenza pandemic.

[0161] The amino acid sequences are shown in Table 7 to emphasize the degree of homology and conservation for 85 years (1917-2002) of the influenza Replikin, for which evidence has first been observed in the 1917 Goose Replikin. No such conservation has ever been observed before. Table 7 also illustrates that the Replikins in the 2003 human SARS virus, in addition to having homologies to the influenza Replikins which first appeared as the 1917 Goose Replikin and the 1918 Human Pandemic influenza Replikin, show certain sequence homologies to both the coronavirus avian bronchitis virus Replikin (*e.g.* "k" in positions 1 and 2,

end in "h") and to the coronavirus acute diarrhea virus Replikin (*e.g.* "k" in positions 1 and 11, "h" at the end of the Replikin). This evidence of relation to both influenza and coronavirus Replikins is of interest because SARS arose in Hong Kong as did several recent influenza epidemics and earlier pandemics, and the SARS virus has been classified as a new coronavirus partly because of its structure, including nucleocapsid, spike, and envelope proteins. Certain epidemiological evidence also is relevant in that SARS made its first appearance in humans as the epidemic pneumonia, which erupted, in a crowded Hong Kong apartment house where there was a severe back-up of fecal sewage, which was airborne by ventilating fans.

[0162] Composition of Replikins in Strains of Influenza Virus B: Of a total of 26 Replikins identified in this strain (Table 3), the following ten Replikins are present in every influenza B isolate examined from 1940-2001. Overlapping Replikin sequences are listed separately. Lysines and histidines are in bold type to demonstrate homology consistent with the "3-point recognition."

kshfanlk (SEQ ID NO. 104)
kshfanlkgtk (SEQ ID NO. 105)
kshfanlkgtktrgklcpk (SEQ ID NO. 106)
hekygglnk (SEQ ID NO. 107)
hekygglnksk (SEQ ID NO. 108)
hekygglnkskpyytgehak (SEQ ID NO. 10)
hakaignepiwvk (SEQ ID NO. 110)
hakaignepiwvvkktplklangtk (SEQ ID NO. 111)
hakaignepiwvktplklangtkyrppak (SEQ ID NO. 112)
hakaignepiwvktplklangtkyrppaklk (SEQ ID NO. 113)

[0163] Tables 3 and 4 indicate that there appears to be much greater stability of the Replikin structures in influenza B hemagglutinins compared with H1N1 Replikins. Influenza B has not been responsible for any pandemic, and it appears not to have an animal or avian reservoirs. (Stuart-Harris et al., Edward Arnold Ltd., London (1985)).

[0164] Influenza H1N1 Replikins: Only one Replikin "hp(v/i)tigecpkyv-(r/k)(s/t)(t/a)k" is present in every H1N1 isolate for which sequences are available from 1918, when the strain first appeared and caused the pandemic of that year, through 2000. (Table 4). ("(v/i)" indicates that the amino acid v or i is present in the same position in different years.) Although H1N1 contains only one persistent Replikin, H1N1 appears to be more prolific than influenza B. There are 95 different Replikin structures in 82 years on H1N1 versus only 31 different Replikins in 62 years of influenza B isolates (Table 4). An increase in the number of new Replikin structures occurs in years of epidemics (Tables 3, 4, 5 and 6) and correlates with increased total Replikin concentration (Figures 7 and 8).

[0165] Influenza H2N2 Replikins: Influenza H2N2 was responsible for the human pandemic of 1957. Three of the 20 Replikins identified in that strain for 1957 were conserved in each of the H2N2 isolates available for examination on PubMed until 1995 (Table 5).

ha(k/q/m)(d/n)ilekthngk (SEQ ID NO. 232)

ha(k/q/m)(d/n)ilekthngklc(k/r) (SEQ ID NO. 233)

kgsnyp(v/i)ak(g/r)synntsgeqmliiwq(v/i)h (SEQ ID No. 238)

[0166] However, in contrast to H1N1, only 13 additional Replikins have been found in H2N2 beginning in 1961. This paucity of appearance of new Replikins correlates with the decline in the concentration of the H2N2 Replikins and the appearance of H2N2 in isolates over the years. (Figure 8).

[0167] Influenza H3N2 Replikins: Influenza H3N2 was responsible for the human pandemic of 1968. Five Replikins which appeared in 1968 disappeared after 1977, but reappeared in the 1990s (Table 6). The only Replikin structure which persisted for 22 years was hcd(g/q)f(q/r)nekwdlf(v/i)er(s/t)k, which appeared first in 1977 and persisted through 1998. The emergence of twelve new H3N2 Replikins in the mid 1990s (Table 6) correlates with the increase in Replikin concentration at the same time (Figure 8), and with the prevalence of the H3N2 strain in recent isolates together with the concurrent disappearance of all Replikins from some of these isolates (Figure 8), this suggests the emergence of the new substrain H3N2(R). The

current epidemic in November - December 2003 of a new strain of H3N2 (Fujian) confirms this prediction made first in the Provisional Application US 60/303,396, filed July 9, 2001.

[0168] Figures 1 and 2 show that influenza epidemics and pandemics correlate with the increased concentration of Replikins in influenza virus, which is due to the reappearance of at least one Replikin from one to 59 years after its disappearance. Also, in the A strain only, there is an emergence of new strain-specific Replikin compositions (Tables 4-6, see also increase in number of new Replikins, pre-epidemic for H5N1 in Figure 11). Increase in Replikin concentration by repetition of individual Replikins within a single protein appears not to occur in influenza virus, but is seen in other organisms.

[0169] It has been believed that changes in the activity of different influenza strains are related to sequence changes in influenza hemagglutinins, which in turn are the products of substitutions effected by one of two poorly understood processes: i) antigenic drift, thought to be due to the accumulation of a series of point mutations in the hemagglutinin molecule, or ii) antigenic shift, in which the changes are so great that genetic reassortment is postulated to occur between the viruses of human and non-human hosts. First, the present data suggests that the change in activity of different influenza strains, rather than being related to non-specific sequence changes, are based upon, or relate to the increased concentration of strain-specific Replikins and strain-specific increases in the replication associated with epidemics. In addition, the data were examined for a possible insight into which sequence changes are due to "drift" or "shift", and which are due to conservation, storage in reservoirs, and reappearance. The data show that the epidemic-related increase in Replikin concentration is not due to the duplication of existing Replikins per hemagglutinin, but is due to the reappearance of at least one Replikin composition from 1 to up to 59 years after its disappearance, plus in the A strains only, the emergence of new strain-specific Replikin compositions (Tables 3-6). Thus the increase in Replikin concentration in the influenza B epidemics of 1951 and 1977 are not associated with the emergence of new Replikin compositions in the year of the epidemic but only with the reappearance of Replikin compositions which had

appeared in previous years then disappeared (Table 3). In contrast, for the A strains, in addition to the reappearance of previously disappeared virus Replikins, new compositions appear (e.g. in HINI in the year of the epidemic of 1996, in addition to the reappearance of 6 earlier Replikins, 10 new compositions emerged). Since the A strains only, not influenza B, have access to non-human animal and avian reservoirs, totally new compositions probably derive from non-human host reservoirs rather than from mutations of existing human Replikins which appear to bear no resemblance to the new compositions other than the basic requirements of "3-point recognition" (Tables 2-5). The more prolific nature of H1N1 compared with B, and the fact that pandemics have been produced by the three A strains only, but not by the B strain, both may also be a function of the ability of the human A strains to receive new Replikin compositions from non-human viral reservoirs.

[0170] Some Replikins have appeared in only one year, disappeared, and not reappeared to date (Tables 3-6). Other Replikins disappear from one to up to 81 years, when the identical Replikin sequence reappears. Key Replikin 'k' and 'h' amino acids, and the spaces between them, are conserved during the constant presence of particular Replikins over many years, as shown in Tables 2 and 3-6 for the following strain-specific Replikins: ten of influenza B, the single Replikin of HIN1, and the single Replikin of H3N2 as well as for the reappearance of identical Replikins after an absence. Despite the marked replacement or substitution activity of other amino acids both inside the Replikin structure and outside it in the rest of the hemagglutinin sequences, influenza Replikin histidine (h) appears never to be, and lysine (k) is rarely replaced. Examples of this conservation are seen in the H1N1 Replikin "hp(v/i)tigecpkyv(r/k)(s/t)(t/a)k," (SEQ ID NO. 135) constant between 1918 and 2000, in the H3N2 Replikin "hed(g/q)f(q,r)nekwdlf(v/i)er(s/t)k" (SEQ ID NO. 277) constant between 1975 and 1998 and in the H3N2 Replikin "hqn(s/e)(e/q)g(t/s)g(q/y)aad(l/q)kstq(a/n)a(i/l)d(q/g)I(n/t)(g/n)k,(l/v)n(r/s) vi(c/c)k" (SEQ ID NO. 276) which first appeared in 1975, disappeared for 25 years, and then reappeared in 2000. While many amino acids were substituted, the basic Replikin structure of 2 Lysines, 6 to 10 residues apart, one histidine, a minimum of 6% lysine in not more than approximately 50 amino acids, was conserved.

[0171] Totally random substitution would not permit the persistence of these H1N1 and H3N2 Replikins, nor from 1902 to 2001 in influenza B the persistence of 10 Replikin structures, nor the reappearance in 1993 of a 1919 18mer Replikin after an absence of 74 years. Rather than a random type of substitution, the constancy suggests an orderly controlled process, or in the least, protection of the key Replikin residues so that they are fixed or bound in some way: lysines, perhaps bound to nucleic acids, and histidines, perhaps bound to respiratory redox enzymes. The mechanisms, which control this conservation, are at present unknown.

CONSERVATION OF REPLIKIN STRUCTURES

[0172] Whether Replikin structures are conserved or are subject to extensive natural mutation also was examined by scanning the protein sequences of various isolates of foot and mouth disease virus (FMDV), where mutations in proteins of these viruses have been well documented worldwide for decades. Protein sequences of FMDV isolates were visually examined for the presence of both the entire Replikin and each of the component Replikin amino acid residues observed in a particular Replikin.

[0173] Rather than being subject to extensive substitution over time as occurs in neighboring amino acids, the amino acids which comprise the Replikin structure are substituted little or not at all, that is the Replikin structure is conserved.

[0174] For example, in the protein VP1 of FMDV type O, the Replikin (SEQ ID NO.: 3) "hkqkivapvk" was found to be conserved in 78% of the 236 isolates reported in PubMed, and each amino acid was found to be conserved in individual isolates as follows: his, 95.6%; lys, 91.8%; gln 92.3%; lys, 84.1%; ile, 90.7%; val, 91.8%; ala, 97.3%; pro, 96.2%; ala, 75.4%; and lys, 88.4%. The high rate of conservation suggests structural and functional stability of the Replikin structure and provides constant targets for treatment.

[0175] Similarly, sequence conservation was found in different isolates of HIV for its Replikins, such as (SEQ ID NO.: 5) "kcfncgkcgh" or (SEQ ID NO.: 6) "kvylawvpahk" in HIV Type 1 and (SEQ ID NO.: 7) "kcwncgkcgh" in HIV Type 2

(Table 2). Further examples of sequence conservation were found in the HIV tat proteins, such as (SEQ ID NO.: 613) "helvckqkkglgisygrkk," wherein the key lysine and histidine amino acids are conserved. (See Table 8).

[0176] Similarly, sequence conservation was observed in plants, for example in wheat, such as in wheat ubiguitin activating enzyme E (SEQ ID NOs. 614-616). The Replikins in wheat even provided a reliable target for stimulation of plant growth as described within. Other examples of conservation are seen in the constant presence of malignin in successive generations, over ten years of tissue culture of glioma cells, and by the constancy of affinity of the glioma Replikin for antimalignin antibody isolated by immunoadsorption from 8,090 human sera from the U.S., U.K., Europe and Asia (e.g., Figure 5 and U.S. Patent 6,242,578 B1).

[0177] Similarly, conservation was observed in trans-activator (Tat) proteins in isolates of HIV. Tat (trans-activator) proteins are early RNA binding proteins regulating lentiviral transcription. These proteins are necessary components in the life cycle of all known lentivirases, such as the human immunodeficiency viruses (HIV). Tat is a transcriptional regulator protein that acts by binding to the trans-activating response sequence (TAR) RNA element and activates transcription Initiation and/or elongation from the LTR promoter. HIV cannot replicate without tat, but the chemical basis of this has been unknown. In the HIV tat protein sequence from 89 to 102 residues, we have found a Replikin that is associated with rapid replication in other organisms. The amino acid sequence of this Replikin is "helvckqkkglgisygrkk." In fact, we found that this Replikin is present in every HIV tat protein. Some tat amino acids are substituted frequently, as shown in Table 9, by alternate amino acids (in small size fonts lined up below the most frequent amino acid (Table 8), the percentage of conservation for the predominant Replikin "helvcfqkkglgisygrkk"). These substitutions have appeared for most of the individual amino acids. However, the key lysine and histidine amino acids within the Replikin sequence, which define the Replikin structure, are conserved 100% in the sequence; while substitutions are common elsewhere in other amino acids, both within and outside the Replikin, none occurs on these key histidine amino acids.

[0178] As shown in Table 8 it is not the case that lysines are not substituted in the tat protein amino acid sequence. From the left side of the table, the very first lysine in the immediate neighboring sequence, but outside the Replikin sequence, and the second lysine (k) in the sequence inside the Replikin, but "extra" in that it is not essential for the Replikin formation, are both substituted frequently. However, the 3rd, 4th and 5th lysines, and the one histidine, in parentheses, which together set up the Replikin structure, are never substituted. Thus, these key amino acid sequences are 100% conserved. As observed in the case of the influenza virus Replikins, random substitution would not permit this selective substitution and selective non-substitution to occur due to chance.

Table 8

% Replikin CONSERVATION of each constituent amino acid in the first 117 different isolates of HIV tat protein as reported in PubMed:

38 (100) 57 86 (100) (100) 66 76 (100) 99 57 49 (100) 94 (100) 97 98 85 97 99 (100)(100)(100)%

Neighboring Amino acids **tat Replikin**

k (c) s y [(h) (c) l v (c) f q k (k) g (l) g i s y g (r) (k) (k)]

below are the amino acid substitutions observed for each amino acid above:

h	**c**	**f**	**q**	**i**	**l**	**h**	**t**	**a**	**a**	**l**	**y**	**h**	**q**	
r	**w**	**p**	**l**	**l**		**i**	**h**	**q**	**v**					
y		**s**		**s**		**l**	**m**	***r***	**s**					
i				**s**		**m**	**s**							
s				**r**		**n**								
v														
a														
f														
p														
q														

[0179] The conservation of the Replikin structure suggests that the Replikin structure has a specific survival function for the HIV virus which must be preserved and conserved, and cannot be sacrificed to the virus 'defense' maneuver of amino acid substitution created to avoid antibody and other 'attack.' These 'defense' functions, although also essential, cannot 'compete' with the virus survival function of HIV replication.

[0180] Further conservation was observed in different isolates of HIV for its Replikins such as "kcfncgkegh" (SEQ ID NO. 5) or "kvylawvpahk" (SEQ ID NO. 6) in HIV Type 1 and "kcwncgkegh" (SEQ ID NO. 7) in HIV Type 2.

[0181] The high rate of conservation observed in FMVD and HIV Replikins suggests that conservation also observed in the Replikins of influenza Replikins is a general property of viral Replikins. This conservation makes them a constant and reliable targeted for either destruction, for example by using specific Replikins such

as for influenza, FMVD or HIV vaccines as illustrated for the glioma Replikin, or stimulation.

[0182] Similarly, as provided in examples found in viruses including influenza viruses, FMDV, and HIV, where high rates of conservation in Replikins suggest that conservation is a general property of viral Replikins and thus making Replikins a constant and reliable target for destruction or stimulation, conservation of Replikin structures occurs in plants. For example, in wheat plants, Replikins are conserved and provide a reliable target for stimulation. Examples of conserved Replikins in wheat plants ubiquitin activating enzyme E include:

E3 **hk**drltkkvvdiarevakvdvpeyrrh (SEQ ID NO. 614)

E2 **hk**erldrkvvdvarevakvevpsyrrh (SEQ ID NO. 615)

E1 **hk**erldrkvvdvarevakmevpsyrrh (SEQ ID NO. 616)

* * * ** *

[0183] Similarly to conservation found in the HIV tat protein, the Replikin in the wheat ubiquitin activating enzyme E is conserved. As with the HIV tat protein, substitutions of amino acids (designated with an '*') adjacent to the Replikin variant forms in wheat ubiquitin activating enzyme E are common. The key k and h amino acids that form the Replikin structure, however, do not vary whereas the 'unessential' k that is only 5 amino acids (from the first k on the left) is substituted.

ANTI-REPLIKIN ANTIBODIES

[0184] An anti-Replikin antibody is an antibody against a Replikin. Data on anti-Replikin antibodies also support Replikin class unity. An anti-Replikin antibody response has been quantified by immunoadsorption of serum antimalignin antibody to immobilized malignin (see Methods in U.S. Patent No. 5,866,690). The abundant production of antimalignin antibody by administration to rabbits of the synthetic version of the 16-mer peptide whose sequence was derived from malignin, absent carbohydrate or other groups, has established rigorously that this peptide alone is an epitope, that is, provides a sufficient basis for this immune response (Figure 3). The 16-mer peptide produced both IgM and IgG forms of the antibody. Antimalignin antibody was found to be increased in concentration in serum in 37% of 79 cases in the U.S. and Asia of hepatitis B and C, early, in the first five years of infection, long before the usual observance of liver cancer, which develops about fifteen to twenty-five years after infection. Relevant to both infectious hepatitis and HIV infections, transformed cells may be one form of safe haven for the virus: prolonging cell life and avoiding virus eviction, so that the virus remains inaccessible to anti-viral treatment.

[0185] Because administration of Replikins stimulates the immune system toproduce antibodies having a cytotoxic effect, peptide vaccines based on the particular influenza virus Replikin or group of Replikins observed to be most concentrated over a given time period provide protection against the particular strain of influenza most likely to cause an outbreak in a given influenza season., e.g., an emerging strain or re-emerging strain For example, analysis of the influenza virus hemagglutinin amino acid sequence on a yearly or bi-yearly basis, provides data which are useful in formulating a specifically targeted influenza vaccine for that year. It is understood that such analysis may be conducted on a region-by-region basis or at any desired time period, so that strains emerging in different areas throughout the world can be detected and specifically targeted vaccines for each region can be formulated.

INFLUENZA

[0186] Currently, vaccine formulations are changed twice yearly at international WHO and CDC meetings. Vaccine formulations are based on serological evidence of the most current preponderance of influenza virus strain in a given region of the world. However, prior to the present invention there has been no correlation of influenza virus strain specific amino acid sequence changes with occurrence of influenza epidemics or pandemics.

[0187] The observations of specific Replikins and their concentration in influenza virus proteins provides the first specific quantitative early chemical correlates of influenza pandemics and epidemics and provides for production and timely administration of influenza vaccines tailored specifically to treat the prevalent emerging or re-emerging strain of influenza virus in a particular region of the world. By analyzing the protein sequences of isolates of strains of influenza virus, such as the hemagglutinin protein sequence, for the presence, concentration and/or conservation of Replikins, influenza virus pandemics and epidemics can be predicted. Furthermore, the severity of such outbreaks of influenza can be significantly lessened by administering an influenza peptide vaccine based on the Replikin sequences found to be most abundant or shown to be on the rise in virus isolates over a given time period, such as about one to about three years.

[0188] An influenza peptide vaccine of the invention may include a single Replikin peptide sequence or may include a plurality of Replikin sequences observed in influenza virus strains. Preferably, the peptide vaccine is based on Replikin sequence(s) shown to be increasing in concentration over a given time period and conserved for at least that period of time. However, a vaccine may include a conserved Replikin peptide(s) in combination with a new Replikin(s) peptide or may be based on new Replikin peptide sequences. The Replikin peptides can be synthesized by any method, including chemical synthesis or recombinant

gene technology, and may include non-Replikin sequences, although vaccines based on peptides containing only Replikin sequences are preferred. Preferably, vaccine compositions of the invention also contain a pharmaceutically acceptable carrier and/or adjuvant.

[0189] The influenza vaccines of the present invention can be administered alone or in combination with antiviral drugs, such as gancyclovir; interferon; interleukin; M2 inhibitors, such as, amantadine, rimantadine; neuraminidase inhibitors, such as zanamivir and oseltamivir; and the like, as well as with combinations of antiviral drugs.

REPLIKIN DECOYS IN MALARIA

[0190] Analysis of the primary structure of a Plasmodium farciparum malaria antigen located at the merozoite surface and/or within the parasitophorous vacuole revealed that this organism, like influenza virus, also contains numerous Replikins (Table 9). However, there are several differences between the observation of Replikins in Plasmodium falciparum and influenza virus isolates. For example, Plasmodium falciparum contains several partial Replikins, referred to herein as "Replikin decoys." These decoy structures contain an abundance of lysine residues, but lack the histidine required of Replikin structures. Specifically, these decoys contain many lysines 6 to 10 residues apart in overlapping fashion, similar to the true malaria recognins but without histidine residues. It is believed that the decoy structure maximizes the chances that an anti-malarial antibody or other agent will bind to the relatively less important structure containing the lysines, i.e., the Replikin decoys, rather than binding to histidine, which is present in Replikin structure, such as Replikins in respiratory enzymes, which could result in destruction of the trypanosome. For example, an incoming antibody, with specificity for Replikin structures, might attach to the Replikin decoy structure, leaving the true Replikin structure remains untouched.

[0191] Therefore, anti-Replikin treatment of malaria requires two phases (dual treatment): i) preliminary treatment with proteolytic enzymes that cleave the Replikin decoys, permitting 'safe passage' of the specific anti-Replikin treatment; and ii) attacking malaria Replikins either with specific antibodies or by cellular immunity engendered by synthetic malaria Replikin vaccines or by organic means targeting the malaria Replikins.

REPETITION AND OVERLAPPING OF REPLIKIN STRUCTURES

[0192] Another difference seen in Plasmodium falciparum is a frequent repetition of individual Replikin structures within a single protein, which was not observed with influenza virus. Repetition may occur by (a) sharing of lysine residues between Replikins, and (b) by repetition of a portion of a Replikin sequence within another Replikin sequence.

[0193] A third significant difference between Replikin structures observed in influenza virus isolates and Plasmodium falciparum is a marked overlapping of Replikin structures throughout malarial proteins, e.g., there are nine overlapping Replikins in the 39 amino acid sequence of SEQ ID NO. 393 (Replikin concentration = 23.1/100 amino acids); and 15 overlapping Replikins in the 41 amino acids of SEQ ID NO. 467 (Replikin concentration = 36.6/100 amino acids). Both of these overlapping Replikin structures occur in blood stage trophozoites and schizonts. In contrast, influenza virus Replikins are more scattered throughout the protein and the maximum Replikin concentration is about 7.5/100 amino acids (Figure 7); and tomato leaf curl gemini virus, which was also observed to have overlapping Replikins.

Replikins of tomato leaf curl gemini virus

[0194] Tomato leaf curl Gemini virus has devastated tomato crops in China and in many other parts of the world. Its replikins reach high counts because of

overlapping replikins as illustrated below in a virus isolated in Japan where the replikin count was 20.7

Replikin Analysis

[0195] Description: The first occurrence of tomato yellow leaf curl virus in tomato (specific host Lycopersicon esculentum Mill.), in Japan
Isolated: 1998 Source:Tomato yellow leaf curl virus-[Aichi]
Strain: Aichi

Protein sequence; amino acid positions 1 to 135
m1 q2 p3 s4 s5 p6 s7 t8 s9 h10 c11 s12 q13 v14 s15 i16 k17 v18 q19 h20 k21 i22 a23 k24 k25 k26 p27 i28 r29 r30 k31 r32 v33 n34 l35 d36 c37 g38 c39 s40 y41 y42 l43 h44 l45 n46 c47 n48 n49 h50 g51 f52 t53 h54 r55 g56 t57 h58 h59 c60 s61 s62 s63 r64 e65 w66 r67 f68 y69 l70 g71 d72 k73 q74 s75 p76 l77 f78 q79 d80 n81 r82 t83 q84 p85 e86 a87 i88 s89 n90 c91 p92 r93 h94 h95 f96 h97 s98 d99 k100 i101 q102 p103 q104 h105 q106 c107 g108 t109 g110 d111 s112 q113 m114 f115 s116 q117 l118 p119 n120 l121 d122 d123 i124 t125 a126 s127 d128 w129 s130 f131 l132 k133 s134 i135

Amino-terminal replikins
h10_c11_s12_q13_v14_s15_i16_k17_v18_q19_h20_k21_i22_a23_k24 (SEQ ID NO. 757)

h10_c11_s12_q13_v14_s15_i16_k17_v18_q19_h20_k21_i22_a23_k24_k25 (SEQ ID NO. 758)

h10_c11_s12_q13_v14_s15_i16_k17_v18_q19_h20_k21_i22_a23_k24_k25_k26 (SEQ ID NO. 759)

h10_c11_s12_q13_v14_s15_i16_k17_v18_q19_h20_k21_i22_a23_k24_k25_k26_p27_i28 r29_r30_k31 (SEQ ID NO. 760)

k17_v18_q19_h20_k21_i22_a23_k24_k25_k26_p27_i28_r29_r30_k31_r32_v33_n34_l35 d36_c37_g38_c39_s40_y41_y42_l43_h44_l45_n46_c47_n48_n49_h50 (SEQ ID NO. 761)

Analysis of this sequence showed that amino acids 1-163 of this "replicating protein" contain five Replikins, namely: (SEQ ID NO.: 13) kfrinaknyfltyph, (SEQ ID NO.: 14) knletpvnklfiricrefh, (SEQ ID NO.: 15) hpniqaaksstdvk, (SEQ ID NO.: 16) ksstdvkaymdkdgdvldh, and (SEQ ID NO.: 17) kasalnilrekapkdfvlqfh.q19_h20 k21_i22_a23_k24_k25_k26_p27_i28_r29_r30_k31_r32_v33_n34_l35_d36_c37_g38_c39 s40_y41_y42_l43_h44_l45_n46_c47_n48_n49_h50_g51_f52_t53_h54 (SEQ ID NO. 762)

k17_v18_q19_h20_k21_i22_a23_k24_k25_k26_p27_i28_r29_r30_k31_r32_v33_n34_l35 d36_c37_g38_c39_s40_y41_y42_l43_h44_l45_n46_c47_n48_n49_h50_g51_f52_t53_h54 r55_g56_t57_h58_(SEQ ID NO. 763)

k17_v18_q19_h20_k21_i22_a23_k24_(SEQ ID NO. 764)

k17_v18_q19_h20_k21_i22_a23_k24_k25_k26_p27_i28_r29_r30_k31_r32_v33_n34_l35 d36_c37_g38_c39_s40_y41_y42_l43_h44_l45_n46_c47_n48_n49_h50_g51_f52_t53_h54 r55_g56_t57_h58_h59_(SEQ ID NO. 765)

k17_v18_q19_h20_k21_i22_a23_k24_k25_(SEQ ID NO. 766)

k17_v18_q19_h20_k21_i22_a23_k24_k25_k26_p27_i28_r29_r30_k31_r32_v33_n34_l35 d36_c37_g38_c39_s40_y41_y42_l43_h44_(SEQ ID NO. 767)

k17_v18_q19_h20_k21_i22_a23_k24_k25_k26_(SEQ ID NO. 768)

h20_k21_i22_a23_k24_k25_k26_p27_i28_r29_r30_k31_(SEQ ID NO. 769)

k21_i22_a23_k24_k25_k26_p27_i28_r29_r30_k31_r32_v33_n34_l35_d36_c37_g38_c39 s40_y41_y42_l43_h44_l45_n46_c47_n48_n49_h50_g51_f52_t53_h54 (SEQ ID NO. 770)

k21_i22_a23_k24_k25_k26_p27_i28_r29_r30_k31_r32_v33_n34_l35_d36_c37_g38_c39 s40_y41_y42_l43_h44_l45_n46_c47_n48_n49_h50_g51_f52_t53_h54_r55_g56_t57_h58 (SEQ ID NO. 771)

k21_i22_a23_k24_k25_k26_p27_i28_r29_r30_k31_r32_v33_n34_l35_d36_c37_g38_c39 s40_y41_y42_l43_h44_l45_n46_c47_n48_n49_h50_g51_f52_t53_h54_r55_g56_t57_h58 h59 (SEQ ID NO. 772)

k21_i22_a23_k24_k25_k26_p27_i28_r29_r30_k31_r32_v33_n34_l35_d36_c37_g38_c39 s40_y41_y42_l43_h44 (SEQ ID NO. 773)

k21_i22_a23_k24_k25_k26_p27_i28_r29_r30_k31_r32_v33_n34_l35_d36_c37_g38_c39 s40_y41_y42_l43_h44_l45_n46_c47_n48_n49_h50 (SEQ ID NO. 774)

k24_k25_k26_p27_i28_r29_r30_k31_r32_v33_n34_l35_d36_c37_g38_c39_s40_y41_y42 l43_h44_l45_n46_c47_n48_n49_h50_g51_f52_t53_h54_r55_g56_t57_h58 (SEQ ID NO. 775)

k24_k25_k26_p27_i28_r29_r30_k31_r32_v33_n34_l35_d36_c37_g38_c39_s40_y41_y42 l43_h44_l45_n46_c47_n48_n49_h50_g51_f52_t53_h54_r55_g56_t57_h58_h59 (SEQ ID NO. 776)

k24_k25_k26_p27_i28_r29_r30_k31_r32_v33_n34_l35_d36_c37_g38_c39_s40_y41_y42 l43_h44 (SEQ ID NO. 777)

k24_k25_k26_p27_i28_r29_r30_k31_r32_v33_n34_l35_d36_c37_g38_c39_s40_y41_y42 l43_h44_l45_n46_c47_n48_n49_h50 (SEQ ID NO. 778)

k24_k25_k26_p27_i28_r29_r30_k31_r32_v33_n34_l35_d36_c37_g38_c39_s40_y41_y42 l43_h44_l45_n46_c47_n48_n49_h50_g51_f52_t53_h54 (SEQ ID NO. 779)

k25_k26_p27_i28_r29_r30_k31_r32_v33_n34_l35_d36_c37_g38_c39_s40_y41_y42_l43 h44_l45_n46_c47_n48_n49_h50 (SEQ ID NO. 780)

k25_k26_p27_i28_r29_r30_k31_r32_v33_n34_l35_d36_c37_g38_c39_s40_y41_y42_l43 h44_l45_n46_c47_n48_n49_h50_g51_f52_t53_h54 (SEQ ID NO. 781)

k25_k26_p27_i28_r29_r30_k31_r32_v33_n34_l35_d36_c37_g38_c39_s40_y41_y42_l43 h44_l45_n46_c47_n48_n49_h50_g51_f52_t53_h54_r55_g56_t57_h58 (SEQ ID NO. 782)

k25_k26_p27_i28_r29_r30_k31_r32_v33_n34_l35_d36_c37_g38_c39_s40_y41_y42_l43 h44_l45_n46_c47_n48_n49_h50_g51_f52_t53_h54_r55_g56_t57_h58_h59 (SEQ ID NO. 783)

k25_k26_p27_i28_r29_r30_k31_r32_v33_n34_l35_d36_c37_g38_c39_s40_y41_y42_l43 h44 (SEQ ID NO. 784)

Mid-molecule: Zero replikins.
Carboxy-terminal: Zero replikins.

Replikin Count = Number of Replikins per 100 amino acids = 28/135 = 20.7

Even higher replikin counts are seen to be achieved by overlapping replikins in malaria.

[0196] The mechanism of lysine multiples is also seen in the Replikins of cancer proteins such as in gastric cancer transforming protein, ktkkgnrvsptmkvth (SEQ ID NO. 88), and in transforming protein P21B (K-RAS 2B) of lung, khkekmskdgkkkkkks (SEQ ID NO. 89).

[0197] The relationship of higher Replikin concentration to rapid replication is also confirmed by analysis of HIV isolates. It was found that the slow-growing low titer strain of HIV (NSI, "Bru," which is prevalent in early stage HIV infection) has a Replikin concentration of 1.1 (+/- 1.6) Replikins per 100 amino acids, whereas the rapidly-growing high titer strain of HIV (SI, "Lai", which is prevalent in late stage HIV infection) has a Replikin concentration of 6.8 (+/- 2.7) Replikins per 100 amino acid residues.

[0198] The high concentration of overlapping Replikins in malaria, influenza virus and cancer cells is consistent with the legendary high and rapid replicating ability of malaria organisms. The multitude of overlapping Replikins in malaria also provides an opportunity for the organism to flood and confuse the immune system of its host and thereby maximize the chance that the wrong antibody will be made and perpetuated, leaving key malaria antigens unharmed.

[0199] As in the case of influenza virus, for example, peptide vaccines based on the Replikin structure(s) found in the malaria organism can provide an effective means of preventing and/or treating malaria. Vaccination against malaria can be achieved by administering a composition containing one or a mixture of Replikin structures observed in Plasmodium falciparum. Furthermore, antibodies to malaria Replikins can be generated and administered for passive immunity or malaria detection.

Replikins in Malaria

[0200] Malaria is a disease which accounts for more than 200 million cases annually world-wide and over 2 million deaths annually, and for which there is as yet no effective vaccine. Replikins have been found to be prominent in Plasmodium falciparum, the most common strain of trypanosome responsible for malaria.

High Replikin Count

[0201] In accord with the legendary high replication rate of trypanosomes, the highest replikin count yet observed in any species has been found in trypanosomes. In the trypanosome plasmodium falciparum, we found that for the ATPase protein recently determined to be the target of the effective anti-malarial artemisinins, in the 1999 3D7 malaria pandemic year, the mean replikin count in the ATPase for all isolates was 57.6, and in one isolate the replikin count reached a record of 111.8 by repeating and overlaping replikins.

Repition and overlapping of replikin structures

[0202] One characteristic seen in Plasmodium falciparum replikins which accounts for the high replikin count compared with influenza replikins, is a frequent repetition of individual replikin structures within a single protein, a feature which was not observed in influenza virus. Repetition may occur a) simply by repeating the entire replikin, (b) by sharing of lysine residues between replikins, and (c) by repetition of a portion of a replikin sequence combined with or within another replikin sequence.

[0203] In addition to repeats of replikin structures, another significant difference between replikin structures observed in influenza virus proteins and those in Plasmodium falciparum proteins is a marked overlap of replikins throughout malarial proteins. For example, overlapping replikin structures occur in blood stage trophozoites and schizonts as seen in the following:

1) in the 39 amino acid sequence of SEQ ID NO. 393, in addition to the four exact repeats of 'ksdhnhk', there are fourteen overlapping replikins (replikin concentration or Replikin Count = 18/39 = 46.2/100 amino acids):

ksdhnhksdhnhksdhnhksdhnhksdpnhkkknnnnnk(SEQIDNO.393)

ksdhnhksdhnhksdhnhksdpnhkkknnnnnk (SEQ ID NO. 394)

ksdhnhksdhnhksdpnhkkknnnnnk (SEQ ID NO. 395)

ksdhnhksdpnhkkknnnnnk (SEQ ID NO. 396)

kkknnnnnkdnksdpnhk (SEQ ID NO. 397)

kknnnnnkdnksdpnhk (SEQ ID NO. 398)

knnnnnkdnksdpnhk (SEQ ID NO. 399)

kdnksdpnhk (SEQ ID NO. 400)

ksdpnhk (SEQ ID NO. 401)

ksdhnhk (SEQ ID NO. 743)

ksdpnhkk (SEQ ID NO. 744)

ksdpnhkkk (SEQ ID NO. 745)

ksdpnhkk (SEQ ID NO. 746)

hkkknnnnnk (SEQ ID NO. 747)

and 2) 15 overlapping Replikins occur in the 41 amino acids of SEQ ID NO. 467 (Replikin concentration or replikin count= 36.6/100 amino acids).

kkdkekkkdsnenrkkkqkedkknpndnklkkieytnkith (SEQ ID NO. 467)

kdkekkkdsnenrkkkqkedkknpndnklkkieytnkith (SEQ ID NO. 468)

kekkkdsnenrkkkqkedkknpndnklkkieytnkith (SEQ ID NO. 469)

kkkdsnenrkkkqkedkknpndnklkkieytnkith (SEQ ID NO. 470)

kkdsnenrkkkqkedkknpndnklkkieytnkith (SEQ ID NO. 471)

kdsnenrkkkqkedkknpndnklkkieytnkith (SEQ ID NO. 472)

kkkqkedkknpndnklkkieytnkith (SEQ ID NO. 473)

kkqkedkknpndnklkkieytnkith (SEQ ID NO. 474)

kqkedkknpndnklkkieytnkith (SEQ ID NO. 475)

kedkknpndnklkkieytnkith (SEQ ID NO. 476)

kknpndnklkkieytnkith (SEQ ID NO. 477)

knpndnklkkieytnkith (SEQ ID NO. 478)

klkkieytnkith (SEQ ID NO. 479)

kkieytnkith (SEQ ID NO. 480)

kieytnkith (SEQ ID NO. 481)

[0204] Both of these overlapping replikin structures occur in blood stage trophozoites and schizonts. In contrast, influenza virus replikins are more scattered throughout the protein and the maximum replikin count (except in 1917-18) is about 7.5 (Figure 7). As described earlier, tomato leaf curl gemini virus was also observed to have overlapping replikins raising the replikin count as high as 20.7.

Table 9

Replikin repeats, overlap, and conservation in one molecule, ATPase

[0205] Overlapping replikins, replikin repeats, and intramolecular conservation of replikin structure, were all found in the single molecule of Plasmodium Falciparum 3D7 ATPase, from amino acid positions 399 through 927. As shown below, identical motifs of each replikin are conserved by the invariant initial '**k**' and terminal '**hk**' (shaded areas), whereas adjacent amino acids 'g/s/n' 'd/g/e' 's/c/n/h' and 's/n' are variable (clear). The last lysine for each replikin is also the first lysine for the next replikin, eg. k411, k417.:

k405	s406	d407	s408	n409	h410	k411	(SEQ ID NO. 785)
k411	s412	d413	s414	n415	h416	k417	
k417	s418	d419	s420	n421	h422	k423	
k423	s424	g425	s426	n427	h428	k429	(SEQ ID NO. 786)
k429	s430	d431	c432	n433	h434	k435	(SEQ ID NO. 787)
k435	s436	g437	s438	n439	h440	k441	
k459	s460	d461	n462	n463	h464	k465	(SEQ ID NO. 788)
k465	s466	d467	s468	s469	h470	k471	(SEQ ID NO. 789)
k477	s478	d479	s480	s481	h482	k483	
k483	s484	g485	s486	n487	h488	k489	
k489	s490	d491	n492	n493	h494	k495	
k495	s496	d497	s498	s499	h500	k501	
k501	s502	g503	s504	n505	h506	k507	
k507	s508	d509	h510	n511	h512	k513	(SEQ ID NO. 790)
k513	s514	d515	s516	n517	h518	k519	
k519	s520	d521	s522	n523	h524	k525	
k525	n526	e527	s528	n529	h530	k531	(SEQ ID NO.791)
k531	n532	e533	s534	n535	h536	k537	
k537	n538	e539	s540	n541	h542	k543	
k543	n544	e545	s546	n547	h548	k549	
k549	n550	e551	s552	n553	h554	k555	
k555	n556	d557	s558	n559	h560	k561	(SEQ ID NO. 792)
k573	s574	d575	n576	n577	h578	k579	(SEQ ID NO.793)
k591	s592	d593	n594	n595	h596	k597	
k609	s610	d611	n612	n613	h614	k615	
k615	s616	d617	n618	n619	h620	k621	
k621	s622	d623	n624	n625	h626	k627	
k639	s640	d641	n642	n643	h644	k645	
k645	s646	d647	h648	n649	h650	k651	
k669	s670	d671	h672	n673	h674	k675	
k675	s676	d677	h678	n679	h680	k681	
k681	s682	d683	n684	n685	h686	k687	
k687	s688	d689	s690	n691	h692	k693	
k693	s694	d695	s696	n697	h698	k699	
k699	s700	d701	h702	n703	h704	k705	
k723	s724	d725	h726	n727	h728	k729	
k729	s730	d731	n732	n733	h734	k735	
k735	s736	d737	s738	n739	h740	k741	
k741	s742	d743	s744	n745	h746	k747	
k747	s748	d749	h750	n751	h752	k753	
k777	s778	d779	h780	n781	h782	k783	
k783	s784	d785	n786	n787	h788	k789	
k789	s790	d791	s792	n793	h794	k795	
k795	s796	d797	s798	n799	h800	k801	
k801	s802	d803	s804	n805	h806	k807	
k807	s808	d809	h810	n811	h812	k813	
k825	s826	d827	n828	n829	h830	k831	
k831	s832	d833	h834	n835	h836	k837	
k837	s838	d839	n840	n841	h842	k843	

k843	s844	d845	h846	n847	h848	k849
k849	s850	d851	s852	n853	h854	k855
k855	s856	d857	s858	n859	h860	k861
k861	s862	d863	s864	n865	h866	k867
k867	s868	d869	s870	n871	h872	k873
k873	s874	d875	n876	n877	h878	k879
k897	s898	d899	h900	n901	h902	k903
k903	s904	d905	h906	n907	h908	k909
k909	s910	d911	n912	n913	h914	k915
k915	s916	d917	n918	n919	h920	k921
k921	s922	d923	h924	n925	h926	k927

Lysine Multiples

[0206] The phenomenon of lysine multiples in replikins, 'kk', 'kkk', etc., producing a high percent lysine, seen in the examples above in malaria, may be related to increased virulence and increased mortality, since it is also seen in the replikins of high mortality cancer proteins, such as in gastric cancer transforming protein, '**k**t**kk**gnrvsptm**k**v**th**' (SEQ ID NO. 88), and in transforming protein P21B (K-RAS 2B) of lung '**kh**k**e**k**n**s**k**dg**kkkkkk**s**' (SEQ ID NO. 89), and in high mortality SARS as in human SARS nucleocapsid protein '**kh**ldayktfppte**p**k**k**d**k**k**k**k', but is less commonly seen in lower mortality influenza. strains as in 1918 H1N1 human influenza '**kk**gssyp**k**l**s**ksyvnn**k**gkevlvlwgv**hh**' (Table 7a).

[0207] The high concentration of overlapping Replikins in malaria, influenza virus and cancer cells is consistent with the legendary high and rapid replicating ability of malaria organisms. The multitude of overlapping Replikins in malaria also provides an opportunity for the organism to flood and confuse the immune system of its host and thereby maximize the chance that the wrong antibody will be made and perpetuated, leaving key malaria antigens unharmed.

[0208] As in the case of influenza virus, for example, peptide vaccines based on the Replikin structure(s) found in the malaria organism can provide an effective means of preventing and/or treating malaria. Vaccination against malaria can be achieved by administering a composition containing one or a mixture of Replikin structures observed in Plasmodium falciparum. Furthermore, antibodies to malaria

Replikins can be generated and administered for passive immunity or malaria detection.

Replikin Decoys:

[0209] In another difference from influenza virus replikins, Plasmodium falciparum contains several repeating partial replikins, referred to herein as "replikin decoys." These decoys contain many lysines 6 to 10 residues apart in overlapping fashion, similar to the true malaria replikins but without histidine residues. It is believed that these decoy structure maximize the chances that an anti-malarial antibody or other agent will bind to the relatively less important structure containing the lysines, i.e., the replikin decoys, rather than the binding including histidine, which is present in replikin structure, such as replikins in respiratory enzymes, which could result in destruction of the trypanosome. For example, an incoming antibody, with specificity for replikin structures, might attach to the replikin decoy structure, leaving the true replikin structure untouched.

[0210] Therefore, anti-replikin treatment of malaria would require two phases (dual treatment): i) preliminary treatment with proteolytic enzymes that cleave theReplikin decoys, permitting 'safe passage' of the specific anti-Replikin treatment; and ii) attacking malaria Replikins either with specific antibodies or by cellular immunity engendered by synthetic malaria Replikin vaccines or by organic means targeting the malaria Replikins.

[0211] Table 10 provides a list of several Plasmodium falciparum Replikin sequences. It should be noted that this list is not meant to be complete. Different isolates of the organism may contain other Replikin structures.

Table 10

Malaria Replikins

a) Primary structure of a Plasmodium falciparum malaria antigen located at the merozoite surface and within the parasitophorous vacuole

i) DECOYS:

(C-Terminal)

keeeekekekekekeekekeekekeekekekeekekekeekeeekk (SEQ ID NO. 293), or

keeeekekekekekeekekeekekeekekekeekekekeekeeekkek (SEQ ID NO. 294), or

keeeekekekekekeekekeekekeekekekeekekekeekeeekk (SEQ ID NO. 295), or

keeeekekek (SEQ ID NO. 296)

ii) Replikins:

hkklikalkkniesiqnkk (SEQ ID NO. 297)

hkklikalkkniesiqnkm (SEQ ID NO. 298)

hkklikalkk (SEQ ID NO. 299)

hkklikalk (SEQ ID NO. 300)

katysfvntkkkiislksqghkk (SEQ ID NO. 301)

katysfvntkkkiislksqghk (SEQ ID NO. 302)

katysfvntkkkiislksqgh (SEQ ID NO. 303)

htyvkgkkapsdpqea dikeeckellkek (SEQ ID NO. 304)

kiislksqghk (SEQ ID NO. 305)

kkkkfeplkngnvsetiklih (SEQ ID NO. 306)

kkkfeplkngnvsetiklih (SEQ ID NO. 307)

kkfeplkngnvsetiklih (SEQ ID NO. 308)

kngnvsetiklih (SEQ ID NO. 309)

klihlgnkdkk (SEQ ID NO. 310)

kvkkigvtlkkfeplkngnvsetiklihlgnkdkkh (SEQ ID NO. 311)

hliyknksynplllscvkkmnmlkenvdyiqnqnlfkelmnqkatysfvntkkkiislk (SEQ ID NO. 312)

hliyknksynplllscvkkmnmlkenvdyiqnqnlfkelmnqkatysfvntk (SEQ ID NO. 313)

hliyknksynplllscvkkmnmlkenvdyiqnqnlfkelmnqk (SEQ ID NO. 314)

hliyknksynplllscvkkmnmlkenvdyiqknqnlfk (SEQ ID NO. 315)

hliyknksynpllllscvkkmnmmlk (SEQ ID NO. 316)

ksannsanngkknnaeemknlvnflqshkklikalkkniesiqnkkh (SEQ ID NO. 317)

kknnaeemknlvnflqshkklikalkkniesiqnkkh (SEQ ID NO. 318)

knlvnflqshkklikalkkniesiqnkkh (SEQ ID NO. 319)

kklikalkkniesiqnkkh (SEQ ID NO. 320)

klikalkkniesiqnkkh (SEQ ID NO. 321)

kkniesiqnkkh (SEQ ID NO. 322)

kniesiqnkkh (SEQ ID NO. 323)

knnaeemknlvnflqsh (SEQ ID NO. 324)

kklikalkkniesiqnkkqghkk (SEQ ID NO. 325)

kknnaeemknlvnflqshk (SEQ ID NO. 326)

knnaeemknlvnflqsh (SEQ ID NO. 327)

klikalkkniesiqnkkqghkk (SEQ ID NO. 328)

kvkkigvtlkkfeplkngnvsetiklih (SEQ ID NO. 329)

kngnvsetiklih (SEQ ID NO. 330)

klihlgnkdkk (SEQ ID NO. 331)

ksannsanngkknnaeemknlvnflqsh (SEQ ID NO. 332)

kknnaeemknlvnflqsh (SEQ ID NO. 333)

kklikalkkniesiqnkkh (SEQ ID NO. 334)

kalkkniesiqnkkh (SEQ ID NO. 335)

kkniesiqnkkh (SEQ ID NO. 336)

kelmnqkatysfvntkkkiislksqgh (SEQ ID NO. 337)

ksqghkk (SEQ ID NO. 338)

kkkiislksqgh (SEQ ID NO. 339)

kkiislksqgh (SEQ ID NO. 340)

kkniesiqnkkh (SEQ ID NO. 341)

kniesiqnkkh (SEQ ID NO. 342)

htyvkgkkapsdpqeadikeeckellkek (SEQ ID NO. 343)

htyvkgkkapsdpqeadikeeckellk (SEQ ID NO. 344)

b) "liver stage antigen-3" gene="LSA-3" Replikins

henvlsaalentqseeekkevidvieevk (SEQ ID NO. 345)

kenvvttilekveettaesvttfsnileeiqentitndtieekleelh (SEQ ID NO. 346)

hylqqmkekfskek (SEQ ID NO. 347)

hylqqmkekfskeknmnvievtnkaekkgnvqvtnktekttk (SEQ ID NO. 348)

hylqqmkekfskeknmnvievtnkaekkgnvqvtnktekttkvdknnk (SEQ ID NO. 349)

hylqqmkekfskeknmnvievtnkaekkgnvqvtnktekttkvdknnkvpkkrrtqk (SEQ ID NO. 350)

hylqqmkekfskeknmnvievtnkaekkgnvqvtnktekttkvdknnkvpkkrrtqksk (SEQ ID NO. 351)

hvdevmkyvqkidkevdkevskalesskndvtnvlkqnqdffskvknfvkkyk (SEQ ID NO. 352)

hvdevmkyvqkidkevdkevskalesskndvtnvlkqnqdffskvknfvkk (SEQ ID NO. 353)

hvdevmkyvqkidkevdkevskalesskndvtnvlkqnqdffsk (SEQ ID NO. 354)

hvdevmkyvqkidkevdkevskalesskndvtnvlk (SEQ ID NO. 355)

hvdevmkyvqkidkevdkevskalesk (SEQ ID NO. 356)

hvdevmkyvqkidkevdkevsk (SEQ ID NO. 357)

hvdevmkyvqkidkevdk (SEQ ID NO. 358)

hvdevmkyvqkidk (SEQ ID NO. 359)

kdevidlivqkckriekvkakkkklekkveegvsglkkh (SEQ ID NO. 360)

kvkakkkklekkveegvsglkkh (SEQ ID NO. 361)

kakkkklekkveegvsglkkh (SEQ ID NO. 362)

kkkklekkveegvsglkkh (SEQ ID NO. 363)

kkklekkveegvsglkkh (SEQ ID NO. 364)

kklekkveegvsglkkh (SEQ ID NO. 365)

klekkveegvsglkkh (SEQ ID NO. 366)

kkveegvsglkkh (SEQ ID NO. 367)

kveegvsglkkh (SEQ ID NO.368)

hveqnvyvdvdvpamkdqflgilneagglkemffnledvfksesdvitveeikdepvqk (SEQ ID NO. 369)

hikgleeddleevddlkgsildmlkgdmelgdmdkesledvtlklgerveslk (SEQ ID NO. 370)

hikgleeddleevddlkgsildmlkgdmelgdmdkesledvtlk (SEQ ID NO. 371)

hikgleeddleevddlkgsildmlkgdmelgdmdk (SEQ ID NO. 372)

hikgleeddleevddlkgsildmlk (SEQ ID NO. 373)

hiisgdadvlssalgmdeeqmktrkkaqrpk (SEQ ID NO. 374)

hdittldevvelkdveedkiek (SEQ ID NO. 375)

kkleevhelk (SEQ ID NO. 376)

kleevhelk (SEQ ID NO. 377)

ktietdileekdkeiekdh (SEQ ID NO. 378)

kkeiekdhfek (SEQ ID NO. 379)

kdhfek (SEQ ID NO. 380)

kfeeeaeeikh (SEQ ID NO. 381)

c) 28 KDA ookinete surface antigen precursor Replikins:

kdgdtkctlecaqgkkcikhksdhnhksdhnhksdpnhkkknnnnnk (SEQ ID NO. 382)

kdgdtkctlecaqgkkcikhksdhnhksdhnhksdpnhkk (SEQ ID NO. 383)

kdgdtkctlecaqgkkcikhksdhnhksdhnhksdpnhk (SEQ ID NO. 384)

kdgdtkctlecaqgkkcikhksdhnhksdhnhk (SEQ ID NO. 385)

kdgdtkctlecaqgkkcikhksdhnhk (SEQ ID NO. 386)

kdgdtkctlecaqgkkcikhk (SEQ ID NO. 387)

kdgdtkctlecaqgkk (SEQ ID NO. 388)

kdgdtkctlecaqgk (SEQ ID NO. 389)

kciqaecnykccgeqkcvwdgih (SEQ ID NO. 390)

kccgeqkcvwdgih (SEQ ID NO. 391)

hieckcnndyvltnryeccpknkctsledtnk (SEQ ID NO. 392)

d) Blood stage trophozoites and schizonts Replikins:

ksdhnhksdhnhksdhnhksdhnhksdpnhkkknnnnnk (SEQ ID NO. 393)

ksdhnhksdhnhksdhnhksdpnhkkknnnnnk (SEQ ID NO. 394)

ksdhnhksdhnhksdpnhkkknnnnnk (SEQ ID NO. 395)

ksdhnhksdpnhkkknnnnnk (SEQ ID NO. 396)

kkknnnnnkdnksdpnhk (SEQ ID NO. 397)

kknnnnnkdnksdpnhk (SEQ ID NO. 398)

knnnnnkdnksdpnhk (SEQ ID NO. 399)

kdnksdpnhk (SEQ ID NO. 400)

ksdpnhk (SEQ ID NO. 401)

hslyalqqneeyqkvknekdqneikkikqlieknk (SEQ ID NO. 402)

hslyalqqneeyqkvknekdqneikkik (SEQ ID NO. 403)

hslyalqqneeyqkvknekdqneikk (SEQ ID NO. 404)

hslyalqqneeyqkvknekdqneik (SEQ ID NO. 405)

hklenleemdk (SEQ ID NO. 406)

khfddntneqk (SEQ ID NO. 407)

kkeddekh (SEQ ID NO. 408)

keennkkeddekh (SEQ ID NO. 409)

ktssgilnkeennkkeddekh (SEQ ID NO. 410)

knihikk (SEQ ID NO. 411)

hikkkegidigyk (SEQ ID NO. 412)

kkmwtcklwdnkgneitknih (SEQ ID NO. 413)

kkgiqwnllkkmwtcklwdnkgneitknih (SEQ ID NO. 414)

kekkdsnenrkkkqkedkknpnklkkieytnkithffkaknnkqqnnvth (SEQ ID NO. 415)

kkdsnenrkkkqkedkknpnklkkieytnkithffkaknnkqqnnvth (SEQ ID NO. 416)

kdsnenrkkkqkedkknpnklkkieytnkithffkaknnkqqnnvth (SEQ ID NO. 417)

kkqkedkknpnklkkieytnkithffkaknnkqqnnvth (SEQ ID NO. 418)

kqkedkknpnklkkieytnkithffkaknnkqqnnvth (SEQ ID NO. 419)

kedkknpnklkkieytnkithffkaknnkqqnnvth (SEQ ID NO. 420)

knpnklkkieytnkithffkaknnkqqnnvth (SEQ ID NO. 421)

kkieytnkithffkaknnkqqnnvth (SEQ ID NO. 422)

kieytnkithffkaknnkqqnnvth (SEQ ID NO. 423)

kithffkaknnkqqnnvth (SEQ ID NO. 424)

hknnedikndnskdikndnskdikndnskdikndnnedikndnskdik (SEQ ID NO. 425)

hknnedikndnskdikndnskdikndnskdikndnnedikndnsk (SEQ ID NO. 426)

hknnedikndnskdikndnskdikndnskdikndnnedik (SEQ ID NO. 427)

hknnedikndnskdikndnskdikndnskdik (SEQ ID NO. 428)

hknnedikndnskdikndnskdikndnsk (SEQ ID NO. 429)

hknnedikndnskdikndnskdik (SEQ ID NO. 430)

hknnedikndnskdikndnsk (SEQ ID NO. 431)

hknnedikndnskdik (SEQ ID NO. 432)

hknnedik (SEQ ID NO. 433)

kkyddlqnkynilnklknsleekneelkkyh (SEQ ID NO. 434)

kyddlqnkynilnklknsleekneelkkyh (SEQ ID NO. 435)

kynilnklknsleekneelkkyh (SEQ ID NO. 436)

klknsleekneelkkyh (SEQ ID NO. 437)

knsleekneelkkyh (SEQ ID NO. 438)

kneelkkyh (SEQ ID NO. 439)

hmgnnqdinenvynikpqefkeeeeedismvntkk (SEQ ID NO. 440)

knsnelkrindnffklh (SEQ ID NO. 441)

kpclykkckisqclykkckisqvwwcmpvkdtfntycmnvlnskicmniekiph (SEQ ID NO. 442)

hinneytnknpknellykneernyndnnikdyinsmnfkk (SEQ ID NO. 443)

hinneytnknpknellykneernyndnnikdyinsmnfk (SEQ ID NO. 444)

hinneytnknpknellyk (SEQ ID NO. 445)

knktnqskgvkgeyekkketngh (SEQ ID NO. 446)

ktnqskgvkgeyekkketngh (SEQ ID NO. 447)

kgvkgeyekkketngh (SEQ ID NO. 448)

kgeyekkketngh (SEQ ID NO. 449)

ksgmytnegnksceesykkkssssnkvh (SEQ ID NO. 450)

ksceesykkkssssnkvh (SEQ ID NO. 451)

kkkssssnkvh (SEQ ID NO. 452)

kkssssnkvh (SEQ ID NO. 453)

kssssnkvh (SEQ ID NO. 454)

himlksgmytnegnksceesykkkssssnk (SEQ ID NO. 455)

himlksgmytnegnksceesykkk (SEQ ID NO. 456)

himlksgmytnegnksceesykk (SEQ ID NO. 457)

himlksgmytnegnksceesyk (SEQ ID NO. 458)

kplaklrkrektqinktkyergdviidnteiqkiiirdyhetlnvhkldh (SEQ ID NO. 459)

krektqinktkyergdviidnteiqkiiirdyhetlnvhkldh (SEQ ID NO. 460)

ktqinktkyergdviidnteiqkiiirdyhetlnvhkldh (SEQ ID NO. 461)

kplaklrkrektqinktkyergdviidnteiqkiiirdyhetlnvh (SEQ ID NO. 462)

kplaklrkrektqinktkyergdviidnteiqkiiirdyh (SEQ ID NO. 463)

klrkrektqinktkyergdviidnteiqkiiirdyh (SEQ ID NO. 464)

krektqinktkyergdviidnteiqkiiirdyh (SEQ ID NO. 465)

ktqinktkyergdviidnteiqkiiirdyh (SEQ ID NO. 466)

kkdkekkkdsnenrkkkqkedkknpndnklkkieytnkith (SEQ ID NO. 467)

kdkekkkdsnenrkkkqkedkknpndnklkkieytnkith (SEQ ID NO. 468)

kekkkdsnenrkkkqkedkknpndnklkkieytnkith (SEQ ID NO. 469)

kkkdsnenrkkkqkedkknpndnklkkieytnkith (SEQ ID NO. 470)

kkdsnenrkkkqkedkknpndnklkkieytnkith (SEQ ID NO. 471)

kdsnenrkkkqkedkknpndnklkkieytnkith (SEQ ID NO. 472)

kkkqkedkknpndnklkkieytnkith (SEQ ID NO. 473)

kkqkedkknpndnklkkieytnkith (SEQ ID NO. 474)

kqkedkknpndnklkkieytnkith (SEQ ID NO. 475)

kedkknpndnklkkieytnkith (SEQ ID NO. 476)

kknpndnklkkieytnkith (SEQ ID NO. 477)

knpndnklkkieytnkith (SEQ ID NO. 478)

klkkieytnkith (SEQ ID NO. 479)

kkieytnkith (SEQ ID NO. 480)

kieytnkith (SEQ ID NO. 481)

hgqikiedvnnenfnneqmknkyndeekmdiskskslksdflek (SEQ ID NO. 482)

hgqikiedvnnenfnneqmknkyndeekmdiskskslk (SEQ ID NO. 483)

hgqikiedvnnenfnneqmknkyndeekmdisksk (SEQ ID NO. 484)

hgqikiedvnnenfnneqmknkyndeekmdisk (SEQ ID NO. 485)

kkyddlqnkynilnklknsleekneelkkyh (SEQ ID NO. 486)

kyddlqnkynilnklknsleekneelkkyh (SEQ ID NO. 487)

kynilnklknsleekneelkkyh (SEQ ID NO. 488)

klknsleekneelkkyh (SEQ ID NO. 489)

knsleekneelkkyh (SEQ ID NO. 490)

kneelkkyh (SEQ ID NO. 491)

hmgnnqdinenvynikpqefkeeeeedismvntkkeddiqenik (SEQ ID NO. 492)

ktnlyniynnknddkdnildnenreglyledvmknsnelkrindnffklh (SEQ ID NO. 493)

knsnelkrindnffklh (SEQ ID NO. 494)

krindnffklh (SEQ ID NO. 495)

hinneytnknpknellykneernyndnnikdyinsmnfkk (SEQ ID NO. 496)

hinneytnknpknellykneernyndnnikdyinsmnfk (SEQ ID NO. 497)

hinneytnknpknellyk (SEQ ID NO. 498)

kpelykkekisqvwwempvkdtfntyernnvlnskienniekiph (SEQ ID NO. 499)

kekisqvwwempvkdtfntyernnvlnskienniekiph (SEQ ID NO. 500)

kienniekiph (SEQ ID NO. 501)

knktngskgvkgeyekkketngh (SEQ ID NO. 502)

ktngskgvkgeyekkketngh (SEQ ID NO. 503)

kgvkgeyekkketngh (SEQ ID NO. 504)

kgeyekkketngh (SEQ ID NO. 505)

ktiekinkskswffeeldeidkplaklrkrektqinktkyergdviidnteiqkiirdyh (SEQ ID NO. 506)

kinkskswffeeldeidkplaklrkrektqinktkyergdviidnteiqkiirdyh (SEQ ID NO. 507)

kplaklrkrektqinktkyergdviidnteiqkiirdyh (SEQ ID NO. 508)

himlksqmytnegnkscecsykkkssssnkvh (SEQ ID NO. 509)

klrkrektqinktkyergdviidnteiqkiirdyh (SEQ ID NO. 510)

krektqinktkyergdviidnteiqkiirdyh (SEQ ID NO. 511)

ktqinktkyergdviidnteiqkiirdyh (SEQ ID NO. 512)

kplaklrkrektqinktkyergdviidnteiqkiirdyhtlnvhkldh (SEQ ID NO. 513)

klrkrektqinktkyergdviidnteiqkiirdyhtlnvhkldh (SEQ ID NO. 514)

krektqinktkyergdviidnteiqkiirdyhtlnvhkldh(SEQ ID NO. 515)

ktqinktkyergdviidnteiqkiirdyhtlnvhkldh (SEQ ID NO. 516)

kplaklrkrektqinktkyergdviidnteiqkiirdyhtlnvh (SEQ ID NO. 517)

klrkrektqinktkyergdviidnteiqkiirdyhtlnvh (SEQ ID NO. 518)

krektqinktkyergdviidnteiqkiirdyhtlnvh (SEQ ID NO. 519)

ktqinktkyergdviidnteiqkiirdyhtlnvh (SEQ ID NO. 520)

himlksqmytnegnkscecsykkkssssnkvh (SEQ ID NO. 521)

ksqmytnegnkscecsykkkssssnkvh (SEQ ID NO. 522)

kscecsykkkssssnkvh (SEQ ID NO. 523)

kkkssssnkvh (SEQ ID NO. 524)

kkssssnkvh (SEQ ID NO. 525)

kssssnkvh (SEQ ID NO. 526)

himlksqmytnegnkscecsykkkssssnk (SEQ ID NO. 527)

himlksqmytnegnkscecsykkk (SEQ ID NO. 528)

himlksqmytnegnkscecsykk (SEQ ID NO. 529)

himlksqmytnegnkscecsyk (SEQ ID NO. 530)

hmnhniqiykdkrinfmnphkvmyhdnmsknertek (SEQ ID NO. 531)

hmnhniqiykdkrinfmnphkvmyhdnmsk (SEQ ID NO. 532)

hmnhniqiykdkrinfmnphk (SEQ ID NO. 533)

hkvmyhdnmsknertek (SEQ ID NO. 534)

hkvmyhdnmsk (SEQ ID NO. 535)

REPLIKINS IN STRUCTURAL PROTEINS

[0212] It has also been determined that some structural proteins include Replikin structures. Structural proteins are molecules involved in tissue and organ

support, such as collagen in skin and connective tissue and in membrane structures, for example amyloid A4 precursor protein (APP) in brain. Overproduction of these proteins is associated with disease; specifically, scleroderma in the case of overproduction of collagen in skin (Table 11) and Alzheimer's Disease in the case of overproduction of APP in the brain (Table 12).

[0213] The association of scleroderma and malignancy has been a source of controversy during recent years. Several mechanisms of interrelationship have been suggested in earlier reports. Recent long-term studies suggest an increased association-ratio of scleroderma and malignancy. However, the underlying mechanisms remain elusive. (Wenzel, J. Eur. J. Dermatol. 20002 May-Jun; 12(3): 296-300).

[0214] Several proteins concerned with the excessive production of proteins in scleroderma have been found to contain Replikin structures. Thus, these provide further examples of unrecognized targets for inhibition or cessation of excessive collagen production. Table 11 provides a list of proteins in scleroderma and the associated Replikins.

[0215] The APP protein is the source of the amyloid beta A4 protein, which in excessive amounts forms plaques in the extracellular spaces in the brain, producing toxic effects associated with nerve cell loss in Alzheimer's Disease. Most studies to date have focused on the inability to clear the excessive deposits of A4, but have not considered that, rather than a waste clearance problem, this may actually be a problem of overproduction of the precursor protein APP. The high concentration of the Replikins in APP (3.3 Replikins per 100 amino acids) strongly suggest that overproduction may well be the cause of Alzheimer's Disease (Table 12). Therefore, the Replikins contained in Table 12 can be blocked or inhibited by the same methods as illustrated in detail for the glioma Replikin.

Table 11

Proteins overproduced in scleroderma and associated Replikins:

PMC1 HUMAN:

hreictiqssggimllkdqvlrcskiagvkvaeitelilk (SEQ ID NO.536)

hreictiqssggimllkdqvlresk (SEQ ID NO.537)

34KD nucleolar scleroderma antigen:

hreictiqssggimllkdqvlrcskiagvkvaeitelikllkalendqk (SEQ ID NO.538)

hreictiqssggimllkdqvlrcskiagvkvaeitelilk (SEQ ID NO.539)

Fibrillarin:

kkmqqenmkqpeqltlepyerdh (SEQ ID NO.540)

kmqqenmkpqeqltlepyerdh (SEQ ID NO. 541)

SPOP HUMAN:

hemeeskknrveindvepevfkemmcfiytgkapnldk (SEQ ID NO.542)

hemeeskknrveindvepevfkemmcfiytgk (SEQ ID NO.543)

Centromere protein C:

khgelkvyk (SEQ ID NO.544)

klilgpqeekgkqh (SEQ ID NO. 545)

hnrihhk (SEQ ID NO. 546)

hhnssrkstkktnqssk (SEQ ID NO. 547)

hnssrkstkktnqssk (SEQ ID NO. 548)

khhnilpktlandkhshkph (SEQ ID NO.549)

hhnilpktlandkhshk (SEQ ID NO. 550)

hnilpktlandkhshk (SEQ ID NO. 551)

hnilpktlandk (SEQ ID NO. 552)

kntpdskkissrnindhh (SEQ ID NO.553)

kntpdskkissrnindh (SEQ ID NO. 554)

kdteiqspskecqkshpksvpvsskkk (SEQ ID NO. 555)

kdteiqspskecqkshpksvpvsskk (SEQ ID NO. 556)

hpksvpvsskkk (SEQ ID NO. 557)

hpksvpvsskk (SEQ ID NO. 558)

hpksvpvssk (SEQ ID NO. 559)

Factor CTCBF, KU antigen:

kalqekveikqlnh (SEQ ID NO. 560)

ktlfplieakkdqvtageifgdnhedgptakklktegggah (SEQ ID NO. 561)

ktlfplieakkkdqvtageifqdnb (SEQ ID NO. 562)

klevfkkierhsih (SEQ ID NO. 563)

klevfkkierh (SEQ ID NO. 564)

kgpsfplkgiteqqkegleivk (SEQ ID NO. 565)

hgpsfplkgiteqqk (SEQ ID NO. 566)

ATP synthase subunit 6:

htllkilstfllfk (SEQ ID NO. 567)

hllgnndknllpsk (SEQ ID NO. 568)

FBRL nuclear protein:

hrhegvficrgkedalvtk (SEQ ID NO. 569)

hegvficrgkedalvtk (SEQ ID NO. 570)

hsggnrgrgrggkrghqsgk (SEQ ID NO. 571)

krgnqsgknvmveph (SEQ ID NO. 572)

krgnqsgknvmvephrh (SEQ ID NO. 573)

kkmqqenmkpqeqltlepyerdh (SEQ ID NO.574)

kmqqenmkpqeqltlepyerdh (SEQ ID NO. 575)

HP1Hs-alpha protein:

haypedaenkeketak (SEQ ID NO. 576)

keanvkepqiviafyeerltwh (SEQ ID NO. 577)

kvldrrvvkgqveyllkwkgfseeh (SEQ ID NO. 578)

kgqveyllkwkgfseeh (SEQ ID NO. 579)

FM/Scl nucleolar protein:

ksevaagvkksglpsaerlenvlfgphdcsh (SEQ ID NO.580)

ksevaagvkksgplpsaerlenvlfgph (SEQ ID NO. 581)

kaaeygkkaksetfrllhakniirpqlk (SEQ ID NO. 582)

kaaeygkkaksetfrllhak (SEQ ID NO. 583)

ksetfrllhak (SEQ ID NO. 584)

hakniirpqlk (SEQ ID NO. 585)

hmnlkiaeelpk (SEQ ID NO. 586)

hsldhllklycnvdsnk (SEQ ID NO. 587)

hllklycnvdsnk (SEQ ID NO. 588)

Table 12

Amyloid beta A4 precursor protein (APP) Replikins:

kakerleakh (SEQ ID NO. 589)

kdrqhtlk (SEQ ID NO. 590)

kdrqhtlkh (SEQ ID NO. 591)

ketcsekstnlh (SEQ ID NO. 592)

kteeisevkmdaefgh (SEQ ID NO. 593)

kteeisevkmdaefghdsgfevrh (SEQ ID NO. 594)

kkyvraeqkdrqhtlkh (SEQ ID NO. 595)

kyvraeqkdrqhtlkh (SEQ ID NO. 596)

kkyvraeqkdrqh (SEQ ID NO. 597)

kyvraeqkdrqht (SEQ ID NO. 598)

hhvfnmlkkyvraeqk (SEQ ID NO. 599)

hvfnmlkkyvraeqk (SEQ ID NO. 600)

hhvfnmlkkyvraeqkdrqhtlkh (SEQ ID NO. 601)

hvfnmlkkyvraeqkdrqhtlkh (SEQ ID NO. 602)

hahfqkakerleakh (SEQ ID NO. 603)

hahfqkakerleak (SEQ ID NO. 604)

hfqkakerleak (SEQ ID NO. 605)

hqermdvcethlhwhtvaketcsekstnlh (SEQ ID NO. 606)

hqermdvcethlhwhtvaketcsek (SEQ ID NO. 607)

hwhtvaketcsek (SEQ ID NO. 608)

htvaketcsek (SEQ ID NO. 609)

hlhwhtvaketcsek (SEQ ID NO. 610)

hmnvqngkwesdpsgtktcigtk (SEQ ID NO. 611)

hmnvqngkwesdpsgtk (SEQ ID NO. 612)

PASSIVE IMMUNITY

[0216] In another embodiment of the invention, isolated Replikin peptides may be used to generate antibodies, which may be used, for example to provide passive immunity in an individual. Passive immunity to the strain of influenza identified by the method of the invention to be the most likely cause of future influenza infections may be obtained by administering antibodies to Replikin sequences of the identified strain of influenza virus to patients in need. Similarly, passive immunity to malaria may be obtained by administering antibodies to Plasmodium falciparum Replikin(s).

[0217] Various procedures known in the art may be used for the production of antibodies to Replikin sequences. Such antibodies include but are not limited to polyclonal, monoclonal, chimeric, humanized, single chain, Fab fragments and fragments produced by an Fab expression library. Antibodies that are linked to a cytotoxic agent may also be generated. Antibodies may also be administered in combination with an antiviral agent. Furthermore, combinations of antibodies to different Replikins may be administered as an antibody cocktail.

[0218] For the production of antibodies, various host animals or plants may be immunized by injection with a Replikin peptide or a combination of Replikin peptides, including but not limited to rabbits, mice, rats, and larger mammals.

[0219] Monoclonal antibodies to Replikins may be prepared by using any technique that provides for the production of antibody molecules. These include but are not limited to the hybridoma technique originally described by Kohler and Milstein, (Nature, 1975, 256:495-497), the human B-cell hybridoma technique (Kosbor et al., 1983, Immunology Today, 4:72), and the EBV hybridoma technique (Cole et al., Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96). In addition, techniques developed for the production of chimeric antibodies (Morrison et al., 1984, Proc. Nat. Acad. Sci USA, 81:6851-6855) or other techniques may be used. Alternatively, techniques described for the production of single chain antibodies (US 4,946,778) can be adapted to produce Replikin-specific single chain antibodies.

[0220] Particularly useful antibodies of the invention are those that specifically bind to Replikin sequences contained in peptides and/or polypeptides of influenza virus. For example, antibodies to any of peptides observed to be present in an emerging or re-emerging strain of influenza virus and combinations of such antibodies are useful in the treatment and/or prevention of influenza. Similarly, antibodies to any Replikins present on malaria antigens and combinations of such antibodies are useful in the prevention and treatment of malaria.

[0221] Antibody fragments which contain binding sites for a Replikin may be generated by known techniques. For example, such fragments include but are not limited to F(ab')2 fragments which can be produced by pepsin digestion of the antibody molecules and the Fab fragments that can be generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries can be generated (Huse et al., 1989, Science, 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity.

[0222] The fact that antimalignin antibody is increased in concentration in human malignancy regardless of cancer cell type (Figure 5), and that this antibody binds to malignant cells regardless of cell type now may be explained by the presence of the Replikin structures herein found to be present in most malignancies (Figure 1 and Table 2). Population studies have shown that antimalignin antibody increases in concentration in healthy adults with age, and more so in high-risk families, as the frequency of cancer increases. An additional two-fold or greater antibody increase, which occurs in early malignancy, has been independently confirmed with a sensitivity of 97% in breast cancers 1-10 mm in size. Shown to localize preferentially in malignant cells in vivo, histochemically the antibody does not bind to normal cells but selectively binds to (Figure 4a,b) and is highly cytotoxic to transformed cells in vitro (Figure 4c-f). Since in these examples the same antibody is bound by several cell types, that is, brain glioma, hematopoietic cells (leukemia), and small cell carcinoma of lung, malignant Replikin class unity is again demonstrated.

[0223] Antimalignin does not increase with benign proliferation, but specifically increases only with malignant transformation and replication in breast in vivo and returns from elevated to normal values upon elimination of malignant cells (Figure 5). Antimalignin antibody concentration has been shown to relate quantitatively to the survival of cancer patients, that is, the more antibody, the longer the survival. Taken together, these results suggest that anti-Replikin antibodies may be a part of a mechanism of control of cell transformation and replication. Augmentation of this immune response may be useful in the control of replication, either actively with synthetic Replikins as vaccines, or passively by the administration of anti-Replikin antibodies, or by the introduction of non-immune based organic agents, such as for example, carbohydrates, lipids and the like, which are similarly designed to target the Replikin specifically.

[0224] In another embodiment of the invention, immune serum containing antibodies to one or more Replikins obtained from an individual exposed to one or more Replikins may be used to induce passive immunity in another individual or animal. Immune serum may be administered via i.v. to a subject in need of treatment. Passive immunity also can be achieved by injecting a recipient with preformed antibodies to one or more Replikins. Passive immunization may be used to provide immediate protection to individuals who have been exposed to an infectious organism. Administration of immune serum or preformed antibodies is routine and the skilled practitioner can readily ascertain the amount of serum or antibodies needed to achieve the desired effect.

SYNTHETIC REPLIKIN VACCINES (ACTIVE IMMUNITY)

[0225] Synthetic Replikin vaccines, based on Replikins such as the glioma Replikin (SEQ ID NO.: 1) "kagvaflhkk" or the hepatitis C Replikin (SEQ ID NO.: 18) "hyppkpgcivpak", or HIV Replikins such as (SEQ ID NO.: 5) "kcfncgkcgh" or (SEQ ID NO.: 6) "kvylawvpahk" or preferably, an influenza vaccine based on conserved and/or emerging or re-emerging Replikin(s) over a given time period may be used to augment antibody concentration in order to lyse the respective virus infected cells and release virus extracellularly where chemical treatment can then be effective. Similarly, a malaria vaccine, based on Replikins observed in Plasmodium

falciparum malaria antigens on the merozoite surface or within the parasitophorous vacuole, for example, can be used to generate cytotoxic antibodies to malaria. Table 7 shows the relation of shortening or compacting of Replikin sequences to mortality rate caused by the organisms which contain these Replikins, to as short as seven amino acids. This correlation has been found by us to be a general phenomenon regardless of the type of organism. We have also found that there may be a progression over time to the shortened Replikin structure, as in influenza and SARS viruses. There is abundant evidence that there are constant evolutionary and competitive pressures for the emergence of constantly increasing "efficacy" of each infectious organism. Based upon these observations, and by projection, it would appear that if evolutionary pressures are towards shorter and shorter Replikins, with higher and higher concentrations of lysine (k), to as high as 70% as in EEL leukemia (Table 7), then the projected theoretical ideal would be the shortest possible Replikin permitted by the algorithm which defines a Replikin, that is six amino acids (two ks six to ten amino acids apart), with the highest possible % k (see Example below in deduced Replikin "kkkkhk", which contains 83.3% k, 5/6, and one obligatory "h"). We have therefore, so-to-speak, taken what appears to be, or might be, the next evolutionary step, not apparently as yet taken by the organisms themselves, and devised the resultant deduced Replikins to use as general vaccines. These Replikins which we have deduced have maximum %'k's, therefore maximum potential binding capacity, plus the constituent 'h' by definition required for the Replikin, giving the potential for 'h' connection to redox energy systems. These devised Replikins are least likely to be cleaved by organisms because of their short length (proteins are cleaved to 6 to 10 amino acids long in processing for presentation to and recognition by immune cells), therefore most likely to present intact to immune–forming apparatuses in the organism to which they are administered, and, because of their high k content, they are most likely to generate a maximum immune response which mimics and may increase the maximum such response which can be generated against short homologous high mortality Replikins. Further, we have found that high % k Replikins generate the highest antibody responses when administered to rabbits. These synthetic peptides, designed by us, are designated as Universal synthetic epitopes, or "UTOPE"s , and the vaccines based upon these UTOPEs, are designated "UVAX"s. UVAXs, deduced synthetic

vaccines, may be used as sole vaccines or as adjuvants when administered with more specific Replikin vaccines or other vaccines. The following are examples of deduced UTOPEs and UVAXs:

DEVISED SYNTHETIC REPLIKIN (UTOPE OR UVAX)	SEQ ID NO:
kkkkhk	732
kkkhkk	733
kkhkkk	734
khkkkk	735
kkkkkkh	736
kkkkkhk	737
kkkkhkk	738
kkkhkkk	739
kkhkkkk	740
khkkkkk	741
hkkkkkk	742

[0226] Recognin and/or Replikin peptides may be administered to a subject to induce the immune system of the subject to produce anti-Replikin antibodies. Generally, a 0.5 to about 2 mg dosage, preferably a 1 mg dosage of each peptide is administered to the subject to induce an immune response. Subsequent dosages may be administered if desired.

[0227] The Replikin sequence structure is associated with the function of replication. Thus, whether the Replikins of this invention are used for targeting sequences that contain Replikins for the purpose of diagnostic identification, promoting replication, or inhibiting or attacking replication, for example, the structure-function relationship of the Replikin is fundamental.

[0228] It is preferable to utilize only the specific Replikin structure when seeking to induce antibodies that will recognize and attach to the Replikin fragment and thereby cause destruction of the cell. Even though the larger protein sequence

may be known in the art as having a "replication associated function," vaccines using the larger protein often have failed or proven ineffective.

[0229] Although the present inventors do not wish to be held to a single theory, the studies herein suggest that the prior art vaccines are ineffective because they are based on the use of the larger protein sequence. The larger protein sequence invariably has one or more epitopes (independent antigenic sequences that can induce specific antibody formation); Replikin structures usually comprise one of these potential epitopes. The presence of other epitopes within the larger protein may interfere with adequate formation of antibodies to the Replikin, by "flooding" the immune system with irrelevant antigenic stimuli that may preempt the Replikin antigens, See, e.g., Webster, R.G., J. Immunol., 97(2):177-183 (1966); and Webster et al., J. Infect. Dis., 134:48-58, 1976; Klenerman et al, Nature 394:421-422 (1998) for a discussion of this well-known phenomenon of antigenic primacy whereby the first peptide epitope presented and recognized by the immune system subsequently prevails and antibodies are made to it even though other peptide epitopes are presented at the same time. This is another reason that, in a vaccine formulation, it is important to present the constant Replikin peptide to the immune system first, before presenting other epitopes from the organism so that the Replikin is not pre-empted but lodged in immunological memory.

[0230] The formation of an antibody to a non-Replikin epitope may allow binding to the cell, but not necessarily lead to cell destruction. The presence of structural "decoys" on the C-termini of malaria proteins is another aspect of this ability of other epitopes to interfere with binding of effective anti-Replikin antibodies, since the decoy epitopes have many lysine residues, but no histidine residues. Thus, decoy epitopes may bind anti-Replikin antibodies, but may keep the antibodies away from histidine-bound respiratory enzymes. Treatment may therefore be most efficacious in two stages: 1) proteases to hydrolyze decoys, then; 2) anti-Replikin antibodies or other anti-Replikin agents.

[0231] It is well known in the art that in the course of antibody production against a "foreign" protein, the protein is first hydrolyzed into smaller fragments.

Usually fragments containing from about six to ten amino acids are selected for antibody formation. Thus, if hydrolysis of a protein does not result in Replikin-containing fragments, anti-Replikin antibodies will not be produced. In this regard, it is interesting that Replikins contain lysine residues located six to ten amino acids apart, since lysine residues are known to bind to membranes.

[0232] Furthermore, Replikin sequences contain at least one histidine residue. Histidine is frequently involved in binding to redox centers. Thus, an antibody that specifically recognizes a Replikin sequence has a better chance of inactivating or destroying the cell in which the Replikin is located, as seen with anti-malignin antibody, which is perhaps the most cytotoxic anti-cancer antibody yet described, being active at picograms per cell.

[0233] One of the reasons that vaccines directed towards a particular protein antigen of a disease causing agent have not been fully effective in providing protection against the disease (such as foot and mouth vaccine which has been developed against the VP1 protein or large segments of the VP1 protein) is that the best antibodies have not been produced, that is – it is likely that the antibodies to the Replikins have not been produced. Replikins have not been produced. That is, either epitopes other than Replikins present in the larger protein fragments may interfere according to the phenomenon of antigenic primacy referred to above, and/or because the hydrolysis of larger protein sequences into smaller sequences for processing to produce antibodies results in loss of integrity of any Replikin structure that is present, e.g., the Replikin is cut in two and/or the histidine residue is lost in the hydrolytic processing. The present studies suggest that for an effective vaccine to be produced, the Replikin sequences, and no other epitope, should be used as the vaccine. For example, a vaccine of the invention can be generated using any one of the Replikin peptides identified by the three-point recognition system.

[0234] Particularly preferred peptides - for example - an influenza vaccine include peptides that have been demonstrated to be conserved over a period of one or more years, preferably about three years or more, and/or which are present in a strain of influenza virus shown to have the highest increase in concentration of

Replikins relative to Replikin concentration in other influenza virus strains, e.g., an emerging strain. The increase in Replikin concentration preferably occurs over a period of at least about six months to one year, preferably at least about two years or more, and most preferably about three years or more. Among the preferred Replikin peptides for use in an influenza virus vaccine are those Replikins observed to "re-emerge" after an absence from the hemagglutinin amino acid sequence for one or more years.

[0235] The Replikin peptides of the invention, alone or in various combinations are administered to a subject, preferably by i.v. or intramuscular injection, in order to stimulate the immune system of the subject to produce antibodies to the peptide. Generally the dosage of peptides is in the range of from about 0.1 μg to about 10 mg, preferably about 10 μg to about 1 mg, and most preferably about 50 μg to about 500 ug. The skilled practitioner can readily determine the dosage and number of dosages needed to produce an effective immune response.

QUANTITATIVE MEASUREMENT EARLY RESPONSE(S) TO REPLIKIN VACCINES

[0236] The ability to measure quantitatively the early specific antibody response in days or a few weeks to a Replikin vaccine is a major practical advantage over other vaccines for which only a clinical response months or years later can be measured.

ADJUVANTS

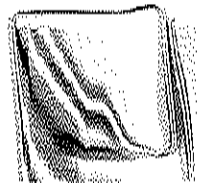
[0237] Various adjuvants may be used to enhance the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels, such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, key limpet hemocyanin, dinitrophenol, and potentially useful human adjuvants such as BCG and Corynebacterium parvum. In addition to the use of synthetic UTOPEs as vaccines in themselves, UTOPEs can be used as adjuvants to other Replikin vaccines and to non-Replikin vaccines.

REPLIKIN NUCLEOTIDE SEQUENCES

[0238] Replikin DNA or RNA may have a number of uses for the diagnosis of diseases resulting from infection with a virus, bacterium or other Replikin encoding agent. For example, Replikin nucleotide sequences may be used in hybridization assays of biopsied tissue or blood, *e.g.*, Southern or Northern analysis, including in situ hybridization assays, to diagnose the presence of a particular organism in a tissue sample or an environmental sample, for example. The present invention also contemplates kits containing antibodies specific for particular Replikins that are present in a particular pathogen of interest, or containing nucleic acid molecules (sense or antisense) that hybridize specifically to a particular Replikin, and optionally, various buffers and/or reagents needed for diagnosis.

[0239] Also within the scope of the invention are oligoribonucleotide sequences, that include antisense RNA and DNA molecules and ribozymes that function to inhibit the translation of Replikin- or recognin-containing mRNA. Both antisense RNA and DNA molecules and ribozymes may be prepared by any method known in the art. The antisense molecules can be incorporated into a wide variety of vectors for delivery to a subject. The skilled practitioner can readily determine the best route of delivery, although generally i.v. or i.m. delivery is routine. The dosage amount is also readily ascertainable.

[0240] Particularly preferred antisense nucleic acid molecules are those that are complementary to a Replikin sequence contained in a mRNA encoding, for example, an influenza virus polypeptide, wherein the Replikin sequence comprises from 7 to about 50 amino acids including (1) at least one lysine residue located six to ten residues from a second lysine residue; (2) at least one histidine residue; and (3) at least 6% lysine residues. More preferred are antisense nucleic acid molecules that are complementary to a Replikin present in the coding strand of the gene or to the mRNA encoding the influenza virus hemagglutinin protein, wherein the antisense nucleic acid molecule is complementary to a nucleotide sequence encoding a Replikin that has been demonstrated to be conserved over a period of six months to one or more years and/or which are present in a strain of influenza virus shown to



have an increase in concentration of Replikins relative to Replikin concentration in other influenza virus strains. The increase in Replikin concentration preferably occurs over a period of at least six months, preferably about one year, most preferably about two or three years or more.

[0241] Similarly, antisense nucleic acid molecules that are complementary to mRNA those that are complementary to a mRNA encoding bacterial Replikins comprising a Replikin sequence of from 7 to about 50 amino acids including (1) at least one lysine residue located six to ten residues from a second lysine residue; (2) at least one histidine residue; and (3) at least 6% lysine residues. More preferred are antisense nucleic acid molecules that are complementary to the coding strand of the gene or to the mRNA encoding a protein of the bacteria.

DIAGNOSTIC APPLICATIONS

[0242] For organisms such as diatom plankton, foot and mouth disease virus, tomato leaf curl gemini virus, hepatitis B and C, HIV, influenza virus and malignant cells, identified constituent Replikins are useful as vaccines, and also may be usefully targeted for diagnostic purposes. For example, blood collected for transfusions may be screened for contamination of organisms, such as HIV, by screening for the presence of Replikins shown to be specific for the contamination organism. Also, screening for Replikin structures specific for a particular pathological organism leads to diagnostic detection of the organism in body tissue or in the environment.

REPLIKIN STIMULATION OF GROWTH

[0243] In another embodiment of the invention, Replikin structures are used to increase the replication rate of cells, tissues or organs. A method is available to increase replication rates by the addition of specific Replikin structures for other cells, tissues or organs that it is desired to replicate more rapidly, together with or without appropriate stimulae to cell division know in the art for said cells, tissues or organs to increase the rate of replication and yield. This may be accomplished, for example, by methods known in the art, by modifying or transforming a gene

encoding for or associated with a protein or enzyme having a replication function in the organism with at least one Replikin structure.

[0244] In another aspect of the invention, Replikin structures are used to increase the replication of organisms. The present invention demonstrates that in influenza virus, for example, increased replication associated with epidemics is associated with increased concentration of Replikins. The increase is due to 1) the reappearance of particular Replikin structures, which were present in previous years, but which then disappeared for one or more years; and/or 2) by the appearance of new Replikin compositions. In addition, in malaria Replikins, repetition of the same Replikin in a single protein occurs. Further, UTOPEs can be used to stimulation growth of an organism or to increase replication of organisms.

[0245] Thus, the present invention provides methods and compositions for increasing the replication of organisms. Similarly, in the manner that Replikins of different organisms can be targeted to inhibit replication of any organism, Replikins can be used to increase the replication of any organism. For example, production of rice, maize, and wheat crops, which are critical to feeding large populations in the world, can be improved, for example, by increasing the concentration (number of Replikins/100 amino acid residues) of any particular strain of rice.

[0246] As an example, in the Oryza sativa strain of rice, catalase isolated from immature seeds was observed to contain the following different Replikins within the 491 amino acid sequence of the protein:

kfpdvihafkpnprsh (SEQ ID NO. 638)

kfpdvihafk (SEQ ID NO. 639)

karyvkfhwk (SEQ ID NO.640)

hpkvspelraiwvnylsqedeslgvkianlnvk (SEQ ID NO. 641)

katihkqndfk (SEQ ID NO. 642)

happtpitprpvvgrrqkatihkqndfk (SEQ ID NO.643)

kfrpsssfdtkttttnagapvwndnealtvgprgpilledyhliekvah (SEQ ID NO. 644)

kfrpsssfdtkttttnagapvwndnealtvgprgpilledyn (SEQ ID NO. 645)

[0247] Thus, by using recombinant gene cloning techniques well known in the art, the concentration of Replikin structures in an organism, such as a food crop plant, can be increased, which will promote increased replication of the organism. For example, inserting additional Replikin sequences, like the Replikins identified above, into the Oryza sativa catalase gene by methods well know in the art will promote this organism's replication.

[0248] Similarly, in the NBS-LRR protein of Oryza sativa (japonica cultivar group), the following Replikins were found:

kvkahfqkh (SEQ ID NO. 647)

kvkahfqk (SEQ ID NO. 648)

kdyeidkddlih (SEQ ID NO. 648)

hmkqefafeavfpk (SEQ ID NO. 650)

hvfwelvwrsffqnvkqigsifqrkvyrygqsdvttskihdlmhdlavh (SEQ ID NO. 651)

kqigsifqrkvrygpsdvttskihdlmhdlavh (SEQ ID NO. 652)

kqigsifqrkvyrygpsdvttskihdlmh (SEQ ID NO. 653)

kqigsifqrkvyrygqsdvttskih (SEQ ID NO. 654)

[0249] Further, for aspartic proteinase oryzasin 1 precursor protein, the following Replikins were found:

khgvsagik (SEQ ID NO.655)

htvfdygkmrvgfak (SEQ ID NO. 657)

hsryksgqsstyqkngk (SEQ ID NO. 658)

[0250] Similarly, in the MADS-box protein FDRMADS3 transcription factor of Oryza sativa (indica cultivar-group), the following Replikins were found:

kqeamvlkqeinllqkglryiygnranch (SEQ ID NO. 659)

kqeinllqkglryiygnraneh (SEQ ID NO. 660)

kskegmlkaaneilqekiveqnglidvgmmvadqqngh (SEQ ID NO.661)

kaaneilqekiveqnglidvgmmvadqqngh (SEQ ID NO. 662)

[0251] Similarly, in LONI MAIZE (ATP-binding redox associated Hydrolase; Serine protease; Multigene family; Mitochondrion), the following Replikins were found:

kvlaahrygik (SEQ ID NO. 663)

klkiamkhliprvleqh (SEQ ID NO.664)

klkiamkh (SEQ ID NO. 665)

ktslassiakalnrkfirislggvkdeadirgh (SEQ ID NO.666)

kalnrkfirislggvkdeadirgh (SEQ ID NO. 667)

kfirislggvkdeadirgh (SEQ ID NO. 668)

kvrlskatelvdrhlqsilvaekitqkvegqlsksqk (SEQ ID NO. 669)

hlqsilvaekitqkvegglsksqk (SEQ ID NO. 670)

kvrlskatelvdrh (SEQ ID NO. 671)

kvggsavesskqdtkngkepihwhskgvaaralh (SEQ ID NO. 672)

kvggsavesskqdtkngkepihwh (SEQ ID NO. 673)

kvggsavesskqdtkngkepih (SEQ ID NO. 674)

kqdtkngkepihwhskgvaaralh (SEQ ID NO. 675)

kqdtkngkepih (SEQ ID NO. 676)

[0252] Similarly, for Glyceraldehyde 3-phospate dehydrogenase A, a chloroplast precursor, the following Replikins are found:

hrdlrraraaalnivptstgaakavslvlpnlk (SEQ ID NO. 677)

kvlddqkfgiikgtmttth (SEQ ID NO. 678)

hiqagakkvlitapgk (SEQ ID NO. 679)

hgrgdaspldviaindtggvkqashllk (SEQ ID NO. 680)

kqashllk (SEQ ID NO. 710)

[0253] Further, examples of rust resistance-like protein RP1-4 (Zea mays) found include the following Replikins:

kvrrvlskdysslkqlmtlmmdddiskhlqiiesgleeredkvwmkeniik (SEQ ID NO. 681)

kvrrvlskdysslkqlmtlmmdddiskh (SEQ ID NO. 682)

hlqiiesgleeredkvwmkeniik (SEQ ID NO.683)

hdlreniimkaddlask (SEQ ID NO. 684)

hvqnlenvigkdealask (SEQ ID NO. 685)

kkqgyelrqlkdlnelggslh (SEQ ID NO. 686)

kqgyelrqlkdlnelggslh (SEQ ID NO. 687)

klylksrlkelilewssengmdamnilh (SEQ ID NO. 688)

hlqllqlngmverlpnkvcnlskhylrgykdqipnigk (SEQ ID NO. 689)

hlqllqlngmverlpnkvcnlskrylrgyk (SEQ ID NO. 690)

hlqllqlngmverlpnkvcnlsk (SEQ ID NO. 691)

hnsnklpksvgelk (SEQ ID NO. 692)

kipkvgelkh (SEQ ID NO. 693)

hlsvrvesmqkhkeiiyk (SEQ ID NO. 694)

khkeiiyk (SEQ ID NO. 695)

klrdilqesqkfllvldlalfkh (SEQ ID NO. 696)

hafsgaeikdqllrmklqdtaeeiakrlgqeplaakvlgsrmerrk (SEQ ID NO. 697)

hafsgaeikdqllrmk (SEQ ID NO. 698)

klqdtaeeiakrlgqelaakvlgsrmerrkdiaewkaadvwfeksh (SEQ ID NO. 699)

kvlgsrmerrkdiaewkaadvwfeksh (SEQ ID NO. 700)

kdiaewkaadvwfeksh (SEQ ID NO. 701)

kaadvwfeksh (SEQ ID NO. 702)

hvptttslptskvfgrnsdrdrivkfllgkttaeasstk (SEQ ID NO. 703)

kailteakqlrdllglph (SEQ ID NO. 704)

kakaksgkgpllredessstattvmkpfh (SEQ ID NO. 705)

ksphrgkleswlrrlkeafydaedlldeh (SEQ ID NO. 706)

ksphrgkleswlrrlk (SEQ ID NO. 707)

hrgkleswlrrlk (SEQ ID NO. 708)

ksphrgk (SEQ ID NO. 709)

[0254] As discussed previously, the Replikin in wheat ubiquitin activating enzyme E (SEQ ID Nos. 614-616) is conserved. This conservation of Replikin structure provides reliable targets for stimulation of plant growth.

[0255] The close relationship of Replikins to redox enzymes is also clearly indicated in this structure in wheat. Thus, this wheat ubiquitin activating enzyme E activates ubiquitin by first adenylating with ATP its carboxy-terminal glycine

residue and, thereafter, linking this residue to the side chain of a cysteine residue in E1 (SEQ ID NO. 614), yielding an ubiquitin-E1 thiolester and free AMP.

[0256] A further example of the relationship of wheat Replikins to redox enzymes was also found in the PSABWheat Protein, Photosystem I P700 chlorophyll A apoprotein A2 (PsaB) (PSI-B) isolated from bread Chinese spring wheat Chloroplast Triticum aestivum. This protein functions as follows: PsaA and PsaB bind 9700, the primary electron donor of photosystem I (PSI), as well as the electron acceptors A0, A1, and FX. PSI functions as a plastocyanin/cytochrome c6-ferredoxin oxidoreductase. Cofactor P700 is a chlorophyll A dimer, A0 is chlorophyll A, A1 is a phylloquinone and FX is a 4Fe- 4S iron-sulfur center. The subunit A psaA/S heterodimer binds the P700 chlorophyll special pair and subsequent electron acceptors. The PSI reaction center of higher plants and algae is composed of one at least 11 subunits. This is an integral membrane protein of the Chloroplast thylakoid membrane. The 4Fe-4S iron-sulfur"center" to which 'h' bind is critical; hence the significance of 'h' in Replikin structure. Next to bacterial Replikins, these wheat Replikins and plant Replikins are the most primitive evolutionary illustrations of the importance of the Replikin structure to replication and the energy source needed for replication. This basic relationship carries through algae, virus Replikins, bacteria, cancer cells, and apparently all organisms with regard to replication.

[0257] Further examples of Replikins were found in the PSAB Wheat protein, which is critical for wheat growth. These include:

hlqpkwkpslswfknaesrlnhh (SEQ ID NO.617)
hlqpkwkpslswfk (SEQ ID NO. 618)
kwkpslswfknaesrlnhh (SEQ ID NO. 619)
kwkpslswfknaesrlnh (SEQ ID NO. 620)
kpslswfknaesrlnhh (SEQ ID NO. 621)
kpslswfknaesrlnh (SEQ ID NO. 622)
hhaialglhtttlilvkgaldargsklmpdkk (SEQ ID NO. 623)
haialglhtttlilvkgaldargsklmpdkk (SEQ ID NO. 624)
hhaialglhtttlilvkgaldargsk (SEQ ID NO. 625)

haialglhtttlilvkgaldargsk (SEQ ID NO..626)

htttlilvkgaldargsklmpdkk (SEQ ID NO.627)

htttlilvkgaldargsklmpdk (SEQ ID NO. 628)

htttlilvkgaldargsk (SEQ ID NO. 629)

[0258] A further example of the relationship of wheat Replikins to redox is provide in the PSAA_WHEAT Photosystem I 9700 chlorophyll A apoprotein A1, that include:

hhhlaiailfliaghmyrtnwgighglkdileahkgpftgqghk (SEQ ID NO. 630)

hhlaiailfliaghmyrtnwgighglkdileahkgpftgqghk (SEQ ID NO. 631)

hlaiailfliaghmyrtnwgighglkdileahkgpftgqghk (SEQ ID NO. 632)

hmyrtnwgighglkdileahkgpftgqghk (SEQ ID NO. 633)

hglkdileahkgpftgqghk (SEQ ID NO. 634)

hdileahkgpftgqghk (SEQ ID NO. 635)

hkgpftgqghk (SEQ ID NO. 636)

kgpftgqghk (SEQ ID NO. 637)

COMPUTER SOFTWARE FOR IDENTIFYING REPLIKINS

[0259] The present invention also provides methods for identifying Replikin sequences in an amino acid or nucleic acid sequence. Visual scanning of over four thousand sequences was performed in developing the present 3-point-recognition methods. However, data banks comprising nucleotide and/or amino acid sequences can also be scanned by computer for the presence of sequences meeting the 3 point recognition requirements.

[0260] According to another embodiment of the invention, three-point recognition methods described herein may be performed by a computer. Figure 6 is a block diagram of a computer available for use with the foregoing embodiments of the present invention. The computer may include a processor, an input/output device and a memory storing executable program instructions representing the 3-point-recognition methods of the foregoing embodiments. The memory may include a static memory, volatile memory and/or a nonvolatile memory. The static memory conventionally may be a read only memory ("ROM") provided on a

magnetic, or an electrical or optical storage medium. The volatile memory conventionally may be a random accessmemory ("RAM") and may be integrated as a cache within the processor or provided externally from the processor as a separate integrated circuit. The non-volatile memory may be an electrical, magnetic or optical storage medium.

[0261] From a proteomic point of view the construction of a "3-point recognition" template based on the new glioma peptide sequence led directly to identification of a biology-wide class of proteins having related structures and functions. The operation of the 3-point-recognition method resembles identification by the use of a "keyword"search; but instead of using the exact spelling of the keyword "kagvaflhkk" (SEQ ID NO.: 1) as in a typical sequence homology search, or in the nucleotide specification of an amino acid, an abstraction of the keyword delimited by the "3-point-recognition" parameters is used. This delimited abstraction, although derived from a single relatively short amino acid sequence leads to identification of a class of proteins with structures that are defined by the same specifications. That particular functions, in this case transformation and replication, in addition to structures, turn out also to be shared by members of the exposed class suggests that these structures and functions are related. Thus, from this newly identified short peptide sequence, a molecular recognition 'language' has been formulated, which previously has not been described. Further, the sharing of immunological specificity by diverse members of the class, as here demonstrated for the cancer Replikins, suggests that B cells and their product antibodies recognize Replikins by means of a similar recognition language.

OTHER USES OF THE THREE POINT RECOGNITION METHOD

[0262] Since "3-point-recognition" is a proteomic method that specifies a particular class of proteins, using three or more different recognition points for other peptides similarly should provide useful information concerning other protein classes. Further, the "3-point- recognition" method is applicable to other recognins, for example to the TOLL 'innate' recognition of lipopolyssacharides of organisms. The three point recognition method may also be modified to identify other useful compounds of covalently linked organic molecules, including other covalently

linked amino acids, nucleotides, carbohydrates, lipids or combinations thereof. In this embodiment of the invention a sequence is screened for subsequences containing three or more desired structural characteristics. In the case of screening compounds composed of covalently linked amino acids, lipids or carbohydrates the subsequence of 7 to about 50 covalently linked units should contain (1) at least one first amino acid, carbohydrate or lipid residue located seven to ten residues from a second of the first amino acid, carbohydrate or lipid residue; (2) encoding at least one second amino acid, lipid or carbohydrate residue; and (3) at least 6% of the first amino acid, carbohydrate or lipid residue. In the case of screening nucleotide sequences, the subsequence of about 21 to about 150 nucleotides should contain (1) at least one codon encoding a first amino acid located within eighteen to thirty nucleotides from a second codon encoding the first amino acid residue; (2) at least one second amino acid residue; and (3) encodes at least 6% of said first amino acid residue.

[0263] Several embodiments of the present invention are specifically illustrated and described herein. However, it will be appreciated that modifications and variations of the present invention are encompassed by the above teachings and within the purview of the appended claims without departing from the spirit and intended scope of the invention.

EXAMPLE 1

PROCESS FOR EXTRACTION, ISOLATION AND IDENTIFICATION OF REPLIKINS AND THE USE OF *REPLIKINS* TO TARGET, LABEL OR DESTROY *REPLIKIN*-CONTAINING ORGANISMS

a) Algae

[0264] The following algae were collected from Bermuda water sites and either extracted on the same day or frozen at -20 degrees C and extracted the next day. The algae were homogenized in a cold room (at 0 to 5 degrees C) in 1 gram aliquots in neutral buffer, for example 100 cc. of 0.005M phosphate buffer solution, pH 7 ("phosphate buffer") for 15 minutes in a Waring blender, centrifuged at 3000 rpm, and the supernatant concentrated by perevaporation and dialyzed against phosphate buffer in the cold to produce a volume of approximately 15 ml. The

volume of this extract solution was noted and an aliquot taken for protein analysis, and the remainder was fractionated to obtain the protein fraction having a pK range between 1 and 4.

[0265] The preferred method of fractionation is chromatography as follows: The extract solution is fractionated in the cold room (4° C) on a DEAE cellulose (Cellex-D) column 2.5x11.0 cm, which has been equilibrated with 0.005M phosphate buffer. Stepwise eluting solvent changes are made with the following solutions:

Solution 1- 4.04 g. NaH2P04 and 0.5g NaH2P04 are dissolved in 15 litres of distilled water (0.005 molar, pH 7);

Solution 2 - 8.57 g. NaH2P04 is dissolved in 2,480 ml. of distilled water;

Solution 3 - 17.1 g. of NaH2P04 is dissolved in 2480 ml of distilled water (0.05 molar, pH 4.7);

Solution 4 - 59.65 g. of NaH2P04 is dissolved in 2470 ml distilled water (0.175 molar);

Solution 5 - 101.6 g. of NaH2P04 is dissolved in 2455 ml distilled water (pH 4.3);

Solution 6 - 340.2 g. of NaH2P04 is dissolved in 2465 of distilled water (1.0 molar, pX-i 4.1);

Solution 7 - 283.63 g. of 80% phosphoric acid (H3P04) is made up in 2460 ml of distilled water (1.0 molar, pH 1.0).

[0266] The extract solution, in 6 to 10 ml volume, is passed onto the column and overlayed with Solution 1, and a reservoir of 300 ml of Solution 1 is attached and allowed to drip by gravity onto the column. Three ml aliquots of eluant are collected and analyzed for protein content at OD 280 until all of the protein to be removed with Solution 1 has been removed from the column. Solution 2 is then applied to the column, followed in succession by Solutions 3, 4, 5, 6 aid 7 until all of the protein which can, be removed with each Solution is removed from the column. The eluates from Solution 7 are combined, dialyzed against phosphate buffer, the protein content determined of both dialysand and dialyzate, and both analyzed by gel electrophoresis. One or two bands of peptide or protein of molecular weight

between 3,000 and 25,000 Daltons are obtained in Solution 7. For example the algae Caulerpa mexicana, Laurencia obtura, Cladophexa prolifera, Sargassum natans, Caulerpa verticillata, Halimeda tuna, and Penicillos capitatus, after extraction and treatment as above, all demonstrated in Solution 7 eluates sharp peptide bands in this molecular weight region with no contaminants. These Solution 7 proteins or their eluted bands are hydrolyzed, and the amino acid composition determined. The peptides so obtained, which have a lysine composition of 6% or greater are Replikin precursors. These Replikin peptide precursors are then determined for amino acid sequence and the Replikins are determined by hydrolysis and mass spectrometry as detailed in U.S. Patent 6,242,578 B1. Those that fulfill the criteria defined by the " 3-point-recognition" method are identified as Replikins. This procedure can also be applied to obtain yeast, bacterial and any plant Replikins.

b) Virus

[0267] Using the same extraction and column chromatography separation methods as above in a) for algae, Replikins in virus-infected cells are isolated and identified.

c) Tumor cells in vivo and in vitro tissue culture

[0268] Using the same extraction and column chromatography separation methods as above in a) for algae, Replikins in tumor cells are isolated and identified. For example, Replikin precursors of Astrocytin isolated from malignant brain tumors, Malignin (Aglyco 1OB) isolated from glioblastoma tumor cells in tissue culture, MCF7 mammary carcinoma cells in tissue culture, and P3J Lymphoma cells in tissue culture each treated as above in a) yielded Replikin precursors with lysine content of 9.1%, 6.7%, 6.7%, and 6.5% respectively. Hydrolysis and mass spectrometry of Aglyco 1OB as described in Example 10 U.S. 6,242,578 B1 produced the amino acid sequence, ykagvaflhkkndiide the 16-mer Replikin.

EXAMPLE 2:

[0269] As an example of diagnostic use of Replikins: Aglyco 1OB or the 16-mer Replikin may be used as antigen to capture and quantify the amount of its

corresponding antibody present in serum for diagnostic purposes are as shown in Figures 2, 3, 4 and 7 of U.S. 6,242,578 B1.

[0270] As an example of the production of agents to attach to Replikins for labeling, nutritional or destructive purposes: Injection of the 16-mer Replikin into rabbits to produce the specific antibody to the 16-mer Replikin is shown in Example 6 and Figures 9A and 9B of U.S. 6,242,578 B1.

[0271] As an example of the use of agents to label Replikins: The use of antibodies to the 16-mer Replikin to label specific cells which contain this Replikin is shown in Figure 5 and Example 6 of U.S. 6,242,578 B1.

[0272] As an example of the use of agents to destroy Replikins: The use of antibodies to the 16-mer Replikin to inhibit or destroy specific cells which contain this Replikin is shown in Figure 6 of U.S. 6,242,578 B1.

EXAMPLE 3

[0273] Analysis of sequence data of isolates of influenza virus hemagglutinin protein or neuraminidase protein for the presence and concentration of Replikins is carried out by visual scanning of sequences or through use of a computer program based on the 3-point recognition system described herein. Isolates of influenza virus are obtained and the amino acid sequence of the influenza hemagglutinin and/or neuraminidase protein is obtained by any art known method, such as by sequencing the hemagglutinin or neuraminidase gene and deriving the protein sequence therefrom. Sequences are scanned for the presence of new Replikins, conservation of Replikins over time and concentration of Replikins in each isolate. Comparison of the Replikin sequences and concentrations to the amino acid sequences obtained from isolates at an earlier time, such as about six months to about three years earlier, provides data that are used to predict the emergence of strains that are most likely to be the cause of influenza in upcoming flu seasons, and that form the basis for seasonal influenza peptide vaccines or nucleic acid based vaccines. Observation of an increase in concentration, particularly a stepwise increase in concentration of Replikins in a given strain of influenza virus for a period of about six months to

about three years or more is a predictor of emergence of the strain as a likely cause of influenza epidemic or pandemic in the future.

[0274] Peptide vaccines or nucleic acid-based vaccines based on the Replikins observed in the emerging strain are generated. An emerging strain is identified as the strain of influenza virus having the highest increase in concentration of Replikin sequences within the hemagglutinin and/or neuraminidase sequence during the time period. Preferably, the peptide or nucleic acid vaccine is based on or includes any Replikin sequences that are observed to be conserved in the emerging strain. Conserved Replikins are preferably those Replikin sequences that are present in the hemagglutinin or neuraminidase protein sequence for about two years and preferably longer. The vaccines may include any combination of Replikin sequences identified in the emerging strain.

[0275] For vaccine production, the Replikin peptide or peptides identified as useful for an effective vaccine are synthesized by any method, including chemical synthesis and molecular biology techniques, including cloning, expression in a host cell and purification therefrom. The peptides are preferably admixed with a pharmaceutically acceptable carrier in an amount determined to induce a therapeutic antibody reaction thereto. Generally, the dosage is about 0.1 μg to about 10 mg.

[0276] The influenza vaccine is preferably administered to a patient in need thereof prior to the onset of "flu season." Influenza flu season generally occurs in late October and lasts through late April. However, the vaccine may be administered at any time during the year. Preferably, the influenza vaccine is administered once yearly, and is based on Replikin sequences observed to be present, and preferably conserved in the emerging strain of influenza virus. Another preferred Replikin for inclusion in an influenza vaccine is a Replikin demonstrated to have re-emerged in a strain of influenza after an absence of one or more years.

EXAMPLE 4

[0277] Analysis of sequence data of isolates of coronavirus nucleocapsid, or spike, or envelope, or other protein for the presence and concentration of Replikins

is carried out by visual scanning of sequences or through use of a computer program based on the 3-point recognition method described herein. Isolates of coronavirus are obtained and the amino acid sequence of the coronavirus protein is obtained by any method known in the art, such as by sequencing the protein's gene and deriving the protein sequence therefrom. Sequences are scanned for the presence of new Replikins, conservation of Replikins over time and concentration of Replikins in each isolate. Comparison of the Replikin sequences and concentrations to the amino acid sequences obtained from isolates at an earlier time, such as about six months to about three years earlier, provides data that are used to predict the emergence of strains that are most likely to be the cause an outbreak or pandemic, and that form the basis for coronavirus peptide vaccines or nucleic acid based vaccines. Observation of an increase in concentration, particularly a stepwise increase in concentration of Replikins in a given class, or strain, of coronavirus for a period of about six months to about three years or more is a predictor of emergence of the strain as a likely cause of an epidemic or pandemic, such as SARS, in the future.

[0278] Peptide vaccines or nucleic acid-based vaccines based on the Replikins observed in the emerging strain of coronaviruses are generated. An emerging strain is identified as the strain of coronovirus having the highest increase in concentration of Replikin sequences within the nucleocapsid sequence during the time period. Preferably, the peptide or nucleic acid vaccine is based on or includes any Replikin sequences that are observed to be conserved in the strain. Conserved Replikins are preferably those Replikin sequences which are present in the nucleocapsid protein sequence for about two years and preferably longer. The vaccines may include any combination of Replikin sequences identified in the emerging strain.

[0279] For vaccine production, the Replikin peptide or peptides identified as useful for an effective vaccine are synthesized by any method, including chemical synthesis and molecular biology techniques, including cloning, expression in a host cell and purification therefrom. The peptides are preferably admixed with a pharmaceutically acceptable carrier in an amount determined to induce a therapeutic antibody reaction thereto. Generally, the dosage is about 0.1 μg to about 10 mg.

[0280] The coronavirus vaccine may be administered to a patient at any time of the year. Preferably, the coronavirus vaccine is administered once and is based on Replikin sequences observed to be present, and preferably conserved, in the classes of coronavirus.

EXAMPLE 5

[0281] Analysis of sequence data of isolates of Plasmodium falciparum antigens for the presence and concentration of Replikins is carried out by visual scanning of sequences or through use of a computer program based on the 3-point recognition method described herein. Isolates of Plasmodium falciparum are obtained and the amino acid sequence of the protein is obtained by any art known method, such as by sequencing the gene and deriving the protein sequence therefrom. Sequences are scanned for the presence of Replikins, conservation of Replikins over time and concentration of Replikins in each isolate. This information provides data that are used to form the basis for anti-malarial peptide vaccines or nucleic acid based vaccines.

[0282] Peptide vaccines or nucleic acid-based vaccines based on the Replikins observed in the malaria causing organism are generated. Preferably, the peptide or nucleic acid vaccine is based on or includes any Replikin sequences that are observed to be present on a surface antigen of the organism. The vaccines may include any combination of Replikin sequences identified in the malaria causing strain.

[0283] For vaccine production, the Replikin peptide or peptides identified as useful for an effective vaccine are synthesized by any method, including chemical synthesis and molecular biology techniques, including cloning, expression in a host cell and purification therefrom. The peptides are preferably admixed with a pharmaceutically acceptable carrier in an amount determined to induce a therapeutic antibody reaction thereto. Generally, the dosage is about 0.1 μg to about 10 mg.

[0284] Then malaria vaccine is preferably administered to a patient in need thereof at any time during the year, and particularly prior to travel to a tropical environment.

[0285] Another embodiment includes an antisense nucleic acid molecule complementary to the coding strand of the gene or the mRNA encoding organism for the replikins in organisms including, but not limited to, viruses, trypanosomes, bacteria, fungi, algae, amoeba, and plants, wherein said antisense nucleic acid molecules is complementary to a nucleotide sequence of a replikin containing organism.

What is claimed is:

Claim 1. An isolated or synthesized coronavirus peptide consisting of 7 to 50 amino acids wherein said peptide is isolated or synthesized by:
identifying a motif consisting of
(1) at least one lysine residue located at a first terminus of said motif and at least one lysine residue or at least one histidine residue located at a second terminus of said motif;
(2) a first lysine residue located six to ten residues from a second lysine residue;
(3) at least one histidine residue; and
(4) at least 6% lysine residues; and
selecting said identified motif; and
isolating or synthesizing said coronavirus peptide comprising said identified motif.

Claim 2. The isolated or synthesized coronavirus peptide of claim 1 wherein said coronavirus peptide amino acid sequence consists of 7 to 25 amino acids residues.

Claim 3. The isolated or synthesized coronavirus peptide of claim 1 wherein said coronavirus peptide amino acid sequence consists of 7 to 16 amino acid residues.

Claim 4. A coronavirus vaccine comprising at least one isolated peptide of claim 1 present in the nucleocapsid, or spike, or envelope, or other protein of an emerging strain of coronavirus and a pharmaceutically acceptable carrier.

Claim 5. The coronavirus vaccine of claim 4 wherein the vaccine comprises a plurality of isolated peptides of claim 1.

Claim 6. A method of preventing or treating SARS comprising administering to a patient a vaccine comprising at least one isolated peptide of claim 1 present in the nucleocapsid, or spike, or envelope, or other protein of an emerging strain of coronavirus and a pharmaceutically acceptable carrier.

Claim 7. A synthesized peptide wherein the synthesized peptide is about 7 to 10 amino acids long comprising all lysines except for one histidine and at least one lysine

is located six to ten residues from a second lysine.

Claim 8. A method of providing passive immunity comprising administering a composition comprising the synthesized peptide of claim 7 and a pharmaceutically acceptable carrier.

Claim 9. The method of claim 8 wherein said composition further comprises an adjuvant.

Claim 10. A method of inducing an immune response against SARS comprising administering a synthesized peptide about 6 to about 10 amino acids long comprising all lysines except for one histidine.

Claim 11. A method of inducing an immune response against SARS comprising administering a synthetic peptide of claim 7.

Claim 12. A method of inducing an immune response against SARS comprising administering a vaccine of claim 4.

Claim 13. A method of treating SARS comprising administering a UTOPE, wherein said UTOPE is a synthesized peptide about 6 to about 10 amino acids long comprising all lysines except for one histidine.

Claim 14. A synthetic peptide consisting of the amino acid sequence of a UTOPE, wherein said UTOPE is a synthesized peptide about 6 to about 10 amino acids long comprising all lysines except for one histidine.

Claim 15. A method of identifying an emerging strain of a virus for diagnostic, preventive or therapeutic purposes comprising:

(1) obtaining at least one isolate of each strain of a plurality of strains of the virus;

(2) analyzing the amino acid sequence of the individual proteins of at least one isolate of each strain of the plurality of strains of the virus for the presence and concentration of Replikin sequences wherein said Replikin sequences comprise:

a) at least one lysine residue located at a first terminus of said Replikin sequence and at least one lysine residue or at least one histidine residue located at a second terminus of said Replikin sequence;

b) a first lysine residue located six to ten residues from a second lysine residue;

c) at least one histidine residue; and

d) at least 6% lysine residues;

(3) comparing the concentration of Replikin sequences in the sum of the amino acid sequences of all of the individual proteins of the at least one isolate of each strain of the plurality of strains of the virus to the concentration of Replikin sequences observed in the amino acid sequences of the sum of all of the individual proteins of the same strain of the virus isolated at at least one earlier time period to provide the concentration of Replikins for at least two time periods in each strain of the plurality of strains, said at least one earlier time period being within about six months to about three years prior to step (1); and

(4) identifying the strain of the virus having the highest increase in concentration of Replikin sequences in the sum of all of the individual proteins of the strain of the virus between the at least two time periods as compared to the concentration of Replikin sequences in the sum of all of the individual proteins of the other strains of the plurality of strains as an emerging strain of virus.

Claim 16. A method according to claim 15 wherein said emerging strain of virus is selected from the group consisting of influenza virus, SARS, and coronavirus.

Claim 17. A method of identifying an emerging strain of a virus for diagnostic, preventive or therapeutic purposes comprising:

(1) obtaining at least one isolate of each strain of a plurality of strains of the virus;

(2) analyzing the amino acid sequence of at least one individual protein of the at least one isolate of each strain of the plurality of strains of the virus for the presence and concentration of Replikin sequences, wherein said Replikin sequences comprise:

a) at least one lysine residue located at a first terminus of said Replikin sequence and at least one lysine residue or at least one histidine residue located at a second terminus of said Replikin sequence;

b) a first lysine residue located six to ten residues from a second lysine residue;

c) at least one histidine residue; and

d) at least 6% lysine residues;

and wherein said at least one individual protein is the same protein in each isolate of each strain of the virus;

(3) comparing the concentration of Replikin sequences in the amino acid sequence of the at least one individual protein of the at least one isolate of each strain of the plurality of strains of the virus to the concentration of Replikin sequences observed in the amino acid sequence of the same individual protein of each of the strains isolated at at least one earlier time period to provide the concentration of Replikins for at least one protein in each strain of the plurality of strains of the virus at at least two time periods, said at least one earlier time period being within about six months to about three years prior to step (1); and

(4) identifying the strain of the virus having the highest increase in concentration of Replikin sequences in the at least one individual protein between the at least two time periods as compared to the increase in concentration of Replikin sequences in the at least one individual protein between the at least two periods in the other strains of the plurality of strains of the virus.

Claim 18. A method according to claim 17 wherein said emerging strain of virus is selected from the group consisting of influenza virus, SARS, and coronavirus.

Claim 19. A method of identifying an emerging strain of an influenza virus for diagnostic, preventive or therapeutic purposes comprising:

(1) obtaining at least one isolate of each strain of a plurality of strains of the influenza virus;

(2) analyzing the amino acid sequence of the hemagglutinin protein of the at least one isolate of each strain of the plurality of strains of the virus for the presence and concentration of Replikin sequences, wherein said Replikin sequences comprise:

a) at least one lysine residue located at a first terminus of said Replikin sequence and at least one lysine residue or at least one histidine residue located at a second terminus of said Replikin sequence;

b) a first lysine residue located six to ten residues from a second lysine residue;

c) at least one histidine residue; and

d) at least 6% lysine residues;

(3) comparing the concentration of Replikin sequences of the hemagglutinin of each strain of the plurality of strains of the virus with the concentration of Replikin sequences of the hemagglutinin amino acid of each strain at at least one earlier time period to provide the concentration of Replikins for at least two time periods, said at least one earlier time period being within about six months to about three years prior to step (1); and

(4) identifying the strain of the influenza virus having the highest increase in concentration of Replikin sequences of the hemagglutinin protein during the at least two time periods as an emerging strain of influenza virus.

Claim 20. A method of identifying an emerging strain of virus according to claims 15, 16, 17, 18 and 19 further comprising identifying the strain of virus having the greatest increase in concentration of new Replikin sequences between said at least two time periods as an emerging strain of virus, wherein said new Replikin sequences are those exact sequence structures that are present in said strain of virus isolated at the later of said two time periods and not present in said strain of virus isolated at the earlier of said two time periods wherein said Replikin sequences comprise at least one lysine residue located at a first terminus of said Replikin sequence and at least one lysine residue or at least one histidine residue located at a second terminus of said Replikin sequence, a first lysine residue located six to ten residues from a second lysine residue, at least one histidine residue; and at least 6% lysine residues.

Claim 21. A method of identifying a non-emerging strain of virus, or a strain of virus that is dormant, for diagnostic, preventive or therapeutic purposes comprising:

(1) obtaining at least one isolate of each strain of a plurality of strains of the virus;

(2) analyzing the amino acid sequence of the at least one isolate of each strain of the plurality of strains of the virus for the presence and concentration of Replikin sequences, wherein said Replikin sequences comprise:

a) at least one lysine residue located at a first terminus of said Replikin sequence and at least one lysine residue or at least one histidine residue located at a second terminus of said Replikin sequence;

b) a first lysine residue located six to ten residues from a second lysine residue;

c) at least one histidine residue; and

d) at least 6% lysine residues;

(3) comparing the concentration of Replikin sequences in the amino acid sequence of the at least one isolate of each strain of the plurality of strains of the virus to the concentration of Replikin sequences observed in the amino acid sequence of each of the strains at at least one earlier time period to provide the concentration of Replikins for at least two time periods, said at least one earlier time period being within about six months to about three years prior to step (1); and

(4) identifying the strain of the virus having the lowest increase in concentration of Replikin sequences during the at least two time periods.