FLAVIVIRUS NS5A PROTEINS FOR THE TREATMENT OF HIV

Inventors: Jack T. Stapleton, Iowa City, IA (US); Jinhua Xiang, Iowa City, IA (US); Qing Chang, Iowa City, IA (US); James McLinden, Coralville, IA (US)

Assignee: University of Iowa Research Foundation, Iowa City, IA (US)

Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 562 days.

Appl. No.: 11/345,662
Filed: Feb. 1, 2006

Prior Publication Data

Related U.S. Application Data
Provisional application No. 60/653,823, filed on Feb. 17, 2005.

Int. Cl.
C12Q 1/70 (2006.01)
A61K 39/21 (2006.01)
A61K 39/29 (2006.01)

U.S. Cl. .................. 435/5; 424/208.1; 424/228.1

Field of Classification Search None

References Cited
U.S. PATENT DOCUMENTS
5,441,936 A 8/1995 Hoogse et al. .............. 414/16
5,824,507 A 10/1998 Kim et al. .............. 435/69.3
5,849,532 A 12/1998 Kim et al. .............. 435/69.3
5,856,134 A 1/1999 Kim et al. .............. 435/69.3
5,874,563 A 2/1999 Kim et al. .............. 536/23.72
6,020,195 A 2/2000 Schmolke et al. ........ 435/399
6,156,495 A 12/2000 Pilot-Matias et al. ... 435/5
2005/013828 A1 1/2005 George et al. ....... 424/189.1

FOREIGN PATENT DOCUMENTS

OTHER PUBLICATIONS

* cited by examiner

Primary Examiner — Stacy B Chen
(74) Attorney, Agent, or Firm — Fulbright & Jaworski

ABSTRACT
GB virus C (GBV-C or hepatitis G virus) is a flavivirus that frequently leads to chronic viremia in humans. The invention provides compositions and methods involving a GBV-C NS5A peptide or polypeptide for inhibiting and treating HIV infections.

6 Claims, 21 Drawing Sheets
X4 HIV-1 or mock control

96 hrs

No Doxy (NS5A positive)

d0

d2

d4

d6

X4 HIV-1 or mock control

96 hrs

1 μg/ml (NS5A negative)

Jurkat Cells passed (x3) in Doxycycline

FIG. 4A
NS5A Inhibits HIV MN (X4) Clade B

HIV p24 Antigen Release

FIG. 11A
NS5A Inhibits HIV JF(X4) Clade B
NS5A Inhibits HIV ELI(X4) Clade D
NS5A (414 aa) - IFN-R
NS5A (414 aa) - IFN-R + GFP
NS5A (414 aa) - IFN-S + GFP
NS5A (363 aa) - IFN-R + GFP
NS5A (250 aa) - IFN-R + GFP
NS5A (167 aa) - IFN-R + GFP
5A (123 aa) - R
NS5A (314 aa) - IFN-S + GFP
NS5A (123-363 aa) - IFN-R + GFP
167-250 - R

- Inhibits HIV
- Does not inhibit HIV

FIG. 13
*Top construct does not include GFP
1

FLAVIVIRUS NS5A PROTEINS FOR THE TREATMENT OF HIV

BACKGROUND OF THE INVENTION

The present application claims benefit of priority to U.S. Provisional Application Ser. No. 60/653,823, filed Feb. 17, 2005, the entire contents of which are hereby incorporated by reference.

The U.S. Government owns rights in this invention pursuant to grant number AI58740 from NIH and merit grants awarded to Jack Stapleton and Jinhua Xiang from the Veterans Administration.

I. Field of the Invention

The present invention relates generally to the fields of molecular biology and virology. More particularly, it concerns methods and compositions to treat, inhibit or prevent HIV infection.

II. Description of Related Art

A. GB Virus Type C

GB virus type C (GBV-C), also known as hepatitis G virus (HGV), is a virus whose genomic organization and nucleotide sequence places it in the Flaviviridae family (Robertson et al., 1998). It is the most closely related human virus to hepatitis C virus (HCV) (Leary et al., 1996; Linnen et al., 1996; Simons et al., 1995). It has been suggested that these viruses should be classified together with non-human GB-hepatitis agents as the hepacivirus genus. Although GBV-C was originally associated with post-transfusion hepatitis in humans (Linnen et al., 1996), subsequent epidemiological studies indicated that it does not cause acute or chronic hepatitis (Alter et al., 1997a; Alter et al., 1997b). In addition, experimental GBV-C infection of chimpanzees was not associated with acute hepatitis (Bokh et al., 1998).

Persistent GBV-C viremia (as detected by RT-PCR) is common, with 0.9% to 3% of healthy U.S. blood donors and approximately 20%-30% of patients with HCV infection persistently infected with GBV-C (Dawson et al., 1996; Feucht et al., 1997; Simons et al., 1995a; Simons et al., 1995b; Tacke et al., 1997). Following infection, about 80% of people clear their viremia, concomitantly developing antibodies to the GBV-C E2 protein (Feucht et al., 1997; Thomas et al., 1998). Thus, it is estimated that approximately 20% of infected people remain viremic for long periods of time. GBV-C appears to be transmitted primarily by parenteral exposure (Simons et al., 1995), although there are data suggesting that sexual and/or household transmission of GBV-C infection may occur (Akiyoshi et al., 1999; de Martino et al., 1998; Nemurak et al., 1998; Tanaka et al., 1997; Wu et al., 1997).

B. GBV-C and HIV

During progressive human immunodeficiency virus type 1 (HIV-1) infection, the virus-specific immune responses of an infected subject gradually deteriorate, leading to the development of acquired immunodeficiency syndrome (AIDS). Most infected patients do not exhibit overt clinical manifestations of the disease for six to ten years following initial infection, however, most individuals infected with HIV eventually die from conditions or infections; that the individuals’ immune system is no longer equipped to fight. While treatment for AIDS has been forthcoming, no effective cure has been reported. Thus, preventative and treatment options against HIV infection and the development of AIDS remain highly desirable.

GBV-C has been investigated in the context of HIV infection. The course of HIV-1 infection is extremely variable among infected individuals, although the reasons for this observation are not fully understood. Individuals whose HIV disease progresses slowly are often called long-term non-progressors (LTNP). The prevalence of LTNP varies from 1% to 25% of infected people, depending upon the definition used (reviewed in Easterbrook, 1999). There are no specific clinical criteria for LTNP. However, non-progression generally implies the absence of HIV-related clinical disease 10 or more years following infection and an absolute CD4 count of 500 cells/mm³ (Easterbrook, 1999). Evaluation of LTNP’s has identified HIV isolates with deletions in key replicative genes (Deacon et al., 1995) and host genetic factors, including specific HLA haplotypes (reviewed in reference Rowland-Jones, 1999). In some individuals, polymorphisms that result in absent or reduced expression of HIV co-receptors have been identified (Huang et al., 1996). However, these findings are uncommon and thought to account for no more than one-third of LTNP’s (Rowland-Jones, 1999).

Persistent GBV-C infection is common in humans, with infection rates of approximately 0.9% to 3% in healthy blood donors, 20-30% in HCV-positive people (Dawson et al., 1996), and 35%-40% in HIV-positive individuals. GBV-C infection can persist for decades in the absence of any clinical morbidity or mortality. Among immune-competent individuals, it is estimated that 60% to 75% of GBV-C-infected people clear the infection, concomitantly developing antibodies to the envelope glycoprotein E2 (Thomas et al., 1998). It is also known that GBV-C can be propagated in cultures of peripheral blood mononuclear cells (PBMC’s) (Fogeda et al., 1999).

In 1998, Toyoda et al. found that hemophiliacs co-infected with HIV and GBV-C had a lower plasma HIV RNA concentration and a lower incidence of AIDS diagnoses compared to those infected with HIV alone (Toyoda et al., 1998), although the differences were not statistically significant. In contrast, Sabin and colleagues found an increased rate of AIDS and death in hemophiliacs “exposed” to GBV-C (Sabin et al., 1998) compared to non-exposed individuals. This study included HIV-positive subjects who were either GBV-C viremic as determined by detection of GBV-C RNA in plasma, or HIV-infected people who were not viremic but were anti-GBV-C E2 antibody-positive. Although the mortality rate was higher among the GBV-C “exposed” individuals, the results were not statistically significant. Looking at HIV-infected persons, Lefrere and colleagues reported a significant delay in the rate of CD4+ T cell decline, development of AIDS, and death in 23 HIV-positive individuals with GBV-C viremia compared to 72 HIV-infected people without GBV-C viremia (Lefrere et al., 1999). In this study, HIV-infected individuals who were also GBV-C-positive were compared to HIV-infected individuals who were GBV-C-negative. When these subjects were matched by age, sex, baseline HIV RNA load, and baseline CD4 T cell count, HIV disease progression appeared to be worse in GBV-C-negative subjects.

The interrelationship between HIV and GBV-C continues to be explored, with possible therapeutic aspects of GBV-C infection being examined.

SUMMARY OF THE INVENTION

A pharmaceutical composition comprising an isolated flavivirus NS5A peptide or polypeptide, or multiple flavivirus NS5A peptides from the same or different NS5A polypeptide. The NS5A polypeptide may be a full length NS5A polypeptide, a fusion polypeptide. The fusion may comprise a targeting signal, such as a nuclear targeting signal or a cell surface receptor (e.g., a CD4 receptor). The NS5A peptide or polypeptide may be formulated in a lipid vehicle, such as a liposome. The NS5A peptide or polypeptide may be formulated with an amphiphilic peptide, an insect peptide, or pyr-
rhoricin. The flavivirus may be selected from the group consisting of DENV4, YFV, TBEV, WN4, CSFV, BDV, GBV-A, GBV-B, GBV-C, HGV, HCV 2a, HCV 3a, HCV 2b, HCV 1a and HCV 1b. The flavivirus NS5A peptide or polypeptide may reside 152-237 of GBV-C NS5A, or the corresponding sequences thereto from other flavivirus NS5A proteins, or may comprise domain II of HCV NS5A, or the corresponding sequences thereto from other flavivirus NS5A proteins. The NS5A peptide or polypeptide may be from an IFN-sensitive flavivirus, or a IFN-resistant flavivirus. The IFN-sensitive flavivirus or IFN-resistant flavivirus may be a GBV-C virus.

In yet another embodiment, there is provided a method for preventing or treating HIV infection comprising administering to a subject a composition comprising a flavivirus NS5A peptide or polypeptide. The flavivirus NS5A peptide or polypeptide comprises multiple flavivirus NS5A peptides, from the same or different NS5A polypeptides. The NS5A polypeptide may be a full length NS5A polypeptide, and may be a fusion polypeptide, for example, comprising a targeting signal, such as a nuclear targeting signal or a targeting signal that targets a cell surface receptor (e.g., the CD4 receptor). The NS5A peptide or polypeptide may be formulated in a lipid vehicle, such as a liposome. The NS5A peptide or polypeptide may be formulated with an amphipathic peptide, an insect peptide, or pyrrothicin. The flavivirus may be selected from the group consisting of DENV4, YFV, TBEV, WN4, CSFV, BDV, GBV-A, GBV-B, GBV-C, HGV, HCV 2a, HCV 3a, HCV 2b, HCV 1a and HCV 1b. The flavivirus NS5A peptide or polypeptide may reside 152-237 of GBV-C NS5A, or the corresponding sequences thereto from other flavivirus NS5A proteins, or may comprise domain II of HCV NS5A, or the corresponding sequences thereto from other flavivirus NS5A proteins. The NS5A peptide or polypeptide may be from an IFN-sensitive flavivirus, or a IFN-resistant flavivirus. The IFN-sensitive flavivirus or IFN-resistant flavivirus may be a GBV-C virus. The method may further comprise administration of at least a second anti-HIV therapy, before or after said flavivirus NS5A peptide or polypeptide. The second anti-HIV therapy may be HAART therapy, or AZT therapy. The method may comprise multiple administrations of the composition.

In yet another embodiment, there is provided a method of reducing HIV replication in an HIV-infected cell comprising contacting said cell with a composition comprising a flavivirus NS5A peptide or polypeptide. The cell may be a T lymphocyte. In yet another embodiment, there is provided a method of inhibiting HIV infection of a cell comprising contacting said cell with a composition comprising a flavivirus NS5A peptide or polypeptide. The cell may be a T lymphocyte.

In yet another embodiment, there is provided a method for preventing or treating HIV infection comprising administering to a subject a composition comprising an expression construct encoding a flavivirus NS5A peptide or polypeptide. The expression construct may be a viral expression construct, such as an adenovirus, a retrovirus, a lentivirus, an adeno-associated virus, a polyoma virus, a herpesvirus, or a pox virus. The expression construct may be a non-viral expression construct, for example, dispersed in a lipid vehicle. The expression construct may encode an NS5A peptide or an NS5A peptide, a full length NS5A polypeptide, a fusion polypeptide, for example, comprising a nuclear targeting signal or a signal that targets a cell surface receptor, e.g., the CD4 receptor. The flavivirus NS5A peptide or polypeptide may comprise residues 152-237 of GBV-C NS5A2, or the corresponding sequences thereto from other flavivirus NS5A proteins, or domain II of HCV NS5A, or the corresponding sequences thereto from other flavivirus NS5A proteins. The flavivirus may be selected from the group consisting of DENV4, YFV, TBEV, WN4, CSFV, BDV, GBV-A, GBV-B, GBV-C, HGV, HCV 2a, HCV 3a, HCV 2b, HCV 1a and HCV 1b.

The use of the word "a" or "an" when used in conjunction with the term "comprising" in the claims and/or the specification may mean "one," but it is also consistent with the meaning of "one or more," "at least one," and "one or more than one." Furthermore, where multiple steps of a method of process are cited, it is understood that the steps are not required to be performed in the particular order recited unless one of skill in the art is not able to practice the method in a different order.

Other objects, features, and advantages of the present invention will become apparent from the following detailed description. It should be understood, however, that the detailed description and the specific examples, while indicating preferred embodiments of the invention, are given by way of illustration only, since various changes and modifications within the spirit and scope of the invention will become apparent to those skilled in the art from this detailed description.

**BRIEF DESCRIPTION OF THE DRAWINGS**

The following drawings form part of the present specification and are included to further demonstrate certain aspects of the present invention. The invention may be better understood by reference to one or more of these drawings in combination with the detailed description of specific embodiments presented herein.

**FIG. 1**—NS5A expression in Jurkat cells. Jurkat cells containing GBV-C NS5A (A11a clone) expressed two immunoreactive proteins when incubated without doxycycline (Doxyc). Rabbit anti-NS5A antiserum was raised against NS5A expressed in E. Coli. This antigen is shown as the positive control on lower panel (+), whereas the positive control [C(+)] top panel represents NS5A expressed in CHO cells; JICR In Press). Growth of cells in low levels of doxycycline for 48 hrs reduced NS5A protein levels to non-detectable levels. Doxycycline concentration in mg/mi.

**FIGS. 2A-C**—HIV replication in Jurkat cells with and without NS5A expression. The amount of HIV p24 antigen released into culture supernatants from two cloned Jurkat cell lines expressing GBV-C NS5A (A11a and A11b) or vector control cell line (VC) 2 to 6 days following infection using an X-4 virus isolate (MOI of 0.1 (FIG. 2A) or 0.5 (FIG. 2B)). Cells were grown without doxycyclone, and data represent the mean of 3 replicate samples. NS5A expression is shown (immunoblot of cell lysates in triplicate) of A11a, A11b VC—vector control cell lysate. Note that A11a has greater levels of NS5A expressed and HIV replication in these cells is lower than in A11b (on day 6). Compared to vector control cells, HIV inhibition was significantly lower on day 6 (p=0.02).

**FIG. 3A**—Experimental Design. Jurkat cells with either vector control (VC) or NS5A were grown without doxycyclone. Cells were divided, and half were maintained without doxycyclone (NS5A was expressed) or with various doses of doxycycline (NS5A expression was differentially suppressed). Cells (VC and NS5A containing) were infected with X4 HIV-1 (clinical isolate) 48 hrs after dividing cells, and cell culture supernatants monitored 2, 4, and 6 days later for HIV p24 antigen.

**FIG. 3B**—Dose-dependent inhibition of HIV by NS5A. Experimental design described above. NS5A expression was
suppressed by cell growth in various concentrations of doxycycline (0.01, 0.1 and 1 mg/ml as indicated) for 2 days prior to infection with HIV-1. HIV replication was monitored by measuring p24 antigen in culture supernatants. The amount of NS5A expression was directly correlated with inhibition of HIV expression in Jurkat-NS5A cells (left panel), and doxycycline had no effect in Jurkat-vector control cells (VC; right panel).

FIG. 4A—Experimental Design. Jurkat cells with NS5A were passaged with doxycycline to suppress NS5A expression. Cells were divided, and half were maintained without doxycycline (NS5A was expressed) or with 1 mg/ml doxycycline (NS5A was not expressed). NS5A containing cells were infected with X4 HIV-1 (clinical isolate) 96 hrs after cells were divided, and cell culture supernatants monitored 2, 4, and 6 days later for HIV p24 antigen.

FIG. 4B—GBV-C NS5A inhibits HIV replication. Data represent HIV replication (p24 antigen in supernatants). HIV replication was not different on days 1 and 2; however, by day 3, Jurkat-NS5A cells had 45% reduction in HIV p24 Ag at 3 days post-infection (T-test, P<0.03), and by 56%-60% on days 4, 5, and 6 (p<0.001 for each day). Infections were performed in triplicate, NS5A and vector control error bars shown. Results were similar for Jurkat cells without any plasmid DNA as well.

FIG. 5—Release of SDF-1 into culture supernatants by Jurkat cells with or without GBV-C NS5A. SDF-1 in supernatants was measured by ELISA in triplicate at time points shown. There were no differences between Jurkat and vector control Jurkat cells; however, SDF-1 increased in NS5A expressing cells on days 5 and 6 (T test, P<0.001 for both days).

FIG. 6—Jurkat cells expressing GBV-C NS5A have lower surface density of CXCR4 compared to Jurkat cells with a vector control. Jurkat cells that express NS5A or the vector control were grown without doxycycline, and analyzed by flow cytometry for surface expression of CXCR4 (shown above). Reproducibly, cells with NS5A demonstrated a 49.6% reduction in CXCR4 mean fluorescence (p=0.003). No difference in CCR5 or CD4 was observed (data not shown). Preliminary microarray data demonstrated decreased levels of CXCR4 25 mRNA in NS5A expressing cells.

FIG. 7—Release of caspase 3/7 into culture supernatants following incubation with anti-Fas antibody (CH11). Jurkat cells with NS5A were relatively resistant to induction of apoptosis by CH11 when compared to Jurkat cells with the vector control. Similarly, spontaneous apoptosis was greater in Jurkat NS5A cells compared to vector control. Preliminary microarray data demonstrated NS5A-related increases in mRNA levels for TGFB1-induced anti-apoptotic factor 1 and chemokine ligand 25 mRNA. Both of these genes enhance resistance to apoptosis, and CC ligand 25 does so in CD4+ T cells.

FIG. 8—Dose-Dependent HIV Inhibition by the NS5A Protein. Left panel—NS5A-expressing Jurkat cells; right panel—vector control Jurkat cells. HIV replication is measured by p24 antigen levels on days 2-6 (D2-D6). Amounts of doxycycline, which inversely relate to NS5A production, are shown in mg/ml.

FIG. 9—Both IFN-R and IFN-S GBV-C NS5A Inhibit HIV Replication. HIV replication is measured by p24 antigen levels on days 2-6 (D2-D6). VC=vector control; GFP=vector expression GFP only; IFN-S=interferon sensitive NS5A; IFN-R=interferon resistant NS5A.

FIG. 10—Anti-SDF-1 Blocks NSF5A-Mediated Inhibition of HIV Replication. Results reported as percent increase in HIV replication when cultures are incubated with anti-SDF-1 neutralizing antibody as compared to isotype control antibody. IFN-R=interferon resistant NS5A; IFN-R-GFP=interferon resistant NS5A linked to GFP; IFN-S-GFP=interferon sensitive NS5A linked to GFP.

FIGS. 11A–C—NS5A Inhibits Multiple Distinct HIV Strains. HIV replication is measured by p24 antigen levels on days 2-6 (D2-D6). VC-GFP=vector control expression GFP; IFN-R-interferon resistant NS5A; IFN-R-GFP=interferon resistant NS5A linked to GFP; IFN-S-GFP=interferon sensitive NS5A linked to GFP.

FIG. 12—Fragments of NS5A That Have Been Stably Expressed in Jurkat Cells. Shaded bars indicate fusions or fragments.

FIG. 13—Fragments of NS5A That Inhibit HIV Replication. Shaded bars inhibit HIV; open bars do not inhibit HIV.

FIG. 14—Doxycycline-Represents Expression of NS5A. Lanes are labeled with amounts of doxycycline (µg/ml); arrows indicate NS5A or mutants (µl NS5A).

FIG. 15—Percent Inhibition of HIV Replication by NS5A Fragments. Various fragments identified by residues of NS5A remaining, and their ability to inhibit HIV replication (by p24 antigen levels) on day 6 after infection.

FIG. 16—Fragments of NS5A That Inhibit HIV Replication. Shaded bars inhibit HIV; open bars do not inhibit HIV; lined bars not yet tested.

FIG. 17—R 152–237 NS5A inhibits HIV when expressed in Jurkat cells. By deduction, it appears that 152-182 is the minimal sequence required for HIV inhibition. Note that the latter sequence is conserved (IFN-R and IFN-S) GBV-C NS5A, and that there are multiple potential phosphorylation sites (SEQ ID: NOS-10-13).

DESCRIPTION OF ILLUSTRATIVE EMBODIMENTS

Anti-retroviral medications suppress viral replication in HIV disease, yet they have failed to eradicate the virus from the body due to the multi-faceted nature of HIV infection, as well as the complexities of the immune system. Methods are being developed that both prevent infection and boost the immune system to keep it functioning at a level where it can assist in fighting HIV infection.

The present inventors have previously reported on methods and compositions for therapeutic and/or prophylactic treatment of HIV infection, including GBV-C envelope proteins, in particular GBV-C envelope protein E2 (E2). More specifically, the inventors have shown that HIV-infected subjects that are co-infected with GB virus C (GBV-C) typically have reduced mortality and slower progression to AIDS as compared to HIV-infected subjects without GBV-C co-infection (PCT/US2004/017706). Infection of peripheral blood mononuclear cells (PBMCs) with GBV-C and HIV results in inhibition of HIV-1 replication. GBV-C infection typically inhibits HIV by inducing β-chemokines and reducing expression of the HIV co-receptor CCR5, explaining part of the beneficial clinical findings of GBV-C on HIV disease progression.

The inventors also described a therapeutic use for antibodies and/or binding agents that bind GBV-C proteins (e.g., envelope proteins), in particular, the E2 protein, and similar antigens used for producing these antibodies or binding agents (PCT/US03/33925).

The inventors now demonstrate a unique role for the NS5A protein of GBV-C, as well as NS5A’s from other flaviviruses, in the inhibition of HIV replication. Various aspects of the invention are described below.
III. Flaviviruses

A. Family

With a total of 69 pathogens in its ranks, Flaviviridae contains a myriad of viruses that cause disease in humans. Foremost among these is Yellow Fever Virus, the type virus of the Flaviviridae, from which the family begets its name (flavus in Latin means "yellow"). Flaviviruses have been subdivided by the ICTV into three genera: Flavivirus, Pestivirus and Hepacivirus.

The Flavivirus genus contains several dangerous viruses including yellow fever virus, dengue fever virus, and Japanese encephalitis (JE) virus. The Pestivirus genus is home to the three serotypes of bovine viral diarrhea, but no known human pathogens. The genus Hepacivirus consists of hepatitis C virus and its relatives.

Flavivirus genomes consist of a monopartite (i.e., one piece of linear, single-stranded, positive sense RNA. Because the RNA is positive sense, the nucleic acid itself is capable of instigating an infection in the appropriate host cells. The total genome can range from 10 to 11 kilobase pairs. The genome 3' terminus is not polyadenylated. The 5' end has a methylated nucleotide cap (allows for translation) or a genome-linked protein (Vpg). Pestivirus genomes are reported to be 12.5 kb in length. Like the Flavivirus genus, no poly-A tail exists on the 3' end of the RNA, however, Pestivirus genus members lack a 5' cap. In both genera, structural genes are found towards the 5' end of the RNA. Both the Pestivirus and Hepacivirus genera contain internal ribosomal entry sites (IRES) that provide a site of translation initiation for host ribosomes. This is in contrast to the Flavivirus genus that uses the technique of ribosomal scanning to commence protein synthesis.

Under the EM, virions appear roughly as spheres (some experts say they're "pleomorphic"), 40-65 nm in diameter. What can be seen under the microscope is the virus's lipid envelope, which it obtains from host cells during egress (leaving the cell). Underneath the envelope can be found an icosahedral capsid coat approximately 25-30 nm in diameter.

All members of the Flavivirus genus are transmitted by arthropods (i.e., mosquitoes and ticks) while Hepatitis C is spread parenterally (i.e., through contaminated bodily fluids). A key feature for viral transmission in Flaviviruses is that they are capable of reproducing in their vector. Without the ability to replicate in the vector, they would not remain viable to be passed from one host to the next.

B. GBV-C

Like other members of the Flaviviridae, GBV-C is a positive-strand RNA virus that encodes a single long open reading frame (I.Leney et al., 1996). GBV-C does not cause acute or chronic hepatitis, yet it is the family member most closely related to HCV, the cause of hepatitis C. Sequences of GBV-C have been previously reported, for example in U.S. Pat. No. 5,874,563, which is specifically incorporated by reference. In particular, an infectious GBV-C clone has been described in the PCT application WO 01/77157, which is incorporated herein by reference.

The GBV-C polyprotein is predicted to be cleaved into two envelope proteins (E1 and E2, referred to collectively as GBV-C envelope protein), an RNA helicase, a trypsin-like serine protease, and an RNA-dependent RNA polymerase. A major difference between GBV-C and HCV is in the amino terminus of the polyprotein. In many isolates, this region is truncated, and no core (or nucleocapsid) protein is present (Simons et al., 1995; Xiang et al., 1999). In vitro translation experiments suggest that the AUG immediately upstream of the putative E1 protein is preferentially used to initiate translation, although there may be as many as four AUG's in frame with the polyprotein upstream of this AUG (Simons et al., 1996).

The site of GBV-C replication has not been clearly identified, but it appears that replication in the hepatocyte, if it occurs, is not the primary source of virus in infected individuals (Lankus et al., 1998; Pessoa et al., 1998; Seipp et al., 1999). Recently, there are reports that human peripheral blood mononuclear cells (PBMC's) and interferon-resistant Daupl cells are permissive for GBV-C replication (Fogeda et al., 1999; Shimizu, 1999). In addition, transient replication of GBV-C was described in MT-2 cells (a human T-cell line), and Pfisch (a human hepatocyte line immortalized with simian virus 40 large T antigen) (Seipp et al., 1999).

C. Other Flavivirus NS5A's

Other Flaviviruses contain NS5A's that can be used in accordance with the present invention. These viruses include DEN1-4, YFV, TBEV, WNV, CSFV, BDV, GBV-A, GBV-B, HGV, HCV2a, HCV3a, HCV2b, HCV1a, HCV1c and HCV1b.

IV. GBV-C Polypeptides

In certain aspects, the invention is directed to the nSS5A polypeptide of GBV-C virus, or a peptide or polypeptide derived there from SEQ ID NO:2 shows the NS5A translated product of SEQ ID NO:1 (cDNA). It is contemplated that the compositions and methods disclosed herein may be utilized to express all or part of SEQ ID NO:2 and derivatives thereof.

In certain embodiments, compositions of the invention may include the nucleic acids encoding the peptides as set forth in SEQ ID NO:1 or 2. Determination of which protein or DNA molecules inhibit HIV may be achieved using functional assays measuring HIV replication and infectivity, which are familiar to those of skill in the art. The structure of the various polypeptides or peptides can be modeled or resolved by computer modeling, NMR, or x-ray crystallography. Such structures may be used to engineer derivatives of the various NS5A protein.

Exemplary accession nos. (incorporated by reference) for other NS5A's are as follows:

<table>
<thead>
<tr>
<th>Virus</th>
<th>Accession No.</th>
<th>Virus</th>
<th>Accession No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>West Nile</td>
<td>DQ018019</td>
<td>Dengaue 1-4</td>
<td>AF063290</td>
</tr>
<tr>
<td>Yellow fever</td>
<td>NC002031</td>
<td>HCV 1a</td>
<td>AF031783</td>
</tr>
<tr>
<td></td>
<td>Y603338</td>
<td></td>
<td>AF333324</td>
</tr>
<tr>
<td>BVDV</td>
<td>AF502399</td>
<td></td>
<td>D43853</td>
</tr>
<tr>
<td>Dengaue 1-4</td>
<td>M257512</td>
<td></td>
<td>D00944</td>
</tr>
<tr>
<td></td>
<td>M14091</td>
<td></td>
<td>D10888</td>
</tr>
<tr>
<td></td>
<td>M20558</td>
<td></td>
<td>AF048666</td>
</tr>
</tbody>
</table>

A. Variants of GBV-C NS5A Polypeptides

Embodiments of the invention include various GBV-C NS5A polypeptides, peptides, and derivatives thereof. Amino acid sequence variants of a polypeptide can be substitutional, insertional or deletion variants. Deletion variants lack one or more residues of the native protein that are not essential for function or immunogenic activity. Insertional mutants typically involve the addition of material at a non-terminal point in the polypeptide. This may include the insertion of an immunoreactive epitope or simply a single residue. Terminal additions, called fusion proteins, are discussed below.

Substitutional variants typically contain the exchange of one amino acid for another at one or more sites within the protein, and may be designed to modulate one or more properties of the polypeptide, such as stability against proteolytic cleavage, without the loss of other functions or properties.
Substitutions of this kind preferably are conservative, that is, one amino acid is replaced with one of similar shape and charge. Conservative substitutions are well known in the art and include, for example, the changes of: alanine to serine; arginine to lysine; asparagine to glutamine or histidine; asparagine to glutamate; cysteine to serine; glutamine to asparagine; glutamate to asparagine; glycine to proline; histidine to asparagine or glutamine; isoleucine to leucine or valine; leucine to valine or isoleucine; lysine to arginine; methionine to leucine or isoleucine; phenylalanine to tyrosine, leucine or methionine; serine to threonine; threonine to serine; tryptophan to tyrosine; tyrosine to tryptophan or phenylalanine; and valine to isoleucine or leucine.

The term "biologically functional equivalent" is well understood in the art and is further defined in detail herein. Accordingly, sequences that have between about 70% and about 80%; or more preferably, between about 81% and about 90%; or even more preferably, between about 91% and about 99%; of amino acids that are identical or functionally equivalent to the amino acids of GBV-C NSSA polypeptides, for example SEQ ID NO.2, provided the biological activity of the protein or peptide is maintained.

The term "functionally equivalent codon" is used herein to refer to codons that encode the same amino acid, such as the six codons for arginine or serine, and also refers to codons that encode biologically equivalent amino acids (see Table 1, below).

Certain embodiments of the invention include various peptides or polypeptides of the NSSA protein. For example, all or part of a GBV-C NSSA protein as set forth in SEQ ID NO.2 may be used in various embodiments of the invention. In certain embodiments, a fragment of the NSSA protein may comprise, but is not limited to about 5, about 6, about 7, about 8, about 9, about 10, about 11, about 12, about 13, about 14, about 15, about 16, about 17, about 18, about 19, about 20, about 21, about 22, about 23, about 24, about 25, about 26, about 27, about 28, about 29, about 30, about 31, about 32, about 33, about 34, about 35, about 36, about 37, about 38, about 39, about 40, about 41, about 42, about 43, about 44, about 45, about 46, about 47, about 48, about 49, about 50, about 51, about 52, about 53, about 54, about 55, about 56, about 57, about 58, about 59, about 60, about 61, about 62, about 63, about 64, about 65, about 66, about 67, about 68, about 69, about 70, about 71, about 72, about 73, about 74, about 75, about 76, about 77, about 78, about 79, about 80, about 81, about 82, about 83, about 84, about 85, about 86, about 87, about 88, about 89, about 90, about 91, about 92, about 93, about 94, about 95, about 96, about 97, about 98, about 99, about 100, about 110, about 120, about 130, about 140, about 150, about 160, about 170, about 180, about 190, about 200, about 210, about 220, about 230, about 240, about 250, about 275, about 300, about 325, about 350, about 375, about 400, about 415, and any range derivable therein.

It also will be understood that amino acid and nucleic acid sequences may include additional residues, such as additional N — or C-terminal amino acids or 5' or 3' sequences, and yet still be essentially as set forth in one of the sequences disclosed herein, so long as the sequence meets the criteria set forth above, including the maintenance of biological activity (e.g., immunogenicity) where protein expression is concerned. The addition of terminal sequences particularly applies to nucleic acid sequences that may, for example, include various non-coding sequences flanking either of the 5' or 3' portions of the coding region.

The following is a discussion based upon changing of the amino acids of an NSSA polypeptide or peptide to create an equivalent, or even an improved, second-generation molecule. For example, certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies or binding sites on substrate molecules. Since it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid substitutions can be made in a protein sequence, and in its underlying DNA or RNA coding sequence, and nevertheless produce a protein with like properties. It is thus contemplated by the inventors that various changes may be made in the DNA or RNA sequences of genes or coding regions without appreciable loss of their biological utility or activity, as discussed herein. Table 1 shows the codons that encode particular amino acids.

<table>
<thead>
<tr>
<th>Table 1: CODON TABLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amino Acids</td>
</tr>
<tr>
<td>Alanine</td>
</tr>
<tr>
<td>Cysteine</td>
</tr>
<tr>
<td>Aspartic acid</td>
</tr>
<tr>
<td>Glutamic acid</td>
</tr>
<tr>
<td>Phenylalanine</td>
</tr>
<tr>
<td>Glycine</td>
</tr>
<tr>
<td>Histidine</td>
</tr>
<tr>
<td>Isoleucine</td>
</tr>
<tr>
<td>Lysine</td>
</tr>
<tr>
<td>Leucine</td>
</tr>
<tr>
<td>Methionine</td>
</tr>
<tr>
<td>Asparagine</td>
</tr>
<tr>
<td>Proline</td>
</tr>
<tr>
<td>Glutamine</td>
</tr>
<tr>
<td>Arginine</td>
</tr>
<tr>
<td>Serine</td>
</tr>
<tr>
<td>Threonine</td>
</tr>
<tr>
<td>Valine</td>
</tr>
<tr>
<td>Tyrosine</td>
</tr>
</tbody>
</table>

In making such changes, the hydrophilic index of amino acids may be considered. The importance of the hydrophilic amino acid index in conferring interactive biologic function on a protein is generally understood in the art (Kyte and Doolittle, 1982). It is accepted that the relative hydrophilic character of the amino acid contributes to the secondary structure of the resultant protein, which in turn defines the interaction of the protein with other molecules, for example, enzymes, substrates, receptors, DNA, antibodies, antigens, and the like.

It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U.S. Pat. No. 4,554,101, incorporated herein by reference, states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with a biological property of the protein.

It is understood that an amino acid substituted for another having a similar hydrophilicity value still produces a biologically equivalent and immunologically equivalent protein.

In certain embodiments, an NSSA polypeptide may be a fusion protein. Fusion proteins may alter the characteristics of a given polypeptide, such as antigenicity or purification characteristics. A fusion protein is a specialized type of insertional variant. This molecule generally has all or a substantial portion of the native molecule, linked at the N— or C-terminus, to all or a portion of a second polypeptide. For example, fusions typically employ leader sequences from other species.
to permit the recombinant expression of a protein in a heterologous host. Another useful fusion includes the addition of an immunologically active domain, such as an antibody epitope, to facilitate purification of the fusion protein. Inclusion of a cleavage site at or near the fusion junction will facilitate removal of the extraneous polypeptide after purification. Other useful fusions include linking of functional domains, such as active sites from enzymes such as a hydrolase, glycosylation domains, cellular targeting signals, or transmembrane regions.

B. In vitro Production of N5S5A Polypeptides or Peptides

Various types of expression vectors are known in the art that can be used for the production of protein products. Following transfection with an expression vector, a cell in culture, e.g., a primary mammalian cell, a recombinant product may be prepared in various ways. A host cell strain may be chosen that modulates the expression of the inserted sequences, or that modifies and processes the gene product in the manner desired. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may be important for the function of the protein. Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of proteins. Appropriate cell lines or host systems can be chosen to insure the correct modification and processing of the foreign protein expressed. In order for the cells to be kept viable while in vitro and in contact with the expression construct, it is necessary to ensure that the cells maintain contact with the correct ratio of oxygen and carbon dioxide and nutrients but are protected from microbial contamination. Cell culture techniques are well documented (for exemplary methods see Freshney, 1992).

Animal cells can be propagated in vitro in two modes: as non-anchorage-dependent cells growing in suspension throughout the bulk of the culture or as anchorage-dependent cells requiring attachment to a solid substrate for their propagation (i.e., a monolayer type of cell growth).

Non-anchorage dependent or suspension cultures from continuous established cell lines are the most widely used means of large-scale production of cells and cell products. However, suspension cultured cells have limitations, such as tumorigenic potential and lower protein production than adherent cells.

In further aspects of the invention, other protein production methods known in the art may be used, including but not limited to prokaryotic, yeast, and other eukaryotic hosts such as insect cells and the like.

Because of their relatively small size, the peptides of the invention can also be synthesized in solution or on a solid support in accordance with conventional techniques. Various automatic synthesizers are commercially available and can be used in accordance with known protocols. See, for example, Stewart and Young, (1984); Tam et al., (1983); Merrifield, (1986); and Barany and Merrifield (1979), each incorporated herein by reference. Short peptide sequences, or libraries of overlapping peptides, usually from about 6 up to about 35 to 50 amino acids, which correspond to the selected regions described herein, can be readily synthesized and then screened in screening assays designed to identify reactive peptides.

C. Protein Purification

It may be desirable to purify N5S5A polypeptides and peptides, or variants and derivatives thereof. Protein purification techniques are well known to those of skill in the art. These techniques involve, at one level, the crude fractionation of the cellular milieu to polypeptide and non-polypeptide fractions. Having separated the polypeptide from other proteins, the polypeptide of interest may be further purified using chromatographic and electophoretic techniques to achieve partial or complete purification (or purification to homogeneity). Analytical methods particularly suited to the preparation of a pure peptide are ion-exchange chromatography, hydrophobic interaction chromatography, exclusion chromatography; polyacrylamide gel electrophoresis; isoelectric focusing. A particularly efficient method of purifying peptides is fast protein liquid chromatography (FPLC).

Certain aspects of the present invention concern the purification, and in particular embodiments, the substantial purification, of an encoded protein or peptide. The term “purified protein or peptide” as used herein, is intended to refer to a composition, isolatable from other components, wherein the protein or peptide is purified to any degree relative to its naturally obtainable state. A purified protein or peptide therefore also refers to a protein or peptide, free from the environment in which it may naturally occur.

Generally, “purified” will refer to a protein or peptide composition that has been subjected to fractionation to remove various other components, and which composition substantially retains its expressed biological activity. Where the term “substantially purified” is used, this designation will refer to a composition in which the protein or peptide forms the major component of the composition, such as constituting about 50%, about 60%, about 70%, about 80%, about 90%, about 95% or more of the proteins in the composition.

Various methods for quantifying the degree of purification of the protein or peptide will be known to those of skill in the art in light of the present disclosure. These include, for example, determining the specific activity of an active fraction, or assessing the amount of polypeptides within a fraction by SDS/PAGE analysis. A preferred method for assessing the purity of a fraction is to calculate the specific activity of the fraction, to compare it to the specific activity of the initial extract, and to thus calculate the degree of purity, herein assessed by a “fold purification number.” The actual units used to represent the amount of activity will, of course, be dependent upon the particular assay technique chosen to follow the purification and whether or not the expressed protein or peptide exhibits a detectable activity.

There is no general requirement that the protein or peptide always be provided in their most purified state. Indeed, it is contemplated that less substantially purified products will have utility in certain embodiments. Partial purification may be accomplished by using fewer purification steps in combination, or by utilizing different forms of the same general purification scheme.

V. Flavivirus Polynucleotides

Certain embodiments of the invention include Flavivirus N5S5A polynucleotides or nucleic acid molecules and fragments thereof. The polynucleotides of the invention may be isolated and purified from Flavivirus or cells infected or transfected with Flavivirus polynucleotides. The term isolated indicating they are free or substantially free from total viral or cellular genomic RNA or DNA, and proteins. It is contemplated that an isolated and purified Flavivirus nucleic acid molecule may take the form of RNA or DNA. A Flavivirus nucleic acid molecule refers to an RNA or DNA molecule that is capable of yielding all or part of a Flavivirus N5S5A from a transfected cell.

When the present application refers to the function or activity of an infectious Flavivirus that is encoded by a Flavivirus polynucleotide, it is meant that the polynucleotide encodes a molecule that has the ability to propagate an infectious Flavivirus virus particle from a cell. It is contemplated that a Flavivirus polynucleotide may refer to a Flavivirus RNA transcript that is able to propagate an infectious Flavivirus
virus particle after introduction to a cell or to a Flavivirus expression construct, clone, or vector composed of double-stranded DNA or DNA/RNA hybrid that is similarly capable, or a doublestranded DNA that is similarly capable following in vitro transcription.

The term “cDNA” is intended to refer to DNA prepared using RNA as a template. The advantage of using a cDNA, as opposed to genomic RNA or an RNA transcript is stability and the ability to manipulate the sequence using recombinant DNA technology (see Maniatis, 1990; Ausubel, 1996). There may be times when the full or partial genomic sequence is preferred.

It is contemplated that a given Flavivirus may be represented by natural variants or strains that have slightly different nucleic acid sequences but, nonetheless, encode the same viral polypeptides (see Table 1 above). Consequently, the present invention also encompasses derivatives of Flaviviruses with minimal amino acid changes in its viral proteins, but that possesses the same activities.

The term “gene” is used for simplicity to refer to the nucleic acid giving rise to a functional protein, polypeptide, or peptide-encoding unit. As will be understood by those in the art, this functional term includes genomic sequences, cDNA sequences, and smaller engineered gene sequences that express, or may be adapted to express, proteins, polypeptides, domains, peptides, fusion proteins, and mutants. The nucleic acid molecule encoding Flavivirus may contain a contiguous nucleic acid sequence encoding one or more Flavivirus genes and regulatory regions and be of the following lengths: about 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, 1090, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, 3000, 3100, 3200, 3300, 3400, 3500, 3600, 3700, 3800, 3900, 4000, 4100, 4200, 4300, 4400, 4500, 4600, 4700, 4800, 4900, 5000, 5100, 5200, 5300, 5400, 5500, 5600, 5700, 5800, 5900, 6000, 6100, 6200, 6300, 6400, 6500, 6600, 6700, 6800, 6900, 7000, 7100, 7200, 7300, 7400, 7500, 7600, 7700, 7800, 7900, 8000, 8100, 8200, 8300, 8400, 8500, 8600, 8700, 8800, 8900, 9000, 9100, 9200, 9300, 9400, 9500, 10,000 or more nucleotides, nucleosides, or base pairs. Such sequences may be identical or complementary to all or part of SEQ ID NO:1.

In particular embodiments, the invention concerns isolated nucleic acid segments and recombinant vectors incorporating DNA sequences that encode Flavivirus NS5A polypeptides or peptides. Such vectors used in the present invention, regardless of the length of the coding sequence itself, may be combined with other DNA or RNA sequences, such as promoters, polyadenylation signals, additional restriction enzyme sites, multiple cloning sites, other coding segments, and the like, such that their overall length may vary considerably. It is therefore contemplated that a nucleic acid fragment of almost any length may be employed, with the total length preferably being limited by the ease of preparation and use in the intended recombinant DNA protocol.

In a non-limiting example, one or more nucleic acid constructs may be prepared that include a contiguous stretch of nucleotides identical to or complementary to a Flavivirus genome. A nucleic acid construct may be about 50, 70, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1,000, 2,000, 3,000, 4,000, 5,000, 6,000, 7,000, 8,000, 9,000, and 9,400, nucleotides in length, as well as constructs of greater size, up to and including chromosomal sizes (including all intermediate lengths and intermediate ranges), given the advent of nucleic acid constructs such as a yeast artificial chromosome are known to those of ordinary skill in the art. It will be readily understood that “intermediate lengths” and “intermediate ranges,” as used herein, means any length or range including or between the quoted values (i.e., all integers including and between such values). Non-limiting examples of intermediate lengths include about 11, about 12, about 13, about 16, about 17, about 18, about 19, etc.; about 21, about 22, about 23, etc.; about 31, about 32, etc.; about 51, about 52, about 53, etc.; about 101, about 102, about 103, etc.; about 151, about 152, about 153, etc.

The nucleic acid segments used in the present invention encompass biologically functional and/or immunogenically equivalent Flavivirus NS5A proteins and peptides. Such sequences may arise as a consequence of codon redundancy and functional equivalency that are known to occur naturally within nucleic acid sequences and the proteins thus encoded. Alternatively, functionally and immunogenically equivalent proteins or peptides may be created via the application of recombinant DNA technology, in which changes in the protein structure may be engineered, based on considerations of the properties of the amino acids being exchanged. Changes designed by human may be introduced through the application of site-directed mutagenesis techniques, e.g., to introduce improvements to the antigenicity of the protein.

A. Vectors Encoding Flavivirus

The present invention encompasses the use of vectors to encode for all or part of one or more Flavivirus NS5A polypeptides, including an infectious Flavivirus. The term “vector” is used to refer to a carrier nucleic acid molecule into which a nucleic acid sequence can be inserted for introduction into a cell where it can be replicated. A nucleic acid sequence can be “exogenous,” which means that it is foreign to the cell into which the vector is being introduced or that the sequence is homologous to a sequence in the cell but in a position within the host cell nucleic acid in which the sequence is ordinarily not found. Vectors include plasmids, cosmids, viruses (bacteriophage, animal viruses, and plant viruses), and artificial chromosomes (e.g., YACs). In particular embodiments, gene therapy or immunization vectors are contemplated. One of skill in the art would be well equipped to construct a vector through standard recombinant techniques, which are described in Maniatis et al., 1990 and Ausubel et al., 1996, both incorporated herein by reference.

The term “expression vector” or “expression construct,” refers to a vector containing a nucleic acid sequence coding for at least part of a gene product capable of being transcribed. In some cases, RNA molecules are then translated into a protein, polypeptide, or peptide. In other cases, these sequences are not translated, for example, in the production of antisense molecules or ribozymes. Expression vectors can contain a variety of “control sequences,” which refer to nucleic acid sequences necessary for the transcription and possibly translation of an operably linked coding sequence in a particular host organism. In addition to control sequences that govern transcription and translation, vectors and expression vectors may contain nucleic acid sequences that serve other functions as well and are described infra. It is contemplated that an infectious Flavivirus particle of the present invention may arise from a vector containing Flavivirus sequence or RNA encoding Flavivirus sequence into a cell. Either of these, or any other nucleic acid molecules of the present invention may be constructed with any of the follow-
ing nucleic acid control sequences. Thus, the full-length RNA transcript may contain the benefit of recombinant DNA technology such that it contains exogenous control sequences or genes.

1. Promoters and Enhancers

A “promoter” is a control sequence that is a region of a nucleic acid sequence at which initiation and rate of transcription are controlled. It may contain genetic elements at which regulatory proteins and molecules may bind such as RNA polymerase and other transcription factors. The phrases “operatively positioned,” “operatively linked,” “under control,” and “under transcriptional control” means that a promoter is in a correct functional location and/or orientation in relation to a nucleic acid sequence to control transcriptional initiation and/or expression of that sequence. A promoter may or may not be used in conjunction with an “enhancer,” which refers to a cis-acting regulatory sequence involved in the transcriptional activation of a nucleic acid sequence.

A promoter may be one naturally associated with a gene or sequence, as may be obtained by isolating the 5’ non-coding sequences located upstream of the coding segment and/or exon. Such a promoter can be referred to as “endogenous.” Similarly, an enhancer may be one naturally associated with a nucleic acid sequence, located either downstream or upstream of that sequence. Alternatively, certain advantages will be gained by positioning the coding nucleic acid segment under the control of a recombinant or heterologous promoter, which refers to a promoter that is not normally associated with a nucleic acid sequence in its natural environment. A recombinant or heterologous promoter refers also to an enhancer not normally associated with a nucleic acid sequence in its natural environment. Such promoters or enhancers may include promoters or enhancers of other genes, and promoters or enhancers isolated from any other prokaryotic, viral, or eukaryotic cell, and promoters or enhancers not “naturally occurring,” i.e., containing different elements of different transcriptional regulatory regions, and/or mutations that alter expression. In addition to producing nucleic acid sequences of promoters and enhancers synthetically, sequences may be produced using recombinant cloning and/or nucleic acid amplification technology, including PCR, in connection with the compositions disclosed herein (see U.S. Pat. Nos. 4,683,202 and 5,928,906, each incorporated herein by reference). Furthermore, it is contemplated the control sequences that direct transcription and/or expression of sequences within non-nuclear organelles such as mitochondria, chloroplasts, and the like, can be employed as well.

Naturally, it will be important to employ a promoter and/or enhancer that effectively directs the expression of the nucleic acid segment in the cell type, organelle, and organism chosen for expression. Those of skill in the art of molecular biology generally know the use of promoters, enhancers, and cell type combinations for protein expression, for example, see Sambrook et al. (2001), incorporated herein by reference. The promoters employed may be constitutive, tissue-specific, inducible, and/or useful under the appropriate conditions to direct high level expression of the introduced DNA segment, such as is advantageous in the large-scale production of recombinant proteins and/or peptides. The promoter may be heterologous or exogenous, i.e., from a different source the Flavivirus sequence. In some examples, a prokaryotic promoter is employed for use with in vitro transcription of a desired sequence. Prokaryotic promoters for use with many commercially available systems include T7, T3, and Sp6.

Table 2 lists several elements/promoters that may be employed, in the context of the present invention, to regulate the expression of a gene. This list is not intended to be exhaustive of all the possible elements involved in the promotion of expression but, merely, to be exemplary thereof. Table 3 provides examples of inducible elements, which are regions of a nucleic acid sequence that can be activated in response to a specific stimulus.

| Table 2 |
|---|---|
| Promoter and/or Enhancer | References |
| Immunoglobulin Heavy Chain | Buermeyer et al., 1983; Gilles et al., 1983; Grosschedl et al., 1985; Athersuch et al., 1986; 1987; Hirtel et al., 1987; Weinberger et al., 1984; Kiledjian et al., 1988; Porteus et al., 1990 |
| Immunoglobulin Light Chain | Queen et al., 1985; Ficaro et al., 1984 |
| HLA DQ α and/or DQ β | Ludin et al., 1987; Wiinoto et al., 1989; Redondo et al., 1990 |
| β-Interferon | Sullivan et al., 1987; Goodbourn et al., 1986; Fujita et al., 1987; Goodbourn et al., 1988 |
| Interleukin-2 | Greene et al., 1989; Lin et al., 1990 |
| Interleukin-2 Receptor | Koch et al., 1989 |
| MIC Class II | Sherman et al., 1989 |
| MIC Class II HLA-DRα | Kawano et al., 1988; Ng et al., 1989 |
| Muscle Creatine Kinase (MCK) | Jaynes et al., 1988; Horlick et al., 1989; Johnson et al., 1989 |
| Prealbumin (Transfthylarin) | Cost et al., 1988 |
| Elastase I | Omura et al., 1987 |
| Metallothionein (MTII) | Katin et al., 1987; Culotta et al., 1989 |
| Collagenase | Pinkert et al., 1987; Angel et al., 1987 |
| Albunin | Pinkert et al., 1987; Torchia et al., 1989, 1990 |
| α-Fetoprotein | Godbourn et al., 1988; Compere et al., 1989 |
| γ-Globin | Bodine et al., 1987; Perez-Stable et al., 1990 |
| β-Globin | Truel et al., 1987 |
| c-fos | Cohen et al., 1987 |
| c-HA-ras | Truesman, 1986; Deschamps et al., 1985 |
| Insulin | Edlund et al., 1985 |
| Neural Cell Adhesion Molecule (NCAM) | Hirsch et al., 1990 |
TABLE 2-continued

<table>
<thead>
<tr>
<th>Promoter/Enhancer</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-Interferon</td>
<td>Latimer et al., 1990</td>
</tr>
<tr>
<td>H2B (TH2B) Histone</td>
<td>Hwang et al., 1990</td>
</tr>
<tr>
<td>Mouse and/or Type I Collagen</td>
<td>Rihe et al., 1989</td>
</tr>
<tr>
<td>Glucose-Regulated Proteins (GPR54 and GRP78)</td>
<td>Chang et al., 1989</td>
</tr>
<tr>
<td>Rat Growth Hormone</td>
<td>Larsen et al., 1986</td>
</tr>
<tr>
<td>Human Serum Amyloid A (SAA)</td>
<td>Elbashir et al., 1989</td>
</tr>
<tr>
<td>Troponin I (TN I)</td>
<td>Yuzoy et al., 1989</td>
</tr>
<tr>
<td>Platelet-Derived Growth Factor (PDGF)</td>
<td>Pech et al., 1989</td>
</tr>
<tr>
<td>Hepatitis B Virus</td>
<td>Buja et al., 1986, Ismoel et al., 1990, Shaad et al., 1987, Spandau et al., 1988, Vannrice et al., 1988</td>
</tr>
<tr>
<td>Cytomegalovirus (CMV)</td>
<td>Weber et al., 1984, Boskliart et al., 1985, Fordeking et al., 1986</td>
</tr>
<tr>
<td>Gibbon Ape Leukemia Virus</td>
<td>Holbrook et al., 1987, Quinn et al., 1989</td>
</tr>
</tbody>
</table>

TABLE 3

<table>
<thead>
<tr>
<th>Element</th>
<th>Inducer</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>MT II</td>
<td>Phorbol Ester (TPA)</td>
<td>Palmier et al., 1982; Haulinger et al., 1985; Searle et al., 1985; Stuart et al., 1985; Imigamata et al., 1987, Karin et al., 1987; Angel et al., 1987b; McNeall et al., 1989</td>
</tr>
<tr>
<td>MMTV (mouse mammary tumor virus)</td>
<td>Glucocorticoids</td>
<td>Haang et al., 1981; Lee et al., 1981, 1983; Majors et al., 1983; Chandler et al., 1983; Lee et al., 1984, Poza et al., 1985, Sukai et al., 1988</td>
</tr>
<tr>
<td>β-Interferon</td>
<td>poly(rI) × poly(rC)</td>
<td>Tavernier et al., 1983</td>
</tr>
<tr>
<td>Adenovirus 5 E2</td>
<td>EIA</td>
<td>Imperiale et al., 1984</td>
</tr>
<tr>
<td>Collagenase</td>
<td>Phorbol Ester (TPA)</td>
<td>Angel et al., 1987b</td>
</tr>
<tr>
<td>Stromelysin</td>
<td>Phorbol Ester (TPA)</td>
<td>Angel et al., 1987b</td>
</tr>
<tr>
<td>SV40</td>
<td>Interferon, Newcastle Disease Virus</td>
<td>Hoo et al., 1988</td>
</tr>
<tr>
<td>Murine MA Gene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GRP78 Gene</td>
<td>A23187</td>
<td>Resende et al., 1988</td>
</tr>
<tr>
<td>α-2-Macroglobulin</td>
<td>IL-5</td>
<td>Kuro et al., 1989</td>
</tr>
<tr>
<td>Vimentin</td>
<td>Serum</td>
<td>Ritting et al., 1989</td>
</tr>
<tr>
<td>MHC Class I Gene H-2kb</td>
<td>Interferon</td>
<td>Binare et al., 1989</td>
</tr>
</tbody>
</table>
The identity of tissue-specific promoters or elements, as well as assays to characterize their activity, is well known to those of skill in the art. Examples of such regions include the human LIMK2 gene (Nomoto et al. 1999), the somatostatin receptor 2 gene (Kraus et al. 1998), murine epididymal retinoic acid-binding gene (Lazure et al. 1999), human CD4 (Zhao-Emonet et al. 1998), mouse alpha2 (XI) collagen (Tsumuki et al. 1998), DIA dopamine receptor gene (Lee et al. 1997), insulin-like growth factor II (Wu et al. 1997), human platelet endothelial cell adhesion molecule-1 (Almendo et al. 1996).

2. Initiation Signals and Internal Ribosome Binding Sites

A specific initiation signal also may be required for efficient translation of coding sequences. These signals include the ATG initiation codon or adjacent sequences. Exogenous translational control signals, including the ATG initiation codon, may need to be provided. One of ordinary skill in the art would readily be capable of determining this 15 and providing the necessary signals. It is well known that the initiation codon must be "in-frame" with the reading frame of the desired coding sequence to ensure translation of the entire insert. The exogenous translational control signals and initiation codons can be either natural or synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements.

In certain embodiments of the invention, the use of internal ribosome entry sites (IRES) elements are used to create multigene or polycistronic, messages. IRES elements are able to bypass the ribosome-scanning model of 5' methylated Cap dependent translation and begin translation at internal sites (Pelletier and Sonenberg, 1988). IRES elements derive from two members of the picornavirus family (polio andencephalomyocarditis) have been described (Pelletier and Sonenberg, 1988), as well as an IRES from a mammalian message (Macejak and Sarnow, 1991). IRES elements can be linked to heterologous open reading frames. Multiple open reading frames can be transcribed together, each separated by an IRES, creating polycistronic messages. By virtue of the IRES element, each open reading frame is accessible to ribosomes for efficient translation. Multiple genes can be efficiently expressed using a single promoter/enhancer to transcribe a single message (see U.S. Pat. Nos. 5,925,565 and 5,935,819, herein incorporated by reference).

3. Multiple Cloning Sites

Vectors can include a multiple cloning site (MCS), which is a nucleic acid region that contains multiple restriction enzyme sites, any of which can be used in conjunction with standard recombinant technology to digest the vector. (See Carbonelli et al., 1999, Levenson et al., 1998, and Coccar, 1997, incorporated herein by reference.) "Restriction enzyme digestion" refers to catalytic cleavage of a nucleic acid molecule with an enzyme that functions only at specific locations in a nucleic acid molecule. Many of these restriction enzymes are commercially available. Use of such enzymes is widely understood by those of skill in the art. Frequently, a vector is linearized or fragmented using a restriction enzyme that cuts within the MCS to enable equogenous sequences to be ligated to the vector. "Ligation" refers to the process of forming phosphodiester bonds between two nucleic acid fragments, which may or may not be contiguous with each other. Techniques involving restriction enzymes and ligation reactions are well known to those of skill in the art of recombinant technology.

4. Termination Signals

The vectors or constructs of the present invention will generally comprise at least one termination signal. A "termination signal" or "terminator" is comprised of the DNA sequences involved in specific termination of an RNA transcript by an RNA polymerase. Thus, in certain embodiments a termination signal that ends the production of an RNA transcript is contemplated. A terminator may be necessary in vivo to achieve desirable message levels.

In eukaryotic systems, the terminator region may also comprise specific DNA sequences that permit site-specific cleavage of the new transcript to expose a polyadenylation site. This signals a specialized endogenous polymerase to add a stretch of about 200 A residues (polyA) to the 3' end of the transcript. RNA molecules modified with this polyA tail appear to be more stable and are translated more efficiently. Thus, in other embodiments involving eukaryotes, it is preferred that that terminator comprises a signal for the cleavage of the RNA, and is more preferred that the terminator signal promotes polyadenylation of the message. The terminator and/or polyadenylation site elements can serve to enhance message levels and/or to minimize read through from the cassette into other sequences.

Terminators contemplated for use in the invention include any known terminator of transcription described herein or known to one of ordinary skill in the art, including but not limited to, for example, the termination sequences of genes, such as for example the bovine growth hormone terminator or viral termination sequences, such as for example the SV40 terminator. In certain embodiments, the termination signal may be a lack of transcribable or translatable sequence, such as due to a sequence truncation.

5. Polyadenylation Signals

For expression, particularly eukaryotic expression, one will typically include a polyadenylation signal to effect proper polyadenylation of the transcript. The nature of the polyadenylation signal is not believed to be crucial to the successful practice of the invention, and/or any such sequence may be employed. Preferred embodiments include the SV40 polyadenylation signal and/or the bovine growth hormone polyadenylation signal, convenient and/or known to function well in various target cells. Polyadenylation may increase the stability of the transcript or may facilitate cytoplasmic transport.

<table>
<thead>
<tr>
<th>Element</th>
<th>Inducer</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSP70</td>
<td>E1A, SV40 Large T</td>
<td>Taylor et al., 1989, 1990a, 1990b</td>
</tr>
<tr>
<td>Proliferin</td>
<td>Antigen</td>
<td></td>
</tr>
<tr>
<td>Tumor Necrosis Factor</td>
<td>Phorbol Ester-TPA</td>
<td>Mordacq et al., 1989</td>
</tr>
<tr>
<td>Thyroid Stimulating Hormone α Gene</td>
<td>Thyroid Hormone</td>
<td>Chatterjee et al., 1989</td>
</tr>
</tbody>
</table>
6. Origins of Replication

In order to propagate a vector in a host cell, it may contain one or more origins of replication sites (often termed "ori"), which is a specific nucleic acid sequence at which replication is initiated. Alternatively, an autonomously replicating sequence (ARS) can be employed if the host cell is yeast.

7. Selectable and Screenable Markers

In certain embodiments of the invention, the cells containing a nucleic acid construct of the present invention may be identified in vitro or in vivo by including a marker in the expression vector. Such markers would confer an identifiable change to the cell permitting easy identification of cells containing the expression vector. Generally, a selectable marker is one that confers a property that allows for selection. A positive selectable marker is one in which the presence of the marker allows for its selection, while a negative selectable marker is one in which its presence prevents its selection. An example of a positive selectable marker is a drug resistance marker.

Usually the inclusion of a drug selection marker aids in the cloning and identification of transformants, for example, genes that confer resistance to neomycin, puromycin, hygromycin, DHFR, GPT, zeocin and histidinol are useful selectable markers. In addition to markers conferring a phenotype that allows for the discrimination of transformants based on the implementation of conditions, other types of markers including screenable markers such as GFP, whose basis is calorimetric analysis, are also contemplated. Alternatively, screenable enzymes such as herpes simplex virus thymidine kinase (tk) or chloramphenicol acetyltransferase (CAT) may be utilized. One of skill in the art would also know how to employ immunologic markers, possibly in conjunction with FACS analysis. The marker used is not believed to be important, so long as it is capable of being expressed simultaneously with the nucleic acid encoding a gene product. Further examples of selectable and screenable markers are well known to one of skill in the art.

B. Host Cells

As used herein, the terms “cell,” “cell line,” and “cell culture” may be used interchangeably. All of these terms also include their progeny, which refers to any and all subsequent generations. It is understood that all progeny may not be identical due to deliberate or inadvertent mutations. In the context of expressing a heterologous nucleic acid sequence, “host cell” refers to a prokaryotic or eukaryotic cell, and it includes any transformable organisms that are capable of replicating a vector and/or expressing a heterologous gene encoded by a vector. A host cell can, and has been, used as a recipient for vectors. A host cell may be “transfected” or “transformed,” which refers to a process by which exogenous nucleic acid is transferred or introduced into the host cell. A transformed cell includes the primary subject cell and its progeny.

Host cells may be derived from prokaryotes or eukaryotes, depending upon whether the desired result is replication of the vector, expression of part or all of the vector-encoded nucleic acid sequences, or production of infectious viral particles. Numerous cell lines and cultures are available for use as a host cell, and they can be obtained through the American Type Culture Collection (ATCC), which is an organization that serves as an archive for living cultures and genetic materials. An appropriate host can be determined by one of skill in the art based on the vector backbone and the desired result. A plasmid or cosmid, for example, can be introduced into a prokaryotic host cell for replication of many vectors. Bacterial cells used as host cells for vector replication and/or expression include DH15a, JM109, and KC8, as well as a number of commercially available bacterial hosts such as SURE® Competent Cells and SOLOPACK™ Gold Cells (STRATEGEN®, La Jolla). Alternatively, bacterial cells such as E. coli LE392 could be used as host cells for phage viruses.

Examples of eukaryotic host cells for replication and/or expression of a vector include HeLa, NIH3T3, Jurkat, 293, COS, C10, Saos, and PC12. Many host cells from various cell types and organisms are available and would be known to one of skill in the art. Similarly, a viral vector may be used in conjunction with either an eukaryotic or prokaryotic host cell, particularly one that is permissive for replication or expression of the vector.

C. Expression Systems

Numerous expression systems exist that comprise at least all or part of the compositions discussed above. Prokaryote- and/or eukaryote-based systems can be employed for use with the present invention to produce nucleic acid sequences, or their cognate polypeptides, proteins and peptides. Many such systems are commercially and widely available.

The insect cell/baculovirus system can produce a high level of protein expression of a heterologous nucleic acid segment, such as described in U.S. Pat. Nos. 5,871,886 and 4,879,236, both herein incorporated by reference, and which can be bought, for example, under the name MAXBAC® 2.0 from INVITROGEN® and BACPACK® Baculovirus Expression System from CLONTECH®.

Other examples of expression systems include STRATEGEN®'s COMPLETE CONTROL® Inducible Mammalian Expression System, which involves a synthetic edosynone-inducible receptor, or its PET Expression System, an E. coli expression system. Another example of an inducible expression system is available from INVITROGEN®, which carries the T-Rex® (tetracycline-regulated expression) System, an inducible mammalian expression system that uses the full-length CMV promoter. The Tet-On® and Tet-Off® systems from CLONTECH® can be used to regulate expression in a mammalian host using tetracycline or its derivatives. The implementation of these systems is described in Gossen et al., 1992 and Gossen et al., 1995, and U.S. Pat. No. 5,650,298, all of which are incorporated by reference.

INVITROGEN® also provides a yeast expression system called the Pichia methanolica Expression System, which is designed for high-level production of recombinant proteins in the methyloptrophic yeast Pichia methanolica. One of skill in the art would know how to express a vector, such as an expression construct, to produce a nucleic acid sequence or its cognate polypeptide, protein, or peptide.

D. Introduction of Nucleic Acids into Cells

In certain embodiments, a nucleic acid may be introduce into a cell in vitro for production of polypeptides or in vivo for immunization purposes. There are a number of ways in which nucleic acid molecules such as expression vectors may be introduced into cells. In certain embodiments of the invention, the expression vector comprises a Flavivirus infectious particle or engineered vector derived from a Flavivirus genome. In other embodiments, an expression vector known to one of skill in the art may be used to express a segment of a Flavivirus nucleic, which may be translated into a Flavivirus polypeptide or peptide. The ability of certain viruses to enter cells via receptor-mediated endocytosis, to integrate into host cell genome and express viral genes stably and efficiently have made them attractive candidates for the transfer of foreign genes into mammalian cells (Ridgeway, 1988; Nicolas and Rubenstein, 1988; Baichwal and Sugden, 1986; Temin, 1986).

"Viral expression vector" is meant to include those vectors containing sequences of that virus sufficient to (a) support
packaging of the vector and (b) to express a polynucleotide that has been cloned therein. In this context, expression may require that the gene product be synthesized. A number of such viral vectors have already been thoroughly researched, including adenovirus, adenov-associated viruses, retroviruses, herpesviruses, and vaccinia viruses.

Delivery may be accomplished in vitro, as in laboratory procedures for transforming cells lines, or in vivo or ex vivo, as in the treatment of certain disease states. One mechanism for delivery is via viral infection where the expression vector is encapsidated in an infectious viral particle. Several non-viral methods for the transfer of expression vectors into cultured mammalian cells also are contemplated by the present invention. These include calcium phosphate precipitation (Graham and Van Der Eb, 1973; Chen and Okayama, 1987; Rippe et al., 1990) DEAE-dextran (Gopal, 1985), electroporation (Tur-Kaspa et al., 1986; Poter et al., 1984), direct microinjection (Harland and Weinstein, 1985), DNA-loaded liposomes (Nicolaou and Sene, 1982; Fraley et al., 1979) and lipofectamine-DNA complexes, cell sonication (Fechheimer et al., 1987), gene bombardment using high velocity microprojectiles (Yang et al., 1990), liposome (Ghosn and Bachhawat, 1991; Kaneda et al., 1989) and receptor-mediated transfection (Wu and Wu, 1987; Wu and Wu, 1988). Some of these techniques may be successfully adapted for in vivo or ex vivo use.

In certain embodiments, the nucleic acid encoding a gene or genes may be stably integrated into the genome of the cell. This integration may be in the cognate location and orientation via homologous recombination (gene replacement) or it may be integrated in a random, non-specific location (gene augmentation). In yet further embodiments, the nucleic acid may be stably maintained in the cell as a separate, episomal segment of DNA. Such nucleic acid segments or “episomes” encode sequences sufficient to permit maintenance and replication independent of or in synchronization with the host cell cycle. How the expression vector is delivered to a cell and where in the cell the nucleic acid remains is dependent on the type of expression vector employed.

Transfer of a nucleic acid molecule may be performed by any of the methods mentioned above which physically or chemically permeabilize the cell membrane. This is particularly applicable for transfer in vitro, but it may be applied to in vivo use as well.

VI. Anti-HIV Therapies

In certain embodiments, therapeutic methods will include administering to a patient or subject a composition comprising an antigen or an antibody derived from a Flavivirus NS5A peptide or polypeptide, such as human or humanized animal derived antibodies. In various embodiments, the treatment methods of the invention may be used in combination with other anti-HIV treatments, such as Flaviviruses infection as a therapeutic or preventative treatment for AIDS. For exemplary compositions and methods see PCT application WO 01/77157, which is incorporated herein by reference.

As a therapeutic measure, a Flavivirus NS5A agent can be used to reduce the severity or progression of AIDS, including the prevention of AIDS in HIV-infected individuals. A reduction in severity or progression of AIDS includes, but is not limited to, prevention of or a reduction in the severity, duration, or discomfort associated with the following conditions: prolonged and unexplained fatigue; swollen glands; prolonged fever; chills; excessive sweating; swollen gums and mouth lesions; sore throat; cough; shortness of breath; constipation; diarrhea; symptoms of well-known opportunistic infections; Kaposi sarcoma; skin rashes or lesions; loss of appetite or weight loss; malaise; headaches; speech impair-
being most preferred. In some situations, it may be desirable to extend the time period for treatment significantly, however, where several days (2, 3, 4, 5, 6, or 7) to several weeks (1, 2, 3, 4, 5, 6, 7, or 8) lapse between the respective administrations.

It also is conceivable that more than one administration of a NS5A-based therapeutic agent will be desired. Various combinations may be employed, where a NS5A is “A” and the other agent is “B,” as exemplified below:

<table>
<thead>
<tr>
<th>A/B</th>
<th>A/B</th>
<th>B/A</th>
<th>B/A</th>
<th>A/A</th>
<th>B/A</th>
<th>A/B</th>
<th>B/B</th>
<th>A/B</th>
<th>B/B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A/B</td>
<td>A/A</td>
<td>B/A</td>
<td>A/A</td>
<td>A/B</td>
<td>A/B</td>
<td>A/B</td>
<td>A/B</td>
<td>B/B</td>
</tr>
</tbody>
</table>

Other combinations are contemplated as well.

A. AZT

A well known, traditional therapy for the treatment of AIDS involves zidovudine (AZT™ available from Burroughs Wellcome). This is one of a class of nucleoside analogues known as dideoxynucleosides which block HIV replication by inhibiting HIV reverse transcriptase. The anti-AIDS drug zidovudine (also known as AZT) may also be used in limited circumstances, mostly in combination with rifampin, as described by Burger et al. (1993).

The compositions and methods disclosed herein will be particularly effective in conjunction with other forms of therapy, such as AZT and/or protease inhibitors that are designed to inhibit viral replication, by maintaining desirable levels of white blood cells. Thus, in effect, buys the patient the time necessary for the anti-viral therapies to work.

B. HAART

New combination drug therapy has shown promising results in the treatment of HIV-infected patients. Treatment with potent anti-HIV drug combinations is referred to as “highly active anti-retroviral therapy” (HAART), and it has provided clinical improvement, longer survival, and improved quality of life for people infected with HIV during all four stages of HIV disease. Examples of HAART include a protease inhibitor (indinavir, nelfinavir, ritonavir, ritonavir/ saquinavir, or saquinavir) combined with two nucleoside analogs (AZT/ddI, d4T/ddC, AZT/ddC, AZT/3TC, or d4T/3TC).

In many instances, it will be desirable to have multiple administrations of the inventive compositions and/or a vaccine, usually not exceeding six administrations or vaccinations, more usually not exceeding four vaccinations. In certain embodiments, one or more, usually at least about three administrations or vaccinations may be provided. The administrations or vaccinations will normally be at from two to twelve weeks intervals, more usually from three to five week intervals. Periodic boosters at intervals of 1-5 years, usually three years, will be desirable to maintain protective levels of the antibodies. The course of the immunization or treatment may be followed by standard antibody assays. The assays may be performed by labeling with conventional labels, such as radionuclides, enzymes, fluorescent, and the like. These techniques are well known and may be found in a wide variety of patents, such as U.S. Pat. Nos. 3,791,952; 4,174,384 and 3,949,064, as illustrative of these types of assays.

The manner of application may be varied widely. Any of the conventional methods for administration of an antibody or vaccine are applicable. These are believed to include oral application on a solid physiologically acceptable base or a physiologically acceptable dispersion, parenterally, by injection or the like. The dosage of the NS5A agent will depend on the route of administration and will vary according to the size of the host.

The NS5A agents and flavivirus nucleic acids of the invention may be formulated into a pharmaceutically acceptable composition, see below, or vaccine as neutral or salt forms. Pharmaceutically acceptable salts include the acid addition salts (formed with the free amino groups of the peptide) and those that are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups may also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, 2-ethylamino ethanol, histidine, procaine, and the like.

The preparation of flavivirus NS5A agents as active ingredients is generally well understood in the art by analogy, as exemplified by U.S. Pat. Nos. 6,479,243, 6,399,763, 5,714,153, 5,582,981, and 4,833,077, all incorporated herein by reference. The preparation of vaccines that contain flavivirus sequences as active ingredients is generally well understood in the art by analogy, as exemplified by U.S. Pat. Nos. 5,958,895, 6,004,799, and 5,620,896, all incorporated herein by reference.

VIII. Pharmaceutical Compositions and Routes of Administration

Pharmaceutical compositions including NS5A peptides and polypeptides will be formulated along the line of typical pharmaceutical drug and biological preparations. A discussion of formulations may be found in Remington’s Pharmaceutical Sciences (1990). The percentage of active compound in any pharmaceutical preparation is dependent upon both the activity of the compound, in this case ability of NS5A agents to inhibit HIV replication. Typically, such compositions should contain at least 0.1% active compound. The percentage of the compositions and preparations may, of course, be varied and may conveniently be between about 2 to about 60% of the weight of the unit. The amount of active compounds in such therapeutically useful compositions is such that a suitable dosage will be obtained.

The pharmaceutical forms suitable for injectable use include sterile aqueous solutions or suspensions and sterile powders for the extemporaneous preparation of sterile injectable solutions or suspensions. In all cases, the form must be sterile and must be fluid to the extent that easy injection is possible. It must be stable under the conditions of manufacture and storage and must be preserved against the contamination of microorganisms, such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and vegetable oils. The proper fluidity can be maintained, for example, by the use of a coating, such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. The prevention of the action of microorganisms can be brought about by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic.
acid, thimerosal, phenylmercuric nitrate, m-cresol, and the like. In many cases, it will be preferable to use isotonic solutions, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate, and gelatin.

Sterile injectable solutions are prepared by incorporating the active compounds in the required amount in the appropriate solvent with various of the other ingredients enumerated above, as required, followed by sterile filtration. Generally, dispersions are prepared by incorporating the various sterilized active ingredients into a sterile vehicle which contains the basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preliminary step of preparing a sterile powder will be vacuum drying and freeze-drying techniques that yield a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

The phrases “pharmacologically acceptable” or “pharmaceutically acceptable” refer to molecular entities and compositions that do not produce an adverse, allergic, or other untoward reaction when administered to an animal, or human, as appropriate. As used herein, “pharmacologically acceptable carrier” includes any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like. The use of such media and agents for pharmaceutical active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active ingredients, its use in the therapeutic compositions is contemplated. Supplementary active ingredients, such as other anti-cancer agents, can also be incorporated into the compositions.

The active compounds of the present invention can be formulated for parenteral administration, e.g., formulated for injection via the intravenous, intramuscular, intrasynovial, subcutaneous, or even intraarteriopetal routes. Administration by i.v. or i.m. are specifically contemplated. Pharmaceutically acceptable salts, include the acid salts and those which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups may also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, 2-ethylamino ethanol, histidine, procaine, and the like.

In certain embodiments, it may be desirable to provide a continuous supply of therapeutic compositions to the patient. For intravenous or intraarteriopetal routes, this is accomplished by drip system. For various approaches, delayed release formulations could be used that provided limited but constant amounts of the therapeutic agent over and extended period of time. For internal application, continuous perfusion may be preferred. This could be accomplished by catheterization followed by continuous administration of the therapeutic agent. The time period for perfusion will be selected by the clinician for the particular patient and situation, but times could range from about 1-2 hours, to 2-6 hours, to about 6-10 hours, to about 10-24 hours, to about 1-2 days, to about 1-2 weeks or longer. Generally, the dose of the therapeutic composition via continuous perfusion will be equivalent to that given by single or multiple injections, adjusted for the period of time over which the injections are administered. It is believed that higher doses may be achieved via perfusion, however.

For parenteral administration in an aqueous solution, for example, the solution should be suitably buffered if necessary and the liquid diluent first rendered isotonic with sufficient saline or glucose. These particular aqueous solutions are especially suitable for intravenous, intramuscular, subcutaneous and intraperitoneal administration. In this connection, sterile aqueous media that can be employed will be known to those of skill in the art in light of the present disclosure. For example, one dosage could be dissolved in 1 mL of isotonic NaCl solution and either added to 1000 mL of hypodermoclysis fluid or injected at the proposed site of infusion, (see for example, Remington’s Pharmaceutical Sciences, 1990). Some variation in dosage will necessarily occur depending on the condition of the subject being treated. The person responsible for administration will, in any event, determine the appropriate dose for the individual subject.

An effective amount of the therapeutic composition is determined based on the intended goal. The term “unit dose” or “dosage” refers to physically discrete units suitable for use in a subject, each unit containing a predetermined quantity of the therapeutic composition calculated to produce the desired responses, discussed above, in association with its administration, i.e., the appropriate route and treatment regimen. The quantity to be administered, both according to number of treatments and unit dose, depends on the protection desired.

Peptides or polypeptides may be administered in a dose that can vary from 0.01, 0.05, 0.1, 0.5, 1, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 mg/kg of weight to 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200 mg/kg of weight in one or more daily, weekly, monthly, or yearly administrations during one or various days, weeks, months, or years. The antibodies can be administered by parenteral injection (intravenous, intraperitoneal, intramuscular, subcutaneous, intracavity or transdermic). For viral vectors, one generally will prepare a viral vector stock. Depending on the kind of virus and the titer attainable, one will deliver 1 to 100, 10 to 50, 100-1000, or up to 1 x 10^8, 1 x 10^9, 1 x 10^10, 1 x 10^11, 1 x 10^12 infectious particles to the patient. Similar figures may be extrapolated for liposomal or other non-viral formulations by comparing relative uptake efficiencies. Formulation as a pharmaceutically acceptable composition is discussed below.

In many instances, it will be desirable to have multiple administrations of the NS5A agent. The compositions of the invention may be administered 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more times. The administrations will normally be at from one to twelve week intervals, more usually from one to four week intervals. Periodic re-administration will be desirable with recurrent exposure to the pathogen (e.g., HIV). For example, an HIV positive mother would be re-inoculated prior to parturition from a second pregnancy.

Precise amounts and delivery regimen for the therapeutic composition also depend on the judgment of the practitioner and are peculiar to each individual. Factors affecting dose include physical and clinical state of the patient, the route of administration, the intended goal of treatment (amelioration of symptoms versus cure) and the potency, stability, and toxicity of the particular therapeutic substance.

In a particular embodiment of the invention, the NS5A agent may be entrapped in a liposome. Liposomes are vesicular structures characterized by a phospholipid bilayer membrane and an inner aqueous medium. Multilamellar liposomes have multiple lipid layers separated by aqueous medium. They form spontaneously when phospholipids are suspended in an excess of aqueous solution. The lipid components undergo self-rearrangement before the formation of
closed structures and entrap water and dissolved solutes between the lipid bilayers (Ghosh and Bachhawat, 1991).

Current in vivo lipid delivery methods use subcutaneous, intradermal, intratumoral, or intracranial injection to avoid the toxicity and stability problems associated with cationic lipids in the circulation. The DOTAP:cholesterol lipid formulation is said to form a unique structure termed a “sandwich liposome.” This formulation is reported to “sandwich” DNA between an invaginated bi-layer or “vase” structure. Beneficial characteristics of these lipid structures include a positive ρ, colloidal stabilization by cholesterol, two dimensional DNA packing and increased serum stability.

The production of lipid formulations often is accomplished by sonication or serial extrusion of liposomal mixtures after (I) reverse phase evaporation (II) dehydration-rehydration (III) detergent dialysis and (IV) thin film hydration. Once manufactured, lipid structures can be used to encapsulate compounds that are toxic (chemotherapeutics) or labile (nucleic acids) when in circulation. Lipid encapsulation has resulted in a lower toxicity and a longer serum half-life for such compounds (Gubizon et al., 1990).

Example 1

Cells: Tet-Off Jurkat cells (Clontech) were transfected with pTRE 2 Hyg plasmids containing full-length GBV-C NS5A (from the full-length GBV-C infectious clone AF121950) or with the vector only (control) by electroporation. Transfectants were selected using hygromycin, and cell lines cloned twice. Cells were grown in RPMI 1640 with or without doxycycline as recommended.

NS5A expression: Cells were lysed in RIPA buffer containing protease and phosphatase inhibitors and examined by immunoblot using rabbit anti-NS5A antisera.

HIV infections: Clinical and laboratory CXCR4 tropic strains of HIV-1 were used to infect Jurkat cell lines expressing NS5A or the vector control. Cells were maintained in doxycycline at various concentrations to inhibit NS5A expression. HIV replication was measured using culture supernatant p24 antigen ELISA.

Results: Jurkat cell lines stably expressing GBV-C NS5A protein were established, and expression of NS5A was regulated by doxycycline. HIV infection of cells expressing NS5A was decreased, as measured by p24 antigen compared to control cells or NS5A-containing cells grown in the presence of doxycycline. NS5A-expressing cells also had lower surface density of CXCR4 and released increased levels of the CXCR4 ligand SDF-1 into culture supernatants, which may account for observations. In addition, NS5A led to increased levels of two genes associated with resistance to apoptosis, and rendered cells resistant to Fas-mediated apoptosis. This latter observation may account for slower decline in CD4 cell counts observed in HIV-infected people.

Summary: Jurkat cell lines stably expressing GBV-C NS5A protein (regulated by doxycycline) were established (FIG. 1). HIV infection of Jurkat cells that expressed GBV-C NS5A resulted in decreased levels of HIV replication (measured by p24 antigen) compared to vector control cells or cells in which NS5A expression was reduced by doxycycline (FIGS. 2A-C). NS5A expressing cells had lower levels of CXCR4 on the surface than vector control cells, and also had reduced release of SDF-1 into culture supernatants, suggesting that NS5A expression may interact with chemokine receptor pathways (FIGS. 3A-B). Although part of the HIV inhibitory effect of GBV-C is mediated by E2 protein interactions with cellular receptor(s), the NS5A protein appears to also contribute to an anti-HIV effect of GBV-C (FIGS. 4A-B). GBV-C NS5A, like HCV, appears to have two phosphorylation states. HCV NS5A hyperphosphorylation has been reported by some to require NS4 expression. These data show that NS4 is not required for hyperphosphorylation of GBV-C NS5A (FIG. 5). NS5A led to relative resistance to Fas-mediated apoptosis, which could slow CD4 cell decline. HCV NS5A has previously been shown to be anti-apoptotic (FIG. 6). GBV-C infected PBMCs did not show decreased CXCR4 expression, although CCR5 expression was reduced. Because active GBV-C replication is required to produce NS5A, and active replication occurs in a small percentage of PBMCs, downregulation of CXCR4 would be missed. In contrast, E2 (structural protein) is released into supernatant and will have more widespread effects leading to decreased CCR5 (FIG. 7).

Example 2

NS5A Inhibition of HIV Replication is Dose Dependent, Independent of PKR and May Involve SDF-1

Using Jurkat cells that are infected with HIV, doxycycline-repressible expression constructs encoding GBV-C NS5A were used to inhibit HIV replication. As shown in FIG. 8, the use of 0.01, 0.1 and 1 μg of doxycycline produced decreasing amounts of HIV inhibition as measured by p24 antigen levels. No such effect was seen in control Jurkat cells.

The inventors previously showed that NS5A’s from interferon resistant (IFN-R) GBV-C inhibited PKR-mediated phosphorylation of eIF2α, which is part of the interferon cascade, whereas interferon sensitive (IFN-S) did not. The NS5A’s from IFN-R and IFN-S strains were compared for their respective abilities to inhibit HIV replication. As shown in FIG. 9, both strains were equally effective at inhibiting HIV replication as measured by p24 antigen concentration. Thus, the effect is apparently not related to PKR-function.

Following treatment with of HIV-infected Jurkat cells with NS5A, SDF-1 is released into culture supernatant (data not shown). In an effort to determine whether there is any link between SDF-1 and HIV inhibition, anti-SDF-1 antibodies were added to the Jurkat cell cultures. As shown in FIG. 10, there was an increase in HIV replication ranging from 50% to almost 60% as compared to untreated controls, thereby suggesting that SDF-1 might play a role.

Example 3

NS5A Inhibits Distinct Strains of HIV

Three different strains of HIV—ELI (X4) Clade D, JF (X4) Clade B and MN (X4) Clade B—were tested for replication in the presence of NS5A. As shown in FIGS. 11A-C, all were sensitive to NS5A-mediated inhibition.
GBV-C NS5A Fragments Inhibit HIV Replication

A variety of NS5A fragments were tested for their ability to inhibit HIV replication.

Methods. Stable Jurkat cell lines expressing PKR-inhibiting and non-inhibiting NS5A proteins were generated, as were a series of C-terminal deletion mutants (Fig. 12). These cell lines expressed NS5A proteins of 123, 161, 181, 250, 314, and 363 amino acids (aa). All constructs had stop codons after NS5A, followed by the EMC IRES and GFP. A control cell line expressing only GFP served as a negative control, and NS5A and GFP expression were regulated by tetracycline. HIV replication was measured by p24 antigen release into culture supernatant or by measuring infectivity.

Results. Jurkat cell lines demonstrated regulated expression of NS5A and deletion mutants as shown by western blot and GFP expression. The NS5A and GFP genes were shown to remain linked by RT-PCR of cellular DNA from recombinant cell lines. Expression of either PKR-inhibiting or non-inhibiting NS5A proteins resulted in HIV inhibition (>95% reduction in p24 antigen), thus the HIV- and PKR-inhibiting functions are independent. All deletion mutants of 250 aa or larger retained HIV-inhibiting effects, whereas those with 181 aa or smaller did not (Fig. 13).

Example 5

HCV and GBV-B Virus NS5A Proteins Inhibit HIV Replication in Jurkat Cells

Hepatitis C virus (HCV), like GBV-C, commonly infect humans. GBV-B is a primate virus that is closely related to both HCV and GBV-C. HCV and GBV-C NS5A proteins both inhibit PKR function. To further characterize the in vitro works showing inhibition of X4-tropic HIV replication, they expressed HCV and GBV-B NS5A's and a series of GBV-C NS5A deletions in Jurkat cells, and measured the effect of these proteins on HIV replication.

Methods. Jurkat cell lines stably expressing HCV, GBV-B, and GBV-C NS5A proteins were generated, as were a series of GBV-C NS5A deletion mutants. All constructs had stop codons after NS5A, followed by the EMC IRES and GFP. A control cell line expressing only GFP served as a negative control, and NS5A and GFP expression were regulated by tetracycline. Plasmids were transfected using the AmoXa nucleofection method; cells were selected for growth in hygromycin and for GFP expression. HIV replication was measured by p24 antigen release into culture supernatant or by measuring infectivity.

Results. Jurkat cell lines demonstrated regulated expression of HCV, GBV-B GBV-C NS5A proteins and GBV-C deletion mutants as determined by GFP expression and when antibodies were available, by immunoblot (Fig. 14). For constructs in which antibodies were not available, the NS5A and GFP coding sequences were linked (detected by PCR of cellular DNA). Expression of HCV, GBV-B and GBV-C NS5A proteins resulted in HIV inhibition (>95% reduction in p24 antigen), thus the inhibitory effect appears conserved between these three flaviviruses (Fig. 15). All deletion mutants containing GBV-C NS5A amino acids (aa) between number 152 and 237 retained HIV-inhibiting effects, whereas those with C-terminal deletions containing ≤152 NS5A aa's did not inhibit HIV replication (Fig. 16). The effect was reduced in NS5A containing cells grown in doxycycline (NS5A expression turned off), but not in control cells grown in doxycycline. Thus, expression of GBV-C, GBV-B, and HCV NS5A proteins resulted in inhibition of CXCR4-tropic HIV replication in Jurkat cells, and the inhibitory effect requires GBV-C aa's 152-237, coinciding with domain II of the HCV NS5A protein. The region may possibly narrowed to 152-182 (Fig. 17).

All of the compositions and methods disclosed and claimed herein can be made and executed without undue experimentation in light of the present disclosure. While the compositions and methods of this invention have been described in terms of preferred embodiments, it will be apparent to those of skill in the art that variations may be applied to the compositions and methods and in the steps or in the sequence of steps of the method described herein without departing from the concept, spirit and scope of the invention. More specifically, it will be apparent that certain agents that are both chemically and physiologically related may be substituted for the agents described herein while the same or similar results would be achieved. All such similar substituents and modifications apparent to those skilled in the art are deemed to be within the spirit, scope and concept of the invention as defined by the appended claims.

REFERENCES

The following references, to the extent that they provide exemplary procedural or other details supplementary to those set forth herein, are specifically incorporated herein by reference.

U.S. Pat. No. 3,791,932
U.S. Pat. No. 3,949,064
U.S. Pat. No. 4,174,384
U.S. Pat. No. 4,554,101
U.S. Pat. No. 4,683,202
U.S. Pat. No. 4,833,077
U.S. Pat. No. 4,879,236
U.S. Pat. No. 5,582,981
U.S. Pat. No. 5,620,896
U.S. Pat. No. 5,650,298
U.S. Pat. No. 5,714,153
U.S. Pat. No. 5,871,986
U.S. Pat. No. 5,874,563
U.S. Pat. No. 5,925,565
U.S. Pat. No. 5,928,906
U.S. Pat. No. 5,935,819
U.S. Pat. No. 5,938,895
U.S. Pat. No. 6,004,799
U.S. Pat. No. 6,399,763
U.S. Pat. No. 6,649,243
PCT Appl. WO 01/77157
PCT Appl. PCT/US03/33925
PCT Appl. PCT/US2004/017706
Angel et al., Cell, 49:729, 1987b.
<400> SEQUENCE: 1

ccccccccc gcacctggtt gcaagccccca gaaccccagc cctactgaaag tagacgtaat 60
 ggcccgccgc gcagacggcc ccgcggcgc cgggcgc gagc ggtgtggtgta caggggttggt 120
 aggtctaa agctgctcagt cctgtcgaccc actataggtg gtccttaagg gcagggctagc 180
 gtcgctcttgg gcagatagga gggagggcgc cggtgctgggtgctcctgggtggt 240
 aactaaggccc cggcgtagct caaccgctgcatt aaccggcgc gttactcccc tgccggcagc 300
 acggcagagc ccgggctcgc ctgoccttcgta aagagctctcttc gcagcccaagtg ggtctcgag 360
 ccggttgacac gcggagagcgc ccgggggggg gcgggggggg gcgggggggg gcgggggggg 420
 tcgggaggag gcggagaatt gtcggggcga cccagctgcc gcggggccta cccagggcgtg 480
 agcggcaaga agcggaaaggt ggccggggag gcagggggtt gctatttttt tctataacgt atcg 536

Met

1

gca gto ctt ctc cta ctc gtc gtt gga gag gcc ggg gct att tta gcc
Ala Val Leu Leu Leu Leu Leu Val Val Glu Ala Gly Ala Ile Leu Ala
5 10 15

cgc gcc acc acc cat gct gtt agc ggc aaa gaa caa tat ttc ctc aca aac
Pro Ala Thr His Ala Cys Ser Ala Lys Gly Gin Tyr Xaa Leu Thr Aam
20 25 30

tgt gtc gcc ctc gga gac ata ggc ttc tgc gta gaa gcc ggg tgc ctg
Cys Cys Ala Leu Ala Asp Ile Gly Phe Cys Leu Glu Gly Gly Leu
35 40 45

gtc gct cgg ggc acc att tgc acc gac cgg ctc tgg cca ctg tat
Val Ala Leu Gly Cys Thr Ile Cys Thr Arg Cys Trp Pro Leu Tyr
50 55 60 65

cag gcc gat tgg gcc gtt ggc ccc ggc aag tcc gcc ggc cag tgg gcc
Gln Ala Gly Leu Ala Val Pro Gly Lys Ser Ala Ala Ala Gin Leu Val
70 75 80

ggg gaa ctc ggt agt ctc tac ggg ctc tca gtc cgg gct gct gct ggc tac
gly Glu Leu Gly Ser Leu Tyr Gly Pro Leu Ser Val Ser Tyr Val
85 90 95

gcc ggg atc cgt cgg cgt ggg ctc tac cgg tgg ggc gtc ctg acc gtc
Ala Gly Ile Leu Leu Gly Val Val Gin Tyr Ser Gly Val Leu Thr Val
100 105 110

ggg gtt gcc ccc ccc gcc gtc ctc gtt ggc ctc cgg gct gcc aac ctg acy
gly Glu Leu Ala Thr Arg Val Arg Val Tyro Pro Val Pro Aam Leu Thr Cys
115 120 125

gca gta gag tgt gta gaa ggt gaa gtt gta gaa aag act
Ala Val Glu Cys Glu Leu Lys Try Glu Ser Phe Thr Arg Arg
130 135 140 145

ga cag ctc gcc tca aac tac cag att ctt cag ctc tgg aag ctc
Glu Gin Leu Ala Ser Aam Tyr Thr Ile Leu Glu Tyr Leu Thr Val
150 155 160

cct ttc gcc ttt tgg cgg gga gtt agt aac ctt act cct ctc tgg gtc
Pro Pro Arg Phe Thr Arg Val Met Ser Leu Thr Pro Leu Leu Val
165 170 175

tgc gtt gcc gtc ctc ctc gtc gtt gca gat cgt att gtc atg ctt tgc
Cys Val Ala Ala Leu Leu Leu Leu Glu Arg Ile Val Met Val Phe
180 185 190

cct cgg ctc ctc ctc ctc cgc cgg tgt cgg gct cgg cgg cgc tca gtc
Leu Leu Val Met Thr Met Ser Gin Glu Gin Arg Pro Ala Ser Val
195 200 205

tgg ggg cta cgg cct tgc gag gcc ggg tgg act tgt cag tct
tgg Leu Gly Ser Arg Pro Phe Glu Ala Gly Leu Thr Trp Gin Ser Cys
210 215 220 225

tgc acc tgc cgg cgc ctc ggc gcc agc ggg gaa agg gtt tgg gaa
Cys Arg Ser Aam Gly Ser Arg Val Pro Thr Gly Glu Arg Val Trp Glu
230 235 240

1256
<table>
<thead>
<tr>
<th>230</th>
<th>235</th>
<th>240</th>
</tr>
</thead>
<tbody>
<tr>
<td>cgt ggg aac gtc aca ctt ttg tgg gac tgc ccc aac ggt cct tgg gtc</td>
<td>245</td>
<td>250</td>
</tr>
<tr>
<td>Arg Gly Aen Val Thr Leu Leu Cys Asp Cys Pro Aem Gly Pro Trp Val</td>
<td>255</td>
<td></td>
</tr>
<tr>
<td>tgg gtc ccg gcc cct ttg cag gca atc gga tgg ggc gac cct atc act</td>
<td>260</td>
<td>265</td>
</tr>
<tr>
<td>Trp Val Pro Ala Leu Cys Glu Ala Ile Gly Trp Gly Arg Pro Ile Thr</td>
<td>270</td>
<td></td>
</tr>
<tr>
<td>cat tgg agc cac gga caa aat cag tgg ccc ttc tgt ccc caa ttg</td>
<td>275</td>
<td>280</td>
</tr>
<tr>
<td>His Trp Ser His Gly Glu Asn Gly Trp Pro Leu Ser Cys Pro Glu Phe</td>
<td>285</td>
<td></td>
</tr>
<tr>
<td>gtc tac ggc gcc gtt tca gtt acc tgc gtt ggg gct gtt gtc tgg val tyr gly ala val ser val thr val cys val trp gly ser trp ser trp</td>
<td>290</td>
<td>295</td>
</tr>
<tr>
<td>Val Tyr Gly Ala Val Ser Val Thr Cys Val Gly Ser Gly Ser Val Ser Trp</td>
<td>300</td>
<td></td>
</tr>
<tr>
<td>ttc gct tcc act ggg gtt ggc gtc ccc aag gtt gat gtt ggc gtt gtt tgg</td>
<td>305</td>
<td>310</td>
</tr>
<tr>
<td>Phe Ala Thr Ser His Gly Arg Arg Ser Lys Val Arg Val Trp Ser Leu</td>
<td>315</td>
<td></td>
</tr>
<tr>
<td>gtt cca gtt ggc tct gcc agc tgc acc ata ggc gca ctc gga gta tct tgg</td>
<td>320</td>
<td>325</td>
</tr>
<tr>
<td>Val Pro Val Gly Ser Ala Ser Cys Thr Ile Ala Leu Gly Ser Ser</td>
<td>330</td>
<td></td>
</tr>
<tr>
<td>gat cgg gac aca ctc gtg gtt gaa ccc tcc ggc gac cct cag gag gga att ccc tgc asp arg aen thr val val glu leu ser glu trp gly ile pro cys ala</td>
<td>335</td>
<td>340</td>
</tr>
<tr>
<td>345</td>
<td></td>
<td></td>
</tr>
<tr>
<td>act tgt atc ctc gac aag cgg cct gcc tgg tgt ggc acc tgt gtt aag</td>
<td>350</td>
<td>355</td>
</tr>
<tr>
<td>Thr Cys Ile Leu Asp Arg Pro Arg Pro Cys Gly Thr Cys Val Arg</td>
<td>360</td>
<td></td>
</tr>
<tr>
<td>gac tgc tgt ccc gag acc gcc gcc gtc gta cgt ttc gcc ctc cac agg tgt</td>
<td>365</td>
<td>370</td>
</tr>
<tr>
<td>Asp Cys Pro Gly Thr Gly Ser Val Arg Arg Pro Phe Pro His Arg Cys</td>
<td>375</td>
<td></td>
</tr>
<tr>
<td>380</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ggc gcc gca ccg agg ctc acc aag gac ctt gag gct gtt ggg ccc ttc gtc</td>
<td>385</td>
<td>390</td>
</tr>
<tr>
<td>Gly Ala Gly Pro Arg Leu Thr Arg Asp Leu Glu Ala Val Pro Phe Val</td>
<td>395</td>
<td></td>
</tr>
<tr>
<td>aat agg aca act ccc ttc acc ata agg ggg ccc ctc gcc gac cag ggg</td>
<td>400</td>
<td>405</td>
</tr>
<tr>
<td>Ann Arg Thr Thr Pro Phe Thr Ile Arg Gly Pro Leu Gly Ann Gly Glu</td>
<td>410</td>
<td></td>
</tr>
<tr>
<td>415</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cga ggc aac ccc ctc gtt ccc ccc tgg ctt ttt ggg ccc tac acc atg</td>
<td>420</td>
<td>425</td>
</tr>
<tr>
<td>Arg Gly Aen Pro Val Arg Ser Gly Gly Phe Gly Phe Ser Tyr Thr Met</td>
<td>430</td>
<td></td>
</tr>
<tr>
<td>435</td>
<td></td>
<td></td>
</tr>
<tr>
<td>acc aag atc cga gac ccc tta ccc ttt gtt gaa aat tgt ccc acc cca gcc</td>
<td>440</td>
<td>445</td>
</tr>
<tr>
<td>Thr Lys Ile Arg Asp Ser Leu Leu Leu Gly Cys Pro Thr Pro Ala</td>
<td>450</td>
<td></td>
</tr>
<tr>
<td>455</td>
<td></td>
<td></td>
</tr>
<tr>
<td>att gag ctt ccc acc agg acg tgg tgg ttc cca gga gtc gcc ccc ccc</td>
<td>460</td>
<td>465</td>
</tr>
<tr>
<td>Ile Glu Pro Pro Thr Gly Thr Phe Gly Phe Gly Pro Gly Val Pro Pro</td>
<td>470</td>
<td></td>
</tr>
<tr>
<td>475</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ctt aac aac tgc atg ctt ctc gcc act gag gtt tca gag gta tgt ggt</td>
<td>480</td>
<td>485</td>
</tr>
<tr>
<td>Leu Asn Asp Leu Met Leu Leu Leu Gly Thr Gly Ser Val Ser Gly Val Gly</td>
<td>490</td>
<td></td>
</tr>
<tr>
<td>ggg ggc ggc ctc act ggg ggg ttt tac gas cct ctt ctt ctt ggg cgg tgt</td>
<td>495</td>
<td>500</td>
</tr>
<tr>
<td>Gly Ala Gly Leu Thr Gly Gly Phe Thr Gly Pro Leu Arg Arg Cys</td>
<td>505</td>
<td></td>
</tr>
<tr>
<td>510</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tca gag ctt atg ggt cgg cag aat ccc tgc gcc cgg ggg ttc gca tgg</td>
<td>515</td>
<td>520</td>
</tr>
<tr>
<td>Ser Glu Leu Met Gly Arg Arg Arg Pro Val Cys Pro Gly Phe Ala Trp</td>
<td>525</td>
<td></td>
</tr>
<tr>
<td>ctc tgt ctc gaa cgg cct gat ggg ttt atc cta cag gag gcc cac tgg</td>
<td>530</td>
<td>535</td>
</tr>
<tr>
<td>Leu Ser Gly Arg Pro Arg Gly Phe Ile His His Val Glu Gly His Leu</td>
<td>540</td>
<td></td>
</tr>
<tr>
<td>545</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cag gag gty gat ggc ggc aac ttc att ccc ccc cca cgc tgg tgt ctc</td>
<td>550</td>
<td>555</td>
</tr>
<tr>
<td>Gly Glu Val Arg Ala Asp Ile Pro Pro Arg Pro Trp Leu Leu</td>
<td>560</td>
<td></td>
</tr>
<tr>
<td>ttc gac ttt gta gta gtc ctt ttc tac ctt atg aag ctt gcc gag gca</td>
<td>565</td>
<td>570</td>
</tr>
<tr>
<td>Leu Asp Phe Val Phe Val Leu Leu Tyr Leu Met Lys Leu Ala Glu Ala</td>
<td>575</td>
<td>580</td>
</tr>
</tbody>
</table>
cgg tgg gtc ccg ctc atc ctc ctc ctc taa tgg tgg gtt gaa cag 
Arg Leu Val Pro Leu Ile Leu Leu Leu Leu Leu Leu Val Val Gly 
550 555 560 2264

htg ggc gtc ctt ggt gtg ggc ctc aam cgc ctc gaa cag 
Leu Ala Val Leu Xaa Val Xaa Ala Ala Val Val Val Gly Gly 
551 556 561 2311

ctg ttt gct ctc cct ctc tgg ttc ctc gtc gta cta ccc ttc 
Val Phe Ala Gly Pro Ala Leu Ser Trp Cys Leu Gly Leu Pro Phe Val 
562 567 572 2360

agt atg atc ctc ggg ctc gtc tgg tgg tac ttc cgc tgg atg 
Ser Met Ile Leu Gly Leu Ala Arg Val Val Val Tyr Arg Trp Met 
573 578 583 2408

agt ctc cca cgc ctc ggc att tcc cgc act cgc ggc cgc 
Gly Pro Gln Arg Leu Met Phe Leu Val Leu Leu Ala Gly Gly 
584 589 594 2456

ctg ttc ccc ctc gca tta ctc gtt ggg att tcc cgc act cgc ggc cgc 
Ala Phe Pro Leu Ala Leu Met Gly Ile Ser Ala Thr Arg Gly Arg 
595 600 605 2504

acc ttc ggt ctt ggg cct cgc gtt gag tgc acc tcc gaa atg 
Thr Ser Val Leu Gly Ala Glu Phe Cys Phe Asp Val Thr Phe Glu Val 
606 611 616 2552

gac acg tca gtc ttc ggt gtt gtt gct gta gtc gtt gct cgc 
Asp Thr Ser Val Leu Gly Trp Val Val Val Leu Leu Trp Ala 
617 622 627 2600

ata cgc ctc ctc agc ttc atg aag cgc ggg gtt cag ctt gtc 
Ile Ala Leu Leu Ser Ser Met Ser Ala Gly Gly Trp Lys His Lys Ala 
628 633 638 2648

ata atc tac agg cgc tgt ctt cag ctt cgc cag cgc 
Ile Ile Tyr Arg Thr Tyr Gly Gly Xaa Leu Arg Arg Arg 
639 644 649 2696

gtg gtg ctt ggt gtt gct gtt gct gat tgt gtc gtt ggt ggc 
Val Val Val Gly Glu Gly Arg Pro Arg Pro Thr Lys Pro Leu Thr 
650 655 660 2744

ata gcc tct tgt ggc tct tac atc tgg cag gac gct gtt atg tgt 
Ile Ala Pro Cys Leu Ala Ser Tyr Ile Trp Pro Asp Val Met Leu 
661 666 671 2792

gtg gtt gtt ggc atg gtt ctc ctc ctc ctc tgg ggc ttc ttc 
Val Val Val Met Val Val Leu Leu Phe Gly Leu Phe Asp Ala Leu Asp 
672 677 682 2840

tgg gcc tgt gag ggc ttc ctt cgg ctc gct ctc ttc 
Val Ala Leu Glu Leu Val Leu Val Ser Arg Ser Leu Arg Arg 
683 688 693 2888

gcc agg tgt ggt ggt gat tgt gtt gct ggg gac gag cag 
Ala Arg Val Val Gly Cys Val Met Ala Gly Gly Lys Ala Thr 
694 699 704 2936

gtc gtt cgg ctc ggg tgg ggg ggc cgg ctc ctc ggc 
Val Arg Leu Val Ser Lys Met Cys Ala Arg Gly Ala Tyr Leu Phe Asp 
705 710 715 2984

cac cag gtg ctt ggg ctt ctc ggc ggc ctc aam cgc ttc tgg tgg 
His Met Gly Ser Phe Ser Arg Ala Val Lys Glu Arg Leu Leu Trp 
716 721 726 3032

gac ggt ggt gct tgg tgg cgg aag ggg goc tat ctc tgg 
Asp Ala Ala Leu Lys Leu Ser Phe Thr Arg Thr Asp Cys Arg Ile 
727 732 737 3080

ata cga gcc ggc cgg gcc acc ctc agc tgc gcc cgg cgg 
Ile Arg Asp Ala Ala Arg Arg Cys Cys Gly Cys Val Met Gly 
738 743 748 3128

ttg ccc cgg gtc ggt ggt gat ggt cgg cgc gat ggg ttc 
Val Pro Val Val Ala Arg Glu Glu Leu Ile Gly Val Phe 
749 754 759 3176
cag gat gtg anc cac tgt gct ccg ggg ttt gyt cct aca ggc cct gtt
Gln Asp Val Amn His Leu Pro Pro Gly Phe Xaa Pro Thr Ala Pro Val 885 890 895 3224

gtc atc gct ggg tgg tgc gaa aag ggc ttc ctc ggg gtc act aag gct gcc
Val Ile Arg Arg Cys Gly Lys Gly Phe Leu Gly Val Thr Lys Ala Ala 900 905 910 3272

ttg act ggt ggg gat cct gac tta cac cca gga aac gtc atg gtt tgg
Leu Thr Gly Arg Asp Pro Arg His Pro Gly Val Met Val Leu 915 920 925 3320

ggg acg got acc tcg cgc agc atg gga acg tgc tta aac ggg tgg ctc
Gly Thr Ala Thr Ser Arg Ser Met Gly Thr Cys Leu Amn Gly Leu Leu 930 935 940 945 3368

ttc acg cca atc cat ggc gtt ctt tcc cga acc att gcc aca cct gtc
Phe Thr Pro Gly Ala Ser Ser Arg Thr Ile Ala Thr Pro Val 950 955 960 3416

ggg gco ctt aac cca agg tgg tgg tcg gcc gat gac gtc agc gcc
Gly Ala Leu Amn Pro Arg Asp Arg Ser Asp Thr Val Thr 965 970 975 3464

tat ccc ctc ccc gat gga gct aac tgg ctc tgg ctc aac gac aag gcc
Tyr Pro Pro Pro Gly Ala Pro Leu Ser Val Pro Cys Ser Gln 980 985 990 3512

gct gag tcc tgt gtc aty cga tcc gat ggg gct ctt tgt tgc cat gcc
Ala Glu Ser Cys Trp Val Val Ser Asp Ser Arg Ala Leu Cys His Gly 995 1000 1005 3560

ttg acg aag ggg gac aag gta gag gta cgg gtc ggt goc atg gag gtt gct
Leu Ser Gly Gly Asp Lys Val Leu Amn Leu Met Gly Val Ala 1010 1015 1020 1025 3608

gac ttt cgt ggg tgg tct ggg tct gtc cta tgc gac gsg gac cac
Amp Phe Arg Gly Ser Ser Gly Pro Val Leu Cys Arg Gly His 1030 1035 1040 3656

gct gta gga atg ctt gtt ctg tgt ctc ctt cat tgt gg aac ggg ggt gcc
Ala Val Gly Met Leu Val Ser Val Leu His Ser Gly Gly Arg Thr 1040 1045 1050 1055 3704

gcg gct cga ttc act cgg ccc tgg acc cca aca gtc cca aca gac gcc
Ala Ala Arg Phe Thr Arg Pro Thr Gln Val Pro Arg Ala Lys 1060 1065 1070 3752

cct ccc tgt cct ctc aag gcc ggg ggg aag gcc aag cca gcc tct
Thr Thr Glu Pro Pro Pro Ala Ala Gly Lys Ser Thr Gly Val Pro 1075 1080 1085 3800

gct cct ctt ctc atg cca aag ggg gcc ggg aag gcc aca cag gcc tct
Ala Pro Leu Phe Pro Thr Gly Ala Gly Lys Ser Thr Gly Val Pro 1090 1095 1100 1105 3848

ttg gag tat gga aac atg ggg cac aac ggt ctc atg att ctc aac cgg tgg
Leu Glu Tyr Gly Amn Met Gly His Lys Val Leu Ile Leu Amn Pro Ser 1110 1115 1120 3896

gtt gcc act gtc agg gcc atg ggc ctt tac atg gag agg ctg gcc ggg
Val Ala Thr Val Arg Ala Met Gly Phe Met Arg Leu Ala Gly 1125 1130 1135 3944

aaa cat cct aag att ttc tgt gga cac gac aca aca gct ttc aca cgg
Lys His Pro Ser Ile Phe Cys Gly His Arg Thr Thr Ala Phe Arg 1140 1145 1150 3992

atc aag gac tct cca tgg ctc aag ctc cct aag ggg ttt cag gcc
Ile Thr Pro Ser Pro Leu Thr Tyr Ser Thr Gly Arg Phe Leu Ala 1155 1160 1165 4040

aac cgy agg cag atg ctt gga gtt tcc gtc atc tgg gat gag
Amp Pro Arg Gln Met Leu Arg Gly Val Ser Val Ile Cys Asp Glu 1170 1175 1180 1185 4088

tgc cac aat cat gac tca act gtt tgt ctt ggt ata ggc agg gtc egg
Cys His Ser His Asp Ser Thr Val Leu Leu Gly Ile Gly Arg Val Arg 4136
gac gtc gcc cgg ggg tgt gga gtt ctc tac act gct gac
Amp Val Ala Arg Gly Cys Gly Val Gln Leu Val Leu Tyr Ala Thr Ala
1190 1195 1200
1205 1210 1215

act ccc cgc tgg cct atg act cag cat cca tcc ata att gac aca
Thr Pro Pro Gly Ser Pro Met Thr Gln His Pro Ser Ile Ile Glu Thr
1220 1225 1230

aag ctc gac gtt ggt gag att ctc ttc tga ttc cat tcc aag ggc
Lys Leu Asp Val Gly Ile Pro Phe Tyr Gly His Gly Ile Pro Leu
1235 1240 1245

1250 1255 1260

gag tgt gaa gaa ggt aag agt ggt cgc cac ctt gta ttc tgc cat tcc aag ggc
Glu Arg Met Arg Thr Gln Arg His Leu Val Phe Cys His Ser Lys Ala
1270 1275 1280

1285 1290 1295

gag tgt gac ggt ggt gac atg gcc cag ttc tcc ggg gtw aat gcc
Glu Cys Glu Arg Leu Ala Gly Phe Ser Ala Arg Gly Val Asn Ala
1300 1305 1310

1315 1320 1325

att gcc tat tgg ggt gaa gac agt tcc act atc ata aaa gac gga gac
Ile Ala Tyr Tyr Arg Gly Lys Ser Arg Ser Ile Lys Arg Gly Asp
1330 1335 1340

1345 1350 1355

cgg gtc gtt tgt ggc cac cag ctc tcc ggg tac aca gag aac
Leu Val Val Cys Ala Thr Asp Ala Leu Ser Thr Gly Tyr Gly Arg
1360 1365 1370

1375 1380 1385

tcc gat ctc gcc acc gac tgt tgc tgg gtt gac gaa gac gtt gac
Phe Asp Ser Val Thr Asp Cys Leu Val Gly Leu Val Gly Val Glu
1390 1395 1400

1405 1410 1415

gtg acc ctt gat ccc acc att acc att tcc tgt cgg act gtc cct gct
Val Thr Asp Pro Thr Ile Thr Ile Ser Leu Arg Thr Val Pro Ala
1420 1425 1430

1435 1440 1445

tcg gtg gaa ttc tgg cat cgg cgc gga cgc cgg aga ggt cgg
Ser Ala Glu Leu Arg Ser Met Gly Arg Gly Arg Thr Gly Arg
1450 1455 1460

1465 1470 1475

tcg gcc cgc tac tac tac gct gcc ggt aag gct ccc gcc ggg ggt
glu Arg Tyr Tyr Tyr Ala Gly Val Gly Lys Ala Pro Ala Gly Val
1480 1485 1490

1495 1500 1505

gtg gcc tct gct gcc gcc gtc gtt ggg cgg gga gtt act gcc
Glu Val Arg Ser Pro Thr Ser Ala Gly Val Gly Val Thr Trp
1510 1515 1520

1525 1530 1535

tat gaa atg gaa cct gac tgt gcc aca acc ctt gta acc ctt ccc
tg ctc cct acc aca gcc gtc gcc gct gcc gct gcc act gcc
Tyr Pro Met Glu Pro Asp Leu Thr Ala Asn Leu Leu Arg Leu Tyr Asp
1540 1545 1550

1555 1560 1565

gtc ggt ggg ggt gcc gcc gcc gcc gct gcc gct gcc gat ggt gac
glu Arg Cys Pro Tyr Thr Ala Val Ala Ala Ala Asp Ile Gly Glu Ala Ala
1570 1575 1580

1585 1590 1595

gtc ctc ccc gcc gcc gcc gcc gcc gcc gct gcc gcc gcc gct ggc
Val Phe Phe Ala Gly Ala Pro Leu Arg Met His Pro Asp Ser Val Ser
1600 1605 1610

1615 1620 1625

tgc ggc aac gtt ggc gcc ggc gcc gcc gcc gcc gct gcc gcc gcc gcc gcc
Arg Val Ala Arg Gly Val Arg Leu Arg Leu Cys Gly Glu Val Val Val
1630 1635 1640

1645 1650 1655

cgg cag atg tgt cgg gaa aca ctc gtc ccc gcc cgg gcc gcc gcc gcc gcc gcc
Arg Thr Met Cys Arg Gly Thr Leu Ser Pro Gly Pro Ser Asp Pro
1660 1665 1670

1675 1680 1685

cgg gcc gcc gcc ggc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc ggc
Glu Trp Ala Ala Gly Leu Gly Pro Asp Pro Val Pro Leu Leu Leu Arg
1690 1695 1700

1705 1710 1715

tgc ttc ccc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc
Val Ala Arg Arg Leu Arg Leu Leu Ala Gly Val Val Val Val Val Val
1720 1725 1730

1735 1740 1745

cgc ggg gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc
Glu Gly Val Val Val Val Val Val Val Val Val Val Val Val Val Val Val
1745 1750 1755

1760 1765 1770

tgc ccc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc
Val Val Val Val Val Val Val Val Val Val Val Val Val Val Val Val Val
1775 1780 1785

1790 1795 1800

tgc ggg gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc
Val Val Val Val Val Val Val Val Val Val Val Val Val Val Val Val Val
1805 1810 1815

1820 1825 1830

cgt gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc
Glu Val Val Val Val Val Val Val Val Val Val Val Val Val Val Val Val Val
1835 1840 1845

1850 1855 1860

tgc ccc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc
Val Val Val Val Val Val Val Val Val Val Val Val Val Val Val Val Val
1865 1870 1875

1880 1885 1890

cgc ggg gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc
1895 1900 1905

1910 1915 1920

tgc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc
Val Val Val Val Val Val Val Val Val Val Val Val Val Val Val Val Val
1925 1930 1935

1940 1945 1950

cgt gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc
Glu Val Val Val Val Val Val Val Val Val Val Val Val Val Val Val Val Val
1955 1960 1965


tgc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc gcc
Val Val Val Val Val Val Val Val Val Val Val Val Val Val Val Val Val
1985 1990 1995

2000 2005 2010


1510  
aaa  
---continued---  
1515  

gct ggr ccc atc ctc atg ggc  
   tg cgc ata gcg ggc atg atc  
   Ala Xaa Pro Ile Leu Met Val Gly  
   Leu Ala Ile Ala Gly Gly Met Ile  
   1520  
   1525  
   1530  
   1535  

tac gcc tct tac act ggg tcc cta  
   gtt gtg ata gca gac tgg gat gtg  
   Tyr Ala Ser Tyr Thr Gly Ser Leu Val thr Thr Asp Asp Val  
   1540  
   1545  
   1550  
   1555  

aag gga ggt ggc aat ccc ccc  
   ctt tat agg agt ggt gac cag gcc acc cct  
   Lys Gly Gly Gly Arg Pro Leu Tyr Arg Ser Gly Asp Gin Ala Thr Pro  
   1560  
   1565  

caaa ccg gtc ctc ccc cgg gta  
   gac cat cgg ccc ggg ggg gag  
   Gin Pro Val Val Gin Val Pro Val Asp Arg Gly Gly Glu  
   1570  
   1575  
   1580  
   1585  

tct gcc cca ggg gat ggc aag aca  
   gta gcc gac gct ggc gca gcc act  
   Ser Ala Pro Arg Asp Ala Lys Thr Val Thr Asp Ala Val Ala Ile  
   1590  
   1595  
   1600  

cag gtg aac tgc gat tgg tct  
   gtt atg acc ctc tgt gcg atc ggg gaa gtc  
   Gin Val Aam Cys Asp Thr Ser Thr Ser Leu Ile Gly Gly Val  
   1605  
   1610  
   1615  

cct acc tgt ggt cag gct  
   aag aca gcc gac ggc tac gca gct act tcc  
   Leu Thr Leu Ala Gin Ala Lys Thr Ala Glu Ala Tyr Ala Thr Ser  
   1620  
   1625  
   1630  

aggg tgg ctc ctc gtc  
   acg ggg acg ggg ggc gcg gcc act gta  
   Arg Thr Leu Ala Gly Cys Thr Thr Thr Asp Thr Val  
   1635  
   1640  
   1645  

tca att gtt gac aag ctc ttc  
   gcc ggg gtt gcc gcc gtc gtt ggt  
   Ser Ile Val Asp Lys Leu Phe Ala Gly Gly Thr Asp Ala Val Val Gly  
   1650  
   1655  
   1660  
   1665  

cac tgt cac gcc gcg ctc cct gcc  
   att gct gcc gct ggg gac cag tcc  
   His Cys His Ser Val Ile Ala Ala Ala Tyr Val Gly Gly Ser  
   1670  
   1675  
   1680  

aag gcg cca tcc ctc gcg  
   ggc gcg gcc gcc act cct  
   Arg Ser Pro Pro Leu Ala Ala Ala Ser Tyr Leu Met Gly Leu Gly  
   1685  
   1690  
   1695  

gtt gcc gcg  
   aac gcg cag ggg gcg tgg gtt ctc  
   ctt ctc tgg gcg  
   Val Gly Gly Arg Ala Gin Ala Arg Leu Ala Ser Ala Leu Leu Gly  
   1700  
   1705  
   1710  

gct ggt gcc gcg  
   acg ctc gcc gct ggg acc cct gtc gtt gcg  
   Ala Gly Thr Ala Leu Gly Thr Pro Val Leu Thr Met Ala  
   1715  
   1720  
   1725  

egg gcg  
   ttc atc gcc gcg gcg gtt gcc tcc ccc ccc  
   gcc gtc act gcc  
   Gin Ala Phe Met Gly Gly Gly Ala Ser Arg Ser Leu Val Thr Val  
   1730  
   1735  
   1740  
   1745  

cat gtt ggg gtt gga gcc gtt ggg gtt ctc  
   gac gcg gct gcc agt  
   Leu Leu Gly Ala Gly Gly Trp Gly Gly Val Val Ala Ser  
   1750  
   1755  
   1760  

ctc gtc gtc gcg ttc aag  
   ccc gcc gta ctc  
   thr ggg aac ccc ctt ctc gaa gcg ctt tgg  
   Leu Val Phe Asp Met Gly Ala Lys Leu Ser Thr Glu Asp Leu Trp  
   1765  
   1770  
   1775  

att gcc atc ccc gta ctc  
   act act gcc cgt ggg gcg ctc gcg gcg atg  
   Tyr Ala Lle Pro Val Leu Thr Ser Pro Pro Xaa Ala Gly Gly Ile  
   1780  
   1785  
   1790  

gcc ctt ggt gcc  
   gtt tac tca gcg aac gcc tct gcc gcg ctc gcc gcg gcc ctc gtc  
   Ala Leu Gly Leu Tyr Ser Asp Ala Ser Ser Asp Gly Thr Thr Thr  
   1795  
   1800  
   1805  

tgg gcg  
   ctt ggt ctt atc gcg aac ttc ggg act gcc  
   Gtc ctc atc act gtg  
  Thr Ser Ala Amn Ser Ala Thr Ser Ala Ser Ile Pro  
   1810  
   1815  
   1820  
   1825  

gac gcg ttc cta ccg cag gcg  
   tac tcc gac aag gcc ctc gca gcc  
   Gin Asp Ser Tyr Phe Gin Gin Ala Asp Tyr Cys Asp Lys Val Ser Ala Ile  
   1830  
   1835  
   1840  
   1845
<table>
<thead>
<tr>
<th>1930</th>
<th>1835</th>
<th>1940</th>
</tr>
</thead>
<tbody>
<tr>
<td>gtt ccg cg ctc aag ctt aca cgg ggt gct gac gtc aag agg</td>
<td>Val Arg Arg Leu Ser Leu Thr Arg Thr Val Val Ala Leu Val Asn Arg</td>
<td>6104</td>
</tr>
<tr>
<td>gag cct aag gtt gat gag gtc cag gtt gtt gag tgc tgg gat ctt tgg</td>
<td>Glu Pro Lys Val Asp Glu Val Gin Val Gly Tyr Val Trp Asp Leu Trp</td>
<td>6152</td>
</tr>
<tr>
<td>gag tgg gtt gtt atg cgc cag gtc atg gtt atg tct ata gtc cgg gcc</td>
<td>Glu Trp Val Met Arg Gin Val Arg Met Val Met Ser Arg Leu Arg Ala</td>
<td>6200</td>
</tr>
<tr>
<td>ctc gtc cct gtt gtt tca ctc ccc tgg tgg cac tgg ggg ggg tgg</td>
<td>Leu Cys Pro Val Val Ser Leu Pro Leu Trp His Cys Gly Glu Gly Trp</td>
<td>6248</td>
</tr>
<tr>
<td>tcc gtt gaa tgg ctt ctc gat ggg cac gtt gaa agt ctt ctt cgg tgg</td>
<td>Ser Gly Glu Trp Leu Leu Asp Gly His Val Glu Ser Arg Cys Leu Cys</td>
<td>6296</td>
</tr>
<tr>
<td>ggg tgg gta atc acc ggc gac gtc ctc aat ggg caa ctc aca gat cca</td>
<td>Gly Cys Val Ile Thr Gly Val Gly Gin Gin Gly Gin Leu Asp Pro</td>
<td>6344</td>
</tr>
<tr>
<td>gtc tac tct acc aag gtt gtc gac tac tgg atg gga act gtt cgg</td>
<td>Val Tyr Ser Thr Lys Leu Cys Arg His Tyr Trp Met Gly Thr Val Pro</td>
<td>6392</td>
</tr>
<tr>
<td>gca ac tgg ctc ggc tac ggg gaa gct cta ctt ctc gcc tct gcc</td>
<td>Val Ala Met Leu Gly Tyr Gly Pro Ser Leu Leu Ala Ser Asp</td>
<td>6440</td>
</tr>
<tr>
<td>acc cgg aag gtt ccc ttc ggg acc tcg ggg tgg gct gag ggt gtt</td>
<td>Pro Lys Val Val Pro Phe Gly Thr Ser Gly Trp Ala Glu Val Val</td>
<td>6488</td>
</tr>
<tr>
<td>gtc acc ctt acc cag gtt gtt atc agg cgc acy tcc tgt tac aca cgg</td>
<td>Val Thr Pro Thr His Val Ile Arg Arg Ser Thr Ser Cys Tyr Leu Leu</td>
<td>6536</td>
</tr>
<tr>
<td>ctt cgc cag caa att ctt tca gca gct gta gct gag ccc tac tac gtt</td>
<td>Leu Arg Gin Gin Ile Leu Ser Ala Val Ala Glu Pro Tyr Tyr Val Val</td>
<td>6584</td>
</tr>
<tr>
<td>gat gcc ac ggc gca ctg ctc gtc aag cgg aga ggc ggc gcc atg</td>
<td>Asp Gly Ile Asp Pro Ser Leu Pro Ala Asp Ala Arg Asp Ala Met</td>
<td>6632</td>
</tr>
<tr>
<td>gtc tac ggt cgc caa agt gtt acc att gat ggg gag cgc tac acc</td>
<td>Val Gly Pro Gly Gin Ser Val Thr Ile Asp Gly Gly Arg Tyr Thr</td>
<td>6680</td>
</tr>
<tr>
<td>ctt cgc cag ctc gtt cgg atg cgg aat gtt ggc ccc tct gat gtt tca</td>
<td>Leu Pro His Gin Leu Arg Met Arg Ala Pro Ser Glu Val Ser</td>
<td>6728</td>
</tr>
<tr>
<td>tcc ggt gtt atc act gtt ggg agg gag act gaa gac tca gaa ctt</td>
<td>Ser Glu Val Ser Ile Gin SerThr Gly Val Trp Glu Asp Ser Glu Arg</td>
<td>6776</td>
</tr>
<tr>
<td>aca ggc aac gat tgg cca cca gcc ggt gct gcc ctc caa ggc ata gag</td>
<td>Thr Glu Ala Asp Leu Pro Ala Ala Ala Leu Gin Ala Ile Gin</td>
<td>6824</td>
</tr>
<tr>
<td>aat gct gcc aga att ctc gca ggg cag act gat gcc cgc atgc aag gag</td>
<td>Arg Ala Arg Arg Ile Leu Gin Pro His Ile Asp Val Xaa Met Glu Asp</td>
<td>6872</td>
</tr>
<tr>
<td>tgc cgg cag ctc ctt tgg cgg atc aag cgg gag atg ctt ctc ggg tgg</td>
<td>Cys Ser Ser Pro Leu Ser Gin Gin Ser Ser Arg Gin Met Pro Val Trp</td>
<td>6920</td>
</tr>
</tbody>
</table>
| gga gag gag atc gcc gag ctc cag gca ctt atc tcc cgc tct gtt aag | Gly Gin Gin Asp Pro Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Gin Git
atc cca ggg ggt ttc cct ctg tcc ccc ccc tac atg ggg gtt gtt cat
Ile Pro Gly Gly Phe Pro Leu Ser Pro Pro Tyr Met Gly Val Val His
2805 2810 2815

cau ttg gat ttc aca gcc cac cag ggt cgc cgg cgg tgg cgg tgc
cys leu phe thr his val val ser arg arg tyr trp leu gly phe
8260 8265 8270 8280

tta ggc ctc ctc act gta gcc ctc ctt ggg tga actaattca tctggtcgg
leu ala leu leu ile val ala leu phe gly
8285 8290 8295

caggttgtgag cggtggtaca ccgctaaagg aggttcccgc ccctcgcgc ccaggggtct
9145

cctgctggg taaaaagggc cccgctttgg gaggcaatgct gtttactaac cccctgcag
9205

ggtaaacgcc tggatgtgtct aatctcactgc cgtctggcg ccggctggct accttatagc
9265

gttaactggt gccctgcag ccctcgcgc ccagggccca gcggactcg
9325

<210> SEQ ID NO 2
<211> LENGTH: 2843
<212> TYPE: DNA
<213> ORGANISM: Hepatitis GB virus C
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (30)....(2823)
<223> OTHER INFORMATION: X = anything

<400> SEQUENCE: 2

Met Ala Val Leu Leu Leu Leu Val Val Glu Ala Gly Ala Ile Leu
1   5   10  15

Ala Pro Ala Thr His Ala Cys Ser Ala Lys Gly Glu Tyr Xaa Leu Thr
20  25  30

Asn Cys Cys Ala Leu Glu Asp Ile Gly Phe Cys Leu Glu Gly Gly Cys
35  40  45

Leu Val Ala Leu Gly Cys Thr Ile Cys Thr Asp Arg Cys Trp Pro Leu
50  55  60

Tyr Gln Ala Gly Leu Val Arg Pro Gly Lys Ser Ala Ala Gln Leu
65  70  75  80

Val Gly Glu Leu Gly Ser Leu Tyr Gly Pro Leu Ser Val Ser Ala Tyr
85  90  95

Val Ala Gly Ile Leu Gly Gly Leu Val Tyr Ser Gly Val Leu Thr
100 105 110

Val Gly Val Ala Leu Thr Arg Val Val Tyr Pro Val Pro Aam Leu Thr
115 120 125

Cys Ala Val Glu Cys Leu Lys Trp Glu Ser Glu Phe Trp Arg Trp
130 135 140

Thr Glu Gln Leu Ala Ser Arg Tyr Trp Ile Leu Gly Trp Lys
145 150 155 160

Val Pro Phe Asp Phe Trp Arg Gly Val Met Ser Leu Thr Pro Leu Leu
165 170 175

Val Cys Val Ala Leu Leu Leu Leu Glu Gin Arg Ile Val Met Val
180 185 190

Phe Leu Leu Val Thr Met Ala Gly Met Ser Glu Gly Ala Pro Ala Ser
195 200 205

Val Leu Gly Ser Arg Pro Phe Glu Ala Gly Leu Thr Trp Glu Ser Cys
210 215 220

Ser Cys Arg Ser Aam Gly Ser Arg Val Pro Thr Gly Glu Arg Val Trp
225 230 235 240
<table>
<thead>
<tr>
<th>Amino Acid Sequence</th>
<th>Start Position</th>
<th>End Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glu Arg Gly Asn Val Thr Leu Leu Cys Asp Cys Pro Asn Gly Pro Trp</td>
<td>245</td>
<td>255</td>
</tr>
<tr>
<td>Val Trp Val Pro Ala Leu Cys Gln Ala Ile Gly Trp Gly Asp Pro Ile</td>
<td>260</td>
<td>265</td>
</tr>
<tr>
<td>Thr His Trp Ser His Gly Gln Asn Gln Trp Pro Leu Ser Cys Pro Gln</td>
<td>275</td>
<td>280</td>
</tr>
<tr>
<td>Phe Val Tyr Gly Ala Val Ser Val Thr Cys Val Trp Gly Ser Val Ser</td>
<td>290</td>
<td>295</td>
</tr>
<tr>
<td>Trp Phe Ala Ser Thr Gly Gly Arg Ser Lys Val Asp Val Trp Ser</td>
<td>305</td>
<td>310</td>
</tr>
<tr>
<td>Leu Val Pro Val Gly Ser Ala Ser Cys Thr Ile Ala Ala Leu Gly Ser</td>
<td>325</td>
<td>330</td>
</tr>
<tr>
<td>Ser Asp Arg Asp Thr Val Val Glu Leu Ser Glu Trp Gly Ile Pro Cys</td>
<td>340</td>
<td>345</td>
</tr>
<tr>
<td>Ala Thr Cys Ile Leu Asp Arg Arg Pro Ala Ser Cys Gly Thr Cys Val</td>
<td>355</td>
<td>360</td>
</tr>
<tr>
<td>Arg Asp Cys Trp Pro Glu Thr Gly Ser Val Arg Phe Pro Pro Phe His Arg</td>
<td>370</td>
<td>375</td>
</tr>
<tr>
<td>Cys Gly Ala Gly Pro Arg Leu Thr Arg Asp Leu Glu Ala Val Pro Phe</td>
<td>385</td>
<td>390</td>
</tr>
<tr>
<td>Val Asn Arg Thr Thr Pro Phe Thr Ile Arg Gly Pro Leu Gly Asn Gln</td>
<td>405</td>
<td>410</td>
</tr>
<tr>
<td>Gly Arg Gly Asn Pro Val Arg Ser Pro Leu Gly Phe Gly Ser Tyr Thr</td>
<td>420</td>
<td>425</td>
</tr>
<tr>
<td>Met Thr Lys Ile Arg Asp Ser Leu His Leu Val Lys Cys Pro Thr Pro</td>
<td>430</td>
<td>440</td>
</tr>
<tr>
<td>Ala Ile Glu Pro Pro Thr Gly Thr Phe Gly Phe Pro Gly Val Pro</td>
<td>455</td>
<td>460</td>
</tr>
<tr>
<td>Pro Leu Asn Asn Cys Met Leu Leu Gly Thr Glu Val Ser Glu Val Leu</td>
<td>465</td>
<td>470</td>
</tr>
<tr>
<td>Gly Gly Ala Gly Leu Thr Gly Gly Phe Tyr Gly Pro Leu Val Arg Arg</td>
<td>485</td>
<td>490</td>
</tr>
<tr>
<td>Cys Ser Gly Leu Met Gly Arg Arg Asn Pro Val Cys Pro Gly Phe Ala</td>
<td>505</td>
<td>510</td>
</tr>
<tr>
<td>Trp Leu Ser Ser Gly Arg Pro Asp Gly Phe Ile His Val Gln Gly His</td>
<td>515</td>
<td>520</td>
</tr>
<tr>
<td>Leu Gln Glu Val Asp Ala Gly Asn Phe Ile Pro Pro Arg Trp Leu</td>
<td>530</td>
<td>535</td>
</tr>
<tr>
<td>Leu Leu Asp Phe Val Phe Val Leu Leu Tyr Leu Met Lys Leu Ala Glu</td>
<td>540</td>
<td>545</td>
</tr>
<tr>
<td>Ala Arg Leu Val Pro Leu Ile Leu Leu Leu Trp Trp Val Asn</td>
<td>550</td>
<td>555</td>
</tr>
<tr>
<td>Gln Leu Ala Val Leu Xaa Val Xaa Ala Xaa Xaa Ala Ala Val Ala Gly</td>
<td>560</td>
<td>565</td>
</tr>
<tr>
<td>Glu Val Phe Ala Gly Pro Ala Leu Ser Trp Cys Leu Gly Leu Pro Phe</td>
<td>570</td>
<td>575</td>
</tr>
<tr>
<td>Val Ser Met Ile Leu Gly Leu Ala Asn Leu Val Leu Tyr Phe Arg Trp</td>
<td>580</td>
<td>585</td>
</tr>
<tr>
<td>Met Gly Pro Gln Arg Leu Met Phe Leu Val Leu Trp Lys Leu Ala Arg</td>
<td>590</td>
<td>595</td>
</tr>
<tr>
<td>Gly Ala Phe Pro Leu Ala Leu Leu Met Gly Ile Ser Ala Thr Arg Gly</td>
<td>600</td>
<td>605</td>
</tr>
<tr>
<td>Arg Thr Ser Val Leu Gly Ala Glu Phe Cys Phe Asp Val Thr Phe Glu</td>
<td>610</td>
<td>615</td>
</tr>
<tr>
<td>1090</td>
<td>1095</td>
<td>1100</td>
</tr>
<tr>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>Glu</td>
<td>Ala</td>
<td>Pro</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1110</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pro</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1115</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ser</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1120</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gly</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1130</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arg</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1135</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1140</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glu</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1150</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arg</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1160</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1170</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thr</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1180</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leu</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1190</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1200</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1205</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1210</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asn</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1215</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1220</th>
</tr>
</thead>
<tbody>
<tr>
<td>Val</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1225</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glu</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1230</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1235</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arg</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1240</th>
</tr>
</thead>
<tbody>
<tr>
<td>Val</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1245</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trp</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1250</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1255</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1260</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ser</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1265</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gln</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1270</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pro</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1275</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arg</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1280</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1285</th>
<th>1290</th>
<th>1295</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>Ile</td>
<td>Ala</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1295</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1300</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thr</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1305</th>
<th>1310</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leu</td>
<td>Glu</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1315</th>
<th>1320</th>
<th>1325</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>Glu</td>
<td>Cys</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1330</th>
<th>1335</th>
<th>1340</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>Ser</td>
<td>Ala</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1340</th>
<th>1345</th>
<th>1350</th>
<th>1355</th>
<th>1360</th>
</tr>
</thead>
<tbody>
<tr>
<td>Val</td>
<td>Val</td>
<td>Arg</td>
<td>Ser</td>
<td>Gly</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1365</th>
<th>1370</th>
<th>1375</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arg</td>
<td>Ser</td>
<td>Gly</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1380</th>
<th>1385</th>
<th>1390</th>
</tr>
</thead>
<tbody>
<tr>
<td>Val</td>
<td>Val</td>
<td>Arg</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1395</th>
<th>1400</th>
<th>1405</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trp</td>
<td>Tyr</td>
<td>Gly</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1410</th>
<th>1415</th>
<th>1420</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp</td>
<td>Asp</td>
<td>Cys</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1425</th>
<th>1430</th>
<th>1435</th>
<th>1440</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>Val</td>
<td>Phe</td>
<td>Phe</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1445</th>
<th>1450</th>
<th>1455</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gln</td>
<td>Arg</td>
<td>Thr</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1460</th>
<th>1465</th>
<th>1470</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pro</td>
<td>Gln</td>
<td>Trp</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1475</th>
<th>1480</th>
<th>1485</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arg</td>
<td>Trp</td>
<td>Gly</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1490</th>
<th>1495</th>
<th>1500</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp</td>
<td>Asp</td>
<td>Leu</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1505</th>
<th>1510</th>
<th>1515</th>
<th>1520</th>
</tr>
</thead>
<tbody>
<tr>
<td>1525</td>
<td>1530</td>
<td>1535</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>Ile Tyr Ala Ser Tyr Thr Gly Ser Leu Val Val Thr Asp Trp Asp</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1540</td>
<td>1545</td>
<td>1550</td>
<td></td>
</tr>
<tr>
<td>Val Lys Gly Gly Gly Asn Pro Leu Tyr Arg Ser Gly Asp Gln Ala Thr</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1555</td>
<td>1560</td>
<td>1565</td>
<td></td>
</tr>
<tr>
<td>Pro Gln Pro Val Val Gln Val Pro Pro Val Asp His Arg Pro Gly Gly</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1570</td>
<td>1575</td>
<td>1580</td>
<td></td>
</tr>
<tr>
<td>Glu Ser Ala Pro Arg Asp Ala Lys Thr Val Thr Asp Ala Val Ala Ala</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1595</td>
<td>1590</td>
<td>1595</td>
<td>1600</td>
</tr>
<tr>
<td>Ile Gln Val Arg Cys Asp Trp Ser Val Met Thr Leu Ser Ile Gly Glu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1605</td>
<td>1610</td>
<td>1615</td>
<td></td>
</tr>
<tr>
<td>Val Leu Thr Leu Ala Gln Ala Lys Thr Ala Gly Tyr Ala Ala Thr</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1620</td>
<td>1625</td>
<td>1630</td>
<td></td>
</tr>
<tr>
<td>Ser Arg Trp Leu Ala Gly Cys Tyr Thr Gly Thr Arg Ala Val Pro Thr</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1635</td>
<td>1640</td>
<td>1645</td>
<td></td>
</tr>
<tr>
<td>Val Ser Ile Val Asp Lys Leu Phe Ala Gly Gly Trp Ala Ala Val Val</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1650</td>
<td>1655</td>
<td>1660</td>
<td></td>
</tr>
<tr>
<td>Gly His Cys His Ser Val Ile Ala Ala Ala Ala Tyr Gly Ala</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1665</td>
<td>1670</td>
<td>1675</td>
<td>1680</td>
</tr>
<tr>
<td>Ser Arg Ser Pro Pro Leu Ala Ala Ala Ala Ser Tyr Leu Met Gly Leu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1685</td>
<td>1690</td>
<td>1695</td>
<td></td>
</tr>
<tr>
<td>Gly Val Gly Gly Asn Ala Gln Ala Arg Leu Ala Ala Ser Leu Leu Leu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1700</td>
<td>1705</td>
<td>1710</td>
<td></td>
</tr>
<tr>
<td>Gly Ala Ala Gly Thr Ala Leu Gly Thr Pro Val Val Gly Leu Thr Met</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1715</td>
<td>1720</td>
<td>1725</td>
<td></td>
</tr>
<tr>
<td>Ala Gly Ala Phe Met Gly Gly Ala Ser Ser Pro Ser Leu Val Thr</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1730</td>
<td>1735</td>
<td>1740</td>
<td></td>
</tr>
<tr>
<td>Val Leu Leu Gly Ala Val Gly Gly Trp Glu Gly Val Asn Ala Ala</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1745</td>
<td>1750</td>
<td>1755</td>
<td>1760</td>
</tr>
<tr>
<td>Ser Leu Val Phe Asp Phe Met Ala Gly Lys Leu Ser Thr Glu Asp Leu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1765</td>
<td>1770</td>
<td>1775</td>
<td></td>
</tr>
<tr>
<td>Trp Tyr Ala Ile Pro Val Leu Thr Ser Pro Xaa Ala Gly Leu Ala Gly</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1780</td>
<td>1785</td>
<td>1790</td>
<td></td>
</tr>
<tr>
<td>Ile Ala Leu Gly Leu Val Leu Tyr Ser Ala Asn Asn Ser Gly Thr Thr</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1795</td>
<td>1800</td>
<td>1805</td>
<td></td>
</tr>
<tr>
<td>Thr Trp Leu Asn Arg Leu Thr Thr Leu Pro Arg Ser Ser Cys Ile</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1810</td>
<td>1815</td>
<td>1820</td>
<td></td>
</tr>
<tr>
<td>Pro Asp Ser Tyr Phe Gln Gln Ala Asp Tyr Cys Asp Lys Val Ser Ala</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1825</td>
<td>1830</td>
<td>1835</td>
<td>1840</td>
</tr>
<tr>
<td>Ile Val Arg Arg Leu Ser Leu Thr Arg Thr Val Val Ala Leu Val Asn</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1845</td>
<td>1850</td>
<td>1855</td>
<td></td>
</tr>
<tr>
<td>Arg Glu Pro Lys Val Asp Glu Val Gln Val Gly Tyr Val Trp Asp Leu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1860</td>
<td>1865</td>
<td>1870</td>
<td></td>
</tr>
<tr>
<td>Trp Glu Trp Val Met Arg Gln Val Arg Met Val Met Ser Arg Leu Arg</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1875</td>
<td>1880</td>
<td>1885</td>
<td>1890</td>
</tr>
<tr>
<td>Ala Leu Cys Pro Val Val Ser Leu Pro Leu Trp His Cys Gly Glu Gly</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1895</td>
<td>1900</td>
<td>1905</td>
<td></td>
</tr>
<tr>
<td>Trp Ser Gly Glu Trp Leu Leu Asp Gly His Val Glu Ser Arg Cys Leu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1905</td>
<td>1910</td>
<td>1915</td>
<td>1920</td>
</tr>
<tr>
<td>Cys Gly Cys Val Ile Thr Gly Asp Val Leu Asn Gly Gln Leu Lys Asp</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1925</td>
<td>1930</td>
<td>1935</td>
<td></td>
</tr>
<tr>
<td>Pro Val Tyr Ser Thr Lys Leu Cys Arg His Tyr Trp Met Gly Thr Val</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1940</td>
<td>1945</td>
<td>1950</td>
</tr>
<tr>
<td>----</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>Pro Val Asn Met Leu Gly Tyr Gly Thr Ser Pro Leu Leu Ala Ser</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp Thr Pro Lys Val Val Pro Phe Gly Thr Ser Gly Trp Ala Glu Val</td>
<td>1970</td>
<td>1975</td>
<td>1980</td>
</tr>
<tr>
<td>Val Val Thr Pro Thr His Val Val Ile Arg Arg Thr Ser Cys Tyr Lys</td>
<td>1995</td>
<td>2000</td>
<td>2000</td>
</tr>
<tr>
<td>Leu Leu Arg Gln Gln Ile Leu Ser Ala Ala Ala Val Glu Pro Tyr Tyr</td>
<td>2015</td>
<td>2015</td>
<td>2015</td>
</tr>
<tr>
<td>Val Asp Gly Ile Pro Val Ser Trp Glu Ala Asp Ala Arg Ala Pro Ala</td>
<td>2020</td>
<td>2025</td>
<td>2030</td>
</tr>
<tr>
<td>Met Val Tyr Gly Pro Gly Gln Ser Val Thr Ile Asp Gly Glu Arg Tyr</td>
<td>2040</td>
<td>2040</td>
<td>2045</td>
</tr>
<tr>
<td>Thr Leu Pro His Gln Leu Arg Met Arg Asn Val Ala Pro Ser Glu Val</td>
<td>2050</td>
<td>2055</td>
<td>2060</td>
</tr>
<tr>
<td>Ser Ser Glu Val Ser Ile Glu Ile Gly Thr Glu Thr Glu Asp Ser Glu</td>
<td>2065</td>
<td>2070</td>
<td>2075</td>
</tr>
<tr>
<td>Leu Thr Glu Ala Asp Leu Pro Pro Ala Ala Ala Ala Leu Gln Ala Ala Ile</td>
<td>2085</td>
<td>2090</td>
<td>2095</td>
</tr>
<tr>
<td>Glu Aen Ala Ala Arg Ile Leu Glu Pro His Ile Asp Val Xaa Met Glu</td>
<td>2100</td>
<td>2105</td>
<td>2110</td>
</tr>
<tr>
<td>Asp Cys Ser Thr Pro Ser Leu Cys Gly Ser Ser Arg Glu Met Pro Val</td>
<td>2115</td>
<td>2120</td>
<td>2125</td>
</tr>
<tr>
<td>Trp Gly Glu Asp Ile Pro Arg Thr Pro Ser Pro Ala Leu Ile Ser Val</td>
<td>2130</td>
<td>2135</td>
<td>2140</td>
</tr>
<tr>
<td>Thr Glu Ser Ser Ser Asp Glu Thr Leu Val Ser Thr Ser Gln</td>
<td>2145</td>
<td>2150</td>
<td>2155</td>
</tr>
<tr>
<td>Glu Asp Thr Pro Ser Ser Asp Ser Phe Glu Val Ile Gln Glu Ser Asp</td>
<td>2165</td>
<td>2170</td>
<td>2175</td>
</tr>
<tr>
<td>Thr Ala Gly Ser Glu Ser Val Phe Asn Val Ala Leu Ser Val Leu</td>
<td>2180</td>
<td>2185</td>
<td>2190</td>
</tr>
<tr>
<td>Lys Ala Leu Phe Pro Gln Ser Asp Ala Thr Arg Lys Leu Thr Val Lys</td>
<td>2195</td>
<td>2200</td>
<td>2205</td>
</tr>
<tr>
<td>Met Ser Cys Cys Val Glu Lys Ser Val Thr Arg Phe Phe Ser Leu Gly</td>
<td>2210</td>
<td>2215</td>
<td>2220</td>
</tr>
<tr>
<td>Leu Thr Val Ala Asp Ala Ser Leu Cys Glu Met Glu Ile Gln Aen</td>
<td>2225</td>
<td>2230</td>
<td>2235</td>
</tr>
<tr>
<td>His Thr Ala Tyr Cys Asp Lys Val Arg Thr Pro Leu Glu Leu Glu Val</td>
<td>2240</td>
<td>2245</td>
<td>2250</td>
</tr>
<tr>
<td>Gly Cys Leu Val Gly Aen Glu Thr Asp Met Lys Cys Gly</td>
<td>2260</td>
<td>2265</td>
<td>2270</td>
</tr>
<tr>
<td>Ala Arg Gln Glu Thr Leu Ala Ser Phe Ser Tyr Ile Trp Ser Gly Val</td>
<td>2275</td>
<td>2280</td>
<td>2285</td>
</tr>
<tr>
<td>Pro Leu Thr Arg Ala Thr Arg Ala Lys Pro Pro Val Val Arg Pro Val</td>
<td>2290</td>
<td>2295</td>
<td>2300</td>
</tr>
<tr>
<td>Gly Ser Leu Leu Val Ala Asp Thr Thr Lys Val Tyr Val Thr Aen Pro</td>
<td>2305</td>
<td>2310</td>
<td>2315</td>
</tr>
<tr>
<td>Asp Aen Val Gly Arg Arg Val Lys Val Thr Phe Trp Arg Ala Pro</td>
<td>2320</td>
<td>2325</td>
<td>2330</td>
</tr>
<tr>
<td>Arg Val His Asp Lys Phe Leu Val Asp Ser Ile Glu Arg Ala Arg Arg</td>
<td>2335</td>
<td>2340</td>
<td>2345</td>
</tr>
<tr>
<td>Ala Ala Gln Gly Cys Leu Ser Met Gly Tyr Thr Tyr Glu Glu Ala Ile</td>
<td>2350</td>
<td>2355</td>
<td>2360</td>
</tr>
<tr>
<td>Ala Ala Gln Gly Cys Leu Ser Met Gly Tyr Thr Tyr Glu Glu Ala Ile</td>
<td>2355</td>
<td>2360</td>
<td>2365</td>
</tr>
<tr>
<td>Sequence</td>
<td>Start</td>
<td>End</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
<td>-----</td>
<td></td>
</tr>
<tr>
<td>Arg Thr Val Arg Pro His Ala Ala Met Gly Trp Gly Ser Lys Val Ser</td>
<td>2370</td>
<td>2380</td>
<td></td>
</tr>
<tr>
<td>Val Lys Asp Leu Ala Thr Pro Ala Gly Lys Met Ala Val His Asp Arg</td>
<td>2395</td>
<td>2400</td>
<td></td>
</tr>
<tr>
<td>Leu Gln Glu Ile Leu Glu Gly Thr Pro Val Pro Phe Thr Leu Thr Val</td>
<td>2405</td>
<td>2415</td>
<td></td>
</tr>
<tr>
<td>Lys Lys Glu Val Phe Phe Lys Asp Arg Lys Glu Glu Lys Ala Pro Arg</td>
<td>2420</td>
<td>2430</td>
<td></td>
</tr>
<tr>
<td>Leu Ile Val Phe Pro Pro Leu Asp Phe Arg Ile Ala Glu Lys Leu Ile</td>
<td>2435</td>
<td>2445</td>
<td></td>
</tr>
<tr>
<td>Leu Gly Asp Pro Gly Arg Val Ala Lys Ala Gly Val Gly Gly Ala Tyr</td>
<td>2450</td>
<td>2460</td>
<td></td>
</tr>
<tr>
<td>Ala Phe Gln Tyr Thr Pro Asn Gln Arg Val Lys Glu Met Leu Lys Leu</td>
<td>2465</td>
<td>2470</td>
<td></td>
</tr>
<tr>
<td>Trp Gly Ser Lys Thr Pro Cys Ala Ile Cys Val Asp Ala Thr Cys</td>
<td>2485</td>
<td>2495</td>
<td></td>
</tr>
<tr>
<td>Phe Asp Ser Ser Ile Thr Glu Asp Val Ala Leu Glu Thr Glu Leu</td>
<td>2500</td>
<td>2510</td>
<td></td>
</tr>
<tr>
<td>Tyr Ala Leu Ala Ser Asp His Pro Glu Trp Val Arg Ala Leu Gly Lys</td>
<td>2515</td>
<td>2525</td>
<td></td>
</tr>
<tr>
<td>Tyr Xaa Ala Ser Gly Thr Met Val Thr Pro Gly Val Gly Val Pro Val Gly</td>
<td>2530</td>
<td>2540</td>
<td></td>
</tr>
<tr>
<td>Glu Arg Tyr Cys Arg Ser Ser Gly Val Leu Thr Thr Ser Ala Ser Asn</td>
<td>2545</td>
<td>2550</td>
<td></td>
</tr>
<tr>
<td>Cys Leu Thr Cys Tyr Ile Lys Val Arg Ala Ala Cys Glu Arg Ile Gly</td>
<td>2565</td>
<td>2575</td>
<td></td>
</tr>
<tr>
<td>Leu Lys Asn Val Ser Leu Leu Ile Ala Gly Asp Asp Cys Leu Ile Val</td>
<td>2580</td>
<td>2590</td>
<td></td>
</tr>
<tr>
<td>Cys Glu Arg Pro Val Cys Asp Pro Cys Glu Ala Leu Gly Arg Thr Leu</td>
<td>2595</td>
<td>2605</td>
<td></td>
</tr>
<tr>
<td>Ala Ser Tyr Gly Tyr Ala Cys Glu Pro Ser Tyr His Ala Ser Leu Asp</td>
<td>2610</td>
<td>2620</td>
<td></td>
</tr>
<tr>
<td>Thr Ala Pro Phe Cys Ser Thr Trp Leu Ala Glu Cys Asn Ala Asp Gly</td>
<td>2625</td>
<td>2630</td>
<td></td>
</tr>
<tr>
<td>Xaa Arg His Phe Phe Leu Thr Thr Arg Phe Arg Arg Pro Leu Ala Arg</td>
<td>2645</td>
<td>2650</td>
<td></td>
</tr>
<tr>
<td>Met Ser Ser Glu Tyr Ser Pro Ala Ser Ala Ile Gly Tyr Ile</td>
<td>2660</td>
<td>2670</td>
<td></td>
</tr>
<tr>
<td>Leu Leu Tyr Pro Trp Xaa Pro Ile Thr Arg Trp Val Ile Ile Pro His</td>
<td>2675</td>
<td>2685</td>
<td></td>
</tr>
<tr>
<td>Val Leu Thr Cys Ala Ser Arg Gly Gly Gly Tyr Xaa Ser Asp Pro</td>
<td>2690</td>
<td>2700</td>
<td></td>
</tr>
<tr>
<td>Val Trp Cys Gln Val His Gly Asn Tyr Tyr Lys Phe Pro Leu Asp Lys</td>
<td>2705</td>
<td>2710</td>
<td></td>
</tr>
<tr>
<td>Leu Pro Asn Ile Ile Val Ala Leu His Gly Pro Ala Ala Leu Arg Val</td>
<td>2725</td>
<td>2730</td>
<td></td>
</tr>
<tr>
<td>Thr Ala Asp Thr Thr Lys Thr Met Glu Ala Gly Lys Val Leu Ser</td>
<td>2740</td>
<td>2750</td>
<td></td>
</tr>
<tr>
<td>Asp Leu Lys Leu Pro Gly Leu Ala Val His Arg Lys Ala Gly Ala</td>
<td>2755</td>
<td>2760</td>
<td></td>
</tr>
<tr>
<td>Leu Arg Thr Arg Met Leu Arg Ser Arg Gly Trp Ala Glu Leu Ala Arg</td>
<td>2770</td>
<td>2780</td>
<td></td>
</tr>
<tr>
<td>Gly Leu Leu Trp His Pro Gly Leu Arg Leu Pro Pro Pro Glu Ile Ala</td>
<td>2785</td>
<td>2790</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
<210> SEQ ID NO 4
<211> LENGTH: 312
<212> TYPE: PRT
<213> ORGANISM: Hepatitis GB virus C
<220> FEATURE:
<221> NAME/KEY: MOD_RES
<222> LOCATION: (253)...(258)
<223> OTHER INFORMATION: X = anything

<400> SEQUENCE: 4
Thr Ile Ala Ala Leu Gly Ser Ser Asp Arg Asp Thr Val Val Glu Leu
1  5     10    15
Ser Glu Trp Gly Ile Pro Cys Ala Thr Cys Ile Leu Asp Arg Arg Pro
 20     25    30
 Ala Ser Cys Gly Thr Cys Val Arg Asp Cys Trp Pro Glu Thr Gly Ser
  35     40    45
Val Arg Phe Pro Phe His Arg Cys Gly Ala Gly Pro Arg Leu Thr Arg
  50     55    60
Asp Leu Glu Ala Val Pro Phe Val Asp Arg Thr Pro Phe Thr Ile
   65     70    75    80
Arg Gly Pro Leu Gly Asn Gln Gly Arg Gly Asn Pro Val Arg Ser Pro
   85     90    95
Leu Gly Phe Gly Ser Tyr Thr Met Thr Lys Ile Arg Asp Ser Leu His
 100    105    110
Leu Val Lys Cys Pro Thr Pro Ala Ile Glu Pro Pro Thr Gly Thr Phe
115    120    125
Gly Phe Phe Pro Gly Val Pro Pro Leu Asn Asn Cys Met Leu Leu Gly
130   135    140
Thr Glu Val Ser Glu Val Leu Gly Gly Ala Leu Thr Gly Gly Phe
145   150    155    160
Tyr Glu Pro Leu Val Arg Arg Cys Ser Glu Leu Met Gly Arg Arg Asn
165   170    175
Pro Val Cys Pro Gly Phe Ala Trp Leu Ser Ser Gly Arg Pro Asp Gly
180   185    190
Phe Ile His Val Gln Gly His Leu Glu Gly Val Asp Ala Gly Asn Phe
195   200    205
Ile Pro Pro Pro Arg Trp Leu Leu Leu Asp Phe Val Phe Val Leu Leu
210   215    220
Tyr Leu Met Lys Leu Ala Glu Ala Arg Leu Val Pro Leu Ile Leu Leu
225   230    235    240
Leu Leu Trp Trp Trp Val Asn Gln Leu Ala Val Leu Xaa Val Xaa Ala
245   250    255
Xaa Xaa Ala Ala Val Ala Gly Glu Val Phe Ala Gly Pro Ala Leu Ser
260       265       270
Trp Cys Leu Gly Leu Pro Phe Val Ser Met Ile Leu Gly Leu Ala Asn
275       280       285
Leu Val Leu Tyr Phe Arg Trp Met Gly Pro Gln Arg Leu Met Phe Leu
290       295       300
Val Leu Trp Lys Leu Ala Arg Gly
305       310

<210> SEQ ID NO 5
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence:
    Synthetic Peptide
<400> SEQUENCE: 5
Leu Thr Gly Phe Tyr Glu Pro Leu
1       5

<210> SEQ ID NO 6
<211> LENGTH: 17
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence:
    Synthetic Peptide
<400> SEQUENCE: 6
Gly Gly Ala Gly Leu Thr Gly Gly Phe Tyr Glu Pro Leu Val Arg Arg
1       5      10      15
Cys

<210> SEQ ID NO 7
<211> LENGTH: 13
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence:
    Synthetic Peptide
<400> SEQUENCE: 7
Leu Thr Gly Phe Tyr Glu Pro Leu Val Arg Arg Cys
1       5      10

<210> SEQ ID NO 8
<211> LENGTH: 9
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence:
    Synthetic Peptide
<400> SEQUENCE: 8
Phe Tyr Glu Pro Leu Val Arg Arg Cys
1       5

<210> SEQ ID NO 9
<211> LENGTH: 415
<212> TYPE: PRT
<213> ORGANISM: Hepatitis GB virus C
<400> SEQUENCE: 9
Val Gly Tyr Val Trp Asp Leu Trp Glu Trp Ile Met Arg Gln Val Arg
15 Met Val Met Ala Arg Leu Arg Ala Leu Cys Pro Val Val Ser Leu Pro Val Trp His Cys Gly Glu Gly Trp Ser Gly Glu Trp Leu Leu Asp Gly
20 25 30
35 His Val Glu Ser Arg Cys Leu Cys Gly Cys Val Ile Thr Gly Asp Val
40 45
50 55 60
65 Leu Asn Gly Gin Leu Lys Glu Pro Val Tyr Ser Thr Lys Leu Cys Arg
70 75 80
85 His Tyr Trp Met Gly Thr Val Pro Val Asn Met Leu Gly Tyr Gly Glu
90 95
100 105 110
115 Thr Ser Pro Leu Leu Ala Ser Asp Thr Pro Lys Val Val Pro Phe Gly
120 125
130 Arg Arg Thr Ser Ala Tyr Lys Leu Arg Gin Gin Ile Leu Ser Ala
135 140
145 Ala Val Ala Glu Pro Tyr Tyr Val Asp Gly Ile Pro Val Ser Trp Asp
150 155 160
165 Ala Asp Ala Arg Ala Pro Ala Met Val Tyr Gly Pro Gly Gin Ser Val
170 175
180 Thr Ile Asp Gly Glu Arg Tyr Thr Leu Pro His Gin Leu Arg Leu Arg
185 190
195 Asn Val Ala Pro Ser Glu Val Ser Ser Glu Val Ser Ile Asp Ile Gly
200 205
210 Thr Glu Thr Glu Asp Ser Glu Leu Thr Glu Ala Asp Leu Pro Pro Ala
215 220
225 Ala Ala Ala Leu Gin Ala Ile Glu Asn Ala Ala Arg Ile Leu Glu Pro
230 235 240
245 His Ile Asp Val Ile Met Glu Asp Cys Ser Thr Pro Ser Leu Cys Gly
250 255
260 Ser Ser Arg Glu Met Pro Val Thr Gly Glu Asp Ile Pro Arg Thr Pro
265 270
270 Ser Ser Pro Ala Leu Ile Ser Val Thr Glu Ser Pro Ser Asp Glu Lys
280 285 290
295 Thr Pro Ser Val Ser Ser Gin Glu Asp Thr Pro Ser Ser Asp Ser
300
305 Phe Glu Val Ile Gin Glu Ser Glu Thr Ala Glu Gly Glu Glu Ser Val
310 315 320
325 Phe Asn Val Ala Leu Ser Val Leu Lys Ala Leu Phe Pro Gin Ser Asp
330 335
340 Ala Thr Arg Lys Leu Thr Val Lys Met Ser Cys Cys Val Glu Lys Ser
345 350
360 Val Thr Arg Phe Phe Ser Leu Gly Leu Thr Val Ala Asp Val Ala Ser
365
370 Leu Cys Glu Met Glu Ile Gin Asn His Thr Ala Tyr Cys Asp Lys Val
375 380
385 Arg Thr Pro Leu Glu Leu Gin Val Gly Cys Leu Val Gly Asn Glu Leu
390 395 400
405 Thr Phe Glu Cys His Asn Cys Glu Ala Arg Gin Glu Thr Leu Ala
410 415
What is claimed is:

1. A method for treating HIV infection comprising administering to a subject a composition comprising HCV or GBV-C NS5A peptide or polypeptide, wherein said NS5A peptide or polypeptide inhibits HIV replication and thereby treats HIV infection, and further wherein the NS5A peptide or polypeptide further comprises a targeting signal.

2. The method of claim 1, wherein the targeting signal is a nuclear targeting signal.

3. The method of claim 1, wherein the targeting signal targets a cell surface receptor.

4. The method of claim 3, wherein the cell surface receptor is the CD4 receptor.

5. A method for treating HIV infection comprising administering to a subject a composition comprising GBV-C NS5A peptide or polypeptide, wherein said NS5A peptide or polypeptide inhibits HIV replication and thereby treats HIV infection.

6. The method of claim 5, wherein the NS5A polypeptide comprises residues 152-237 of GBV-C NS5A.

* * * * *

* * * * *
In column 1, lines 10-13, delete
“The U.S. Government own rights in this invention pursuant to grant number AI58740
from NIH and merit grants awarded to Jack Stapleton and Jinhua Xiang from the
Veterans Administration.” and insert
--This invention was made with government support under grant number AI58740
awarded by the National Institutes of Health and merit grants awarded to Jack Stapleton
and Jinhua Xiang from the Veterans Administration. The government has certain rights
in the invention--.