Antiviral Therapy

The present invention relates to treatments for improving antiviral therapies and to method for determining whether or not antiviral therapies will be effective.

Viral infections represent a serious threat to health and are known to be a major cause of morbidity to animals and man. For instance, Hepatitis C virus (HCV) infection is a major cause of chronic liver disease worldwide. An important and striking feature of hepatitis C is its tendency towards chronicity. In over 70% of infected individuals, HCV establishes a persistent infection over decades that may lead to cirrhosis and hepatocellular carcinoma.

An interesting hypothesis in HCV biology proposes that the viral NS3-4A protease not only processes the viral proteins but also cleaves and inactivates components of the intracellular sensory pathways that detect viral infection and induce the transcriptional activation of type I interferons (IFN). One of the targets of NS3-4A is TRIF (TIR domain-containing adapter inducing IFNβ), an essential link between dsDNA detection in endosomes by TLR3 (toll-like receptor 3) and the induction of IFNβ. More recently, retinoic acid inducible gene-I (RIG-I) and MDA5 (helicard) were identified as intracellular sensors of dsRNA. Both sensors signal through Cardif (also known as IPS-1, MAVS, VISA) to induce IFNβ production. Cardif can be cleaved and inactivated by HCV NS3-4A. Two RNA helicases, RIG-I and MDA5, identified as intracellular sensors of dsRNA act through Cardif to induce IFNβ production.

Type I IFNs are not only crucial components of the innate immune system, but are also the most important components of current therapies against CHC. The current standard therapy consists of pegylated IFNα (pegIFNα) injected once weekly subcutaneously and daily intake of the oral antiviral agent ribavirin. This regimen achieves an overall sustained virological response (SVR) in about 55% of the patients, with significant differences between genotypes. An SVR is defined as the loss of detectable HCV RNA during treatment and its continued absence for at least 6 months after stopping therapy. Several studies of long-term follow-up on patients who achieve an SVR demonstrate that this response is durable in over 95% of patients. The probability of a SVR strongly depends on the early response to treatment. Patients who do not show an early virological response (EVR), defined as a decline of the viral load by at least 2 $\log_{10}$ after 12 weeks of therapy, are highly unlikely to develop an SVR, and treatment can be stopped in these patients. On the other hand, patients with an EVR have a good chance to be cured, with 65% of them achieving a SVR. The prognosis is even better for patients who have a rapid virological response (RVR), defined as serum HCV RNA undetectable after 4 weeks of treatment. Over 85% of them will achieve a SVR. Unfortunately, less than 20% of patients with genotype 1 and about 60% of patients with genotypes 2 or 3 show a RVR. The host factors that are important for an early response to therapy are currently not known.

Type I IFNs achieve their potent antiviral effects through the regulation of hundreds of genes (ISG, interferon stimulated genes). The transcriptional activation of ISGs induces proteins that are usually not synthesized in resting cells and that establish a non-virus-specific antiviral state within the cell. Interferons induce their synthesis by activating the Jak-STAT pathway, a paradigm of cell signaling used by many cytokines and growth factors. All type I IFNs bind to the same cell surface receptor (IFNAR) and activate the receptor-associated Janus kinase family members Jak1 and Tyk2. The kinases then phosphorylate and activate signal transducer and activator of transcription 1 (STAT1) and STAT2. The activated STATs translocate into the nucleus where they bind specific DNA elements in the promoters of ISGs. Many of the ISGs have antiviral activity but others are involved in lipid metabolism, apoptosis, protein degradation and inflammatory cell responses. As is the case with many viruses, HCV interferes with the IFN system, probably at multiple levels. IFN induced Jak-STAT signaling is inhibited in cells and transgenic mice that express HCV proteins, and in liver biopsies of patients with CHC. *In vitro*, HCV proteins NS5A and E2 bind and inactivate protein kinase R (PKR), an important non-specific antiviral protein. However, the molecular mechanisms that are important for the response to therapeutically applied IFN in patients with CHC are currently unknown.

The capacity of HCV to interfere with the IFN pathway at many different levels is a likely mechanism underlying HCV success to establish a chronic infection (2). However, quite paradoxically, in chimpanzees acutely or chronically infected with HCV hundreds of ISGs are induced in the liver (16, 17). Nevertheless, despite the activation of the endogenous IFN system, the virus is not cleared from chronically infected animals (18). The results obtained with chimpanzees are difficult to extrapolate to humans since there are some important differences in the pathobiology of HCV infection between these species. Whereas most chimpanzees acutely infected with HCV clear the virus spontaneously, infections in men mostly become chronic. On the other hand, chronically infected chimpanzees can rarely be cured with IFN, whereas more than half of the patients with CHC are successfully treated (19).

Induction of ISGs was also found in pre-treatment liver biopsies of many patients with CHC, again demonstrating that HCV infection can lead to activation of the endogenous IFN system (20). Notably, patients with pre-elevated expression of ISGs tended to respond poorly to therapy when compared to patients having low initial expression (20). The cause of this differential response to therapy is not understood.

The present invention is based upon studies in which the inventors investigated IFN induced signaling and ISG induction in paired samples of liver biopsies and peripheral blood mononuclear cells (PBMCs) of patients with chronic hepatitis before and during therapy with pegIFNα. They further

correlated biochemical and molecular data with the response to treatment. Their work is set out in more detail in the accompanying Example.

The inventors established that some subjects with a viral infection of the liver are in a state of "pre-activation", such that the IFN signalling pathway is in a state of stimulation with activated ISGs. The inventors have found that such individuals, when subsequently treated with IFN and an antiviral agent, had a poor, or no, response to the antiviral treatment. In contrast, another group of infected subjects appeared to have no prior stimulation of IFN receptors (and stimulation of ISGs) and this group responded well to the antiviral therapy (i.e. they had a rapid virological response (RVR)). Moreover it is possible to determine whether a subject would be a poor responder to treatment or have a RVR according to the expression level of a number of specific genes, some of which are ISG genes. In other words, the inventors identified a set of genes that are prognostic genetic markers, the expression levels of which predict whether a subject will respond to antiviral treatment.

This lead the inventors to realise that a method could be developed to help a clinician decide on a treatment regimen for subjects suffering from a viral infection of the liver. Gene expression from an infected individual can be compared with gene expression from a control (i.e. a subject without viral infection).

Infected subjects with altered gene expression (compared to the control) would be unlikely to benefit from the use of IFN in a treatment regimen (i.e. these individuals would not be expected to have an RVR) whereas infected subjects for whom gene expression was mostly unaltered, compared to control expression, are likely to benefit from IFN therapy and have an RVR. The inventors were surprised to make these correlations because a skilled person would expect activation of ISGs to be associated with better viral clearance and not with a subset of subjects who respond poorly to treatment.

While there have been previous studies of gene expression levels in "responder" and "non-responder" subjects with a viral infection of the liver, e.g. Chen *et al* (2005) Gastroenterology 128, 1437-1444, the research conducted as part of the present invention studied a very much broader set of genes in order to determine which would act as prognostic markers. Moreover, the inventors also analysed gene expression levels in samples taken before *and* after antiviral treatment, and used this information to identify prognostic genetic markers, while previous studies only attempted to correlate treatment outcome to gene expression levels present in samples taken *before* treatment. Thus the data set which lead to the identification of the prognostic genetic markers set out below are considered to be much more complete and robust than that in previous studies.

Accordingly in the first aspect of the invention there is provided a method for determining the likelihood that a subject having a viral infection of the liver will be responsive to antiviral therapy that includes stimulation of Interferon (IFN) activity, the method comprising:

(a) analysing a sample from the subject for expression of at least one gene from each of the following groups of genes:

(i) KYNU; PAH; LOC129607; DDC; FOLH1; YBX1; BCHE; ACADL; ACSM3; NARF; SLPI; RPS5; RPL3; RPLP0; TRIM5 and HERC5;

(ii) HTATIP2; CDK4; IFI44L; and KLHDC3;

(iii) C7; IF; IFI27; IFIT1; OAS2; G1P2; OAS1; IRF7; RSAD2; IFI44; OAS3; SIGIRR; and IFIT2;

(iv) RAB4A; PPP1R1A; PPM1E; ENPP2; CAP2; ADCY1; CABYR; EVI1; PTGFRN; TRIM55; and IL28RA;

(v) MME; KCNN2; SLC16A10; AMOTL1; SPP2; LRCH4; HIST1H2BG; TSPYL5; HIST1H2AC; HIST1H2BD; PHTF1; ZNF684; GSTM5; FLJ20035; FIS; PARP12; C14orf21; PNPT1; FLJ39051; GALNTL1; OSBPL1A; LGALS3BP; TXNRD2; LOC201725,TOMM7; SRPX2; DCN; PSMAL; MICAL-L2; FLJ30046; SAMD9; ANKRD35; LOC284013; LOC402560; and LOC147646; and,

(b) comparing expression of the genes in the sample to expression of the same genes in a control sample.

One embodiment of the invention is wherein altered expression of the genes in the sample compared to expression of the same genes in the control sample indicates that the subject is not likely to be responsive to said antiviral therapy.

An alternative embodiment of the invention is wherein unaltered expression of the genes in the sample compared to expression of the same genes in the control sample indicates that the subject is likely to be responsive to said antiviral therapy.

Further information regarding each of the genes assessed in the first aspect of the invention is provided in Table 2 in the accompanying Example. In particular, we provide the Affimetrix identification number for each of the genes, thus allowing each gene to be specifically identified.

It will be appreciated that the method of the first aspect of the invention will be of great benefit to clinicians. IFN is a protein growth factor and pharmaceutical preparations containing IFN are expensive to manufacture. It is therefore very important for a clinician to be confident that IFN is being used in an appropriate and cost-effective way. Furthermore, independent of the cost, it is often desirable to eliminate a viral infection of the liver as quickly as possible. It is therefore wasting time (which could be spent utilising alternative therapies) if a clinician administers IFN and subsequently discovers that it has no beneficial effects. The method of the first aspect of the invention is therefore of great assistance to a clinician because he can identify two populations of subjects. One population will show altered expression of the genes listed above and in table 2 and will therefore not benefit from treatment with IFN. The other population, with expression of the genes listed above and in table 2 that do not significantly differ from control samples, will benefit from therapy with IFN.

In an alternative embodiment, it is considered that the expression of the genes of Table 3 (differentially expressed 4 hours after treatment) can be used in a similar way.

By "antiviral therapy" we mean any treatment regimen for reducing viral infection that involves the stimulation of IFN activity. Such a regimen may involve the use of compounds that stimulate Type I IFN activity and/or induce IFN stimulated genes (ISGs). The therapy may involve treatment with IFN *per se* or other IFN receptor agonists. For example the therapy may utilise pegylated IFNα (pegIFNα).

The therapy may involve the stimulation of IFN activity alone. However the inventors have found that the method according to the first aspect of the invention is particularly useful for predicting the effectiveness of an antiviral therapy that comprises the use of a combination therapy comprising a stimulator of IFN activity in conjunction with a known antiviral agent. Many antiviral agents are known to the art and the method of the invention can be used to evaluate the effectiveness of a number of different combination therapies. However the inventors have found that the method of the first aspect of the invention has particular value for predicting the effectiveness of therapy with a stimulator of IFN activity used in conjunction with the antiviral agent ribavirin.

It is most preferred that the method of the first aspect of the invention is used to predict the usefulness of pegIFNα and ribavirin as an antiviral therapy.

The method of the first aspect of the invention may be utilised to evaluate the effectiveness of treatments for a number of different viral infections of the liver, including Hepatitis B virus and Hepatitis C virus infections. It is most preferred that the method is utilised to evaluate the effectiveness of therapies for Hepatitis C Virus (HCV) infection. The inventors have found that the method of the invention is particularly useful for distinguishing between subjects that will be expected to have a rapid virological response (RVR) and those which will not (non-RVR).

Samples representative of gene expression in a subject that may be used in accordance with the present invention encompass any sample that may provide information as to genes being expressed by the subject.

Examples of suitable samples include biopsies, samples excised during surgical procedures, blood samples, urine samples, sputum samples, cerebrospinal fluid samples, and swabbed samples (such as saliva swab samples). It will be appreciated that the source of the sample will depend upon which type of viral infection the subject may have.

It is most preferred samples are from liver tissue. Liver samples have been found to be particularly instructive when the method is applied to assessing subjects with HCV infection. The inventors were surprised to find that RVR could be distinguished from non-RVR subjects by analysing gene expression from liver samples whereas peripheral blood leukocytes exhibited no significant changes in gene expression before or after exposure to IFN.

Suitable samples may include tissue sections such as histological or frozen sections. Methods by which such sections may be prepared in such a way as to be able to provide information representative of gene expression in the subject from which the section is derived will be well known to those skilled in the art, and should be selected with reference to the technique that it is intended to use when investigating gene expression.

Although the use of samples comprising a portion of tissue from the subject is contemplated, it may generally be preferred that the sample representative of gene expression comprise a suitable extract taken from such a tissue, said extract being capable of investigation to provide information regarding gene expression in the subject. Suitable protocols which may be used for the production of tissue extracts capable of providing information regarding gene expression in a subject will be well known to those skilled in the art. Preferred protocols may be selected with reference to the manner in which gene expression is to be investigated.

In the case of samples derived from liver suitable preparation steps are disclosed in 1.1.1 and 1.1.3 of the Example.

By "control sample" we mean a sample, equivalent to that from the subject, that has been derived from an individual that is not suffering from a viral infection of the liver. Although equivalent tissue or organ samples, constituting control samples, or extracts from such samples, may be used directly as the source of information regarding gene expression in the control sample, it will be appreciated, and generally be preferred, that information regarding the expression of the selected gene (or genes) in an

"ideal" control sample be provided in the form of reference data. Such reference data may be provided in the form of tables indicative of gene expression in the chosen control tissue. Alternatively, the reference data may be supplied in the form of computer software containing retrievable information indicative of gene expression in the chosen control tissue. The reference data may, for example, be provided in the form of an algorithm enabling comparison of expression of at least one selected gene(s) from each groups of genes in the subject with expression of the same genes in the control tissue sample.

In the event that expression of genes listed above and in Table 2 in a control sample is to be investigated via processing of a tissue or organ sample constituting the control sample, it is preferred that such processing is conducted using the same methods used to process the sample from the subject. Such parallel processing of subject samples and control samples allows a greater degree of confidence that comparisons of gene expression in these tissues will be normalised relative to one another (since any artefacts associated with the selected method by which tissue is processed and gene expression investigated will be applied to both the subject and control samples).

The method according to the first aspect of the invention may involve the analysis of gene expression of at least one gene, selected from each of the groups of genes. The finding that altered expression of the genes listed above and in Table 2 or 3 may be used in determining the effectiveness of an antiviral therapy is surprising, since although the expression of certain genes (such as those encoding STAT1) has been linked to HCV infection, most of the genes listed above and in Table 2 had never previously been identified as being associated IFN regulated gene expression or with the likelihood of a therapy being effective for treating viral infections. Furthermore, irrespective of the association of these genes with IFN activity, it was total unexpected that increased expression of ISGs would be associated with poor response to subsequent IFN treatment.

The inventors have identified a total of 83 different genes, the expression levels of which can be prognostic markers for the outcome of antiviral therapy. These genes have been distributed into five different groups according to their function: group (i) are considered to be involved in cell metabolism; group (ii) are considered to be involved in cell cycle; group (iii) are considered to be involved in immune response; group (iv) are considered to be involved in signal transduction; group (v) are each unassigned to any particular group set out above. This distribution is shown in the method of the invention, in which the expression level of at least one gene from each of the groups of genes is assessed in order to determine the likelihood that the subject will be responsive to antiviral therapy. The inventors have further found that these subsets of the genes have particular value and can be effective for that purpose when the expression level of at least one member of each of those groups is analysed.

It is preferred that the method is based on the analysis of at least five, six, seven, eight, nine, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 71, 72, 73, 74, 75, 76, 77 or 78 genes listed above.

Expression of genes listed above and in Table 2 or 3 may be investigated by analysis of target molecules representative of gene expression in the sample. The presence or absence of target molecules in a sample will generally be detected using suitable probe molecules. Such detection will provide information as to gene expression, and thereby allow comparison between gene expression occurring in the subject and expression occurring in the control sample. Probes will generally be capable of binding specifically to target molecules directly or indirectly representative of gene expression. Binding of such probes may then be assessed and correlated with gene expression to allow an effective prognostic comparison between gene expression in the subject and in the control.

By "altered expression" we include where the gene expression is both elevated or reduced in the sample when compared to the control, as discussed above.

Conversely by "unaltered expression" we include where the gene expression is not elevated or reduced in the sample when compared to the control, as discussed above.

An assessment of whether a gene expression is altered or unaltered can be made using routine methods of statistical analysis.

The target molecule may be peptide or polypeptide. The amount of peptide or polypeptide can be determined using a specific binding molecule, for instance an antibody. In a preferred instance, the amount of certain target proteins present in a sample may be assessed with reference to the biological activity of the target protein in the sample. Assessment and comparison of expression in this manner is particularly suitable in the case of protein targets having enzyme activity. Suitable techniques for the measurement of the amount of a protein target present in a sample include, but are not limited to, aptamers and antibody-based techniques, such as radio-immunoassays (RIAs), enzyme-linked immunoassays (ELISAs) and Western blotting.

Nucleic acids represent preferred target molecules for assaying gene expression according to the third aspect of the invention.

It will be understood that "nucleic acids" or "nucleic acid molecules" for the purposes of the present invention refer to deoxyribonucleotide or ribonucleotide polymers in either single-or double-stranded form. Furthermore, unless the context requires otherwise, these terms should be taken to encompass

known analogues of natural nucleotides that can function in a similar manner to naturally occurring nucleotides.

Furthermore it will be understood that target nucleic acids suitable for use in accordance with the invention need not comprise "full length" nucleic acids (e.g. full length gene transcripts), but need merely comprise a sufficient length to allow specific binding of probe molecules.

It is preferred that the nucleic acid target molecule is a mRNA gene transcript and artificial products of such transcripts. Preferred examples of artificial target molecules generated from gene transcripts include cDNA and cRNA, either of which may be generated using well known protocols or commercially available kits or reagents.

In a preferred embodiment of the method of the first aspect of the invention, samples may be treated to isolate RNA target molecules by a process of lysing cells taken from a suitable sample (which may be achieved using a commercially available lysis buffer such as that produced by Qiagen Ltd.) followed by centrifugation of the lysate using a commercially available nucleic acid separation column (such as the RNeasy midi spin column produced by Qiagen Ltd). Other methods for RNA extraction include variations on the phenol and guanidine isothiocyanate method of Chomczynski, P. and Sacchi, N. (1987) *Analytical Biochemistry* 162, 156. "Single Step Method of RNA Isolation by Acid Guanidinium Thiocyanate-Phenol-Chloroform Extraction." RNA obtained in this manner may constitute a suitable target molecule itself, or may serve as a template for the production of target molecules representative of gene expression.

It may be preferred that RNA derived from a subject or control sample may be used as substrate for cDNA synthesis, for example using the Superscript System (Invitrogen Corp.). The resulting cDNA may then be converted to biotinylated cRNA using the BioArray RNA Transcript labelling Kit (Enzo Life Sciences Inc.) and this cRNA purified from the reaction mixture using an RNeasy mini kit (Qiagen Ltd).

mRNA, representative of gene expression, may be measured directly in a tissue derived from a subject or control sample, without the need for mRNA extraction or purification. For example, mRNA present in, and representative of gene expression in, a subject or control sample of interest may be investigated using appropriately fixed sections or biopsies of such a tissue. The use of samples of this kind may provide benefits in terms of the rapidity with which comparisons of expression can be made, as well as the relatively cheap and simple tissue processing that may be used to produce the sample. *In situ* hybridisation techniques represent preferred methods by which gene expression may be investigated and compared in tissue samples of this kind. Techniques for the processing of tissues of interest that maintain the availability of RNA representative of gene expression in the subject or control sample are well known to those of skill in the art.

However, techniques by which mRNAs representative of gene expression in a subject or control sample may be extracted and collected are also well known to those skilled in the art, and the inventors have found that such techniques may be advantageously employed in accordance with the present invention. Samples comprising extracted mRNA from a subject or control sample may be preferred for use in the method of the third aspect of the invention, since such extracts tend to be more readily investigated than is the case for samples comprising the original tissues. For example, suitable target molecules allowing for comparison of gene expression may comprise the total RNA isolated from a sample of tissue from the subject, or a sample of control tissue.

Furthermore, extracted RNA may be readily amplified to produce an enlarged mRNA sample capable of yielding increased information on gene expression in the subject or control sample. Suitable examples of techniques for the extraction and amplification of mRNA populations are well known, and are considered in more detail below.

By way of example, methods of isolation and purification of nucleic acids to produce nucleic acid targets suitable for use in accordance with the invention are described in detail in Chapter 3 of Laboratory Techniques in Biochemistry and Molecular Biology: Hybridization With Nucleic Acid Probes, Part I. Theory and Nucleic Acid Preparation, P. Tijssen, ed. Elsevier, N.Y. (1993).

In a preferred method, the total nucleic acid may be isolated from a given sample using, the techniques described in the Example.

In the event that it is desired to amplify the nucleic acid targets prior to investigation and comparison of gene expression it may be preferred to use a method that maintains or controls for the relative frequencies of the amplified nucleic acids in the subject or control tissue from which the sample is derived.

Suitable methods of "quantitative" amplification are well known to those of skill in the art. One well known example, quantitative PCR, involves simultaneously co-amplifying a control sequence whose quantities are known to be unchanged between control and subject samples. This provides an internal standard that may be used to calibrate the PCR reaction.

In addition to the methods outlined above, the skilled person will appreciate that any technology coupling the amplification of gene-transcript specific product to the generation of a signal may also be suitable for quantitation. A preferred example employs convenient improvements to the polymerase chain reaction (US 4683195 and 4683202) that have rendered it suitable for the exact quantitation of specific mRNA transcripts by incorporating an initial reverse transcription of mRNA to cDNA. Further

key improvements enable the measurement of accumulating PCR products in real-time as the reaction progresses.

In many cases it may be preferred to assess the degree of gene expression in subject or control samples using probe molecules capable of indicating the presence of target molecules (representative of one or more of the genes listed above and in Table 2) in the relevant sample.

Probes for use in the method of the invention may be selected with reference to the product (direct or indirect) of gene expression to be investigated. Examples of suitable probes include oligonucleotide probes, antibodies, aptamers, and binding proteins or small molecules having suitable specificity.

Oligonucleotide probes constitute preferred probes suitable for use in accordance with the method of the invention. The generation of suitable oligonucleotide probes is well known to those skilled in the art (Oligonucleotide synthesis: Methods and Applications, Piet Herdewijn (ed) Humana Press (2004).). Oligonucleotide and modified oligonucleotides are commercially available from numerous companies.

For the purposes of the present invention an oligonucleotide probe may be taken to comprise an oligonucleotide capable of hybridising specifically to a nucleic acid target molecule of complementary sequence through one or more types of chemical bond. Such binding may usually occur through complementary base pairing, and usually through hydrogen bond formation. Suitable oligonucleotide probes may include natural (ie., A, G, C, or T) or modified bases (7-deazaguanosine, inosine, etc.). In addition, a linkage other than a phosphodiester bond may be used to join the bases in the oligonucleotide probe(s), so long as this variation does not interfere with hybridisation of the oligonucleotide probe to its target. Thus, oligonucleotide probes suitable for use in the methods of the invention may be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages.

The phrase "hybridising specifically to" as used herein refers to the binding, duplexing, or hybridising of an oligonucleotide probe preferentially to a particular target nucleotide sequence under stringent conditions when that sequence is present in a complex mixture (such as total cellular DNA or RNA). In one embodiment, a probe may bind, duplex or hybridise only to the particular target molecule.

The term "stringent conditions" refers to conditions under which a probe will hybridise to its target subsequence, but minimally to other sequences. In some embodiments, a probe may hybridise to no sequences other than its target under stringent conditions. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridise specifically at higher temperatures.

In general, stringent conditions may be selected to be about 5°C lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH. The Tm is the temperature (under defined ionic strength, pH, and nucleic acid concentration) at which 50% of the oligonucleotide probes complementary to a target nucleic acid hybridise to the target nucleic acid at equilibrium. As the target nucleic acids will generally be present in excess, at Tm, 50% of the probes are occupied at equilibrium. By way of example, stringent conditions will be those in which the salt concentration is at least about 0.01 to 1.0 M Na^+ ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30° C for short probes (e.g., 10 to 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide.

Oligonucleotide probes may be used to detect complementary nucleic acid sequences (i.e., nucleic acid targets) in a suitable representative sample. Such complementary binding forms the basis of most techniques in which oligonucleotides may be used to detect, and thereby allow comparison of, expression of particular genes. Preferred technologies permit the parallel quantitation of the expression of multiple genes and include technologies where amplification and quantitation of species are coupled in real-time, such as the quantitative reverse transcription PCR technologies and technologies where quantitation of amplified species occurs subsequent to amplification, such as array technologies.

Array technologies involve the hybridisation of samples, representative of gene expression within the subject or control sample, with a plurality of oligonucleotide probes wherein each probe preferentially hybridises to a disclosed gene or genes. Array technologies provide for the unique identification of specific oligonucleotide sequences, for example by their physical position (e.g., a grid in a two-dimensional array as commercially provided by Affymetrix Inc.) or by association with another feature (e.g. labelled beads as commercially provided by Illumina Inc or Luminex Inc). Oligonuleotide arrays may be synthesised *in situ* (e.g by light directed synthesis as commercially provided by Affymetrix Inc) or pre-formed and spotted by contact or ink-jet technology (as commercially provided by Agilent or Applied Biosystems). It will be apparent to those skilled in the art that whole or partial cDNA sequences may also serve as probes for array technology (as commercially provided by Clontech).

Oligonucleotide probes may be used in blotting techniques, such as Southern blotting or northern blotting, to detect and compare gene expression (for example by means of cDNA or mRNA target molecules representative of gene expression). Techniques and reagents suitable for use in Southern or northern blotting techniques will be well known to those of skill in the art. Briefly, samples comprising DNA (in the case of Southern blotting) or RNA (in the case of northern blotting) target molecules are separated according to their ability to penetrate a gel of a material such as acrylamide or agarose. Penetration of the gel may be driven by capillary action or by the activity of an electrical field. Once separation of the target molecules has been achieved these molecules are transferred to

a thin membrane (typically nylon or nitrocellulose) before being immobilized on the membrane (for example by baking or by ultraviolet radiation). Gene expression may then be detected and compared by hybridisation of oligonucleotide probes to the target molecules bound to the membrane.

In certain circumstances the use of traditional hybridisation protocols for comparing gene expression may prove problematic. For example blotting techniques may have difficulty distinguishing between two or more gene products of approximately the same molecular weight since such similarly sized products are difficult to separate using gels. Accordingly, in such circumstances it may be preferred to compare gene expression using alternative techniques, such as those described below.

Gene expression in a sample representing gene expression in a subject may be assessed with reference to global transcript levels within suitable nucleic acid samples by means of high-density oligonucleotide array technology. Such technologies make use of arrays in which oligonucleotide probes are tethered, for example by covalent attachment, to a solid support. These arrays of oligonucleotide probes immobilized on solid supports represent preferred components to be used in the methods and kits of the invention for the comparison of gene expression. Large numbers of such probes may be attached in this manner to provide arrays suitable for the comparison of expression of large numbers of genes selected from those listed above and in Table 2. Accordingly it will be recognised that such oligonucleotide arrays may be particularly preferred in embodiments of the methods of the invention where it is desired to compare expression of more than one gene selected from each of the groups of genes listed above and in Table 2.

Other suitable methodologies that may be used in the comparison of nucleic acid targets representative of gene expression include, but are not limited to, nucleic acid sequence based amplification (NASBA); or rolling circle DNA amplification (RCA).

It is usually desirable to label probes in order that they may be easily detected. Examples of detectable moieties that may be used in the labelling of probes or targets suitable for use in accordance with the invention include any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Suitable detectable moieties include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, radioactive materials and colourimetric materials. These detectable moieties are suitable for incorporation in all types of probes or targets that may be used in the methods of the invention unless indicated to the contrary.

Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, beta-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone,

fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride, phycoerythrin, texas red, rhodamine, green fluorescent protein, and the like; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin; examples of suitable radioactive material include ^{125}I, ^{131}I, ^{35}S, ^{3}H, ^{14}C, or ^{32}P; examples of suitable colorimetric materials include colloidal gold or coloured glass or plastic (e.g., polystyrene, polypropylene, latex, etc.) beads.

Means of detecting such labels are well known to the skilled person. For example, radiolabels may be detected using photographic film or scintillation counters; fluorescent markers may be detected using a photodetector to detect emitted light. Enzymatic labels are typically detected by providing the enzyme with a substrate and detecting the reaction product produced by the action of the enzyme on the substrate, and colorimetric labels are detected by simply visualizing the coloured label.

In a preferred embodiment of the invention fluorescently labelled probes or targets may be scanned and fluorescence detected using a laser confocal scanner.

In the case of labelled nucleic acid probes or targets suitable labelling may take place before, during, or after hybridisation. In a preferred embodiment, nucleic acid probes or targets for use in the methods of the invention are labelled before hybridisation. Fluorescence labels are particularly preferred and, where used, quantification of the hybridisation of the nucleic acid probes to their nucleic acid targets is by quantification of fluorescence from the hybridised fluorescently labelled nucleic acid. Quantitation may be from a fluorescently labelled reagent that binds a hapten incorporated into the nucleic acid.

In a preferred embodiment of the invention analysis of hybridisation may be achieved using suitable analysis software, such as the Microarray Analysis Suite (Affymetrix Inc.).

Effective quantification may be achieved using a fluorescence microscope which can be equipped with an automated stage to permit automatic scanning of the array, and which can be equipped with a data acquisition system for the automated measurement, recording and subsequent processing of the fluorescence intensity information. Suitable arrangements for such automation are conventional and well known to those skilled in the art.

In a preferred embodiment, the hybridised nucleic acids are detected by detecting one or more detectable moieties attached to the nucleic acids. The detectable moieties may be incorporated by any of a number of means well known to those of skill in the art. However, in a preferred embodiment, such moieties are simultaneously incorporated during an amplification step in the preparation of the sample nucleic acids (probes or targets). Thus, for example, polymerase chain reaction (PCR) using

primers or nucleotides labelled with a detectable moiety will provide an amplification product labelled with said moiety. In a preferred embodiment, transcription amplification using a fluorescently labelled nucleotide (e.g. fluorescein-labelled UTP and/or CTP) incorporates the label into the transcribed nucleic acids.

Alternatively, a suitable detectable moiety may be added directly to the original nucleic acid sample (e.g., mRNA, polyA mRNA, cDNA, etc. from the tissue of interest) or to an amplification product after amplification of the original nucleic acid is completed. Means of attaching labels such as fluorescent labels to nucleic acids are well known to those skilled in the art and include, for example nick translation or end-labelling (e.g. with a labeled RNA) by kinasing of the nucleic acid and subsequent attachment (ligation) of a nucleic acid linker joining the sample nucleic acid to a label (such as a suitable fluorophore).

Although the method of the first aspect of the invention is most suitable for use in association with human subjects it will be appreciated that it may also be useful in determining a course of treatment of viral infection in non-human animals (e.g. horses, dogs, cattle).

An alternative method of the invention comprises a method for determining the likelihood that a subject having a viral infection of the liver will be responsive to antiviral therapy that includes stimulation of Interferon (IFN) activity, the method comprising:

(a) analysing a sample from the subject for expression of at least one gene selected from the genes listed in Table 3 below.

(b) comparing expression of the genes in the sample to expression of the same genes in a control sample;

One embodiment of the method is wherein altered expression of the genes in the sample compared to expression of the same genes in the control sample indicates that the subject is not likely to be responsive to said antiviral therapy.

An alternative embodiment of the method is wherein unaltered expression of the genes in the sample compared to expression of the same genes in the control sample indicates that the subject is likely to be responsive to said antiviral therapy.
Techniques used for performing this aspect of the invention are provided above in relation to the first aspect of the invention. While the specific genes are different, the skilled person would appreciate and be able to identify target molecules to be assessed according to this method, as well as identify specific binding agents that can be used.

The inventors expanded their work investigating the differences between infected subjects that respond well to IFN treatment with those that do not, to examine IFN Induced Jak-STAT signalling.

IFN binds to interferon receptors and activates the Jak-STAT pathway. A central event in this activation is the phosphorylation of STAT1. The inventors found that STAT1 phosphorylation was induced in most subjects when they were treated with pegIFNα2b. However there seemed to be no correlation between STAT1 phosphorylation and the responsiveness of a subject to IFN treatment in an antiviral therapy. However the inventors were surprised to find that there were differences in responders and non-responders with regards the location of STAT1 when examined in samples. STAT1 is known to translocate into the nucleus and bind as a dimer to specific response elements in the promoters of ISGs. All rapid responding subjects had an IFN induced shift in STAT1 location following treatment with pegIFNα2b. In contrast, the non-responsive subjects (i.e. those with pre-activated IFN signalling) had no detectable STAT1 shifts; rather a large proportion of hepatocytes already had appreciable nuclear staining.

Therefore according to a second aspect of the invention, there is provided a method for determining the likelihood that a subject having a viral infection of the liver will be responsive to antiviral therapy that includes stimulation of Interferon (IFN) activity, the method comprising, examining a sample from the subject to identify the subcellular location of STAT1.

As set out in the accompanying examples, the inventors have determined that the location of STAT1 in liver cells is a prognostic marker for the responsiveness of a subject to antiviral therapy that includes stimulation of Interferon (IFN) activity. In that data it is shown that a large proportion of hepatocytes in liver samples taken from non-RVR subjects (i.e. non responders to antiviral therapy) already had an appreciable nuclear staining for STAT1 prior to antiviral therapy, whereas hepatocytes in liver samples from RVR subjects only have minimal nuclear staining. This totally unexpected finding is neither disclosed nor suggested in the art.

Thus, if a majority of hepatocytes in liver samples have nuclear staining for STAT1, then that subject is likely to be non-responsive to antiviral therapy that includes stimulation of Interferon (IFN) activity. Conversely, if a minimal number of the hepatocytes in liver samples have nuclear staining for STAT1 is likely to be responsive to antiviral therapy that includes stimulation of Interferon (IFN) activity.

In some embodiments, the sample is a liver sample. Also In some embodiments, the method examines the subcellular location of STAT1 in hepatocyte cells.

Methods for determining the location of STAT1 protein in liver samples are routine in the art. An example of such a method is standard immunohistochemistry using commercially available anti-STAT antibodies or other specific binding entities. The accompanying example provides a detailed method for determining the location of STAT1 protein in a liver sample. In some embodiments, the STAT1 protein examined in the method of the invention is phospho-STAT1.

By "subject" we include those subjects defined above in relation to the first aspect of the invention. In some embodiments, the subject is human.

As set out above, the present invention is based upon studies in which the inventors investigated IFN induced signaling and ISG induction in paired samples of liver biopsies and peripheral blood mononuclear cells (PBMCs) of patients with chronic hepatitis before and during therapy with pegIFNα; this is described in more detail in the accompanying Example.

The inventors established that the endogenous IFN system is constantly activated in many infected patients. Moreover, the inventors were surprised to correlate patients with a pre-activated IFN system with a poor response to IFN therapy. This finding is counter-intuitive, because one would expect that an active innate immune system would help to eliminate the virus during IFNα therapy.

The inventors analysed ISG expression in liver biopsies and further concluded that there are patients where HCV surprisingly induces (does not block) the endogenous IFN system, and there are patients where HCV does not induce (may be by cleaving TRIF and/or Cardif) the endogenous IFN system, but that this difference has no impact on the ability of HCV to maintain a chronic infection.

In patients without a pre-activation of the IFN system, the inventors found that pegIFNα2b induced within 4 hours a robust (sub-) maximal up-regulation of many ISGs in the liver. Surprisingly, such high ISG expression levels were already present in the pretreatment biopsies of patients that later did not show a rapid virological response at week 4.

It was also found that the pre-activation of the endogenous IFN system was found more frequently in liver biopsies of patients infected with HCV genotype 1 (and 4) than with genotype 2 or 3. This is intriguing because it is well known that genotype 2 and 3 infections can be cured in over 80% of the patients, compared to less than 50% of the patients with genotype 1. The inventors finding that the frequency and degree of pre-activation of the endogenous IFN system depends on the HCV genotype provides an explanation for the different treatment susceptibility of HCV genotypes.

The inventors realised that these data establish that HCV interferes with IFN signalling and thereby impairs the response to therapy. Moreover, an inhibition of IFNα signalling by HCV explains why the

strong pre-activation of the endogenous IFN system does not lead to a spontaneous elimination of HCV. The inventors do not wish to be bound by any hypothesis but believe this means that IFNα would not induce an antiviral state in the hepatocytes that are infected with HCV. The up-regulation of ISGs observed in the liver of many patients would then occur only in the non-infected hepatocytes. The strong induction of ISGs found in liver biopsies is compatible with such a model, because there are more non-infected that infected hepatocytes. IFNβ production would occur in the hepatocytes infected with a virus that is not successful in cleaving Cardif and/or TRIF. Because of the HCV induced inhibition of the Jak-STAT pathway, the secreted IFNβ would not induce an antiviral state in these infected hepatocytes, but only in non-infected neighbor cells.

The inventor realised that their new understanding of the interaction between HCV and the immune system was highly relevant to the design and selection of treatment regimens for viral infections such as HCV infection. It is therefore an aim of certain embodiments of the invention to provide novel means of treating viral infections.

According to a third aspect of the present invention, there is provided a use of an agent that reduces the activation of the IFN system for the prevention or treatment of a viral infection of the liver.

According to a fourth aspect of the present invention, there is provided an agent that reduces the activation of the IFN system in the manufacture of a medicament for the prevention or treatment of a viral infection of the liver.

The inventors, as explained above and in the Example, have realised that some subjects with a viral infection have activation of the IFN system (and associated upregulation of ISGs) and this is associated with a poor response to subsequent antiviral therapy with IFN. This lead them to realise that agents according to the third or fourth aspect of the invention, which will prevent such preactiviation, are useful for reducing the activity of the IFN pathway and will effectively "prime" a subject such that they will respond better to subsequent antiviral therapies which utilise IFN. The inventors were surprised to make these correlations because a skilled person would expect increased IFN activity in a subject to be associated with better viral clearance and not with a subset of subjects who respond poorly to treatment.

It is therefore preferred that the agents are used according to the third or fourth aspects of the invention are used to treat subjects with viral infections that also have increased (relative to uninfected control subjects) activation of IFN system.

By "reduces" we mean that agent is effective for reducing the stimulation of ISGs such that the expression levels of ISGs are not significantly different to expression levels in control tissues.

The agents may be used in the treatment of a number of different viral infections of the liver, including Hepatitis B virus and Hepatitis C virus infections. It is most preferred that the agents are used to prevent or reduce Hepatitis C Virus (HCV) infection.

Examples of agents which may be used according to the invention include where the agent may bind to the IFNα polypeptide and prevent IFN functional activity, e.g. antibodies and fragments and derivatives thereof (e.g. domain antibodies or Fabs). Alternatively the agent may act as a competitive inhibitor to IFN system by acting as an antagonist at IFNα receptors (e.g. IFNAR1, IFNAR2a, b, or c). Alternatively the agent may inhibit enzymes or other molecules in the IFN pathway. Alternatively the agent may bind to mRNA encoding IFNα polypeptide in such a manner as to lead to a reduction in that mRNA and hence a reduction in IFNα polypeptide. Alternatively the agent may bind to a nucleic sequence encoding IFNα in such a manner that it leads to a reduction in the amount of transcribed mRNA encoding IFNα polypeptide. For instance the agent may bind to coding or non-coding regions of the IFNα gene or to DNA 5' or 3' of the IFN and thereby reduce expression of the protein.

It is preferred that the agent of the third or fourth aspect of the invention binds to IFNα polypeptide, an IFNα receptor or to a nucleic acid encoding IFNα polypeptide.

There are a number of different human Interferon α polypeptide sequences. An alignment of these sequences is shown in Figure 8. From this alignment the following consensus sequence has been determined. This information can be used by the skilled person to develop a binding agent to IFNα polypeptide.

When the agents binds to IFNα polypeptide, it is preferred that the agent binds to an epitope defined by the protein that has been correctly folded into its native form. It will be appreciated, that there can be some sequence variability between species and also between genotypes. Accordingly other preferred epitopes will comprise equivalent regions from variants of the gene. Equivalent regions from further IFN polypeptides can be identified using sequence similarity and identity tools, and database searching methods, outlined above in the first aspect of the invention.

It is most preferred that the agent binds to a conserved region of the IFNα polypeptide or a fragment thereof. As can be seen from the alignment of IFNα polypeptide sequences in Figure 8, there are a number of regions of amino acid sequence which are conserved between the different polypeptides. An example of such a conserved region would be positions 161 to 174 of the "consensus" sequence shown in that figure.

Agents which bind to such a region have a particularly dramatic effect on IFNα activity and are therefore particularly effective for preventing pre-activation of the IFN system and thereby improving elimination of HCV from subjects receiving antiviral therapy.

When the agents binds to an IFN receptor, it is preferred that the agent binds to and inhibits the binding of IFNα to the IFN receptor.

There are a number of different Interferon receptors. The amino acid sequences of these are shown in Figure 9. This information can be used by the skilled person to develop a binding agent to IFN receptor polypeptide.

It is preferred that the agent binds to an epitope on the receptor defined by the IFN receptor protein that has been correctly folded into its native form. It will be appreciated, that there can be some sequence variability between species and also between genotypes. Accordingly other preferred epitopes will comprise equivalent regions from variants of the receptor gene. Equivalent regions from further IFN polypeptides can be identified using sequence similarity and identity tools, and database searching methods, outlined above in the first aspect of the invention.

An embodiment of the third or fourth aspects of the invention is wherein the agent is an antibody or fragment thereof.

The use of antibodies as agents to modulate polypeptide activity is well known. Indeed, therapeutic agents based on antibodies are increasingly being used in medicine. As set out above, the inventors realised that an antibody may be used to neutralise IFN system by binding thereto or may act as an inhibitor of an IFN receptor. It is therefore apparent that such agents have great utility as medicaments for the improving the treatment of HCV infections. Moreover, such antibodies can be used in the prognostic methods set out above in further aspects of the invention.

Antibodies, for use in treating human subjects, may be raised against:

(a) IFNα polypeptide *per se* or a number of peptides derived from the IFNα polypeptide, or peptides comprising amino acid sequences corresponding to those found in the IFNα polypeptide; or

(b) the IFN receptor or a number of peptides derived from the IFN receptor, or peptides comprising amino acid sequences corresponding to those found in the IFN receptor.

It is preferred that the antibodies are raised against antigenic structures from human IFNα polypeptide, the human IFN receptor and peptide derivatives and fragments thereof.

Antibodies may be produced as polyclonal sera by injecting antigen into animals. Preferred polyclonal antibodies may be raised by inoculating an animal (e.g. a rabbit) with antigen (e.g. all or a fragment of the IFNα polypeptide) using techniques known to the art.

Alternatively the antibody may be monoclonal. Conventional hybridoma techniques may be used to raise such antibodies. The antigen used to generate monoclonal antibodies for use in the present invention may be the same as would be used to generate polyclonal sera.

In their simplest form, antibodies or immunoglobulin proteins are Y-shaped molecules usually exemplified by the γ-immunoglobulin (IgG) class of antibodies. The molecule consists of four polypeptide chains two identical heavy (H) chains and two identical (L) chains of approximately 50kD and 25kD each respectively. Each light chain is bound to a heavy chain (H-L) by disulphide and non-covalent bonds. Two identical H-L chain combinations are linked to each other by similar non-covalent and disulphide bonds between the two H chains to form the basic four chain immunoglobulin structure $(H\text{-}L)_2$.

Light chain immunoglobulins are made up of one V-domain (V_L) and one constant domain (C_L) whereas heavy chains consist of one V-domain and, depending on H chain isotype, three or four C-domains (C_H1, C_H2, C_H3 and C_H4).

At the N-terminal region of each light or heavy chain is a variable (V) domain that varies greatly in sequence, and is responsible for specific binding to antigen. Antibody specificity for antigen is actually determined by amino acid sequences within the V-regions known as hypervariable loops or Complementarity Determining Regions (CDRs). Each H and L chain V regions possess 3 such CDRs, and it is the combination of all 6 that forms the antibody's antigen binding site. The remaining V-region amino acids which exhibit less variation and which support the hypervariable loops are called frameworks regions (FRs).

The regions beyond the variable domains (C-domains) are relatively constant in sequence. It will be appreciated that the characterising feature of antibodies according to the invention is the V_H and V_L domains. It will be further appreciated that the precise nature of the C_H and C_L domains is not, on the whole, critical to the invention. In fact preferred antibodies for use in the invention may have very different C_H and C_L domains. Furthermore, as discussed more fully below, preferred antibody functional derivatives may comprise the Variable domains without a C-domain (e.g. scFV antibodies).

Preferred antibodies considered to be agents according to the third or fourth aspect of the invention may have the V_L (first domain) and V_H (second domain) domains. A derivative thereof may have 75%

sequence identity, for isntance 90% sequence identity or at least 95% sequence identity. It will be appreciated that most sequence variation may occur in the framework regions (FRs) whereas the sequence of the CDRs of the antibodies, and functional derivatives thereof, should be most conserved.

A number of preferred embodiments of the agent of the third or fourth aspects of the invention relate to molecules with both Variable and Constant domains. However it will be appreciated that antibody fragments (e.g. scFV antibodies or FAbs) are also encompassed by the invention that comprise essentially the Variable region of an antibody without any Constant region.

An scFV antibody fragment considered to be an agent of the third or fourth aspect of the invention may comprise the whole of the V_H and V_L domains of an antibody raised against IFN polypeptide. The V_H and V_L domains may be separated by a suitable linker peptide.

Antibodies, and particularly mAbs, generated in one species are known to have several serious drawbacks when used to treat a different species. For instance when murine antibodies are used in humans they tend to have a short circulating half-life in serum and may be recognised as foreign proteins by the immune system of a patient being treated. This may lead to the development of an unwanted human anti-mouse antibody (HAMA) response. This is particularly troublesome when frequent administration of an antibody is required as it can enhance its clearance, block its therapeutic effect, and induce hypersensitivity reactions. These factors limit the use of mouse monoclonal antibodies in human therapy and have prompted the development of antibody engineering technology to generate humanised antibodies.

Therefore, where the antibody capable of reducing IFN activity is to be used as a therapeutic agent for treating HCV infections in a human subject, then it is preferred that antibodies and fragments thereof of non-human source are humanised.

Humanisation may be achieved by splicing V region sequences (e.g. from a monoclonal antibody generated in a non-human hybridoma) with C region (and ideally FRs from V region) sequences from human antibodies. The resulting 'engineered' antibodies are less immunogenic in humans than the non-human antibodies from which they were derived and so are better suited for clinical use.

Humanised antibodies may be chimaeric monoclonal antibodies, in which, using recombinant DNA technology, rodent immunoglobulin constant regions are replaced by the constant regions of human antibodies. The chimaeric H chain and L chain genes may then be cloned into expression vectors containing suitable regulatory elements and induced into mammalian cells in order to produce fully glycosylated antibodies. By choosing an appropriate human H chain C region gene for this process,

the biological activity of the antibody may be pre-determined. Such chimaeric molecules may be used to treat or prevent cancer according to the present invention.

Further humanisation of antibodies may involve CDR-grafting or reshaping of antibodies. Such antibodies are produced by transplanting the heavy and light chain CDRs of a non-human antibody (which form the antibody's antigen binding site) into the corresponding framework regions of a human antibody.

Humanised antibody fragments represent preferred agents for use according to the invention. Human FAbs recognising an epitope on IFNα polypeptide or an IFN receptor may be identified through screening a phage library of variable chain human antibodies. Techniques known to the art (e.g as developed by Morphosys or Cambridge Antibody Technology) may be employed to generate Fabs that may be used as agents according to the invention. In brief a human combinatorial Fab antibody library may be generated by transferring the heavy and light chain variable regions from a single-chain Fv library into a Fab display vector. This library may yield 2.1 x 10^{10} different antibody fragments. The peptide may then be used as "bait" to identify antibody fragments from then library that have the desired binding properties.

Domain antibodies (dAbs) represent another preferred agent that may be used according to this embodiment of the invention. dAbs are the smallest functional binding unit of antibodies and correspond to the variable regions of either the heavy or light chains of human antibodies. Such dAbs may have a molecule weight of around 13kDa (corresponding to about 1/10 (or less) the size of a full antibody).

According to another embodiment of the third and fourth aspects of the invention, peptides may be used to reduce IFNα polypeptide activity. Such peptides represent other preferred agents for use according to the invention. These peptides may be isolated, for example, from libraries of peptides by identifying which members of the library are able to reduce the activity or expression of IFN α polypeptide. Suitable libraries may be generated using phage display techniques.

Aptamers represent another preferred agent of the third or fourth aspect of the invention. Aptamers are nucleic acid molecules that assume a specific, sequence-dependent shape and bind to specific target ligands based on a lock-and-key fit between the aptamer and ligand. Typically, aptamers may comprise either single- or double-stranded DNA molecules (ssDNA or dsDNA) or single-stranded RNA molecules (ssRNA). Aptamers may be used to bind both nucleic acid and non-nucleic acid targets. Accordingly aptamers may be generated that recognise and so reduce the activity or expression of IFNα. Suitable aptamers may be selected from random sequence pools, from which specific aptamers may be identified which bind to the selected target molecules with high affinity.

Methods for the production and selection of aptamers having desired specificity are well known to those skilled in the art, and include the SELEX (systematic evolution of ligands by exponential enrichment) process. Briefly, large libraries of oligonucleotides are produced, allowing the isolation of large amounts of functional nucleic acids by an iterative process of *in vitro* selection and subsequent amplification through polymerase chain reaction.

Antisense molecules represent another preferred agent for use according to the third or fourth aspects of the invention. Antisense molecules are typically single-stranded nucleic acids, which can specifically bind to a complementary nucleic acid sequence produced by a gene and inactivate it, effectively turning that gene "off". The molecule is termed "antisense" as it is complementary to the gene's mRNA, which is called the "sense" sequence, as appreciated by the skilled person. Antisense molecules are typically are 15 to 35 bases in length of DNA, RNA or a chemical analogue. Antisense nucleic acids have been used experimentally to bind to mRNA and prevent the expression of specific genes. This has lead to the development of "antisense therapies" as drugs for the treatment of cancer, diabetes and inflammatory diseases. Antisense drugs have recently been approved by the US FDA for human therapeutic use. Accordingly, by designing an antisense molecule to polynucleotide sequence encoding IFN polypeptide it would be possible to reduce the expression of IFNα polypeptide in a cell and thereby reduce in IFNα activity and reduce the preactiviation seen in HCV infection. A polynucleotide sequence encoding an IFNα polypeptide is provided in Figure 8.

Small interfering RNA (siRNA), sometimes known as short interfering RNA or silencing RNA, represent further preferred agents for use according to the third or fourth aspects of the invention. As set out above, the inventors realised that preactivation of the IFN system is associated with a resistance to antiviral therapy. It is therefore apparent that siRNA molecules that can reduce IFNα expression have great utility in the preparation of medicaments for the treatment of HCV infection. siRNA are a class of 20-25 nucleotide-long RNA molecules are involved in the RNA interference pathway (RNAi), by which the siRNA can lead to a reduction in expression of a specific gene, or specifically interfere with the translation of such mRNA thereby inhibiting expression of protein encoded by the mRNA. siRNAs have a well defined structure: a short (usually 21-nt) double-strand of RNA (dsRNA) with 2-nt 3' overhangs on either end. Each strand has a 5' phosphate group and a 3' hydroxyl (-OH) group. *In vivo* this structure is the result of processing by Dicer, an enzyme that converts either long dsRNAs or hairpin RNAs into siRNAs. siRNAs can also be exogenously (artificially) introduced into cells by various transfection methods to bring about the specific knockdown of a gene of interest. Essentially any gene of which the sequence is known can thus be targeted based on sequence complementarity with an appropriately tailored siRNA. Given the ability to knockdown essentially any gene of interest, RNAi via siRNAs has generated a great deal of interest in both basic and applied biology. There is an increasing number of large-scale RNAi screens that are designed to identify the important genes in various biological pathways. As disease processes also

depend on the activity of multiple genes, it is expected that in some situations turning off the activity of a gene with a siRNA could produce a therapeutic benefit. Hence their discovery has led to a surge in interest in harnessing RNAi for biomedical research and drug development. Recent phase I results of therapeutic RNAi trials demonstrate that siRNAs are well tolerated and have suitable pharmacokinetic properties. siRNAs and related RNAi induction methods therefore stand to become an important new class of drugs in the foreseeable future. siRNA molecules designed to nucleic acid encoding IFNα polypeptide can be used to reduce the expression of IFNα and therefore result in a reduction in the preactivation of the IFN system. Hence an embodiment of this aspect of the invention is wherein the agent is a siRNA molecule having complementary sequence to IFNα polynucleotide.

A polynucleotide sequence encoding an IFNα polypeptide is provided in Figure 8.

Using such information it is straightforward and well within the capability of the skilled person to design siRNA molecules having complementary sequence to IFNα polynucleotide. For example, a simple internet search yields many websites that can be used to design siRNA molecules.

By "siRNA molecule" we include a double stranded 20 to 25 nucleotide-long RNA molecule, as well as each of the two single RNA strands that make up a siRNA molecule.

It is most preferred that the siRNA is used in the form of hairpin RNA (shRNA). Such shRNA may comprise two complementary siRNA molecules that are linked by a spacer sequence (e.g. of about 9 nueclotides). The complementary siRNA molecules may fold such that they bind together.

A ribozyme capable of cleaving RNA or DNA encoding IFNα polypeptide represent another preferred agent of the third or fourth aspect of the invention.

It is preferred that the agent of the third or fourth aspect of the invention is able to reduce the activation of the IFN system in a subject to be treated but not to reduce the activity of subsequent antiviral therapy supplied to the subject.

For example, where the agent of the third or fourth aspect of the invention is an antibody or fragment thereof, then it is preferred that the agent can bind to and reduce the activity of endogenous IFNα polypeptide but not exogenously supplied IFNα polypeptide. It is possible to derive such antibodies using methods routine in the art, and the information provided previously in this aspect of the invention.

It will be appreciated that the amount of an agent needed according to the invention is determined by biological activity and bioavailability which in turn depends on the mode of administration and the physicochemical properties of the agent. The frequency of administration will also be influenced by the abovementioned factors and particularly the half-life of the agent within the target tissue or subject being treated.

Known procedures, such as those conventionally employed by the pharmaceutical industry (e.g. *in vivo* experimentation, clinical trials etc), may be used to establish specific formulations of the agents and precise therapeutic regimes (such as daily doses and the frequency of administration).

Generally, a daily dose of between 0.01μg/kg of body weight and 0.1g/kg of body weight of an agent may be used in a treatment regimen for treating HCV infection; for instance the daily dose is between 0.01mg/kg of body weight and 100mg/kg of body weight.

By way of example a suitable dose of an antibody according to the invention is 10μg/kg of body weight – 100mg/kg of body weight, for instance about 01mg/kg of body weight – 10mg/kg of body weight and in some embodiments about 6mg/kg of body weight.

Daily doses may be given as a single administration (e.g. a single daily injection or a single dose from an inhaler). Alternatively the agent (e.g. an antibody or aptamer) may require administration twice or more times during a day.

Medicaments according to the invention should comprise a therapeutically effective amount of the agent and a pharmaceutically acceptable vehicle.

A "therapeutically effective amount" is any amount of an agent according to the invention which, when administered to a subject inhibits or prevents cancer growth or metastasis.

A "subject" may be a vertebrate, mammal, domestic animal or human being. It is preferred that the subject to be treated is human. When this is the case the agents may be designed such that they are most suited for human therapy (e.g. humanisation of antibodies as discussed above). However it will also be appreciated that the agents may also be used to treat other animals of veterinary interest (e.g. horses, dogs or cats).

A "pharmaceutically acceptable vehicle" as referred to herein is any physiological vehicle known to those skilled in the art as useful in formulating pharmaceutical compositions.

In one embodiment, the medicament may comprise about 0.01 µg and 0.5 g of the agent. The amount of the agent in the composition can be between 0.01 mg and 200 mg, for instance, between approximately 0.1mg and 100 mg, or between about 1mg and 10mg. Hence, the composition can comprise between approximately 2mg and 5mg of the agent.

In some embodiments,, the medicament comprises approximately 0.1% (w/w) to 90% (w/w) of the agent, and in some embodiments,, 1% (w/w) to 10% (w/w). The rest of the composition may comprise the vehicle.

Nucleic acid agents can be delivered to a subject by incorporation within liposomes, Alternatively the "naked" DNA molecules may be inserted into a subject's cells by a suitable means e.g. direct endocytotic uptake. Nucleic acid molecules may be transferred to the cells of a subject to be treated by transfection, infection, microinjection, cell fusion, protoplast fusion or ballistic bombardment. For example, transfer may be by ballistic transfection with coated gold particles, liposomes containing the DNA molecules, viral vectors (e.g. adenovirus) and means of providing direct DNA uptake (e.g. endocytosis) by application of the DNA molecules directly to the target tissue topically or by injection.

The antibodies, or functional derivatives thereof, may be used in a number of ways. For instance, systemic administration may be required in which case the antibodies or derivatives thereof may be contained within a composition which may, for example, be ingested orally in the form of a tablet, capsule or liquid. It is preferred that the antibodies, or derivatives thereof, are administered by injection into the blood stream. Injections may be intravenous (bolus or infusion) or subcutaneous (bolus or infusion). Alternatively the antibodies may be injected directly to the liver.

Nucleic acid or polypeptide therapeutic entities may be combined in pharmaceutical compositions having a number of different forms depending, in particular on the manner in which the composition is to be used. Thus, for example, the composition may be in the form of a powder, tablet, capsule, liquid, ointment, cream, gel, hydrogel, aerosol, spray, micelle, transdermal patch, liposome or any other suitable form that may be administered to a person or animal. It will be appreciated that the vehicle of the composition of the invention should be one which is well tolerated by the subject to whom it is given, and can enable delivery of the therapeutic to the target cell, tissue, or organ.

In a preferred embodiment, the pharmaceutical vehicle is a liquid and the pharmaceutical composition is in the form of a solution. In another embodiment, the pharmaceutical vehicle is a gel and the composition is in the form of a cream or the like.

Compositions comprising such therapeutic entities may be used in a number of ways. For instance, systemic administration may be required in which case the entities may be contained within a

composition that may, for example, be ingested orally in the form of a tablet, capsule or liquid. Alternatively, the composition may be administered by injection into the blood stream. Injections may be intravenous (bolus or infusion) or subcutaneous (bolus or infusion). The entities may be administered by inhalation (e.g. intranasally).

Therapeutic entities may also be incorporated within a slow or delayed release device. Such devices may, for example, be inserted on or under the skin, and the compound may be released over weeks or even months. Such devices may be particularly advantageous when long term treatment with an entity is required and which would normally require frequent administration (e.g. at least daily injection).

The agents of the first aspect of the invention are particularly useful for pretreating patients about to undergo treatment with antiviral therapy with IFN (e.g. pegIFN) and an antiviral agent such as ribavirin. It is therefore preferred that the agent is administered to a virally infected individual before therapy with IFN and ribavirin is initiated.

The length of time between the pre-treatment with the agents defined in relation to the third and fourth aspect of the invention and the antiviral therapy can depend on the agents used. For example, where the agent is able to reduce the activation of the IFN system in a subject to be treated but not to reduce the activity of subsequent antiviral therapy supplied to the subject, then the length of time can be very short. For example, the subject could be treated concurrently, or even with a combined treatment regime.

If the agent is not distinguishing, then the length of time can depend on the nature of the agent. For example, it is known that exogenously supplied antibody takes around 4 to 6 weeks in order to be cleared from the human body. Therefore, where the agent is an antibody to the IFNα polypeptide or receptor, or other such member of the IFN system, then the subsequent antiviral therapy can be supplied to the patient 4 to 6 weeks later, for example at least 6 weeks.

The various elements required for a technician to perform the method of the first aspect of the invention may be incorporated in to a kit.

Thus, according to a fifth aspect of the invention there is provided a kit for determining the likelihood that a subject having a viral infection of the liver will be responsive to antiviral therapy that includes stimulation of Interferon (IFN) activity, comprising:

(i) means for analysing in a sample from a subject the expression of at least one gene from each of the groups of genes listed above and shown in Table 2; and, optionally,

(ii) means for comparing expression of the genes in the sample to expression of the same genes in a control sample.

By "means for analysing in a sample from a subject the expression of at least one gene from each of the groups of genes listed above and shown in Table 2" we include the specific binding molecules given in the first aspect of the invention that can target molecules representative of gene expression in the sample. In some embodiments, the specific binding molecule is an oligonucleotide probe, antibody, aptamers, or binding proteins or small molecules mentioned above.

By "means for comparing expression of the genes in the sample to expression of the same genes in a control sample" we include the control samples mentioned above in the first aspect of the invention. We also include the control reference data mentioned therein.

The kit of the fifth aspect of the invention may also comprise:

(iii) relevant buffers and regents for analysing the expression of said genes.

The buffers and regents provided with the kit may be in liquid form and in some embodiments, provided as pre-measured aliquots. Alternatively, the buffers and regents may be in concentrated (or even powder form) for dilution.

The various elements required for a technician to perform the method of the second aspect of the invention may be incorporated in to a kit.

Thus, according to a sixth aspect of the invention there is provided a kit for determining the likelihood that a subject having a viral infection of the liver will be responsive to antiviral therapy that includes stimulation of Interferon (IFN) activity, comprising means for examining a sample from the subject to identify the subcellular location of STAT1.

By "means for examining a sample from the subject to identify the subcellular location of STAT1" we include the specific binding molecules given in the second aspect of the invention that can identify the subcellular location of STAT1. In some embodiments, said specific binding molecule is an anti-STAT antibody; in some embodiments, an anti-phospho-STAT1 antibody.

The kit of the sixth aspect of the invention may also comprise:

(iii) relevant buffers and regents for identifying the subcellular location of STAT1.

The buffers and reagents provided with the kit may be in liquid form and in some embodiments, provided as pre-measured aliquots. Alternatively, the buffers and regents may be in concentrated (or even powder form) for dilution.

All of the features described herein (including any accompanying claims, abstract and drawings), and/or all of the steps of any method or process so disclosed, may be combined with any of the above aspects in any combination, except combinations where at least some of such features and/or steps are mutually exclusive.

The invention will now be further described with reference to the following Example and figures in which:

Figure 1. PegIFN-α2b induced regulation of gene expression in liver and PBMCs.
(A) Rapid responders up- or down-regulate significantly more genes in the liver in response to pegIFN-α2b than non-RR patients. Shown are the mean (+SEM) number of genes changed more than 2 fold at significance levels $p<0.01$ (lanes 1,2, 5 and 6) and $p<0.05$ (lanes 3,4,7 and 8) in >75% of patients in liver biopsies and PBMCs. The differences between RR and non-RR patient groups are significant in liver biopsies but not in PBMCs (p values of Mann Whitney tests shown in figure).
(B) Venn diagram of genes significantly ($p < 0.05$) up- or down-regulated >2-fold in response to pegIFN-α in > 50% of the 6 non-RR and 6 randomly selected RR biopsy samples.
(C) Venn diagram of genes significantly ($p < 0.05$) up- or down-regulated >2-fold in response to pegIFN-α in biopsy and PBMC samples of >50% of 6 randomly selected RR patients.

Figure 2. PegIFN-α2b induced gene regulation in HCV-infected patients shows major differences between livers of RVR and non-RVR patients and between liver and PBMCs.
(A) Five ISGs (Mx1, viperin, Mda5/helicard, OAS1, USP18) were chosen from the list of genes significantly ($p<0.05$) regulated >2-fold between B-1 and B-2 in RR patients. In the liver of non-RR patients, expression of these genes is already high before treatment (lanes 25-30), and does not further increase after pegIFNα (lanes 31-36). In RR patients, pre-treatment expression (lanes 5-14) is similar to controls (lanes 1-4), and pegIFNα induces a strong upregulation (lanes 15-24). No pre-activation is found in PBMCs (lanes 37-46 and 57-62), and pegIFNα strongly induces these genes in both RR and non-RR patients (lanes 47-56 and 63-68). The y-axes display absolute expression values.
(B) An example of a gene (CCL8) upregulated in liver in response to pegIFN-α2b in both RR and non-RR patients. The expression values in PBMCs are shown in lanes 37-68).

Figure 3. RT-qPCR analysis of selected ISGs and of the catalytic subunit of PP2A.

(A) RT-qPCR analysis of the USP18 mRNA corroborates the array data. Depicted is the fold induction of USP18 mRNA between B-1 and B-2 in individual patients.

(B) The expression level of selected ISGs in pre-treatment biopsies is lower in patients with early virological response (EVR = more than 2 log drop of viral load at week 12) than in patients with primary non-response (PNR = less than 2 log drop of viral load at week 12).

(C) Both within the group of patients with genotype 1 and 4 ("difficult"-to-treat) and the group with genotype 2 and 3 ("easy"-to-treat) the PNR patients have higher pre-treatment expression levels of USP18 and IFI27.

In panels B and C the y axis shows expression relative to that of GAPDH. Statistical significance was tested with the Mann-Whitney test. *N* = number of patients in each group.

(D) Patients with sustained virological response (SVR = undetectable HCV RNA 6 months after end of treatment) or end-of-treatment response (EoTR) show significantly lower expression of USP18 and IFI27 than patients with PNR or no EoTR.

Figure 4. Analysis of Jak-STAT signaling in liver biopsies.

(A) STAT1 phosphorylation in extracts of liver biopsies collected before (B-1) and after (B-2) pegIFN-α2b injection. Extracts were analyzed by Western blot using antibodies specific for PY(701)-STAT1. The signals were quantified using the Odyssey Imaging Software to calculate the integrated intensity (kilo counts x mm^2). The values represent the fold increase of phosphorylation in B-2 samples. RR-patient numbers are shown in blue, non-RR patients in red. Blots were stripped and reprobed for total STAT1 used as a loading control for each pair of samples.

(B) Representative examples of B-1 and B-2 of RR and non-RR patients. No nuclear staining is evident in pre-treatment biopsies of RR patients (Pat. 4). The light blue color of nuclei originates from the counterstaining with haematoxilin. 4 h after pegIFNα, most hepatocytes show a strong nuclear staining. In non-RR patients (Pat. 12), a weak nuclear staining is already present in pre-treatment biopsies, and pegIFNα induces little changes in hepatocytes. The visible increased nuclear staining is confined to Kupffer cells.

Figure 5. The predominant pattern of gene expression in all patient biopsy samples is shown as a heat map.

The map was generated using a list of 252 genes that are altered > 2 fold in > 50% of all RRs with a p value of <0.05. The color coding of the raw expression values is shown on the left. Many genes have a low expression level in the control patients and the pre-treatment biopsies of the RR patients (B-1).

In RR patients, pegIFNα induces an upregulation (B-2). In non-RR patients, many of the genes are already strongly induced in the pre-treatment biopsy samples (B-1), and no further induction is then found after pegIFNα (B-2).

Figure 6. Supervised classifier prediction in liver biopsy samples and PBMCs with response to treatment at week 4 as grouping criterion.
(A) Supervised classifier prediction using the B-1 biopsies of the two response groups resulted in a list of 29 genes (33 transcripts) as best predictors of treatment outcome with a misclassification rate of 4.3%.
(B) Supervised classifier prediction using the B-2 biopsies of the two response groups revealed a list of 16 genes (16 transcripts) as best predictors of treatment outcome with a misclassification rate of 19.5%.
(C and D) Supervised classifier prediction of PBMC-1 and PBMC-2 samples did not generate a useful list of predictive genes with any of the 4 statistical tests used (Support Vector Machine, Sparse Linear Discriminant Analysis, Fisher Linear Discriminant Analysis, K Nearest Neighbors). The misclassification rates were 38.5% for PBMC-1 and 42.6% for PBMC-2.

Figure 7
(A) **Semiquantitative assessment of immunohistochemical staining of phospho-STAT1 in liver biopsies.** Nuclear staining of hepatocytes was quantified by repeated counting (5 times) in 200 hepatocytes in B-1 (blue) and B-2 (red) samples of the indicated patients (patient numbers correspond to the numbers in table 1). In five out of six non-RR patients, a considerable proportion of hepatocytes had a weak but clear nuclear staining already in the pre-treatment biopsies. All the RR patients had no phospho-STAT1 signals in the nuclei before treatment, but showed a strong induction after pegIFNα.

(B) **The induction of STAT-DNA binding in response to pegIFNα2b is impaired in most of the non-RR patients.** Nuclear extracts from B-1 and B-2 samples were analyzed with EMSAs using the radiolabeled SIE-m67 oligonucleotide probe. The asterisk (*) depicts the signal of the activated STAT1 dimers that have bound the oligonucleotide sequence. The numbers above the gel shift panels represent the patient numbers as already used in table 1. The upper panel shows the 10 patients with a rapid response at week 4 (numbers 1-10). The lower panel shows the 6 non-RR patients (numbers 11-16).

Figure 8: Amino acid and nucleotide sequences of human Interferon α.

Figure 9: Amino acid and nucleotide sequence of human Interferon Receptor 1.

Figure 10: Amino acid and nucleotide sequence of human Interferon Receptor 2.

Figure 11: Amino acid and nucleotide sequence of human Interferon Receptor 2b.

Figure 12: Amino acid and nucleotide sequence of human Interferon Receptor 2c.

Example 1

1.1 Methods

1.1.1 Subject samples and treatment

From January 2006 to April 2007 patients with CHC referred to the outpatient liver clinic of the University Hospital Basel were asked for permission to use part of their diagnostic liver biopsy (B-1) for research purposes. The patients who then were treated with pegylated-IFNα2b (PegIntron) and ribavirin (Rebetol, both from Essex Chemie AG, Switzerland) were asked to participate in this study. 16 patients agreed to undergo a second liver biopsy (B-2) 4 hours after the first injection of 1.5µg/kg body weight pegIFNα2b (PegIntron). All of them were Caucasians. The first dose of ribavirin was given after this second biopsy to avoid further confounding factors. The protocol was approved by the Ethics Committee of the University Hospital Basel. Written informed consent was obtained from all patients. Blood for PBMC isolation was collected before treatment and 4 h after the first pegIFNα2b injection. The patients were treated with pegIFNα2b (1.5 µg/kg body weight) and ribavirin (weight based dosing: <65 kg: 800 mg/d; 65-85 kg: 1 g/d; >85 kg: 1.2 g/d). HCV RNA was quantified before treatment initiation and at weeks 4 and 12 of the treatment. Treatment duration was 24 weeks for patients with genotypes 2/3 and 48 weeks for genotype 1. As non-CHC controls, 4 patients who underwent ultrasound-guided liver biopsies of focal lesions gave informed consent for a biopsy from the normal liver tissue outside the focal lesion. Pre-treatment liver biopsies from 96 additional patients (all but one of them were Caucasians) with CHC were used for RT-qPCR for selected ISGs.

Paired human liver biopsy samples from the16 chronically infected HCV subjects were obtained. From January 2006 to April 2007, all subjects with chronic hepatitis C referred to the outsubject liver clinic of the University Hospital Basel were asked for their permission to use part of their diagnostic liver biopsy for research purposes. Liver biopsies were obtained by ultrasound-guided technique using a coaxial needle.

After removal of two 20- to 25-mm long biopsy specimens for routine histopathological workup for grading and staging of the liver disease according to the Metavir scoring system, the remaining 5- to 20-mm long biopsy cylinders were labeled as B1 (for biopsy 1) and stored as pretreatment samples of future study participants. Pegylated-IFNα2b (Essex Chemie AG, Switzerland) was prescribed to all

subjects participating in this study. Second biopsy (B2) was performed 4 hours following the first pegIFNα2b injection. The first dose of ribavirin was given after this second biopsy to avoid further confounding factors. The protocol was approved by the Ethics Committee of the University Hospitals in Basel. Written informed consent was obtained from all subjects.

In addition, blood for peripheral blood mononuclear cell (PBMC) isolation was collected before treatment and 4 hours after the first pegIFNα2b injection.

The HCV subjects underwent a standard combination treatment with pegIFNα2b (1.5 µg/kg body weight) and ribavirin (weight based dosing: <65 kg: 800 mg/d; 65-85 kg: 1 g/d; <85 kg: 1,2 g/d). HCV-RNA was quantified before treatment initiation, at week 4 and week 12 of the treatment (Table 1). Treatment duration is 24 weeks for subjects with genotypes 2/3 and 48 weeks for genotype 1. From the 16 subjects included in the study, 2 subjects (Nr. 10 and 16) had a primary non-response and treatment was stopped at week 12. From the remaining 9 subjects, 2 (Nr. 1, 2) have accomplished the therapy with an end of treatment response.

As non-HCV controls, two subjects that underwent ultrasound-guided liver biopsies of focal lesions (metastasis of carcinomas) gave informed consent for a biopsy from the normal liver tissue outside the focal lesion. Again, a part of the biopsy was used for routine histopathological diagnosis, and the remaining tissue for the extraction of RNA, as described later. Both control samples showed confirmed absence of liver disease in the routine histopathological workup.

1.1.2 Measurement of IFN alpha serum concentrations

Pretreatment hIFNα serum levels and the serum concentration of pegIFNα2b 4 h after the first injection were measured using the human interferon alpha ELISA kit from PBL Biomedical Laboratories according to manufacturer's instructions. This ELISA kit has previously been shown to recognize both unpegylated and pegylated human IFNα [34].

1.1.3 Preparation of extracts from human liver biopsies

Liver biopsy samples were used for the preparation of whole cell, cytoplasmic and nuclear extracts. For whole cell extracts, samples were dounce homogenized in 100 µl of lysis buffer containing 100 mmol/l NaCl, 50 mmol/l Tris pH 7.5, 1 mmol/l EDTA, 0.1% Triton X-100, 10 mmol/l NaF, 1 mmol/l phenylmethyl sulfonyl fluoride, and 1 mmol/l vanadate. Lysates were centrifuged at 14,000 rpm at 4°C for 5 minutes. Protein concentration was determined by Lowry (BioRad Protein Assay).

For nuclear and cytoplasmic extracts, livers were lysed in a low-salt buffer containing 200 mmol/l Hepes pH 7.6, 10 mmol/l KCl, 1 mmol/l EDTA, 1 mmol/l EGTA, 0.2% NP-40, 10% glycerol, and 0.1 mmol/l vanadate. After centrifugation at 15,000 rpm for 5 minutes, the pellet was resuspended in high-

salt buffer (low-salt buffer supplemented with 420 mmol/L NaCl). After centrifugation, aliquots of nuclear extracts were made for electrophoretic mobility shift assays (EMSAs).

1.1.4 Western blots and Electrophoretic Mobility Shift Assays

10 µg of total protein from human liver lysates was loaded for sodium dodecyl sulphate–polyacrylamide gel electrophoresis and transfered onto a nitrocellulose membrane (Schleicher & Schuell, Bottmingen, Switzerland). The membranes were blocked in 3% BSA/milk (1:1)-0.1% Triton X-100 for 1 hour, washed with Tris-buffered saline Tween-20 (TBST), and incubated with the primary antibody overnight at 4°C.

Proteins were detected with primary antibodies specific to phosphorylated STAT1 (PY(701)-STAT1; Cell Signaling, Bioconcept, Allschwil, Switzerland) and STAT1 (carboxy-terminus; Transduction Laboratories, BD Biosciences, Pharmingen). After 3 washes with TBST, the membranes were incubated with infrared fluorescent secondary goat anti-mouse (IRDye 680) or anti-rabbit (IRDye 800) antibodies (both from LI-COR Biosciences) for 1 hour at room temperature. Blots were analyzed by the Odyssey Infrared Imaging System from LI-COR. The infrared image was obtained in a single scan and the signal was quantified using the integrated intensity.

For loading controls, the membranes were stripped and incubated with anti-β-Actin antibody (Sigma).

EMSAs were performed using 2 µg of nuclear extracts and ^{32}P-radiolabeled DNA-oligonucleotide serum inducible element (SIE)-m67 corresponding to STAT response element sequences [25].

1.1.5 Immunohistochemistry

Standard indirect immunoperoxidase procedures were used for immunohistochemistry (ABC-Elite, Vectra Laboratories). 4-mm-thick sections were cut from paraffin blocks, rehydrated, pretreated (20' in ER2 solution) incubated with a monoclonal rabbit antibody against phospho-STAT1 (dilution 1:200, #9167 Cell Signaling) and counterstained with haematoxilin. The whole staining procedure (dehydration, pre-treatment, incubation, counterstaining and mounting) was performed with an automated stainer (Bond®, Vision BioSystems Europe, Newcastle-upon-Tyne, UK). For quantification of nuclear phospho-STAT1 staining, 5 times 200 hepatocytes were counted for each B-1 and B-2 sample of each patient. In supplementary Fig. 3, the mean values with the standard deviations are shown.

1.1.6 RNA isolation and Microarray analysis

Total RNA was extracted from liver and PBMC samples using the RNeasy Mini Kit (Qiagen) according to manufacturer's instructions. RNA was aliquoted and stored at - 80°C. Gene expression was assessed in liver and PBMCs by microarray analysis using Affymetrix Human Genome U133 Plus 2.0

arrays representing over 56,000 transcripts and variants with 11 perfect-match/mis-match probe pairs per transcript. The microarray hybridizations were performed at the functional genomics facility of the Friedrich Miescher Institute for Biomedical Research in Basel. Total RNA (1-2 µg) from each sample was reverse transcribed and biotinylated using the Affymetrix 1-cycle amplification kit as per manufacturer's instructions. Biotinylated cRNA (20 µg) was fragmented by heating with magnesium (as per Affymetrix's instructions) and 15 µg of fragmented cRNA was hybridized to Human U133 Plus 2.0 GeneChips according to the manufacturer's instructions. Quality control and background normalization was performed using Refiner 4.1 from Genedata AG (Basel, Switzerland). Expression value estimates were obtained using the GC-RMA implementation in Refiner 4.1. LOWESS-normalization and median scaling of the genes called present (detection P-value < 0.04) to a value of 500 was performed in Genedata's Analyst 4.1 package. The LOWESS-normalized data are referred to as 'raw' expression values in this paper. We also performed a point-wise division on the genes by dividing each gene by its median in order to centre its expression level at 1.0. This scaled data shows only the magnitude and direction of change but not of absolute expression level. Scaled data were used for clustering analyses. Unless noted, all other analyses were performed using the raw data.

Data analysis was performed using Expressionist® Analyst 4.1 from Genedata AG. Genes were required to pass a t-test with a $P < 0.05$ and have a median fold change of 1.3, 1.5, 2 and 5 or greater between the paired patient samples in at least 60% of patients within each group. For the supervised classifier prediction of liver biopsy samples and PBMCs using the response at week 4 as a grouping criterion 4 statistical tests were used (Support Vector Machine, Sparse Linear Discriminant Analysis, Fisher Linear Discriminant Analysis, K Nearest Neighbors). The misclassification rates could be determined for every test used and the one with the lowest rate was selected.

1.1.7 RNA isolation, reverse transcription and SYBR-PCR

The array data were validated by quantitative real-time RT-PCR analysis of several IFN regulated genes including STAT1, IP10, USP18, IFI27 SOCS1 and SOCS3.

Total RNA was extracted from liver using the RNeasy Mini Kit (Qiagen) according to manufacturer's instructions. The RNA was reverse transcribed by Moloney murine leukemia virus reverse transcriptase (Promega Biosciences, Inc., Wallisellen, Switzerland) in the presence of random hexamers (Promega) and deoxynucleoside triphosphate. The reaction mixture was incubated for 5 min at 70°C and then for 1 h at 37°C. The reaction was stopped by heating at 95°C for 5 min. SYBR-PCR was performed based on SYBR green fluorescence (SYBR green PCR master mix; Applied Biosystems, Foster City, CA). Primers for GAPDH (glyceraldehyde-3-phosphate dehydrogenase), STAT1, inducible protein 10 (IP10), SOCS1, SOCS3, USP18, IFI27 and PP2Ac were designed across exon-intron junctions. The primer sequences are shown in Table 4. The difference in the cycle threshold (ΔC_T) value was derived by subtracting the C_T value for GAPDH, which served as an

internal control, from the C_T value for STAT1 or other transcripts of interest. All reactions were run in duplicate by using an ABI 7000 sequence detection system (Applied Biosystems). mRNA expression levels of the transcripts were calculated relative to GAPDH from the ΔC_T values using the formula 2-ΔCT. The change of expression in paired liver biopsy samples was calculated as a fold change according to the formula $2^{(\Delta C_T B\text{-}1 - \Delta C_T B\text{-}2)}$.

Box plot diagrams, unpaired t-tests and Mann Whitney tests were performed using GraphPad Prism version 4.00 for Macintosh, GraphPad Software, San Diego California USA, www.graphpad.com.

1.2 Results

Patients and Response to Treatment

16 patients included in this study, 6 women and 10 men, were treated with a weight-adjusted combination of subcutaneously injected pegIFNα2b once weekly and oral ribavirin twice daily. All of them had two liver biopsies, the pre-treatment biopsy (B-1) and the second biopsy (B-2) obtained 4 h after the first injection of pegIFNα2b. We have chosen to analyze gene expression 4 h after pegIFNα2b injection since kinetics of the induction of ISGs by pegIFNα in liver of chimpanzees was maximal at this time and was followed by rapid down-regulation of many genes (22). We realize that we probably missed the up-regulation of some late induced ISGs, but because of rapid down-regulation, we would have missed more ISGs when using later time-points.

Seven of the patients were infected with HCV genotype (GT)1, two with GT 4, four with GT 3 and three with GT 2. Eight patients who had negative serum HCV RNA after 4 weeks of treatment and 2 patients with > 3 log drop of viral titer within the first 4 weeks were classified as rapid responders (RRs), whereas 6 patients showed a viral load reduction of less than 1.5 log and were classified as non-RRs (Table 1).

Serum IFNα concentrations were below the limit of detection in all patients before treatment, and, in accordance with previously published pharmacokinetic data (24), between 34 and 360 pg/ml in samples obtained at 4 h after the pegIFNα2b injection (data not shown). There was no significant correlation between the virological response at week 4 and the serum IFNα concentration at 4 h post-injection. Furthermore, despite the differences in the serum IFNα levels, all patients showed similar ISG induction in PMBCs (see below).

IFN-induced regulation of target genes

Gene expression was analyzed with Affymetrix U133plus2.0 arrays in B-1 and B-2 samples, and also in PBMCs isolated from blood obtained before (PBMC-1) and 4 h after the first pegIFNα2b injection

(PBMC-2). For each patient, the genes that were up- or down-regulated > 2-fold in post-treatment samples (compared to pre-treatment) were identified and saved in gene lists. We then generated 7 and 3 groups of 4 patients randomly selected from the 10 RR and the 6 non-RR patients, respectively. In each group, the genes significantly ($p < 0.05$ or $p < 0.01$) changed in at least 3 out of 4 patients were identified and counted. In liver biopsies of the 7 RR groups, the mean number (± SEM) of regulated genes was 76.71 (± 17.46) and 196.7 (± 31.55) at significance levels $p < 0.01$ and $p < 0.05$, respectively. In the 3 non-RR groups, these numbers were 11.67 (± 3.76) and 28.33 (± 6.12) for $p < 0.01$ and $p < 0.05$, respectively. The difference between RR and non-RR groups was statistically significant (Fig. 1A). There was an overlap in the significantly regulated genes found in RR samples and non-RR samples. For example, 30 of the 36 genes that changed >2-fold between B1 and B2 in more than 50% of the non-RR biopsy samples were also present among 177 genes changed in more than 50% of 6 randomly selected RR patients (Fig. 1B).

Not surprisingly, many of the regulated genes represent known ISGs. However, contrary to our expectations, expression levels of these ISGs were not higher in post pegIFNα2b treatment biopsies from RR patients as compared to non-RRs. Rather, non-RR patient samples had a higher level of ISG expression already in B-1, and the fold change in the B-2 samples was therefore only minor. This is illustrated in Fig. 2A at the example of five ISGs. The genes show a very low expression in biopsies from individuals without hepatitis C and in B-1's of RR patients. The 6 non-RR patients had high expression of these genes before treatment, and pegIFNα2a administration not or only minimally increased their expression. There were very few exceptions to this rule (an example is shown in Fig. 2B). These genes had low expression in the pre-treatment biopsies, and pegIFNα2b induced them in all patients. Nevertheless, the predominant pattern of gene expression resembled this shown in Fig. 2A. A list and a heat map of the expression of 252 genes significantly ($p<0.05$) changed > 2 fold between B-1 and B-2 in the RR group is shown for all biopsy samples in supplementary information (SI) Table 2 and SI Fig. 6.

There was a considerable overlap of pegIFNα2b-regulated genes in liver and PBMCs (Fig. 1C). Interestingly, in all patients pegIFNα2b regulated more genes in PBMCs than in liver. However, the difference in the upregulation of ISGs in PBMCs between RRs and non-RRs was not significant (Fig. 1A). No pre-activation of ISGs was found in PBMCs, and pegIFNα2b treatment had the same effect on ISG regulation in RR and non-RR patients (SI Fig. 5). This indicates that chronic HCV infection has strong local effects on the IFN system in liver, but little effect in PBMCs.

A subset of genes that predicts response to treatment

Supervised classifier analysis of array data allows the identification of a subset of genes that best predicts the outcome, in our case rapid response versus non-response at week 4. All liver biopsy and PBMC data sets were subjected to supervised classifier prediction using the response at 4 weeks of

treatment as grouping criteria. For PBMC samples the analysis did not identify a subset of genes that could predict the treatment outcome. In contrast, a subset of 16 genes was identified in the liver B-2 samples that predicted response to treatment with an error rate of 19.5%. Even better prediction was possible with a subset of 29 genes in the pre-treatment biopsies B-1 where the error rate was 4.3%. In this set there were 22 genes upregulated by pegIFNα2b (Table 2). Therefore, 76% best predictor genes represent ISGs.

Contrary to the predominance of ISGs in the best predictor set from pre-treatment biopsies, only 3 (19%) of the 16 best predictor genes derived from an analysis of the B-2 biopsies were ISGs (Table 3). These results support the findings shown in Fig. 2 that expression levels of ISGs in B-2 do not differ between RR and non-RR samples and therefore are not suited for the discrimination of responders from non-responders. Among the non-ISGs present in the B-1 and B-2 liver biopsy lists discussed above are genes having functions in signal transduction, cell cycle regulation, apoptosis, and amino acid metabolism.

RT-qPCR analysis of ISG expression in liver biopsies

Array analysis of the paired liver biopsies emphasized the importance of ISG expression in B-1 biopsies for the outcome of therapy. To confirm these data, we measured by real time quantitative PCR (RT-qPCR) the expression of selected ISGs (USP18, Stat1, IP10, IFI27) in 16 patients with B1 and B2 biopsies, and in pre-treatment biopsies of 96 additional patients with CHC. In the 16 patients with the paired biopsies, the RT-qPCR values matched well the array expression, validating the quality of the array data (Fig. 3A, and data not shown). The expression of all four ISGs in pre-therapy biopsies was significantly different between the EVR and PNR groups (Fig. 3C), further supporting the conclusion that there is an inverse correlation between the pre-treatment expression of ISGs in liver and the response to IFNα therapy. A significant upregulation of ISGs correlated also with non-response at week 12 and with final treatment outcome.

Pre-treatment ISG expression levels correlate with HCV genotype

We also analyzed the expression of ISGs with regard to the HCV genotype (GT). Interestingly, the investigated ISGs showed significantly higher expression in patients infected with the "difficult-to-treat" GTs 1 and 4 than with GTs 2 and 3, which can be successfully treated in over 80% of patients. Importantly, the expression levels of ISGs were higher in non-RR than RR patients independently from the HCV GT. Therefore, the increased ISG expression level in non-RR patients cannot simply be explained by the fact that GT 1 is over-represented in the non-RR group. Rather, the fact that patients with HCV GT 1 and 4 more frequently have an increased expression of ISGs in their liver provides a plausible explanation for the poor response of these patients to IFN therapy.

Non-responders have higher expression of PP2Ac

We have previously shown that the catalytic subunit of PP2A (PP2Ac) is over-expressed in liver of patients with CHC compared to controls, and that over-expression of PP2Ac inhibits IFNα signaling (14, 25). We therefore analyzed the PP2Ac mRNA levels in a group of patients with known treatment responses at week 12. Patients of the EVR group expressed significantly less PP2Ac mRNA than PNR patients (Fig. 3B).

IFN-induced Jak-STAT signaling

The injected pegIFNα2b binds to IFN receptors and activates the Jak-STAT pathway. A central event in this activation is the phosphorylation of STAT1 on tyrosine 701 (26). We analyzed extracts from all B-1 and B-2 biopsies by Western blot using a phospho-specific STAT1 antibody (Fig. 4A). A semi-quantitative analysis of the phospho-STAT1 bands revealed a median induction of 3.6 fold in RR patients and 1.6 fold in non-RR patients ($p = 0.03$).

Phosphorylated STAT1 translocates into the nucleus and binds as a dimer to specific response elements of ISG promoters (26). Assessment of nuclear translocation by immunohistochemistry, using anti-phospho-STAT1 antibodies, should potentially allow to discriminate between STAT1 activation in hepatocytes and other cells present in the biopsy material. Analysis of paired biopsies of RVR patients revealed a minimal nuclear staining in B-1 samples and a strong staining in most hepatocyte nuclei in B-2 samples, following injection of pegIFNα (Fig. 4B). In contrast, all but one (number 11) non-RVR patients showed a remarkably different staining pattern. In the pre-treatment biopsies, a large proportion of hepatocytes already had an appreciable nuclear staining, which did not increase in B-2 samples. The visible increase in nuclear staining in B-2 samples of non-RVR patients originated from nuclear translocation of STAT1 in Kupffer cells (liver macrophages), and not hepatocytes (Fig. 4B). Activation of STAT1 in Kupffer cells, and possibly contaminating blood cells, may have contributed to the increased STAT1 phosphorylation observed in Western blotting (Fig. 4A).

The next step in the signaling pathway is the binding of nuclear phospho-STAT1 to promoter elements of ISGs. We therefore assessed the STAT1 DNA-binding in extracts of B-1 and B-2 biopsies by performing electrophoretic mobility shift assays (EMSAs). All rapid responders showed a marked increase in the STAT1 DNA binding in the B-2 samples. In contrast, most non-RVR patients showed a minimal or no increase of the gel shift signal upon pegIFNα application.

These data indicate that results of immunohistochemistry and EMSA assays correlate better with the therapy outcome than results of Western analysis for phospo-STAT1. Taken together, the data demonstrate substantial differences in the IFN-induced Jak-STAT signaling between RVR and non-RVR patients.

1.3. Discussion

To learn more about possible mechanisms underlying differential response of HCV-infected patients to IFN therapy, we investigated the IFN-induced signaling and ISG induction in paired liver biopsies collected from patients with CHC before and during therapy pegIFNα. Comparison of IFN signaling in two liver samples obtained from the same patient, and comparison with the ISG induction in matching PBMC samples originating from the same patient, allowed us to obtain unequivocal evidence that patients who respond poorly to the therapy show pre-activation of their IFN system, and that the pre-activation is confined to the liver and is not evident in PBMCs. Importantly, in patients with a low initial ISG expression, representing future responders to therapy, activation of the IFN system in response to pegIFNα did not exceed that seen in non-responders, either before or after therapy. This could suggest that patients with the initial pre-activation of the IFN system, future non-responders, have some defects at steps downstream of ISG expression, making them refractory to both endogenous IFN and IFN therapy.

IFNα treatment induced the STAT1 phosphorylation in all but one patient. There was a tendency for stronger STAT1 activation in RVR compared to non-RVR samples. However, the immunohistochemical analysis revealed a more pronounced difference. In non-RVR samples, pegIFNα strongly induced the nuclear STAT1 translocation in Kupffer cells, contrary to RVR samples, where nuclear STAT1 accumulation was induced predominantly in hepatocytes. Interestingly, non-RVR patients (with one exception) had nuclear phospho-STAT1 already present in pre-treatment biopsies. This is consistent with the observation that ISG transcripts are up-regulated in pre-treatment biopsies of later non-responders. How this preactivation of the Jak-STAT pathway is connected to the refractoriness of the IFN system in non-RVR patients requires further investigations.

Over the last few years, important insights into the interference of HCV with the innate immune system have been gained. Foremost, a series of elegant papers demonstrated the ability of HCV to inhibit both TLR3-TRIF-IRF3 and the RIG-I/MDA5-Cardif signaling pathways of IFNβ induction (27-33). This capacity of HCV could help to explain why the virus often establishes a chronic infection. However, our data and previously published results (20) demonstrate that the endogenous IFN system is constantly activated in many patients. Moreover, patients with a pre-activated IFN system seem to respond poorly to IFN therapy. This finding is counter-intuitive (one would expect that an active innate immune system would help to eliminate the virus during IFNα therapy), but it is largely supported by other published data from chimpanzees and human patients (16, 17, 20). From the analyses of ISG expression in liver biopsies it is apparent that in some patients HCV induces (or at least does not block) the endogenous IFN system, while in others it successfully represses it, possibly by cleaving TRIF and/or Cardif. Paradoxically, this difference has no apparent impact on the ability of HCV to maintain a chronic infection.

In patients without pre-activated IFN system, pegIFNα2b induced a robust up-regulation of many ISGs in the liver within 4 h. Similar high ISG expression was already present in the pre-treatment biopsies of patients that later did not show a rapid virological response at week 4. It is somewhat perplexing why the latter patients do not resolve the chronic HCV infection spontaneously despite the strong activation of the IFN system. One possibility is that ISG proteins that are up-regulated in both cases possess different post-transcriptional modifications. In an alternative scenario, non-response to both endogenous and exogenous IFNα may be caused by the lack of induction of a few critical ISGs that are specifically required for the elimination of HCV. We cannot exclude this possibility, but an array analysis performed on paired liver samples did not reveal ISGs that were specifically up-regulated in rapid responders. Furthermore, this model cannot explain why pre-activation of the endogenous IFN system is so closely linked to later non-response to treatment.

Alternatively, the kinetics of induction of the interferon response could be decisive. In the patients without pre-activated IFN system, the injection of exogenous IFNα during treatment should induce an antiviral state very rapidly in most liver cells, and HCV would not have "enough" time to escape from the IFN-induced defense. On the other hand, the build-up of the antiviral state could be slow in the other group of patients which would give HCV enough time to adapt to and evade the intracellular antiviral defense system, making it also resistant to the subsequent IFN therapy.

How could the induction of the endogenous IFN system compromise the success of IFNα therapy? Clearly, the activation of negative feedback loops that inhibit IFN signaling could play a role. Prominent candidates amongst the negative regulators are: suppressors of cytokine signaling 1 (SOCS1) and SOCS3 (34), two IFN induced proteins that bind to the IFN receptor and inhibit the activity of Jak1 and Tyk2; and the more recently described regulator Ubp43, an IFN-stimulated protein that binds to IFNα receptor 2 (IFNAR2) and blocks the access of Jak1 to it (35). However, we could not find a significant difference in the expression levels of these negative regulators in the pegIFNα2b stimulated liver biopsies of RVR compared to non-RVR patients (data not shown). Moreover, a general up-regulation of negative regulators such as SOCSs and Ubp43 is not compatible with the observed strong constitutive expression of a large number of ISGs in the subset of patients that poorly respond to IFN therapy. If IFNα signaling were indeed inhibited by the induction of SOCSs and Ubp43 in the majority of liver cells, then one should not observe such a pronounced pre-activation of ISGs in pre-treatment livers.

Notably, the pre-activation of tested ISGs occurred more frequently in liver biopsies of patients infected with HCV genotype 1 and 4 than with genotypes 2 or 3. It is well known that genotype 2 and 3 infections can be cured in over 80% of patients, compared to less than 50% of infections with genotype 1 (4). Our finding that the frequency and degree of pre-activation of the endogenous IFN system depends on the HCV genotype could provide an explanation for this differential susceptibility.

Perhaps HCV genotypes 2 and 3 are more successful in preventing the activation of innate immunity in the liver by a more effective cleavage of Cardif and/or TRIF. The success of the virus in preventing the induction of the endogenous IFN system would however come at the cost of being more susceptible to IFNα therapies. Of note, a single chimpanzee infected with the genotype 3 HCV has been shown to have lower ISG expression levels than animals infected with genotype 1 (17).

We have shown previously that HCV inhibits the IFNα-induced signaling via the Jak-STAT pathway by up-regulating a protein phosphatase PP2A (12, 14, 25, 36). PP2A is a heterotrimeric complex of a scaffolding A, a regulatory B, and a catalytic C subunits. The PP2Ac subunit expression is significantly higher in livers of patients infected with genotype 1 than genotype 3 (25). As shown in this work, the expression of PP2Ac mRNA is higher in biopsies of later non-responders than responders. These data support a model where HCV interference with the IFN signaling impairs the response to therapy. Moreover, inhibition of the IFNα signaling by HCV could also explain why the strong pre-activation of the endogenous IFN system does not lead to a spontaneous elimination of HCV. If one assumes that not all hepatocytes are infected by HCV, but rather a minority, then the induction of ISGs observed in pre-treatment biopsies of non-RVR patients could occur predominantly in non-infected hepatocytes. In the infected cells, IFN would be ineffective because of the inhibition of the Jak-STAT signaling pathway. The IFN responsible for the pre-activation of the system would be secreted by hepatocytes that are infected with a virus that is not successful in cleaving Cardif and/or TRIF. Because of the HCV-induced inhibition of the Jak-STAT pathway, the secreted IFNβ would not induce an antiviral state in the infected hepatocytes, but rather in non-infected neighbor cells. To gain further insights into the pathobiology of CHC, future studies should focus on analysis at the single-cell level. Unfortunately, the detection of HCV infected hepatocytes in liver biopsies is still unsatisfactory, making such studies difficult.

Although the precise mechanism of the HCV escape from the immune defense system still remains to be elucidated, the impairment of the hepatitis C therapy by pre-activation of the endogenous IFN system is now well established. It would be interesting to investigate if this pre-activation is a reversible process. The injection of neutralizing anti-IFNα/β antibodies or other factors blocking the IFN response before treatment could return the endogenous IFN system to a "naive" state, and potentially enhance the response to IFNα based therapies.

Table 1

Patient Nr.	4 week response	Sex	Age	HCV GT	Viral Load log IU/ml			12 week response	Follow up	Metavir	Weight (kg)
					Base-line	4 week	12 week				
1	RVR	m	52	3a	7.14	neg.	neg.	EVR	SVR	A2/F2	75
2	RVR	m	37	3a	4.90	neg.	neg.	EVR	EoTR	A1/F2	73
3	RVR	m	38	1a	6.91	neg.	neg.	EVR	EoTR	A2/F1	85
4	RVR	m	33	2b	6.27	neg.	neg.	EVR	EoTR	A1/F2	57
5	RVR	m	48	2b	6.67	neg.	neg.	EVR	EoTR	A3/F4	110
6	RVR	f	53	2a/c	4.95	neg.	neg.	EVR	EoTR	A3/F3	74
7	RVR	m	56	3	5.25	neg.	ongoing	ongoing	ongoing	A3/F4	61
8	RVR	m	38	4	4.08	neg.	ongoing	ongoing	ongoing	A2/F2	69
9	RR	f	50	1b	7.22	3.52	ongoing	ongoing	ongoing	A1/F2	47
10	RR	f	48	1	6.49	3.31	ongoing	ongoing	ongoing	A3/F4	60
11	Non-RVR	f	54	3a	4.52	4.08	1.3	EVR	No EoTR	A3/F4	69
12	Non-RVR	m	64	1b	6.24	4.83	3.46	EVR	No EoTR	A3/F4	74
13	Non-RVR	m	49	4	6.91	5.87	5.22	PNR	-	A3/F4	102
14	Non-RVR	m	56	1b	6.89	6.76	6.01	PNR	-	A2/F3	60
15	Non-RVR	f	50	1a	7.11	6.58	6.35	PNR	-	A1/F2	77
16	Non-RVR	f	47	1a	6.16	5.99	5.52	PNR	-	A2/F2	81

Table 2

Analysis of gene expression in pre-treatment biopsies (B-1). List of 29 genes best predicting treatment outcome at week 4 (IFN stimulated genes shaded in grey; genes that differ between RR and non-RR but are not regulated by IFN are not shaded).

Gene Symbol	Description	Affy-ID	Mean (SEM) expression in 10 RR's	Mean (SEM) expression in 6 non-RR's	Non-RR / RR	Function
IFI44L	interferon-induced protein 44-like	204439_at	306 (50)	3392 (903)	11.10	cell cycle
RSAD2	radical S-adenosyl methionine domain containing 2	242625_at	272 (41)	2405 (425)	8.83	innate immune response
G1P2	interferon, alpha-inducible protein (clone IFI-15K)	205483_s_at	2238 (397)	17375 (2762)	7.76	innate immune response
IFI27	interferon, alpha-inducible protein 27	202411_at	3320 (714)	24927 (2441)	7.51	innate immune response
LAMP3	lysosomal-associated membrane protein 3	205569_at	96 (22)	665 (108)	6.97	cell proliferation
OAS3	2'-5'-oligoadenylate synthetase 3, 100kDa	218400_at	319 (46)	1842 (296)	5.77	innate immune response
HERC6	hect domain and RLD 6	219352_at	144 (32)	795 (93)	5.53	immune response
HIST1H2BD	Histone 1, H2bd	235456_at	40 (4)	202 (33)	5.02	DNA packaging
IFIT1	interferon-induced protein with tetratricopeptide repeats 1	203153_at	2209 (165)	9995 (1706)	4.53	innate immune response

LOC129607	hypothetical protein LOC129607	226702_at	970 (143)	4135 (520)	4.26	amino acid metabolism
IFI44	interferon-induced protein 44	214453_s_at	1101 (153)	4183 (405)	3.80	innate immune response
HERC5	hect domain and RLD 5	219863_at	963 (93)	3144 (552)	3.26	protein ubiquitination
LGALS3BP	lectin, galactoside-binding, soluble, 3 binding protein	200923_at	1537 (238)	4960 (475)	3.23	response to stress
SAMD9	sterile alpha motif domain containing 9	228531_at	323 (32)	997 (166)	3.08	unknown
IFIT2	interferon-induced protein with tetratricopeptide repeats 2	226757_at	951 (63)	2744 (510)	2.89	innate immune response
LOC286208	hypothetical protein LOC286208	1560089_at	39 (3)	110 (18)	2.86	-
IRF7	interferon regulatory factor 7	208436_s_at	240 (16)	679 (107)	2.82	innate immune response
FLJ20035	hypothetical protein FLJ20035	218986_s_at	1191 (84)	3341 (378)	2.80	Helicase
IFIT3	Interferon-induced protein with tetratricopeptide repeats 3	229450_at	2760 (213)	7703 (1327)	2.79	innate immune response
RALGPS1	Ral GEF with PH domain and SH3 binding motif 1	204199_at	22 (2)	61 (8)	2.70	signal transduction
PARP12	poly (ADP-ribose) polymerase family, member 12	218543_s_at	437 (36)	1158 (60)	2.65	poly (ADP-ribose) polymerase family
HIST1H2B	histone 1, H2bg	210387_at	13 (1)	29 (3)	2.13	DNA

G						packaging
PARP9	poly (ADP-ribose) polymerase family, member 9	223220_s_at	1386 (111)	2868 (263)	2.07	poly (ADP-ribose) polymerase family
PNPT1	polyribonucleotide nucleotidyltransferase 1	225291_at	518 (23)	971 (105)	1.88	RNA catabolism
CCDC75	Coiled-coil domain containing 75	213294_at	776 (40)	1382 (105)	1.78	-
CNP	2',3'-cyclic nucleotide 3' phosphodiesterase	208912_s_at	682 (24)	1054 (70)	1.55	nucleotide metabolism
HTATIP2	HIV-1 Tat interactive protein 2, 30kDa	209448_at	3317 (130)	4451 (74)	1.34	Apoptosis
RPLP0	ribosomal protein, large, P0, ribosomal protein, large, P0	211720_x_at	16980 (404)	13294 (439)	0.78	protein biosynthesis
LOC402560	Hypothetical LOC401384	227554_at	562 (95)	90 (20)	0.16	-

Table 3

Analysis of gene expression in biopsies obtained 4 hours after pegIFNα (B-2). List of 16 genes best predicting treatment outcome at week 4 (IFN stimulated genes shaded in grey; genes that differ between RR and non-RR but are not regulated by IFN are not shaded).

Gene Symbol	**Description**	**Affy-ID**	**Mean (SEM) expression in 10 RR's**	**Mean (SEM) expression in 6 non-RR's**	**Non-RR /RR**	**Function**

IFI27	interferon, alpha-inducible protein 27	202411_at	4540 (735)	25501 (2372)	5.62	innate immune response
LGALS3BP	lectin, galactoside-binding, soluble, 3 binding protein	200923_at	1728 (282)	5223 (386)	3.02	response to stress
ZFP3	zinc finger protein 3 homolog (mouse)	235728_at	33 (4)	75 (7)	2.28	ion binding
MYH14	myosin, heavy polypeptide 14	234290_x_at	33 (2)	64 (5)	1.98	cell morphogenesis
PARP6	poly (ADP-ribose) polymerase family, member 6	219639_x_at	149 (8)	285 (23)	1.92	poly (ADP-ribose) polymerase family
GPR143	G protein-coupled receptor 143	206696_at	18 (1)	28 (2)	1.54	signal transduction
BRUNOL5	bruno-like 5, RNA binding protein (Drosophila)	232416_at	13 (0.3)	19 (1)	1.43	RNA binding
ATP5A1	ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit, isoform 1, cardiac muscle	213738_s_at	14472 (290)	17712 (487)	1.22	cellular metabolism
CHMP4A	chromatin modifying protein 4A	228764_s_at	348 (11)	234 (6)	0.67	protein localization
IDS	iduronate 2-sulfatase (Hunter syndrome)	206342_x_at	185 (7)	122 (5)	0.66	carbohydrate metabolism

	gb:AI341383 /DB_XREF=gi:4078310 /DB_XREF=qx91a06.x1 /CLONE=IMAGE:2009842 /FEA=EST /CNT=52 /TID=Hs.112751.2 /TIER=Stack /STK=42 /UG=Hs.112751 /LL=23383 /UG_GENE=KIAA0892 /UG_TITLE=KIAA0892 protein	227092_at	475 (24)	254 (17)	0.54	
VISA	Virus-induced signaling adapter	229741_at	167 (7)	77 (4)	0.46	innate immune response
PCOLCE	procollagen C-endopeptidase enhancer	202465_at	1297 (135)	539 (56)	0.42	development
IRF1	Interferon regulatory factor 1	238725_at	1106 (109)	449 (30)	0.41	innate immune response
PSMAL	prostate-specific membrane antigen-like	211303_x_at	793 (76)	278 (68)	0.35	Unknown
LOC402560	Hypothetical LOC401384	227554_at	578 (85)	83 (17)	0.14	-

(c) **1.4 References**

1. Li, K., *et al.* Immune evasion by hepatitis C virus NS3/4A protease-mediated cleavage of the Toll-like receptor 3 adaptor protein TRIF. *Proc Natl Acad Sci U S A* **102**, 2992-2997 (2005).

2. Yoneyama, M., *et al.* The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nat Immunol* **5**, 730-737 (2004).

3. Kang, D.C., *et al.* mda-5: An interferon-inducible putative RNA helicase with double-stranded RNA-dependent ATPase activity and melanoma growth-suppressive properties. *Proc Natl Acad Sci U S A* **99**, 637-642 (2002).

4. Kawai, T., *et al.* IPS-1, an adaptor triggering RIG-I- and Mda5-mediated type I interferon induction. *Nat Immunol* **6**, 981-988 (2005).

5. Meylan, E., *et al.* Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus. *Nature* **437**, 1167-1172 (2005).

6. Seth, R.B., Sun, L., Ea, C.K. & Chen, Z.J. Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF-kappaB and IRF 3. *Cell* **122**, 669-682 (2005).

7. Xu, L.G., *et al.* VISA is an adapter protein required for virus-triggered IFN-beta signaling. *Mol Cell* **19**, 727-740 (2005).

8. Gale, M., Jr. & Foy, E.M. Evasion of intracellular host defence by hepatitis C virus. *Nature* **436**, 939-945 (2005).

9. Bigger, C.B., Brasky, K.M. & Lanford, R.E. DNA microarray analysis of chimpanzee liver during acute resolving hepatitis C virus infection. *J Virol* **75**, 7059-7066 (2001).

10. Bigger, C.B., *et al.* Intrahepatic gene expression during chronic hepatitis C virus infection in chimpanzees. *J Virol* **78**, 13779-13792 (2004).

11. Chen, L., *et al.* Hepatic gene expression discriminates responders and nonresponders in treatment of chronic hepatitis C viral infection. *Gastroenterology* **128**, 1437-1444 (2005).

12. Strader, D.B., Wright, T., Thomas, D.L. & Seeff, L.B. Diagnosis, management, and treatment of hepatitis C. *Hepatology* **39**, 1147-1171 (2004).

13. Manns, M.P., *et al.* Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomised trial. *Lancet* **358**, 958-965 (2001).

14. Fried, M.W., *et al.* Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. *N Engl J Med* **347**, 975-982 (2002).

15. Hadziyannis, S.J., *et al.* Peginterferon-alpha2a and ribavirin combination therapy in chronic hepatitis C: a randomized study of treatment duration and ribavirin dose. *Ann Intern Med* **140**, 346-355 (2004).

16. Lau, D.T., *et al.* 10-Year follow-up after interferon-alpha therapy for chronic hepatitis C. *Hepatology* **28**, 1121-1127 (1998).

17. Marcellin, P., *et al.* Long-term histologic improvement and loss of detectable intrahepatic HCV RNA in subjects with chronic hepatitis C and sustained response to interferon-alpha therapy. *Ann Intern Med* **127**, 875-881 (1997).

18. Ferenci, P., *et al.* Predicting sustained virological responses in chronic hepatitis C subjects treated with peginterferon alfa-2a (40 KD)/ribavirin. *J Hepatol* **43**, 425-433 (2005).

19. Davis, G.L., *et al.* Early virologic response to treatment with peginterferon alfa-2b plus ribavirin in subjects with chronic hepatitis C. *Hepatology* **38**, 645-652 (2003).

20. Yu, J.W., Wang, G.Q., Sun, L.J., Li, X.G. & Li, S.C. Predictive value of rapid virological response and early virological response on sustained virological response in HCV subjects treated with pegylated interferon alpha-2a and ribavirin. *Journal of gastroenterology and hepatology* **22**, 832-836 (2007).

21. Sen, G.C. Viruses and interferons. *Annual review of microbiology* **55**, 255-281 (2001).

22. Darnell, J.E., Jr., Kerr, I.M. & Stark, G.R. Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science* **264**, 1415-1421 (1994).

23. de Veer, M.J., *et al.* Functional classification of interferon-stimulated genes identified using microarrays. *Journal of leukocyte biology* **69**, 912-920 (2001).

24. Blindenbacher, A., *et al.* Expression of hepatitis c virus proteins inhibits interferon alpha signaling in the liver of transgenic mice. *Gastroenterology* **124**, 1465-1475 (2003).

25. Heim, M.H., Moradpour, D. & Blum, H.E. Expression of hepatitis C virus proteins inhibits signal transduction through the Jak-STAT pathway. *J Virol* **73**, 8469-8475 (1999).

26. Duong, F.H., Filipowicz, M., Tripodi, M., La Monica, N. & Heim, M.H. Hepatitis C virus inhibits interferon signaling through up-regulation of protein phosphatase 2A. *Gastroenterology* **126**, 263-277 (2004).

27. Gale, M., Jr., *et al.* Control of PKR protein kinase by hepatitis C virus nonstructural 5A protein: molecular mechanisms of kinase regulation. *Mol Cell Biol* **18**, 5208-5218 (1998).

28. Wang, C., *et al.* Alpha interferon induces distinct translational control programs to suppress hepatitis C virus RNA replication. *J Virol* **77**, 3898-3912 (2003).

29. Taylor, D.R., Shi, S.T., Romano, P.R., Barber, G.N. & Lai, M.M. Inhibition of the interferon-inducible protein kinase PKR by HCV E2 protein. *Science* **285**, 107-110 (1999).

30. Silva, M., *et al.* A randomised trial to compare the pharmacokinetic, pharmacodynamic, and antiviral effects of peginterferon alfa-2b and peginterferon alfa-2a in subjects with chronic hepatitis C (COMPARE). *J Hepatol* **45**, 204-213 (2006).

31. Shuai, K., Stark, G.R., Kerr, I.M. & Darnell, J.E., Jr. A single phosphotyrosine residue of Stat91 required for gene activation by interferon-gamma. *Science* **261**, 1744-1746 (1993).

32. Darnell, J.E., Jr. STATs and gene regulation. *Science* **277**, 1630-1635 (1997).

33. Duong, F.H., Christen, V., Filipowicz, M. & Heim, M.H. S-adenosylmethionine and betaine correct hepatitis C virus induced inhibition of interferon signaling in vitro. *Hepatology* **43**, 796-806 (2006).

34. Lanford, R.E., *et al.* Genomic response to interferon-alpha in chimpanzees: implications of rapid downregulation for hepatitis C kinetics. *Hepatology* **43**, 961-972 (2006).

35. Krebs, D.L. & Hilton, D.J. SOCS proteins: negative regulators of cytokine signaling. *Stem Cells* **19**, 378-387 (2001).

36. Malakhova, O.A., *et al.* UBP43 is a novel regulator of interferon signaling independent of its ISG15 isopeptidase activity. *Embo J* **25**, 2358-2367 (2006).

37. McHutchison, J.G., *et al.* Interferon alfa-2b alone or in combination with ribavirin as initial treatment for chronic hepatitis C. Hepatitis Interventional Therapy Group. *N Engl J Med* **339**, 1485-1492 (1998).

38. Christen, V., Treves, S., Duong, F.H. & Heim, M.H. Activation of endoplasmic reticulum stress response by hepatitis viruses up-regulates protein phosphatase 2A. *Hepatology* (2007).

CLAIMS

1. A method for determining the likelihood that a subject having a viral infection of the liver will be responsive to antiviral therapy that includes stimulation of Interferon (IFN) activity, the method comprising:

(a) analysing a sample from the subject before antiviral therapy for the level of expression of at least one gene from each of the following groups of genes:

(i) LOC129607; RPLP0; and HERC5;

(ii) HTATIP2; and IFI44L;

(iii) IFI27; IFIT1; G1P2; IRF7; RSAD2; IFI44; OAS3; and IFIT2;

(iv) LAMP3; HERC6; LOC286208; IFIT3; RALGPS1; PARP9; CCDC75; and CNP;

(v) HIST1H2BG; HIST1H2BD; FLJ20035; PARP12; PNPT1; LGALS3BP; SAMD9; and LOC402560 and,

(b) comparing the level of expression of the genes in the sample to level of expression of the same genes in a control sample, wherein an increase in the level of expression of the genes in the sample from the subject before antiviral therapy relative to the level of expression of the same genes in the control sample indicates that the subject is not likely to be responsive to said antiviral therapy.

2. The method of Claim 1 wherein the antiviral therapy includes pegylated IFNα.

3. The method of Claim 2 wherein the antiviral therapy includes ribavirin.

4. The method of any of the previous claims wherein the viral infection is Hepatitis B virus or Hepatitis C virus infection.

5. The method of Claim 2 wherein the virus is Hepatitis C virus.

6. The method of any of the previous claims wherein the sample from the subject comprises liver tissue.

7. The method of any of the previous claims wherein the level of expression of six, seven, eight, nine, 10, 11 , 12, 13, 14, 15, 20, 25, 26, 27, 28 or 29 genes of claim 1 is analysed.

8. The method of any of the previous claims wherein the level of expression is determined by measuring the amount of mRNA gene transcript in the sample from the subject, or the amount of cDNA derived from said mRNA.

9. The method of any of claims 1 to 7 wherein the level of expression is determined by measuring the amount of peptide or polypeptide encoded by the gene in the sample from the subject.

10. The method of claim 9 wherein the amount of peptide or polypeptide is determined using a specific binding molecule.

11. The method of any of the previous claims wherein the subject is human.

SEQUENCE LISTING

<110> Novartis Forschungsstiftung, Zweigniederlassung Friedrich Miescher Institute for Biomedical Research
University Hospital Basel

<120> Antiviral Therapy

<130> 52434

<160> 23

<170> PatentIn version 3.3

<210> 1
<211> 557
<212> PRT
<213> Homo sapiens

<400> 1

```
Met Met Val Val Leu Leu Gly Ala Thr Thr Leu Val Leu Val Ala Val
1               5                  10                  15

Ala Pro Trp Val Leu Ser Ala Ala Ala Gly Gly Lys Asn Leu Lys Ser
            20                  25                  30

Pro Gln Lys Val Glu Val Asp Ile Ile Asp Asp Asn Phe Ile Leu Arg
        35                  40                  45

Trp Asn Arg Ser Asp Glu Ser Val Gly Asn Val Thr Phe Ser Phe Asp
    50                  55                  60

Tyr Gln Lys Thr Gly Met Asp Asn Trp Ile Lys Leu Ser Gly Cys Gln
65                  70                  75                  80

Asn Ile Thr Ser Thr Lys Cys Asn Phe Ser Ser Leu Lys Leu Asn Val
                85                  90                  95

Tyr Glu Glu Ile Lys Leu Arg Ile Arg Ala Glu Lys Glu Asn Thr Ser
            100                 105                 110

Ser Trp Tyr Glu Val Asp Ser Phe Thr Pro Phe Arg Lys Ala Gln Ile
        115                 120                 125

Gly Pro Pro Glu Val His Leu Glu Ala Glu Asp Lys Ala Ile Val Ile
    130                 135                 140
```

```
His Ile Ser Pro Gly Thr Lys Asp Ser Val Met Trp Ala Leu Asp Gly
145                 150                 155                 160

Leu Ser Phe Thr Tyr Ser Leu Val Ile Trp Lys Asn Ser Ser Gly Val
                165                 170                 175

Glu Glu Arg Ile Glu Asn Ile Tyr Ser Arg His Lys Ile Tyr Lys Leu
            180                 185                 190

Ser Pro Glu Thr Thr Tyr Cys Leu Lys Val Lys Ala Ala Leu Leu Thr
        195                 200                 205

Ser Trp Lys Ile Gly Val Tyr Ser Pro Val His Cys Ile Lys Thr Thr
    210                 215                 220

Val Glu Asn Glu Leu Pro Pro Pro Glu Asn Ile Glu Val Ser Val Gln
225                 230                 235                 240

Asn Gln Asn Tyr Val Leu Lys Trp Asp Tyr Thr Tyr Ala Asn Met Thr
                245                 250                 255

Phe Gln Val Gln Trp Leu His Ala Phe Leu Lys Arg Asn Pro Gly Asn
            260                 265                 270

His Leu Tyr Lys Trp Lys Gln Ile Pro Asp Cys Glu Asn Val Lys Thr
        275                 280                 285

Thr Gln Cys Val Phe Pro Gln Asn Val Phe Gln Lys Gly Ile Tyr Leu
    290                 295                 300

Leu Arg Val Gln Ala Ser Asp Gly Asn Asn Thr Ser Phe Trp Ser Glu
305                 310                 315                 320

Glu Ile Lys Phe Asp Thr Glu Ile Gln Ala Phe Leu Leu Pro Pro Val
                325                 330                 335

Phe Asn Ile Arg Ser Leu Ser Asp Ser Phe His Ile Tyr Ile Gly Ala
            340                 345                 350

Pro Lys Gln Ser Gly Asn Thr Pro Val Ile Gln Asp Tyr Pro Leu Ile
        355                 360                 365

Tyr Glu Ile Ile Phe Trp Glu Asn Thr Ser Asn Ala Glu Arg Lys Ile
```

370 375 380

Ile Glu Lys Lys Thr Asp Val Thr Val Pro Asn Leu Lys Pro Leu Thr
385 390 395 400

Val Tyr Cys Val Lys Ala Arg Ala His Thr Met Asp Glu Lys Leu Asn
405 410 415

Lys Ser Ser Val Phe Ser Asp Ala Val Cys Glu Lys Thr Lys Pro Gly
420 425 430

Asn Thr Ser Lys Ile Trp Leu Ile Val Gly Ile Cys Ile Ala Leu Phe
435 440 445

Ala Leu Pro Phe Val Ile Tyr Ala Ala Lys Val Phe Leu Arg Cys Ile
450 455 460

Asn Tyr Val Phe Phe Pro Ser Leu Lys Pro Ser Ser Ser Ile Asp Glu
465 470 475 480

Tyr Phe Ser Glu Gln Pro Leu Lys Asn Leu Leu Leu Ser Thr Ser Glu
485 490 495

Glu Gln Ile Glu Lys Cys Phe Ile Ile Glu Asn Ile Ser Thr Ile Ala
500 505 510

Thr Val Glu Glu Thr Asn Gln Thr Asp Glu Asp His Lys Lys Tyr Ser
515 520 525

Ser Gln Thr Ser Gln Asp Ser Gly Asn Tyr Ser Asn Glu Asp Glu Ser
530 535 540

Glu Ser Lys Thr Ser Glu Glu Leu Gln Gln Asp Phe Val
545 550 555

<210> 2
<211> 6099
<212> DNA
<213> Homo sapiens

<400> 2
aggcggcgcg tgcgtagagg ggcggtgaga gctaagaggg gcagcgcgtg tgcagagggg 60

cggtgtgact taggacgggg cgatggcggc tgagaggagc tgcgcgtgcg cgaacatgta 120

actggtggga tctgcggcgg ctcccagatg atggtcgtcc tcctgggcgc gacgacccta 180

gtgctcgtcg ccgtggcgcc atgggtgttg tccgcagccg caggtggaaa aaatctaaaa 240

tctcctcaaa aagtagaggt cgacatcata gatgacaact ttatcctgag gtggaacagg 300

agcgatgagt ctgtcgggaa tgtgactttt tcattcgatt atcaaaaaac tgggatggat 360

aattggataa aattgtctgg gtgtcagaat attactagta ccaaatgcaa cttttcttca 420

ctcaagctga atgtttatga agaaattaaa ttgcgtataa gagcagaaaa agaaaacact 480

tcttcatggt atgaggttga ctcatttaca ccatttcgca aagctcagat tggtcctcca 540

gaagtacatt tagaagctga agataaggca atagtgatac acatctctcc tggaacaaaa 600

gatagtgtta tgtgggcttt ggatggttta agctttacat atagcttagt tatctggaaa 660

aactcttcag gtgtagaaga aaggattgaa aatatttatt ccagacataa aatttataaa 720

ctctcaccag agactactta ttgtctaaaa gttaaagcag cactacttac gtcatggaaa 780

attggtgtct atagtccagt acattgtata aagaccacag ttgaaaatga actacctcca 840

ccagaaaata tagaagtcag tgtccaaaat cagaactatg ttcttaaatg ggattataca 900

tatgcaaaca tgacctttca agttcagtgg ctccacgcct ttttaaaaag gaatcctgga 960

aaccatttgt ataaatggaa acaaatacct gactgtgaaa atgtcaaaac tacccagtgt 1020

gtctttcctc aaaacgtttt ccaaaaagga atttaccttc tccgcgtaca agcatctgat 1080

ggaaataaca catcttttg gtctgaagag ataaagtttg atactgaaat acaagctttc 1140

ctacttcctc cagtctttaa cattagatcc cttagtgatt cattccatat ctatatcggt 1200

gctccaaaac agtctggaaa cacgcctgtg atccaggatt atccactgat ttatgaaatt 1260

atttttgggg aaaacacttc aaatgctgag agaaaaatta tcgagaaaaa aactgatgtt 1320

acagttccta atttgaaacc actgactgta tattgtgtga aagccagagc acacaccatg 1380

gatgaaaagc tgaataaaag cagtgttttt agtgacgctg tatgtgagaa aacaaaacca 1440

ggaaatacct ctaaaatttg gcttatagtt ggaatttgta ttgcattatt tgctctcccg 1500

tttgtcattt atgctgcgaa agtcttcttg agatgcatca attatgtctt ctttccatca 1560

cttaaacctt cttccagtat agatgagtat ttctctgaac agccattgaa gaatcttctg 1620

ctttcaactt ctgaggaaca aatcgaaaaa tgtttcataa ttgaaaatat aagcacaatt 1680

gctacagtag aagaaactaa tcaaactgat gaagatcata aaaaatacag ttcccaaact 1740

agccaagatt caggaaatta ttctaatgaa gatgaaagcg aaagtaaaac aagtgaagaa 1800

ctacagcagg actttgtatg accagaaatg aactgtgtca agtataaggt ttttcagcag 1860

gagttacact gggagcctga ggtcctcacc ttcctctcag taactacaga gaggacgttt 1920

ccctgtttag ggaaagaaaa aacatcttca gatcataggt cctaaaaata cgggcaagct 1980

cttaactatt taaaaatgaa attacaggcc cgggcacggt ggctcacacc tgtaatccca 2040

gcactttggg aggctgaggc aggcagatca tgaggtcaag agatcgagac cagcctggcc 2100

aacgtggtga aaccccatct ctactaaaaa tacaaaaatt agccgggtgt ggtggcgcgc 2160

gcctgttgtc ttagctactc aggaggctga ggcaggagaa tcgcttgaaa acaggaggtg 2220

gaggttgcag tgagccgaga tcacgccact gcactccagc ctggtgacag cgtgagactc 2280

tttaaaaaaa gaaattaaaa gagttgagac aaacgtttcc tacattcttt tccatgtgta 2340

aaatcatgaa aaagcctgtc accggacttg cattggatga gatgagtcag accaaaacag 2400

tggccacccg tcttcctcct gtgagcctaa gtgcagccgt gctagctgcg caccgtggct 2460

aaggatgacg tctgtgttcc tgtccatcac tgatgctgct ggctactgca tgtgccacac 2520

ctgtctgttc gccattccta acattctgtt tcattcttcc tcgggagata tttcaaacat 2580

ttggtctttt cttttaacac tgagggtagg cccttaggaa atttatttag gaaagtctga 2640

acacgttatc acttggtttt ctggaaagta gcttacccta gaaaacagct gcaaatgcca 2700

gaaagatgat ccctaaaaat gttgagggac ttctgttcat tcatcccgag aacattggct 2760

tccacatcac agtatctacc cttacatggt ttaggattaa agccaggcaa tcttttacta 2820

tgcattaaga cctctgattc aaaacttatt agaacagtag cttctgctgg aatttgcaat 2880

cactgaagtc atagaaaata ggtaactatc taattagaga aataattgtt gtattttaag 2940

atctgagagt gtgtacaagt tttagtatac atgccatgcc agaagatagt gtatgcaaga 3000

agtcttggga ccagaaaatg gcaatgatag gagactgaca tagaagaaga atgcttccct 3060

aggaaaaagg tcgctggctt tggtgcaaga ggaagaagaa tgttccactg gaagcctgag 3120

cacctaatca gctctcagtg atcaacccac tcttgttatg ggtggtctct gtcactttga 3180

atgccaggct ggcttctcgt ctagcagtat tcagataccc cttctgctca gcctgcttgg 3240

cgttaaaata caaatcattg aactgagggg gaaaaatgta actaggaaga aaaacccaat 3300

ttaagaaatt acataatgct ttccaaaggc acctacaact tagttttaaa ttacttgcta 3360

ctggggatta cccatggata tccttaatag gcaggaagtc tgggaattct ggtggcctct 3420

agggcagtgt tctcacagca ccgttccgac agggaccagt gaaagaaaag agacaaagtt 3480

agaacgtgct ggggagcggc catttctaag gccagtctgg tttaagtagt catttctgct 3540

gaaaaaacag atgatcctgg tggaagaaaa ggttgaaggc agctgccctc gggagggctg 3600

tgatgctcgg cacatcctgc ctggcacata cacgtgtctg caggccacac cgtgcatgtc 3660

cccagacctg ccgcctggct tctggagtgc ttcaagcaga gcatggtggg tcattgagga 3720

gacccaggaa tctcatctga gaacccactc tctgccggag aaccccatgg tgacacattt 3780

tcatctttct gaccagaggc tgttttttt ttttttgag acagtctcat tctgttgccc 3840

aggctggagt gcagtggctt gatctcggct cactgcaacc tcgcctcccg ggttcaagca 3900

attctctgcc gcagcctcca gagtagctgg gataacaggt gcccaccacc acaccccact 3960

aatttttgta tttgtatttt tagtagagat ggggtttcac catgttggtc aggctggtct 4020

tggactcctg acctcatgct ccacccgctt cggcctccca aagttctggg attacaggtg 4080

tgagccaccg tgcacggccg gcctgacctt tggaaaagcc ttgtcacttt ggacgtttgc 4140

gtctttgaag aggcgatggg agcatatcat gactgcctgc caccattgct tttcagacta 4200

ccacaactca atcatgctgt ccaggacttc tggccctgtg ttcaccactg ggaaaacgta 4260

cttcagactg gatagcctaa aaaggagcaa tgcccttgta ggatgtggag aagggaaaat 4320

acggacatta acattaaaag acaccagtga aattgttagg tctctaggaa gttggagcac 4380

aaggcttcac gctttaagac catctgtggt tttcagtgaa caagcgctga gcaccagcag 4440

cagaaaacaa caacaaaaaa acacctcgtt tttaccttgt cttctagaca tgaaaaggca 4500

gttgcattcc actctgcatt atgttctaca tgttgcttta tcagtatatg cttagctgta 4560

agtgacaagt attttttctg aacagaagtt tacttagaaa taccatgcac ttgggggtac 4620

caattaaccg cctgaaaatt agcatattga tagttcttag agagaccaga tataatctaa 4680

gaatttatat gaaagatttg tatcattaga gccagaaata attttatatt aatatataat 4740

acagattaac attatatata atatgtacct gtgtcacttc tgacatgagc ctgtaaacat 4800

atattcatat atgtacctgc acatgtaccc acctgatgta ggtcttattc ctttagtatg 4860

gacttaaagt acttattcat ataccttgta actaaaaatt agaacagctc cctagaattg 4920

tgaactttta agagtctgac tagaaatttg caacttataa aaaagttact tttaaaaata 4980

taagttaggg ctaggcacag tggctcatgc ctataatctc agcacttttg ggaggccaag 5040

acaggaggat cacttcaggc caggagttca agatcaacca acctgggtaa catggccaga 5100

ccccatctct atttatatat atatatataa aacttagagt ttttatcttc ccctaaaaga 5160

ggccgtgata tttgcagcag cctcaaattg ctcttaaggg gtttaggtgt gcagaagctt 5220

tcctttccct acccagtaac catgtgacta ctaacgtggt atattgattt attttgtttg 5280

ctgtctgtct cccctgcccc actgctggaa cagaggctcc aagaaaacag ggaccttatt 5340

attcattact gcatccccag taatgaaagt acttagaaaa taattattga atgaatgaaa 5400

tctaaactgt gaacctgagg gtgtttgtgg cagtgtttgt tttactgaat tgtagaagga 5460

cataaccgtg ttttcagtgt ttctatggaa caaacttgta cattttattt cacttgtgtt 5520

ttgtcttaaa ccctactgct ggaaacaatt ttatgtaata agcaatgggc ccaaaagtct 5580

aggagttttt ttgtacttag tgaatttgta tgcaacagag atgctgcagc tgatgccttt 5640

aaaaggtatt catcatggaa gagctgaggc ctgtgcttgg tgttccagag cccagggttg 5700

agcatcctga aggagccact gcagccgtca ctgtccccag agcctgtgga gatagagcct 5760

gtttgctgct ttttcttccc gctcttaaga catggctgga gctcagtctt cattgaatga 5820

agtttgctgt ggtattgcat agccttgctt tcttgaacta aactgtttgc ccttcacaag 5880

tagttcttct ttcaggatta gttcgttcca aggaggctct tcagtctcac agataagtag 5940

atctctcctg ctgtctggac acatttcact cggaaattga atacaatttg tattcaggct 6000

gggaacctga acacacactt gtgtttttaa gcttcccttt tttacagtgg acaaggacac 6060

aaataataaa taaatcatcc ctaatgccca agaaaaaaa 6099

<210> 3
<211> 515
<212> PRT
<213> homo sapiens

<400> 3

Met Leu Leu Ser Gln Asn Ala Phe Ile Phe Arg Ser Leu Asn Leu Val
1 5 10 15

Leu Met Val Tyr Ile Ser Leu Val Phe Gly Ile Ser Tyr Asp Ser Pro
20 25 30

Asp Tyr Thr Asp Glu Ser Cys Thr Phe Lys Ile Ser Leu Arg Asn Phe
35 40 45

Arg Ser Ile Leu Ser Trp Glu Leu Lys Asn His Ser Ile Val Pro Thr
50 55 60

His Tyr Thr Leu Leu Tyr Thr Ile Met Ser Lys Pro Glu Asp Leu Lys
65 70 75 80

```
Val Val Lys Asn Cys Ala Asn Thr Thr Arg Ser Phe Cys Asp Leu Thr
                85                  90                  95

Asp Glu Trp Arg Ser Thr His Glu Ala Tyr Val Thr Val Leu Glu Gly
            100                 105                 110

Phe Ser Gly Asn Thr Thr Leu Phe Ser Cys Ser His Asn Phe Trp Leu
        115                 120                 125

Ala Ile Asp Met Ser Phe Glu Pro Pro Glu Phe Glu Ile Val Gly Phe
    130                 135                 140

Thr Asn His Ile Asn Val Met Val Lys Phe Pro Ser Ile Val Glu Glu
145                 150                 155                 160

Glu Leu Gln Phe Asp Leu Ser Leu Val Ile Glu Glu Gln Ser Glu Gly
                165                 170                 175

Ile Val Lys Lys His Lys Pro Glu Ile Lys Gly Asn Met Ser Gly Asn
            180                 185                 190

Phe Thr Tyr Ile Ile Asp Lys Leu Ile Pro Asn Thr Asn Tyr Cys Val
        195                 200                 205

Ser Val Tyr Leu Glu His Ser Asp Glu Gln Ala Val Ile Lys Ser Pro
    210                 215                 220

Leu Lys Cys Thr Leu Leu Pro Pro Gly Gln Glu Ser Glu Ser Ala Glu
225                 230                 235                 240

Ser Ala Lys Ile Gly Gly Ile Ile Thr Val Phe Leu Ile Ala Leu Val
                245                 250                 255

Leu Thr Ser Thr Ile Val Thr Leu Lys Trp Ile Gly Tyr Ile Cys Leu
            260                 265                 270

Arg Asn Ser Leu Pro Lys Val Leu Asn Phe His Asn Phe Leu Ala Trp
        275                 280                 285

Pro Phe Pro Asn Leu Pro Pro Leu Glu Ala Met Asp Met Val Glu Val
    290                 295                 300

Ile Tyr Ile Asn Arg Lys Lys Lys Val Trp Asp Tyr Asn Tyr Asp Asp
```

305 310 315 320

Glu Ser Asp Ser Asp Thr Glu Ala Ala Pro Arg Thr Ser Gly Gly Gly
325 330 335

Tyr Thr Met His Gly Leu Thr Val Arg Pro Leu Gly Gln Ala Ser Ala
340 345 350

Thr Ser Thr Glu Ser Gln Leu Ile Asp Pro Glu Ser Glu Glu Glu Pro
355 360 365

Asp Leu Pro Glu Val Asp Val Glu Leu Pro Thr Met Pro Lys Asp Ser
370 375 380

Pro Gln Gln Leu Glu Leu Leu Ser Gly Pro Cys Glu Arg Arg Lys Ser
385 390 395 400

Pro Leu Gln Asp Pro Phe Pro Glu Glu Asp Tyr Ser Ser Thr Glu Gly
405 410 415

Ser Gly Gly Arg Ile Thr Phe Asn Val Asp Leu Asn Ser Val Phe Leu
420 425 430

Arg Val Leu Asp Asp Glu Asp Ser Asp Asp Leu Glu Ala Pro Leu Met
435 440 445

Leu Ser Ser His Leu Glu Glu Met Val Asp Pro Glu Asp Pro Asp Asn
450 455 460

Val Gln Ser Asn His Leu Leu Ala Ser Gly Glu Gly Thr Gln Pro Thr
465 470 475 480

Phe Pro Ser Pro Ser Ser Glu Gly Leu Trp Ser Glu Asp Ala Pro Ser
485 490 495

Asp Gln Ser Asp Thr Ser Glu Ser Asp Val Asp Leu Gly Asp Gly Tyr
500 505 510

Ile Met Arg
515

<210> 4
<211> 2898

<212> DNA
<213> Homo sapiens

<400> 4
gcccgcgctt ccgtatcgct cctcgtaggc cggggctcgg cgcgcgcacc cgcactaaag 60

acgcttcttc ccggcgggta ggaatcccgc cggcgagccg aacagttccc cgagcgcagc 120

ccgcggacca ccacccggcc gcacgggccg cttttgtccc ccgcccgccg cttctgtccg 180

agaggccgcc cgcgaggcgc atcctgaccg cgagcgtcgg gtcccagagc cgggcgcggc 240

tggggcccga ggctagcatc tctcgggagc cgcaaggcga gagctgcaaa gatgtaaaag 300

tcaagagaag actctaaaaa tagcaaagat gcttttgagc cagaatgcct tcatcttcag 360

atcacttaat ttggttctca tggtgtatat cagcctcgtg tttggtattt catatgattc 420

gcctgattac acagatgaat cttgcacttt caagatatca ttgcgaaatt tccggtccat 480

cttatcatgg gaattaaaaa accactccat tgtaccaact cactatacat tgctgtatac 540

aatcatgagt aaaccagaag atttgaaggt ggttaagaac tgtgcaaata ccacaagatc 600

attttgtgac ctcacagatg agtggagaag cacacacgag gcctatgtca ccgtcctaga 660

aggattcagc gggaacacaa cgttgttcag ttgctcacac aatttctggc tggccataga 720

catgtctttt gaaccaccag agtttgagat tgttggtttt accaaccaca ttaatgtgat 780

ggtgaaattt ccatctattg ttgaggaaga attacagttt gatttatctc tcgtcattga 840

agaacagtca gagggaattg ttaagaagca taaacccgaa ataaaaggaa acatgagtgg 900

aaatttcacc tatatcattg acaagttaat tccaaacacg aactactgtg tatctgttta 960

tttagagcac agtgatgagc aagcagtaat aaagtctccc ttaaaatgca ccctccttcc 1020

acctggccag gaatcagaat cagcagaatc tgccaaaata ggaggaataa ttactgtgtt 1080

tttgatagca ttggtcttga caagcaccat agtgacactg aaatggattg gttatatatg 1140

cttaagaaat agcctcccca aagtcttgaa ttttcataac tttttagcct ggccatttcc 1200

taacctgcca ccgttggaag ccatggatat ggtggaggtc atttacatca acagaaagaa 1260

gaaagtgtgg gattataatt atgatgatga aagtgatagc gatactgagg cagcgcccag 1320

gacaagtggc ggtggctata ccatgcatgg actgactgtc aggcctctgg gtcaggcctc 1380

tgccacctct acagaatccc agttgataga cccggagtcc gaggaggagc ctgacctgcc 1440

tgaggttgat gtggagctcc ccacgatgcc aaaggacagc cctcagcagt tggaactctt 1500

gagtgggccc tgtgagagga gaaagagtcc actccaggac ccttttcccg aagaggacta 1560

cagctccacg gaggggtctg ggggcagaat taccttcaat gtggacttaa actctgtgtt 1620

tttgagagtt cttgatgacg aggacagtga cgacttagaa gcccctctga tgctatcgtc 1680

tcatctggaa gagatggttg acccagagga tcctgataat gtgcaatcaa accatttgct 1740

ggccagcggg gaagggacac agccaacctt tcccagcccc tcttcagagg gcctgtggtc 1800

cgaagatgct ccatctgatc aaagtgacac ttctgagtca gatgttgacc ttggggatgg 1860

ttatataatg agatgactcc aaaactattg aatgaacttg gacagacaag cacctacagg 1920

gttctttgtc tctgcatcct aacttgctgc cttatcgtct gcaagtgttc tccaagggaa 1980

ggaggaggaa actgtggtgt tcctttcttc caggtgacat cacctatgca cattccccagt 2040

atggggacca tagtatcatt cagtgcattg tttacatatt caaagtggtg cactttgaag 2100

gaagcacatg tgcacctttc ctttacacta atgcacttag gatgtttctg catcatgtct 2160

accagggagc agggttcccc acagtttcag aggtggtcca ggaccctatg atatttctct 2220

tctttcgttc ttttttttttt ttttttttga gacagagtct cgttctgtcg cccaagctgg 2280

agcgcaatgg tgtgatcttg gctcactgca acatccgcct cccaggttca agtgattctc 2340

ctgcctcagc ctccctcgca agtagctggg attacaggcg cctgccacca tgcctagcaa 2400

atttttgtat ttttagtaga gacaggattt taccatgttg gccaggctgg tctcaaactc 2460

ctgacctcaa gtgatctgcc ctcctcagcc tcgtaaagtg ctgggattac aggggtgagc 2520

cgctgtgcct ggctggccct gtgatatttc tgtgaaataa attgggccag ggtgggagca 2580

gggaaagaaa aggaaaatag tagcaagagc tgcaaagcag gcaggaaggg aggaggagag 2640

ccaggtgagc agtggagaga agggggggccc tgcacaagga aacagggaag agccatcgaa 2700

gtttcagtcg gtgagccttg ggcacctcac ccatgtcaca tcctgtctcc tgcaattgga 2760

attccacctt gtccagccct ccccagttaa agtggggaag acagacttta ggatcacgtg 2820

tgtgactaat acagaaagga aacatggcgt cggggagagg gataaaacct gaatgccata 2880

ttttaagtta aaaaaaaa 2898

<210> 5
<211> 331
<212> PRT
<213> Homo sapiens

<400> 5

Met Leu Leu Ser Gln Asn Ala Phe Ile Phe Arg Ser Leu Asn Leu Val
1 5 10 15

```
Leu Met Val Tyr Ile Ser Leu Val Phe Gly Ile Ser Tyr Asp Ser Pro
            20                  25                  30

Asp Tyr Thr Asp Glu Ser Cys Thr Phe Lys Ile Ser Leu Arg Asn Phe
        35                  40                  45

Arg Ser Ile Leu Ser Trp Glu Leu Lys Asn His Ser Ile Val Pro Thr
    50                  55                  60

His Tyr Thr Leu Leu Tyr Thr Ile Met Ser Lys Pro Glu Asp Leu Lys
65                  70                  75                  80

Val Val Lys Asn Cys Ala Asn Thr Thr Arg Ser Phe Cys Asp Leu Thr
                85                  90                  95

Asp Glu Trp Arg Ser Thr His Glu Ala Tyr Val Thr Val Leu Glu Gly
            100                 105                 110

Phe Ser Gly Asn Thr Thr Leu Phe Ser Cys Ser His Asn Phe Trp Leu
        115                 120                 125

Ala Ile Asp Met Ser Phe Glu Pro Pro Glu Phe Glu Ile Val Gly Phe
    130                 135                 140

Thr Asn His Ile Asn Val Met Val Lys Phe Pro Ser Ile Val Glu Glu
145                 150                 155                 160

Glu Leu Gln Phe Asp Leu Ser Leu Val Ile Glu Glu Gln Ser Glu Gly
                165                 170                 175

Ile Val Lys Lys His Lys Pro Glu Ile Lys Gly Asn Met Ser Gly Asn
            180                 185                 190

Phe Thr Tyr Ile Ile Asp Lys Leu Ile Pro Asn Thr Asn Tyr Cys Val
        195                 200                 205

Ser Val Tyr Leu Glu His Ser Asp Glu Gln Ala Val Ile Lys Ser Pro
    210                 215                 220

Leu Lys Cys Thr Leu Leu Pro Pro Gly Gln Glu Ser Glu Ser Ala Glu
225                 230                 235                 240

Ser Ala Lys Ile Gly Gly Ile Ile Thr Val Phe Leu Ile Ala Leu Val
```

245 250 255

Leu Thr Ser Thr Ile Val Thr Leu Lys Trp Ile Gly Tyr Ile Cys Leu
260 265 270

Arg Asn Ser Leu Pro Lys Val Leu Arg Gln Gly Leu Ala Lys Gly Trp
275 280 285

Asn Ala Val Ala Ile His Arg Cys Ser His Asn Ala Leu Gln Ser Glu
290 295 300

Thr Pro Glu Leu Lys Gln Ser Ser Cys Leu Ser Phe Pro Ser Ser Trp
305 310 315 320

Asp Tyr Lys Arg Ala Ser Leu Cys Pro Ser Asp
325 330

<210> 6
<211> 1428
<212> DNA
<213> Homo sapiens

<400> 6
gcccgcgctt ccgtatcgct cctcgtaggc cggggctcgg cgcgcgcacc cgcactaaag 60

acgcttcttc ccggcgggta ggaatcccgc cggcgagccg aacagttccc cgagcgcagc 120

ccgcggacca ccacccggcc gcacgggccg cttttgtccc ccgcccgccg cttctgtccg 180

agaggccgcc cgcgaggcgc atcctgaccg cgagcgtcgg gtcccagagc cgggcgcggc 240

tggggcccga ggctagcatc tctcgggagc cgcaaggcga gagctgcaaa gtttaattag 300

acacttcaga attttgatca cctaatgttg atttcagatg taaaagtcaa gagaagactc 360

taaaaatagc aaagatgctt ttgagccaga atgccttcat cttcagatca cttaatttgg 420

ttctcatggt gtatatcagc ctcgtgtttg gtatttcata tgattcgcct gattacacag 480

atgaatcttg cactttcaag atatcattgc gaaatttccg gtccatctta tcatgggaat 540

taaaaaacca ctccattgta ccaactcact atacattgct gtatacaatc atgagtaaac 600

cagaagattt gaaggtggtt aagaactgtg caaataccac aagatcattt tgtgacctca 660

cagatgagtg gagaagcaca cacgaggcct atgtcaccgt cctagaagga ttcagcggga 720

acacaacgtt gttcagttgc tcacacaatt tctggctggc catagacatg tcttttgaac 780

caccagagtt tgagattgtt ggttttacca accacattaa tgtgatggtg aaatttccat 840

ctattgttga ggaagaatta cagtttgatt tatctctcgt cattgaagaa cagtcagagg 900

gaattgttaa gaagcataaa cccgaaataa aaggaaacat gagtggaaat ttcacctata 960

tcattgacaa gttaattcca aacacgaact actgtgtatc tgtttattta gagcacagtg 1020

atgagcaagc agtaataaag tctcccttaa aatgcaccct ccttccacct ggccaggaat 1080

cagaatcagc agaatctgcc aaaataggag gaataattac tgtgtttttg atagcattgg 1140

tcttgacaag caccatagtg acactgaaat ggattggtta tatatgctta agaaatagcc 1200

tccccaaagt cttgaggcaa ggtctcgcta agggctggaa tgcagtggct attcacaggt 1260

gcagtcataa tgcactacag tctgaaactc ctgagctcaa acagtcgtcc tgcctaagct 1320

tccccagtag ctgggattac aagcgtgcat ccctgtgccc cagtgattaa gttttattat 1380

gtagaaaata aagagcaaac agtacagctg aaaaaaaaaa aaaaaaaa 1428

<210> 7
<211> 331
<212> PRT
<213> Homo sapiens

<400> 7

Met Leu Leu Ser Gln Asn Ala Phe Ile Phe Arg Ser Leu Asn Leu Val
1 5 10 15

Leu Met Val Tyr Ile Ser Leu Val Phe Gly Ile Ser Tyr Asp Ser Pro
20 25 30

Asp Tyr Thr Asp Glu Ser Cys Thr Phe Lys Ile Ser Leu Arg Asn Phe
35 40 45

Arg Ser Ile Leu Ser Trp Glu Leu Lys Asn His Ser Ile Val Pro Thr
50 55 60

His Tyr Thr Leu Leu Tyr Thr Ile Met Ser Lys Pro Glu Asp Leu Lys
65 70 75 80

Val Val Lys Asn Cys Ala Asn Thr Thr Arg Ser Phe Cys Asp Leu Thr
85 90 95

Asp Glu Trp Arg Ser Thr His Glu Ala Tyr Val Thr Val Leu Glu Gly
100 105 110

Phe Ser Gly Asn Thr Thr Leu Phe Ser Cys Ser His Asn Phe Trp Leu

```
        115                 120                 125

Ala Ile Asp Met Ser Phe Glu Pro Pro Glu Phe Glu Ile Val Gly Phe
    130                 135                 140

Thr Asn His Ile Asn Val Met Val Lys Phe Pro Ser Ile Val Glu Glu
145                 150                 155                 160

Glu Leu Gln Phe Asp Leu Ser Leu Val Ile Glu Glu Gln Ser Glu Gly
                165                 170                 175

Ile Val Lys Lys His Lys Pro Glu Ile Lys Gly Asn Met Ser Gly Asn
            180                 185                 190

Phe Thr Tyr Ile Ile Asp Lys Leu Ile Pro Asn Thr Asn Tyr Cys Val
        195                 200                 205

Ser Val Tyr Leu Glu His Ser Asp Glu Gln Ala Val Ile Lys Ser Pro
    210                 215                 220

Leu Lys Cys Thr Leu Leu Pro Pro Gly Gln Glu Ser Glu Ser Ala Glu
225                 230                 235                 240

Ser Ala Lys Ile Gly Gly Ile Ile Thr Val Phe Leu Ile Ala Leu Val
                245                 250                 255

Leu Thr Ser Thr Ile Val Thr Leu Lys Trp Ile Gly Tyr Ile Cys Leu
            260                 265                 270

Arg Asn Ser Leu Pro Lys Val Leu Arg Gln Gly Leu Ala Lys Gly Trp
        275                 280                 285

Asn Ala Val Ala Ile His Arg Cys Ser His Asn Ala Leu Gln Ser Glu
    290                 295                 300

Thr Pro Glu Leu Lys Gln Ser Ser Cys Leu Ser Phe Pro Ser Ser Trp
305                 310                 315                 320

Asp Tyr Lys Arg Ala Ser Leu Cys Pro Ser Asp
                325                 330
```

<210> 8
<211> 1382

<212> DNA
<213> Homo sapiens

<400> 8
gcccgcgctt ccgtatcgct cctcgtaggc cggggctcgg cgcgcgcacc cgcactaaag 60

acgcttcttc ccggcgggta ggaatcccgc cggcgagccg aacagttccc cgagcgcagc 120

ccgcggacca ccacccggcc gcacgggccg cttttgtccc ccgcccgccg cttctgtccg 180

agaggccgcc cgcgaggcgc atcctgaccg cgagcgtcgg gtcccagagc cgggcgcggc 240

tggggcccga ggctagcatc tctcgggagc cgcaaggcga gagctgcaaa gatgtaaaag 300

tcaagagaag actctaaaaa tagcaaagat gcttttgagc cagaatgcct tcatcttcag 360

atcacttaat ttggttctca tggtgtatat cagcctcgtg tttggtattt catatgattc 420

gcctgattac acagatgaat cttgcacttt caagatatca ttgcgaaatt tccggtccat 480

cttatcatgg gaattaaaaa accactccat tgtaccaact cactatacat tgctgtatac 540

aatcatgagt aaaccagaag atttgaaggt ggttaagaac tgtgcaaata ccacaagatc 600

attttgtgac ctcacagatg agtggagaag cacacacgag gcctatgtca ccgtcctaga 660

aggattcagc gggaacacaa cgttgttcag ttgctcacac aatttctggc tggccataga 720

catgtctttt gaaccaccag agtttgagat tgttggtttt accaaccaca ttaatgtgat 780

ggtgaaattt ccatctattg ttgaggaaga attacagttt gatttatctc tcgtcattga 840

agaacagtca gagggaattg ttaagaagca taaacccgaa ataaaaggaa acatgagtgg 900

aaatttcacc tatatcattg acaagttaat tccaaacacg aactactgtg tatctgttta 960

tttagagcac agtgatgagc aagcagtaat aaagtctccc ttaaaatgca ccctccttcc 1020

acctggccag gaatcagaat cagcagaatc tgccaaaata ggaggaataa ttactgtgtt 1080

tttgatagca ttggtcttga caagcaccat agtgacactg aaatggattg gttatatatg 1140

cttaagaaat agcctcccca aagtcttgag gcaaggtctc gctaagggct ggaatgcagt 1200

ggctattcac aggtgcagtc ataatgcact acagtctgaa actcctgagc tcaaacagtc 1260

gtcctgccta agcttcccca gtagctggga ttacaagcgt gcatccctgt gccccagtga 1320

ttaagtttta ttatgtagaa aataaagagc aaacagtaca gctgaaaaaa aaaaaaaaaa 1380

aa 1382

<210> 9
<211> 189
<212> PRT
<213> Homo sapiens

<400> 9

Met Ala Leu Ser Phe Ser Leu Leu Met Ala Val Leu Val Leu Ser Tyr
1 5 10 15

Lys Ser Ile Cys Ser Leu Gly Cys Asp Leu Pro Gln Thr His Ser Leu
20 25 30

Gly Asn Arg Arg Ala Leu Ile Leu Leu Ala Gln Met Gly Arg Ile Ser
35 40 45

His Phe Ser Cys Leu Lys Asp Arg His Asp Phe Gly Phe Pro Glu Glu
50 55 60

Glu Phe Asp Gly His Gln Phe Gln Lys Ala Gln Ala Ile Ser Val Leu
65 70 75 80

His Glu Met Ile Gln Gln Thr Phe Asn Leu Phe Ser Thr Glu Asp Ser
85 90 95

Ser Ala Ala Trp Glu Gln Ser Leu Leu Glu Lys Phe Ser Thr Glu Leu
100 105 110

Tyr Gln Gln Leu Asn Asp Leu Glu Ala Cys Val Ile Gln Glu Val Gly
115 120 125

Val Glu Glu Thr Pro Leu Met Asn Glu Asp Ser Ile Leu Ala Val Arg
130 135 140

Lys Tyr Phe Gln Arg Ile Thr Leu Tyr Leu Thr Glu Lys Lys Tyr Ser
145 150 155 160

Pro Cys Ala Trp Glu Val Val Arg Ala Glu Ile Met Arg Ser Leu Ser
165 170 175

Phe Ser Thr Asn Leu Gln Lys Arg Leu Arg Arg Lys Asp
180 185

<210> 10
<211> 189
<212> PRT
<213> Homo sapiens

<400> 10

```
Met Ala Leu Ser Phe Ser Leu Leu Met Ala Val Leu Val Leu Ser Tyr
1               5                   10                  15

Lys Ser Ile Cys Ser Leu Gly Cys Asp Leu Pro Gln Thr His Ser Leu
            20                  25                  30

Gly Asn Arg Arg Ala Leu Ile Leu Leu Ala Gln Met Gly Arg Ile Ser
        35                  40                  45

His Phe Ser Cys Leu Lys Asp Arg His Asp Phe Gly Phe Pro Glu Glu
    50                  55                  60

Glu Phe Asp Gly His Gln Phe Gln Lys Ala Gln Ala Ile Ser Val Leu
65                  70                  75                  80

His Glu Met Ile Gln Gln Thr Phe Asn Leu Phe Ser Thr Glu Asp Ser
                85                  90                  95

Ser Ala Ala Trp Glu Gln Ser Leu Leu Glu Lys Phe Ser Thr Glu Leu
            100                 105                 110

Tyr Gln Gln Leu Asn Asp Leu Glu Ala Cys Val Ile Gln Glu Val Gly
        115                 120                 125

Val Glu Glu Thr Pro Leu Met Asn Glu Asp Ser Ile Leu Ala Val Arg
    130                 135                 140

Lys Tyr Phe Gln Arg Ile Thr Leu Tyr Leu Thr Glu Lys Lys Tyr Ser
145                 150                 155                 160

Pro Cys Ala Trp Glu Val Val Arg Ala Glu Ile Met Arg Ser Leu Ser
                165                 170                 175

Phe Ser Thr Asn Leu Gln Lys Arg Leu Arg Arg Lys Asp
            180                 185
```

<210> 11
<211> 189
<212> PRT
<213> Homo sapiens

<400> 11

```
Met Ala Leu Ser Phe Ser Leu Leu Met Ala Val Leu Val Leu Ser Tyr
```

```
1               5                   10                  15

Lys Ser Ile Cys Ser Leu Gly Cys Asp Leu Pro Gln Thr His Ser Leu
            20                  25                  30

Gly Asn Arg Arg Ala Leu Ile Leu Leu Ala Gln Met Gly Arg Ile Ser
        35                  40                  45

Pro Phe Ser Cys Leu Lys Asp Arg His Asp Phe Gly Leu Pro Gln Glu
    50                  55                  60

Glu Phe Asp Gly Asn Gln Phe Gln Lys Thr Gln Ala Ile Ser Val Leu
65                  70                  75                  80

His Glu Met Ile Gln Gln Thr Phe Asn Leu Phe Ser Thr Glu Asp Ser
                85                  90                  95

Ser Ala Ala Trp Glu Gln Ser Leu Leu Glu Lys Phe Ser Thr Glu Leu
            100                 105                 110

Tyr Gln Gln Leu Asn Asn Leu Glu Ala Cys Val Ile Gln Glu Val Gly
        115                 120                 125

Met Glu Glu Thr Pro Leu Met Asn Glu Asp Ser Ile Leu Ala Val Arg
    130                 135                 140

Lys Tyr Phe Gln Arg Ile Thr Leu Tyr Leu Thr Glu Lys Lys Tyr Ser
145                 150                 155                 160

Pro Cys Ala Trp Glu Val Val Arg Ala Glu Ile Met Arg Ser Leu Ser
                165                 170                 175

Phe Ser Thr Asn Leu Gln Lys Ile Leu Arg Arg Lys Asp
            180                 185
```

<210> 12
<211> 189
<212> PRT
<213> Homo sapiens

<400> 12

```
Met Ala Arg Ser Phe Ser Leu Leu Met Val Val Leu Val Leu Ser Tyr
1               5                   10                  15
```

```
Lys Ser Ile Cys Ser Leu Gly Cys Asp Leu Pro Gln Thr His Ser Leu
            20                  25                  30

Arg Asn Arg Arg Ala Leu Ile Leu Leu Ala Gln Met Gly Arg Ile Ser
        35                  40                  45

Pro Phe Ser Cys Leu Lys Asp Arg His Glu Phe Arg Phe Pro Glu Glu
    50                  55                  60

Glu Phe Asp Gly His Gln Phe Gln Lys Thr Gln Ala Ile Ser Val Leu
65                  70                  75                  80

His Glu Met Ile Gln Gln Thr Phe Asn Leu Phe Ser Thr Glu Asp Ser
                85                  90                  95

Ser Ala Ala Trp Glu Gln Ser Leu Leu Glu Lys Phe Ser Thr Glu Leu
            100                 105                 110

Tyr Gln Gln Leu Asn Asp Leu Glu Ala Cys Val Ile Gln Glu Val Gly
        115                 120                 125

Val Glu Glu Thr Pro Leu Met Asn Glu Asp Phe Ile Leu Ala Val Arg
    130                 135                 140

Lys Tyr Phe Gln Arg Ile Thr Leu Tyr Leu Met Glu Lys Lys Tyr Ser
145                 150                 155                 160

Pro Cys Ala Trp Glu Val Val Arg Ala Glu Ile Met Arg Ser Phe Ser
                165                 170                 175

Phe Ser Thr Asn Leu Lys Lys Gly Leu Arg Arg Lys Asp
            180                 185
```

<210> 13
<211> 189
<212> PRT
<213> Homo sapiens

<400> 13

```
Met Ala Leu Ser Phe Ser Leu Leu Met Ala Val Leu Val Leu Ser Tyr
1               5                   10                  15

Lys Ser Ile Cys Ser Leu Gly Cys Asp Leu Pro Gln Thr His Ser Leu
```

```
            20                  25                  30

Gly Asn Arg Arg Ala Leu Ile Leu Leu Ala Gln Met Gly Arg Ile Ser
        35                  40                  45

Pro Phe Ser Cys Leu Lys Asp Arg His Asp Phe Gly Phe Pro Gln Glu
    50                  55                  60

Glu Phe Asp Gly Asn Gln Phe Gln Lys Ala Gln Ala Ile Ser Val Leu
65                  70                  75                  80

His Glu Met Ile Gln Gln Thr Phe Asn Leu Phe Ser Thr Lys Asp Ser
                85                  90                  95

Ser Ala Thr Trp Glu Gln Ser Leu Leu Glu Lys Phe Ser Thr Glu Leu
            100                 105                 110

Asn Gln Gln Leu Asn Asp Leu Glu Ala Cys Val Ile Gln Glu Val Gly
        115                 120                 125

Val Glu Glu Thr Pro Leu Met Asn Val Asp Ser Ile Leu Ala Val Lys
    130                 135                 140

Lys Tyr Phe Gln Arg Ile Thr Leu Tyr Leu Thr Glu Lys Lys Tyr Ser
145                 150                 155                 160

Pro Cys Ala Trp Glu Val Val Arg Ala Glu Ile Met Arg Ser Phe Ser
                165                 170                 175

Leu Ser Lys Ile Phe Gln Glu Arg Leu Arg Arg Lys Glu
            180                 185
```

<210> 14
<211> 189
<212> PRT
<213> Homo sapiens

<400> 14

```
Met Ala Leu Ser Phe Ser Leu Leu Met Ala Val Leu Val Leu Ser Tyr
1               5                   10                  15

Lys Ser Ile Cys Ser Leu Gly Cys Asp Leu Pro Gln Thr His Ser Leu
            20                  25                  30
```

```
Gly Asn Arg Arg Ala Leu Ile Leu Leu Ala Gln Met Gly Arg Ile Ser
        35                  40                  45

His Phe Ser Cys Leu Lys Asp Arg Tyr Asp Phe Gly Phe Pro Gln Glu
    50                  55                  60

Val Phe Asp Gly Asn Gln Phe Gln Lys Ala Gln Ala Ile Ser Ala Phe
65                  70                  75                  80

His Glu Met Ile Gln Gln Thr Phe Asn Leu Phe Ser Thr Lys Asp Ser
                85                  90                  95

Ser Ala Ala Trp Asp Glu Thr Leu Leu Asp Lys Phe Tyr Ile Glu Leu
            100                 105                 110

Phe Gln Gln Leu Asn Asp Leu Glu Ala Cys Val Thr Gln Glu Val Gly
        115                 120                 125

Val Glu Glu Ile Ala Leu Met Asn Glu Asp Ser Ile Leu Ala Val Arg
    130                 135                 140

Lys Tyr Phe Gln Arg Ile Thr Leu Tyr Leu Met Gly Lys Lys Tyr Ser
145                 150                 155                 160

Pro Cys Ala Trp Glu Val Val Arg Ala Glu Ile Met Arg Ser Phe Ser
                165                 170                 175

Phe Ser Thr Asn Leu Gln Lys Gly Leu Arg Arg Lys Asp
            180                 185
```

<210> 15
<211> 169
<212> PRT
<213> Homo sapiens

<400> 15

```
Ser Leu Gly Cys Asp Leu Pro Gln Thr His Ser Leu Gly Asn Arg Arg
1               5                   10                  15

Ala Leu Ile Leu Leu Gly Gln Met Gly Arg Ile Ser Pro Phe Ser Cys
            20                  25                  30

Leu Lys Asp Arg His Asp Phe Arg Ile Pro Gln Glu Glu Phe Asp Gly
```

```
        35                  40                  45

Asn Gln Phe Gln Lys Ala Gln Ala Ile Ser Val Leu His Glu Met Ile
    50                  55                  60

Gln Gln Thr Phe Asn Leu Phe Ser Thr Glu Asp Ser Ser Ala Ala Trp
65                  70                  75                  80

Glu Gln Ser Leu Leu Glu Lys Phe Ser Thr Glu Leu Tyr Gln Gln Leu
                85                  90                  95

Asn Asp Leu Glu Ala Cys Val Ile Gln Glu Val Gly Val Glu Glu Thr
            100                 105                 110

Pro Leu Met Asn Glu Asp Ser Ile Leu Ala Val Arg Lys Tyr Phe Gln
        115                 120                 125

Arg Ile Thr Leu Tyr Leu Ile Glu Arg Lys Tyr Ser Pro Cys Ala Trp
    130                 135                 140

Glu Val Val Arg Ala Glu Ile Met Arg Ser Leu Ser Phe Ser Thr Asn
145                 150                 155                 160

Leu Gln Lys Arg Leu Arg Arg Lys Asp
                165
```

<210> 16
<211> 189
<212> PRT
<213> Homo sapiens

<400> 16

```
Met Ala Leu Pro Phe Val Leu Leu Met Ala Leu Val Val Leu Asn Cys
1               5                   10                  15

Lys Ser Ile Cys Ser Leu Gly Cys Asp Leu Pro Gln Thr His Ser Leu
            20                  25                  30

Ser Asn Arg Arg Thr Leu Met Ile Met Ala Gln Met Gly Arg Ile Ser
        35                  40                  45

Pro Phe Ser Cys Leu Lys Asp Arg His Asp Phe Gly Phe Pro Gln Glu
    50                  55                  60
```

```
Glu Phe Asp Gly Asn Gln Phe Gln Lys Ala Gln Ala Ile Ser Val Leu
65                  70                  75                  80

His Glu Met Ile Gln Gln Thr Phe Asn Leu Phe Ser Thr Lys Asp Ser
                85                  90                  95

Ser Ala Thr Trp Asp Glu Thr Leu Leu Asp Lys Phe Tyr Thr Glu Leu
            100                 105                 110

Tyr Gln Gln Leu Asn Asp Leu Glu Ala Cys Met Met Gln Glu Val Gly
        115                 120                 125

Val Glu Asp Thr Pro Leu Met Asn Val Asp Ser Ile Leu Thr Val Arg
    130                 135                 140

Lys Tyr Phe Gln Arg Ile Thr Leu Tyr Leu Thr Glu Lys Lys Tyr Ser
145                 150                 155                 160

Pro Cys Ala Trp Glu Val Val Arg Ala Glu Ile Met Arg Ser Phe Ser
                165                 170                 175

Leu Ser Ala Asn Leu Gln Glu Arg Leu Arg Arg Lys Glu
            180                 185
```

<210> 17
<211> 189
<212> PRT
<213> Homo sapiens

<400> 17

```
Met Ala Leu Pro Phe Ala Leu Leu Met Ala Leu Val Val Leu Ser Cys
1               5                   10                  15

Lys Ser Ser Cys Ser Leu Asp Cys Asp Leu Pro Gln Thr His Ser Leu
            20                  25                  30

Gly His Arg Arg Thr Met Met Leu Leu Ala Gln Met Arg Arg Ile Ser
        35                  40                  45

Leu Phe Ser Cys Leu Lys Asp Arg His Asp Phe Arg Phe Pro Gln Glu
    50                  55                  60

Glu Phe Asp Gly Asn Gln Phe Gln Lys Ala Glu Ala Ile Ser Val Leu
```

65 70 75 80

His Glu Val Ile Gln Gln Thr Phe Asn Leu Phe Ser Thr Lys Asp Ser
85 90 95

Ser Val Ala Trp Asp Glu Arg Leu Leu Asp Lys Leu Tyr Thr Glu Leu
100 105 110

Tyr Gln Gln Leu Asn Asp Leu Glu Ala Cys Val Met Gln Glu Val Trp
115 120 125

Val Gly Gly Thr Pro Leu Met Asn Glu Asp Ser Ile Leu Ala Val Arg
130 135 140

Lys Tyr Phe Gln Arg Ile Thr Leu Tyr Leu Thr Glu Lys Lys Tyr Ser
145 150 155 160

Pro Cys Ala Trp Glu Val Val Arg Ala Glu Ile Met Arg Ser Phe Ser
165 170 175

Ser Ser Arg Asn Leu Gln Glu Arg Leu Arg Arg Lys Glu
180 185

<210> 18
<211> 189
<212> PRT
<213> Homo sapiens

<400> 18

Met Ala Ser Pro Phe Ala Leu Leu Met Val Leu Val Val Leu Ser Cys
1 5 10 15

Lys Ser Ser Cys Ser Leu Gly Cys Asp Leu Pro Glu Thr His Ser Leu
20 25 30

Asp Asn Arg Arg Thr Leu Met Leu Leu Ala Gln Met Ser Arg Ile Ser
35 40 45

Pro Ser Ser Cys Leu Met Asp Arg His Asp Phe Gly Phe Pro Gln Glu
50 55 60

Glu Phe Asp Gly Asn Gln Phe Gln Lys Ala Pro Ala Ile Ser Val Leu
65 70 75 80

```
His Glu Leu Ile Gln Gln Ile Phe Asn Leu Phe Thr Thr Lys Asp Ser
                85                  90                  95

Ser Ala Ala Trp Asp Glu Asp Leu Leu Asp Lys Phe Cys Thr Glu Leu
            100                 105                 110

Tyr Gln Gln Leu Asn Asp Leu Glu Ala Cys Val Met Gln Glu Glu Arg
        115                 120                 125

Val Gly Glu Thr Pro Leu Met Asn Ala Asp Ser Ile Leu Ala Val Lys
    130                 135                 140

Lys Tyr Phe Arg Arg Ile Thr Leu Tyr Leu Thr Glu Lys Lys Tyr Ser
145                 150                 155                 160

Pro Cys Ala Trp Glu Val Val Arg Ala Glu Ile Met Arg Ser Leu Ser
                165                 170                 175

Leu Ser Thr Asn Leu Gln Glu Arg Leu Arg Arg Lys Glu
            180                 185
```

<210> 19
<211> 188
<212> PRT
<213> Homo sapiens

<400> 19

```
Met Ala Leu Thr Phe Ala Leu Leu Val Ala Leu Leu Val Leu Ser Cys
1               5                   10                  15

Lys Ser Ser Cys Ser Val Gly Cys Asp Leu Pro Gln Thr His Ser Leu
            20                  25                  30

Gly Ser Arg Arg Thr Leu Met Leu Leu Ala Gln Met Arg Arg Ile Ser
        35                  40                  45

Leu Phe Ser Cys Leu Lys Asp Arg His Asp Phe Gly Phe Pro Gln Glu
    50                  55                  60

Glu Phe Gly Asn Gln Phe Gln Lys Ala Glu Thr Ile Pro Val Leu His
65                  70                  75                  80

Glu Met Ile Gln Gln Ile Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser
```

```
                85                  90                  95

Ala Ala Trp Asp Glu Thr Leu Leu Asp Lys Phe Tyr Thr Glu Leu Tyr
            100                 105                 110

Gln Gln Leu Asn Asp Leu Glu Ala Cys Val Ile Gln Gly Val Gly Val
        115                 120                 125

Thr Glu Thr Pro Leu Met Lys Glu Asp Ser Ile Leu Ala Val Arg Lys
    130                 135                 140

Tyr Phe Gln Arg Ile Thr Leu Tyr Leu Lys Glu Lys Lys Tyr Ser Pro
145                 150                 155                 160

Cys Ala Trp Glu Val Val Arg Ala Glu Ile Met Arg Ser Phe Ser Leu
                165                 170                 175

Ser Thr Asn Leu Gln Glu Ser Leu Arg Ser Lys Glu
            180                 185
```

<210> 20
<211> 182
<212> PRT
<213> Homo sapiens

<400> 20

```
Leu Leu Val Ala Leu Leu Val Leu Ser Cys Lys Ser Ser Cys Ser Val
1               5                   10                  15

Gly Cys Asp Leu Pro Gln Thr His Ser Leu Gly Ser Arg Arg Thr Leu
            20                  25                  30

Met Leu Leu Ala Gln Met Arg Arg Ile Ser Leu Phe Ser Cys Leu Lys
        35                  40                  45

Asp Arg His Asp Phe Gly Phe Pro Gln Glu Glu Phe Gly Asn Gln Phe
    50                  55                  60

Gln Lys Ala Glu Thr Ile Pro Val Leu His Glu Met Ile Gln Gln Ile
65                  70                  75                  80

Phe Asn Leu Phe Ser Thr Lys Asp Ser Ser Ala Ala Trp Asp Glu Thr
                85                  90                  95
```

```
Leu Leu Asp Lys Phe Tyr Thr Glu Leu Tyr Gln Gln Leu Asn Asp Leu
            100                 105                 110

Glu Ala Cys Val Ile Gln Gly Val Gly Val Thr Glu Thr Pro Leu Met
        115                 120                 125

Lys Glu Asp Ser Ile Leu Ala Val Arg Lys Tyr Phe Gln Arg Ile Thr
    130                 135                 140

Leu Tyr Leu Lys Glu Lys Lys Tyr Ser Pro Cys Ala Trp Glu Val Val
145                 150                 155                 160

Arg Ala Glu Ile Met Arg Ser Phe Ser Leu Ser Thr Asn Leu Gln Glu
                165                 170                 175

Ser Leu Arg Ser Lys Glu
            180
```

<210> 21
<211> 189
<212> PRT
<213> Homo sapiens

<400> 21

```
Met Ala Leu Pro Phe Ala Leu Met Met Ala Leu Val Val Leu Ser Cys
1               5                   10                  15

Lys Ser Ser Cys Ser Leu Gly Cys Asn Leu Ser Gln Thr His Ser Leu
            20                  25                  30

Asn Asn Arg Arg Thr Leu Met Leu Met Ala Gln Met Arg Arg Ile Ser
        35                  40                  45

Pro Phe Ser Cys Leu Lys Asp Arg His Asp Phe Glu Phe Pro Gln Glu
    50                  55                  60

Glu Phe Asp Gly Asn Gln Phe Gln Lys Ala Gln Ala Ile Ser Val Leu
65                  70                  75                  80

His Glu Met Met Gln Gln Thr Phe Asn Leu Phe Ser Thr Lys Asn Ser
                85                  90                  95

Ser Ala Ala Trp Asp Glu Thr Leu Leu Glu Lys Phe Tyr Ile Glu Leu
```

```
            100                 105                 110

Phe Gln Gln Met Asn Asp Leu Glu Ala Cys Val Ile Gln Glu Val Gly
        115                 120                 125

Val Glu Glu Thr Pro Leu Met Asn Glu Asp Ser Ile Leu Ala Val Lys
    130                 135                 140

Lys Tyr Phe Gln Arg Ile Thr Leu Tyr Leu Met Glu Lys Lys Tyr Ser
145                 150                 155                 160

Pro Cys Ala Trp Glu Val Val Arg Ala Glu Ile Met Arg Ser Leu Ser
                165                 170                 175

Phe Ser Thr Asn Leu Gln Lys Arg Leu Arg Arg Lys Asp
            180                 185
```

<210> 22
<211> 189
<212> PRT
<213> Homo sapiens

<400> 22

```
Met Ala Leu Thr Phe Tyr Leu Leu Val Ala Leu Val Val Leu Ser Tyr
1               5                   10                  15

Lys Ser Phe Ser Ser Leu Gly Cys Asp Leu Pro Gln Thr His Ser Leu
            20                  25                  30

Gly Asn Arg Arg Ala Leu Ile Leu Leu Ala Gln Met Arg Arg Ile Ser
        35                  40                  45

Pro Phe Ser Cys Leu Lys Asp Arg His Asp Phe Glu Phe Pro Gln Glu
    50                  55                  60

Glu Phe Asp Asp Lys Gln Phe Gln Lys Ala Gln Ala Ile Ser Val Leu
65                  70                  75                  80

His Glu Met Ile Gln Gln Thr Phe Asn Leu Phe Ser Thr Lys Asp Ser
                85                  90                  95

Ser Ala Ala Leu Asp Glu Thr Leu Leu Asp Glu Phe Tyr Ile Glu Leu
            100                 105                 110
```

```
Asp Gln Gln Leu Asn Asp Leu Glu Ser Cys Val Met Gln Glu Val Gly
        115                 120                 125

Val Ile Glu Ser Pro Leu Met Tyr Glu Asp Ser Ile Leu Ala Val Arg
    130                 135                 140

Lys Tyr Phe Gln Arg Ile Thr Leu Tyr Leu Thr Glu Lys Lys Tyr Ser
145                 150                 155                 160

Ser Cys Ala Trp Glu Val Val Arg Ala Glu Ile Met Arg Ser Phe Ser
                165                 170                 175

Leu Ser Ile Asn Leu Gln Lys Arg Leu Lys Ser Lys Glu
            180                 185
```

<210> 23
<211> 189
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<222> (103)..(103)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (106)..(106)
<223> Xaa can be any naturally occurring amino acid

<220>
<221> misc_feature
<222> (189)..(189)
<223> Xaa can be any naturally occurring amino acid

<400> 23

```
Met Ala Leu Ser Phe Ser Leu Leu Met Ala Leu Leu Val Leu Ser Tyr
1               5                   10                  15

Lys Ser Ile Cys Ser Leu Gly Cys Asp Leu Pro Gln Thr His Ser Leu
            20                  25                  30

Gly Asn Arg Arg Ala Leu Ile Leu Leu Ala Gln Met Gly Arg Ile Ser
        35                  40                  45

Pro Phe Ser Cys Leu Lys Asp Arg His Asp Phe Gly Phe Pro Gln Glu
```

```
    50                  55                  60

Glu Phe Asp Gly Asn Gln Phe Gln Lys Ala Gln Ala Ile Ser Val Leu
65                  70                  75                  80

His Glu Met Ile Gln Gln Thr Phe Asn Leu Phe Ser Thr Lys Asp Ser
                85                  90                  95

Ser Ala Ala Trp Asp Glu Xaa Leu Leu Xaa Lys Phe Tyr Thr Glu Leu
            100                 105                 110

Tyr Gln Gln Leu Asn Asp Leu Glu Ala Cys Val Ile Gln Glu Val Gly
        115                 120                 125

Val Glu Glu Thr Pro Leu Met Asn Glu Asp Ser Ile Leu Ala Val Arg
    130                 135                 140

Lys Tyr Phe Gln Arg Ile Thr Leu Tyr Leu Thr Glu Lys Lys Tyr Ser
145                 150                 155                 160

Pro Cys Ala Trp Glu Val Val Arg Ala Glu Ile Met Arg Ser Phe Ser
                165                 170                 175

Phe Ser Thr Asn Leu Gln Lys Arg Leu Arg Arg Lys Xaa
            180                 185
```