HEPATITIS C VIRUS NS5B POLYMERASE INHIBITOR BINDING POCKET

FIELD OF THE INVENTION

The present invention relates to Hepatitis C Virus NS5B polymerase (HCV NS5B), and in particular, to a novel HCV NS5B inhibitor binding pocket. The crystal structure of HCV NS5B is provided including the coordinates that define the novel inhibitor-binding pocket and 3-dimensional models for use in inhibitor screening. Methods of designing and screening NS5B inhibitors are also provided utilizing the crystal structure and inhibitor-binding information.

BACKGROUND OF THE INVENTION

Hepatitis C virus (HCV) is the major etiological agent of post-transfusion and community-acquired non-A non-B hepatitis worldwide. It is estimated that over 200 million people worldwide are infected by the virus. A high percentage of carriers become chronically infected and many progress to chronic liver disease, so called chronic hepatitis C. This group is in turn at high risk for serious liver disease such as liver cirrhosis, hepatocellular carcinoma and terminal liver disease leading to death.

The mechanism by which HCV establishes viral persistence and causes a high rate of chronic liver disease has not been thoroughly elucidated. It is not known how HCV interacts with and evades the host immune system. In addition, the roles of cellular and humoral immune responses in protecting against HCV infection and disease have yet to be established.

HCV is an enveloped positive strand RNA virus in the *Flaviviridae* family. The single strand HCV RNA genome is of positive polarity and comprises one open reading frame (ORF) of approximately 9600 nucleotides in length, which encodes a linear polyprotein of approx. 3010 amino acids. In infected cells, this polyprotein is cleaved at multiple sites by cellular and viral proteases to produce structural and non-structural (NS) proteins. The structural proteins (C, E1, E2 and E2-p7) comprise polypeptides that constitute the virus particle. The non-structural proteins (NS2, NS3, NS4A, NS4B, NS5A, NS5B) encode for enzymes or accessory factors that catalyze and regulate the replication of the HCV RNA genome. Processing of the structural proteins is catalyzed by host cell proteases. The generation of the mature non-structural proteins is catalyzed by two virally encoded proteases. The first is the

NS2/3 zinc-dependent metalloprotease which auto-catalyzes the release of the NS3 protein from the polyprotein. The released NS3 protein contains an N-terminal serine protease domain and catalyzes the remaining cleavages from the polyprotein. The released NS4A protein has at least two roles. The first role is forming a stable complex with NS3 protein and assisting in the membrane localization of the NS3/NS4A complex; the second is acting as a cofactor for NS3 protease activity. This membrane-associated complex, in turn catalyzes the cleavage of the remaining sites on the polyprotein, thus effecting the release of NS4B, NS5A and NS5B. The C-terminal segment of the NS3 protein also harbors nucleoside triphosphatase and RNA helicase activity. The function of the protein NS4B is unknown. NS5A is a highly phosphorylated protein that appears to be responsible for the interferon resistance of various HCV genotypes. NS5B is an RNA-dependent RNA polymerase (RdRp) that is involved in the replication of HCV.

The open reading frame of the HCV RNA genome is flanked on its 5' end by a non-translated region (NTR) of approx. 340 nucleotides that functions as the internal ribosome entry site (IRES), and on its 3' end by an NTR of approximately 230 nucleotides. Both the 5' and 3' NTRs are important for RNA genome replication. The genomic sequence variance is not evenly distributed over the genome and the 5'NTR and parts of the 3'NTR are the most highly conserved portions.

The cloned and characterized partial and complete sequences of the HCV genome have been analyzed with regard to appropriate targets for a prospective antiviral therapy. The following four viral enzyme activities provide possible targets: (1) the NS2/3 protease; (2) the NS3/4A protease complex, (3) the NS3 helicase and (4) the NS5B RNA-dependent RNA polymerase (NS5B RdRp). The NS5B RNA dependent RNA polymerase has also been crystallized to reveal a structure reminiscent of other nucleic acid polymerases (Ago *et al.* **1999**; Bressanelli *et al.* **1999**; Lesburg *et al.* **1999**) with an enclosed active site.

Virus-specific functions essential for replication are the most attractive targets for drug development. The absence of RNA dependent RNA polymerases in mammals, and the fact that this enzyme appears to be essential to viral replication, would suggest that the HCV NS5B polymerase is an ideal target for anti-HCV therapeutics. It has recently been demonstrated that mutations destroying NS5B activity abolish

infectivity of HCV RNA in a chimp model (Kolykhalov, A.A. *et al.* **2000**). The initial step of viral RNA replication is recognition of the 3'-end of RNA template by NS5B (RdRp), which may occur directly or indirectly with the help of cellular proteins (Lai, **1998**; Strauss *et al.*, **1999**). HCV polymerase then proceeds to elongate this template and form a complementary RNA product.

Several groups have described the crystal structure of HCV NS5B polymerase (Ago *et al.* **1999** *supra*; Bressanelli *et al.* **1999** *supra*; Lesburg *et al.* **1999** *supra*). It resembles a flattened sphere with the approximate dimensions 70Å x 60Å x 40Å. The polypeptide chain encircles the active site, forming a cavity at the center of the molecule, and resulting in an appearance that is very different from other U-shaped polymerases. The domain organization of NS5B is consistent with other polymerases in that it is subdivided into finger, palm and thumb domains in which the palm domain, i.e. residues 188-225 and 287-370, is conserved. In contrast to other polymerases, extensive thumb and finger domain contacts result in a globular-shaped HCV polymerase. These contacts are mediated, in part, by loops that extend from the finger to the thumb domain. Knowledge of the crystal structure of NS5B is useful for structure-based drug design and, indeed, structures of NS5B polymerase/inhibitor complexes have been reported recently (Wang *et al.* **2003**; Love et al, **2003**; EP 1 256 628). Non-nucleoside analogue inhibitors were found to bind in a wedge-like fashion to a hydrophobic binding pocket located near the C-terminal region of the polymerase thumb domain. In this study, the enzyme was determined to undergo only minor conformational changes in the enzyme/inhibitor complex. At least two NTP binding sites have been characterized on NS5B, one in the active site palm and a second potential allosteric site on the thumb (O'Farrell *et al.* **2003**; Bressanelli *et al.* **2002**).

Interestingly, Labonté *et al.* **2002** have reported that a mutation of Leu30 in the N-terminal finger loop of the NS5B affects its polymerase activity and speculate that a local alteration in the structure of the Leu30 mutant is responsible for this decrease in activity. However, the authors are silent on the presence of a binding pocket that is "masked" by the finger loop in its native state and becomes exposed by a mutation or displacement of the Leu30 residue. The discovery of this peculiar binding pocket is the subject-matter of the present invention.

Accordingly, the effort to develop effective treatments to HCV infection can be facilitated by increased knowledge of the structure of enzymes critical to HCV replication, most notably, the NS5B polymerase. An increased knowledge of enzyme structure, particularly when complexed with specific inhibitors, will lead to a means of identifying binding sites in the enzyme, as well as the conformation of enzyme/inhibitor complexes and susceptible residues in the enzyme, knowledge of each of which is critical to the process of drug design and optimization.

SUMMARY OF THE INVENTION

It has now been found that the HCV polymerase, when complexed with certain inhibitors, adopts a conformation in which the finger loop region defined by amino acid residues 18 to 35 is displaced to expose a binding pocket defined generally by amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503. This is in contrast to NS5B crystal structures disclosed in the prior art in which the finger loop defined by amino acid residues 18 to 35 conceals this binding pocket in its native state. This newly "exposed" binding pocket defines a novel target in the search of further chemical entities which are capable of binding to HCV NS5B and modulating, or preferably inhibiting, the polymerase activity of HCV NS5B.

Thus, in one aspect, the present invention provides an isolated and purified polypeptide comprising an HCV NS5B inhibitor binding pocket defined by at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 of native HCV NS5B, or defined by a functionally equivalent analog thereof, wherein said binding pocket is exposed by displacement of a finger loop chain defined by at least amino acid residues 18 to 35 and said binding pocket retains its native functional configuration.

In a second aspect of the present invention, there is provided an isolated and purified HCV NS5B polypeptide analog comprising an HCV NS5B binding pocket defined by at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 of native HCV NS5B, or defined by a functionally equivalent analog thereof, wherein said binding pocket retains its native functional configuration and wherein said binding pocket is exposable.

In a further aspect, there is provided an isolated and purified HCV NS5B polypeptide consisting of an HCV NS5B binding pocket defined by at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 of native HCV NS5B, or defined by a functionally equivalent analog thereof, wherein said binding pocket retains its native functional configuration.

In another aspect, there is provided an HCV NS5B polypeptide variant comprising at least one amino acid mutation within a finger loop defined by amino acid residues 18 to 35 of an HCV NS5B, wherein said mutation provokes displacement of said finger loop to expose a binding pocket essentially defined by amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 of native HCV NS5B, or defined by a functionally equivalent analog thereof, and wherein said binding pocket retains its native functional configuration.

In another aspect of the present invention, there is provided an HCV NS5B polypeptide, or a functionally equivalent analog thereof, characterized by displacement of amino acid residues 18 to 35.

In another aspect of the present invention, there is provided an HCV NS5B polypeptide, or a functionally equivalent analog thereof, in which at least amino acid residues 18 to 35 have been deleted.

In another aspect of the invention, there is provided an HCV NS5B crystal structure comprising a binding pocket defined by the structural coordinates of at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 of a native HCV NS5B, or defined by the structural coordinates of a functionally equivalent analog thereof, wherein a native finger loop chain defined by the structural coordinates of at least amino acids 18 to 35 is displaced to expose said binding pocket.

In another aspect of the present invention, there is provided a complex comprising an HCV NS5B polypeptide, polypeptide variant or polypeptide analog as defined above and a compound, wherein the compound associates with a binding pocket within the NS5B polypeptide, polypeptide variant or polypeptide analog, said binding pocket being defined by amino acid residues 392, 393, 395, 396, 399, 424, 425,

428, 429, 492, 493, 494, 495, 500 and 503 of native HCV NS5B, or defined by a functionally equivalent analog thereof.

In further aspects of the present invention, a method of producing a crystallized HCV NS5B complex is provided, comprising an NS5B polypeptide, polypeptide variant or polypeptide analog as defined above and a compound binding to said polypeptide, polypeptide variant or polypeptide analog, wherein said compound is associated with an NS5B binding pocket defined by at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 of a native HCV NS5B, or defined by a functionally equivalent analog thereof. This method comprises the steps of:

a) incubating a purified HCV NS5B polypeptide in a crystallization buffer to obtain a crystallized NS5B polypeptide;
b) solubilizing the compound; and
c) soaking the crystallized NS5B polypeptide with the solubilized compound in a soaking buffer for a suitable soaking period to generate the HCV NS5B complex.

In an alternative method of preparing a crystallized HCV NS5B complex as defined above, the compound is added to a crystallization buffer containing crystallized HCV NS5B.

Another alternative method of preparing a crystallized HCV NS5B complex as defined above comprises the steps of:

a) combining purified HCV NS5B with the compound in solubilized form to form an NS5B complex; and
b) crystallizing the complex in a crystallization buffer.

In another aspect of the present invention, there is provided X-ray crystal structure coordinates of a complex comprising an HCV NS5B polypeptide, polypeptide variant or polypeptide analog as defined above and a compound, wherein the compound associates with a binding pocket within the NS5B polypeptide, polypeptide variant or polypeptide analog, said binding pocket being defined by amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 of native HCV NS5B, or defined by a functionally equivalent analog thereof.

In another aspect of the present invention, there is provided a computer-readable

storage medium having stored thereon a model of the crystal structure of a complex comprising an HCV NS5B polypeptide, polypeptide variant or polypeptide analog as defined above and a compound, wherein said compound associates with an NS5B binding pocket defined by the structural coordinates of at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 of a native HCV NS5B, or defined by structural coordinates of a functionally equivalent analog thereof.

In a further aspect of the present invention, there is provided a method of identifying a compound that may bind to HCV NS5B, comprising the steps of:

a) applying a 3-dimensional molecular modeling algorithm to the structural coordinates of an HCV NS5B binding pocket defined by the structural coordinates of at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 of native HCV NS5B, or defined by the structural coordinates of a functionally equivalent analog thereof, to determine the spatial coordinates of the binding pocket of HCV NS5B; and

b) electronically screening stored spatial coordinates of the compound against the spatial coordinates of the HCV NS5B binding pocket to determine whether the compound may bind within the HCV NS5B binding pocket.

In another aspect of the present invention, there is provided a virtual screening method to identify potential HCV inhibitors comprising the steps of:

a) constructing a computer model of an HCV NS5B binding pocket defined by the structural coordinates of at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 of a native HCV NS5B, or defined by the structural coordinates of a functionally equivalent analog thereof;

b) employing computational means to perform a fitting program operation between computer models of the compound to be evaluated and the NS5B binding pocket to provide an energy-minimized configuration of the compound in the binding pocket; and

c) evaluating the results of the fitting operation to quantify the association between the compound and the binding pocket, wherein a compound that associates with the binding pocket to yield a low energy, stable complex is a potential NS5B inhibitor.

In yet another aspect of the present invention, a method of screening candidate HCV NS5B inhibitor compounds is provided comprising the steps of:

a) incubating an HCV NS5B polypeptide, polypeptide variant or polypeptide analog as defined above with a candidate inhibitor compound under conditions suitable for binding; and

b) determining whether or not the candidate inhibitor compound binds to the polypeptide, wherein a compound that binds to the polypeptide is a potential HCV NS5B inhibitor.

In another aspect of the invention, a method of designing a compound which binds to an NS5B polypeptide, polypeptide variant or polypeptide analog as defined above is provided which comprises the step of: assessing the complementarity, i.e. the "fit", between the compound and a binding pocket in the NS5B polypeptide defined by at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 of a native HCV NS5B, or defined by a functionally equivalent analog thereof.

In another aspect of the invention, a method of producing a drug which inhibits RNA replication activity of HCV NS5B is provided which comprises identifying or designing a compound which fits into an NS5B binding pocket as defined by at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 of a native HCV NS5B, or defined by a functionally equivalent analog thereof, wherein said binding pocket is exposed by displacement of a finger loop chain defined by at least amino acid residues 18 to 35.

Aspects and embodiments of the present invention are described in more detail herein by reference to the accompanying drawings in which:

BRIEF DESCRIPTION OF THE DRAWINGS

Figure **1A** depicts the NS5B apo-enzyme structure without compound;

Figure **1B** depicts compound **A** bound to the NS5B where amino acids 18-35 are displaced and not seen in the structure;

Figure **1C** depicts compound **A** bound within a solvent accessible surface representation of the binding pocket of the present invention;

Figure **2** depicts compound **B** bound within a solvent accessible surface

representation of the binding pocket of the present invention;

Figure **3** depicts a docked model of the NMR derived bound conformation of compound **C** bound within the accessible representation of the binding pocket of the present invention;

Figure **4** shows the atomic structure coordinates of the HCV NS5B of SEQ ID No. 1 complexed with the compound **A** ;

Figure **5** shows the atomic structure coordinates of the HCV NS5B of SEQ ID No.1 complexed with the compound **B**

Figure **6** shows the atomic structure coordinates of the HCV NS5B of SEQ ID No.1 complexed with the compound **C**

DETAILED DESCRIPTION OF THE INVENTION

Definitions

Unless defined otherwise, the scientific and technological terms and nomenclature used herein have the same meaning as commonly understood by a person of ordinary skill to which this invention pertains. Generally, the procedures for cell culture, infection, molecular biology methods and the like are well-known to those of skill in the art. Such standard techniques can be found in reference manuals such as for example Sambrook *et al.* (1989) and Ausubel *et al.* (1994).

The term "HCV NS5B" refers to the RNA-dependent RNA polymerase (RdRp) of a Hepatitis C virus (HCV) that is involved in the replication of HCV. As one of skill in the art is well aware, the term HCV encompasses a viral family including many different strains, isolates and subtypes. Moreover, the NS5B polymerase of each member of this family, although functionally equivalent and highly homologous structurally, will vary somewhat in amino acid sequence. The amino acid sequence of one HCV NS5B polymerase, NS5B from HCV genotype 1b, is shown in SEQ ID No: 1. The sequences of other isolated HCV NS5B polymerases may be accessed in publicly available sequence databases.

The term "binding pocket", as used herein, refers to a region of a molecule or molecular complex, that, as a result of its configuration, favorably associates with or is occupied by an entity or region of the same molecule or molecular complex, or an entity or region of a different molecule, molecular complex, chemical compound or other compound. In accordance with the present invention, the NS5B binding pocket defined herein becomes exposed by displacement of the finger loop domain of

residues 18-35 thereby allowing binding of a compound that is capable of affecting NS5B activity, for example, inhibiting NS5B activity. Typically, a binding pocket, or at least a portion thereof, comprises a cavity which is the site of interaction with an entity of the same or different molecule. As will be appreciated by those of skill in the art, the nature of the cavity within a binding pocket will vary from molecule to molecule.

The term "isolated and purified" as it is used with respect to polypeptides according to the present invention refers to a polypeptide that is substantially free from other components.

The term "native HCV NS5B", as it is used herein with respect to the amino acid sequence of the binding pocket of the present invention, refers to the natural amino acid sequence of the binding pocket in a given HCV NS5B.

The term "native functional configuration" as it is used with respect to the binding pocket of the present invention refers to the natural arrangement, including spatial arrangement, of amino acids that form a pocket that can associate with or be occupied by certain compounds/entities.

As used herein the term "complex" refers to the combination of two or more entities, at least one of which is a protein or enzyme. In particular, complexes in accordance with the present invention are formed between an NS5B protein, including analogs thereof which may include amino acid substitutions, truncations or insertions, and another compound. The combination or "complexing" of a compound or chemical entity with a protein refers to the nature of the association/binding between the compound or chemical entity and the protein. The association between the components of the complex is the condition of proximity therebetween and may be non-covalent in nature, wherein the juxtaposition is energetically favored by hydrogen bonding, van der Waals forces or electrostatic interactions, or it may be covalent.

The term "analog" as used herein denotes, with respect to a molecular compound, a sequence of amino acids modified from the native or natural sequence of amino acids that retains the biological activity (either functional or structural) of the native

sequence of amino acids. This analog may be from the same or different species and may be a natural analog or may be prepared synthetically. Such analogs include amino acid sequences having substitutions, deletions, or additions of one or more amino acids, provided that the biological activity of the protein is conserved. In particular, the term "conservative analog" denotes an analog having amino acid modifications that retain biological activity. Analogs including amino acid substitutions may include substitutions having either strong or weak similarity (see, for example, Dayhoff, M.O., (1978), Atlas of Protein Sequence and Structure, **5**, suppl. 3, National Biomedical Research Foundation, Washington, D.C.) as defined according to the following "Table of Amino Acid Similarity":

Amino acid	Strong	Weak
A	G, S	C, T, V
C		A, S
D	E	G, H, K, N, Q, R, S
E	D	H, K, N, Q, R, S
F	W, Y	H, I, L, M
G	A	D, N, S
H	Y	D, E, F, K, N, Q, R
I	L, M, V	F
K	R	D, E, H, N, Q, S, T
L	I, M, V	F
M	I, L, V	F
N	Q	D, E, G, H, K, R, S, T
P		S, T
Q	N	D, E, H, K, R, S
R	K	D, E, H, N, Q
S	A, T	C, D, E, G, K, N, P, Q
T	S	A, K, N, P, V
V	I, L, M	A, T
W	F, Y	
Y	F, H, W	

The meaning of the term "functionally equivalent analog" as it is used herein with respect to the binding pocket polypeptide of the invention refers to substitutions,

deletions or insertions of one or more of the amino acids of the binding pocket. Substitutions may be made as set out above, for example, with an appropriate conservative amino acid or conservative synthetic amino acid analog. In essence, the term "analog" corresponds with the foregoing definition. The phrase "functionally equivalent" indicates that the analog retains the biological activity of the native molecule.

The term "side chain" with reference to an amino acid or amino acid residue means a group attached to the α-carbon atom of the α-amino acid. For example, the R-group side chain for glycine is hydrogen, for alanine it is methyl, for valine it is isopropyl. For the specific R-groups or side chains of the α-amino acids reference is made to A.L. Lehninger's text on Biochemistry (see chapter 4).

The term "truncation" refers to any shortened or abbreviated segment of a molecule which, for the purposes of the present invention, retains its biological activity. Truncation may refer to the shortening of a native protein molecule, or to an analog thereof.

The term "root mean square deviation" or "rms deviation" or "rmsd" means the square root of the arithmetic mean of the square of the deviations from the mean. In the context of atomic objects, the numbers are given in angstroms (Å). It is a way to express the deviation or variation from a trend or object.

The following abbreviations are used herein:

DLB: differential line broadening;
DMSO: dimethyl sulfoxide;
DTT: dithiothreitol;
EDTA: ethylenediaminetetraacetic acid;
FID: free induction decay
IPTG: isopropyl-beta-D-thiogalactopyranoside
HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
MES: 2-(N-morpholino)ethanesulfonic acid,
MPD: 2-methyl-2,4-pentanediol
NMR: nuclear magnetic resonance;

NOESY: Nuclear Overhauser Effect Spectroscopy
PEG: polyethylene glycol;
PEG5k mme: monomethyl ether polyethylene glycol 5000;
Tris: tris(hydroxymethyl)aminomethane;
TSP: sodium 3-trimethylsilyltetradeuteriopropionate.

Preferred embodiments

HCV NS5B Polypeptide

In a first aspect, the present invention provides an isolated and purified polypeptide comprising a functional HCV NS5B binding pocket defined by at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 (and optionally one or more of: amino acid residues 37 and 496) of a native HCV NS5B, wherein said binding pocket is exposed by displacement of a finger loop chain defined by at least amino acid residues 18 to 35 and wherein said binding pocket retains its native functional configuration.

Within this aspect of this invention, there is also provided an isolated and purified HCV NS5B polypeptide consisting of an exposed HCV NS5B binding pocket defined by at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 (and optionally one or more of: amino acid residues 37 and 496) of a native HCV NS5B, wherein said binding pocket retains its native functional configuration.

In this aspect of the invention, it is important that the present polypeptide incorporates a binding pocket that assumes its native configuration, i.e. the natural configuration that it assumes in the native HCV NS5B polymerase, in order for it to properly mimic the binding pocket and be a useful HCV NS5B inhibitor screening tool.

Further, in order for the polypeptide to function effectively as an HCV NS5B inhibitor screening tool, it is also important that the inhibitor binding pocket be exposed, or at least exposable, in order to allow inhibitor access to the binding pocket. In HCV NS5B, the inventors have determined that the binding pocket defined herein is exposed when at least amino acid residues 18 to 35 of an NS5B finger loop chain are displaced. Accordingly, in this aspect, a polypeptide is provided in which the

binding pocket is in an exposed configuration.

The binding pocket of the present invention is defined herein by reference to amino acids by their position in the HCV NS5B protein due to expected sequence similarity from one HCV genotype or strain to another. The displaced or displaceable finger loop region is also defined by reference to amino acid positions, i.e. amino acids 18 to 35. It will be readily apparent to those of skill in the art that the numbering of amino acids in all HCV NS5B may be slightly different from that of the HCV NS5B exemplified herein due to amino acid insertions or deletions. The amino acid numbering presented herein is based on native HCV 1b NS5B polymerase sequence as shown **SEQ ID NO: 1**. However, corresponding amino acids in other HCV NS5B can be identified by visual inspection of the amino acid sequences or by using commercially available homology software programs such as Vector NTI (provided by InfoMax Inc.). In this regard, in order to identify the NS5B protein, it is notable that the first four amino acids of the HCV NS5B sequence are typically, -SMSY- **(SEQ ID NO: 2)**, which are conserved from one variant to the next.

In one embodiment of the present invention, the binding pocket is defined by amino acid sequence as follows:

(37)*	392-393	395-396	399	424-425	428-429	492-495	(496)*	500	503
(V)	L-A	A-A	T	I-L	H-F	L-G-V-P	(P)	W	R

(SEQ ID NO: 3)

*(the brackets identifying optional amino acid residues)

or, alternatively,
(V37), L392, A393, A395, A396, T399, I424, L425, H428, F429, L492, G493, V494, P495, (P496), W500 and R503.

As will be appreciated by one of skill in the art, a functionally equivalent analog of the polypeptide is also within the scope of this aspect of the invention. One or more of the amino acid residues of the present polypeptide, either within the binding pocket or outside of the binding pocket domain, may be substituted with a functionally

equivalent amino acid, generally a conservative amino acid replacement as set out in the 'Table of Amino Acid Similarity' above, a synthetic amino acid analog thereof or a naturally occurring amino acid substitution as found in other HCV genotypes while still retaining the binding pocket in functional form or configuration. Amino acid deletions and/or insertions may also be made to the polypeptide. Such amino acid substitutions, insertions or deletions may render a polypeptide that is more practical for use in a screening assay, or a polypeptide that is more readily prepared. Cursory examples of naturally occurring amino acids substitutions within the binding pocket among different HCV genotypes include, but are in no way limited to, HCV genotype 1b NS5B has a T499V substitution; HCV genotype 1a NS5B has M36L, I424V and T499A substitutions; HCV genotype 3a NS5B has M36L, I424V, L425M, and V494C substitutions; HCV genotype 2b NS5B has M36K, L392I, A393S, I424V, L425I and V494A substitutions; HCV genotype 2a, 2k, 6b have a V494A substitution; HCV genotype 3b has a V494I substitution; HCV genotype 6a has a P495L substitution; and HCV genotype 4a has an A396V substitution.

It is notable that a binding pocket is generally defined by atoms of any amino acid of a polypeptide, e.g. NS5B, that are within 5 Å of any atom of an inhibitor when the inhibitor is complexed with the polypeptide. Various computational analyses may be used to determine whether a polypeptide comprising a binding pocket as defined herein is sufficiently similar to the HCV NS5B binding pocket described above so as to be functional. Such analyses may be carried out in well known software applications, such as the Molecular Similarity applications of QUANTA [Molecular Simulations Inc, San Diego, Calif.], Sybyl [Tripos Associates, St. Louis, Mo.], InsightII [Accelrys], and MOE [Chemical Computing Group Inc., Montreal, Quebec, Canada].

There are a number of embodiments that stem from this aspect of the invention. For example, in addition to the amino acids set out above to define the binding pocket of the present invention, the binding pocket may additionally include one or more of amino acid residues 36, 426, 498 or 499 of the HCV NS5B. Preferably, each of these positions is occupied as follows: M36, M426, R498 and T/V499.

In another embodiment of this aspect, the polypeptide may contain the amino acid cluster of residues 36 and 37 of HCV NS5B in their native configuration. It is preferred that positions 36 and 37 are occupied by the following amino acids: M36,

and V37, or has the amino acid sequence M-V.

In another embodiment, the polypeptide may contain the amino acid cluster of residues 392 to 399 of HCV NS5B in their native configuration. It is preferred that positions 392 to 399 are occupied by the following amino acids: L392, A393, R394, A395, A396, W397, E398 and T399, or has the amino acid sequence L-A-R-A-A-W-E-T (**SEQ ID NO: 4**).

In another embodiment, the polypeptide may contain the amino acid cluster of residues 424 to 429 of the HCV NS5B in their native configuration. It is preferred that positions 424 to 429 are occupied by the following amino acids: I424, L425, M426, T427, H428 and F429, or has the amino acid sequence, I-L-M-T-H-F (**SEQ ID NO: 5**).

In another embodiment, the polypeptide may contain the amino acid cluster of residues 492 to 503 of the HCV NS5B in their native NS5B configuration. In this regard, it is preferred that positions 492 to 503 are occupied by the following amino acids: L492, G493, V494, P495, P496, L497, R498, T499, W500, R501, H502 and R503, or has the amino acid sequence, L-G-V-P-P-L-R-T-W-R-H-R (**SEQ ID NO: 6**).

HCV NS5B Polypeptide Analog

In a second aspect, the present invention provides an isolated and purified HCV NS5B polypeptide analog comprising an HCV NS5B binding pocket defined by at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 (and optionally one of: amino acid residues 37 and 496) of a native HCV NS5B, or defined by a functionally equivalent analog thereof, wherein said binding pocket retains its native functional configuration and wherein said binding pocket is exposable.

An NS5B polypeptide analog as defined above which retains the functional configuration of the native binding pocket disclosed herein advantageously provides a mimic of the native NS5B polypeptide which can be designed to be superior to the native NS5B polypeptide for use in inhibitor screening assays. For example, the polypeptide analog can be designed such that it is more readily made and used, or is more stable. It can also be designed in order to provide a binding pocket that is

more readily accessible.

The NS5B polypeptide may be altered to provide an NS5B polypeptide analog by amino acid substitution, deletion or insertion as described in more detail in the definitions set out above, and in particular, the definition of the term "analog". In this regard, it may be particularly desirable to modify the finger loop chain defined by amino acid residues 18 to 35 to provide a finger loop chain that is more readily displaced to expose the binding pocket of the present invention, or to provide a finger loop chain that is displaced or deleted to expose the binding pocket.

As set out above, the present HCV NS5B binding pocket is exposed when at least amino acid residues 18 to 35 of an NS5B finger loop chain are displaced. It has also been determined that certain compounds are capable of displacing the finger loop in order to access the binding pocket. Accordingly, a polypeptide analog is provided in which the binding pocket is exposable, i.e. concealed in its native NS5B configuration by a segment of the NS5B protein, such as the finger loop defined by amino acids 18 to 35, which can be displaced to expose the binding pocket, for example, by a compound with a propensity for the binding pocket.

There are a number of embodiments that stem from this aspect of the invention. For example, in addition to the amino acids set out above to define the binding pocket of the present invention, the binding pocket may additionally include one or more of amino acids residues 36, 426, 498 or 499 of the HCV NS5B. Preferably, the binding pocket includes all of these amino acid residues. Also preferably, each of these positions is occupied as follows: M36, M426, R498 and T/V499.

In another embodiment of this aspect, the binding pocket of the polypeptide may contain amino acid residues 36 and 37 of HCV NS5B in their native configuration. It is preferred that positions 36 and 37 are occupied by the following amino acids: M36, and V37, or has the amino acid sequence M-V.

In another embodiment of this aspect, the binding pocket of the polypeptide may contain amino acid residues 392 to 399 of HCV NS5B in their native configuration. It is preferred that positions 392 to 399 are occupied by the following amino acids: L392, A393, R394, A395, A396, W397, E398 and T399, or has the amino acid

sequence L-A-R-A-A-W-E-T (**SEQ ID NO: 4**).

In another embodiment, the binding pocket of the polypeptide may contain the amino acid residues 424 to 429 of the HCV NS5B in their native configuration. It is preferred that positions 424 to 429 are occupied by the following amino acids: I424, L425, M426, T427, H428 and F429, or has the amino acid sequence, I-L-M-T-H-F (**SEQ ID NO: 5**).

In another embodiment, the binding pocket of the polypeptide may contain the amino acid residues 492 to 503 of the HCV NS5B in their native NS5B configuration. In this regard, it is preferred that positions 492 to 503 are occupied by the following amino acids: L492, G493, V494, P495, P496, L497, R498, T499, W500, R501, H502 and R503, or has the amino acid sequence, L-G-V-P-P-L-R-T-W-R-H-R (**SEQ ID NO: 6**).

In further embodiments, the preferred sequence of the binding pocket residues is in accordance with the sequence set out in **SEQ ID NO: 1.**

As set out above, the binding pocket and displaceable/displaced finger loop region are defined herein by reference to amino acid positions based on the HCV genotype 1b NS5B sequence shown in **SEQ ID NO: 1** due to the high level of sequence homology that exists between HCV genotypes. However, one of skill in the art will appreciate that a slight variation in the position of one or more of the amino acid residues of the binding pocket or finger loop, for example a shift in position of each of the amino acids in the pocket by 1 or 2 (which may occur due to the insertion or deletion of one or more N-terminal amino acids) or a shift in a single amino acid elsewhere in the NS5B protein (e.g. a region which has no impact on binding pocket configuration), may still yield a binding pocket in accordance with the present invention and a finger loop that functions to expose the binding pocket. Accordingly, such position discrepancies are within the scope of the present invention.

HCV NS5B Polypeptide Variants

In another aspect, there is provided an HCV NS5B polypeptide variant comprising at least one amino acid mutation within a finger loop defined by amino acid residues 18 to 35, wherein said mutation provokes a displacement of said finger loop to expose a binding pocket essentially defined by amino acid residues 392, 393, 395, 396, 399,

424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 (and optionally one of: amino acid residues 37 and 496) of native HCV NS5B, or defined by a functionally equivalent analog thereof, wherein said binding pocket retains its native functional configuration.

An NS5B variant in accordance with this aspect of the invention may include any mutation which causes the finger loop defined by amino acids 18 to 35 to become displaced. For example, mutation of an amino acid which forms an association with one or more residues within the binding pocket as defined herein, may prevent such an association and thereby trigger the displacement or "opening" of the finger loop to expose the binding pocket.

In one embodiment of the present invention, at least one of the amino acid residues at positions 30 and 31 of HCV NS5B is mutated to provoke displacement of the finger loop. In a preferred embodiment, amino acid residue 30 is mutated to an amino acid residue other than leucine. More preferably, amino acid residue 30 is selected from: P, F, W, M, G, S, T, C, Y, N, Q, D, E, K, R and H.

Embodiments with respect to the binding pocket, and the additional amino acid residues that may comprise the binding pocket, as well as the specific sequences of the pocket residues, are set out above with respect to the NS5B polypeptides and analogs.

In another aspect of the present invention, there is provided an HCV NS5B polypeptide, or functionally equivalent analog thereof, characterized by displacement of amino acid residues 18 to 35.

In another aspect of the present invention, there is provided an HCV NS5B polypeptide, or functionally equivalent analog thereof, in which at least amino acid residues 18 to 35 have been deleted.

As previously set out, displacement or deletion of amino acid residues 18 to 35 exposes a novel binding pocket as defined herein which has significance in the development of HCV therapeutics.

HCV NS5B Crystal Structure

In a further aspect of the invention, there is provided an HCV NS5B crystal structure comprising a binding pocket defined by the structural coordinates of at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 (and optionally one of: amino acid residues 37 and 496) of a native HCV NS5B, or defined by the structural coordinates of a functionally equivalent analog thereof, wherein a native finger loop chain defined by the structural coordinates of at least amino acids 18 to 35 is displaced to expose said binding pocket. Preferably, the structural coordinates are as set forth in any one of Figures 4, 5 or 6.

A structure in which an inhibitor binding pocket is exposed is a valuable tool for the design and development of candidate NS5B inhibitors because it provides a means to more clearly understand the configuration and overall nature of the binding pocket, knowledge that is crucial to guide the development of therapeutic NS5B inhibitors.

In a preferred embodiment, an HCV NS5B crystal structure is provided in which the binding pocket is additionally defined by amino acid residues 36, 426, 498 and 499. In a further preferred embodiment, an HCV NS5B crystal structure is provided in which the binding pocket is defined by clusters of amino acid residues 36-37, 392-399, 424-429 and 492-503.

HCV NS5B Complex

In another aspect of the present invention, there is provided a complex comprising an HCV NS5B polypeptide and a compound, wherein the compound associates with a binding pocket within the NS5B polypeptide which is defined by amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 (and optionally one of: amino acid residues 37 and 496) of native HCV NS5B, or defined by a functionally equivalent analog thereof.

The NS5B inhibitor-binding pocket of the present invention is exposed on displacement of the finger loop region defined by amino acid residues 18 to 35. The binding pocket itself is defined by at least the amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 (and optionally one of: amino acid residues 37 and 496) of the HCV NS5B. However, it can further be

defined by the addition of one or more of amino acid residues 36, 426, 498 and 499, or even further defined by the following clusters of amino acid residues: 36-37, 392-399, 424-429 and 492-503.

In accordance with this aspect of the invention, thus, the HCV NS5B polypeptide may be a native NS5B polypeptide, or it may be an HCV NS5B polypeptide, variant or analog selected from the group consisting of:

i) an isolated and purified polypeptide comprising a functional HCV NS5B binding pocket defined by at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 (and optionally one of: amino acid residues 37 and 496) of a native HCV NS5B, or defined by a functionally equivalent analog thereof, wherein said binding pocket is exposed by displacement of a finger loop chain defined by at least amino acid residues 18 to 35 and wherein said binding pocket retains its native functional configuration;

ii) an isolated and purified HCV NS5B polypeptide consisting of an HCV NS5B binding pocket defined by at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 (and optionally one of: amino acid residues 37 and 496) of a native HCV NS5B, or defined by a functionally equivalent analog thereof, wherein said binding pocket retains its native functional configuration;

iii) an isolated and purified HCV NS5B polypeptide analog comprising an HCV NS5B binding pocket defined by at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 (and optionally one of: amino acid residues 37 and 496) of a native HCV NS5B, or defined by a functionally equivalent analog thereof, wherein said binding pocket retains its native functional configuration and wherein said binding pocket is exposable;

iv) an HCV NS5B polypeptide variant comprising at least one amino acid mutation within a finger loop defined by amino acid residues 18 to 35, wherein said mutation provokes a displacement of said finger loop to expose a binding pocket essentially defined by amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 (and optionally one of: amino acid residues 37 and 496) of native HCV NS5B, or defined by a functionally equivalent analog thereof, wherein said binding pocket retains its native functional configuration;

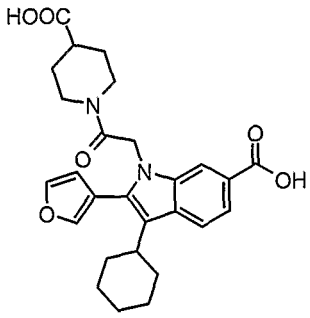
v) an HCV NS5B polypeptide, or a functionally equivalent analog thereof, characterized by displacement of amino acid residues 18 to 35; and

vi) an HCV NS5B polypeptide, or a functionally equivalent analog thereof, in which at least amino acid residues 18 to 35 have been deleted.

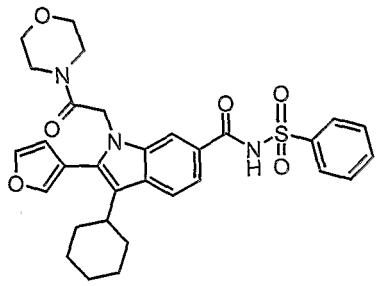
In a preferred embodiment of this aspect, the complex comprises an HCV NS5B polypeptide, analog or variant as defined above, associated with a compound selected from the compound families described in the following patent documents: WO 01/047883, WO 02/004425, WO 03/000254, WO 03/007945, WO 03/010140, WO 03/010141 and WO 03/026587. In an alternative preferred embodiment of this aspect, the complex comprises an HCV NS5B polypeptide, analog or variant as defined above, associated with a compound selected from the compound families described in the US co-pending applications 10/755,256, 10/755,544 and 60/546,213, herein incorporated by reference. Such compounds have not previously been shown to associate with the presently claimed binding pocket of NS5B to form an NS5B complex in accordance with the present invention.

In more preferred embodiments, the HCV NS5B complex comprises NS5B polypeptide, analog or variant associated with one of compounds **A**, **B** or **C** as set out below:

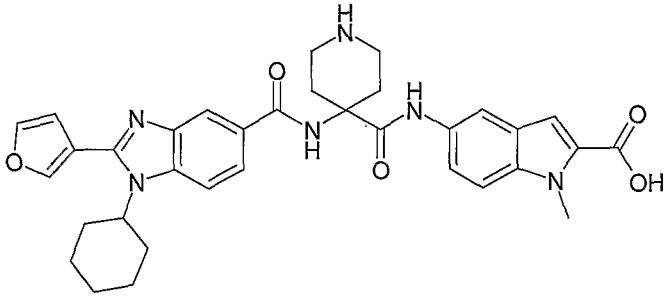
Compound A



Compound B



Compound C



Figures 4, 5 and 6 show the structure coordinates of the HCV NS5B of **SEQ ID No: 1** complexed with the compounds **A**, **B** and **C**, respectively. The manner of obtaining these structure coordinates, interpretation of the coordinates and their utility in understanding the protein structure and specifically the binding pocket as described herein, will be understood by those of skill in the art. Reference may also be made to standard texts such as Crystal Structure Analysis, Jenny Pickworth Glusker and Kenneth N. Trueblood, 2nd Ed. Oxford University Press, 1985, New York; and Principles of Protein Structure, G. E. Schulz and R. H. Schirmer, Springer-Verlag, 1985, New York which provide further guidance in this regard.

Moreover, as is appreciated by one of skill in the art, a set of structure coordinates for an enzyme-complex, as set out in Fig. 4, 5 and 6, is a relative set of points that define a three-dimensional shape. It is possible, thus, that an entirely different set of coordinates could define a similar or identical shape, i.e. a functionally equivalent analog of the native NS5B binding pocket, and thus, be within the scope of the present invention. Moreover, slight variations in the individual coordinates will have little effect on overall shape. In terms of binding pockets, these variations would not be expected to significantly alter the nature of compounds that could associate with

those pockets.

It is also noteworthy that modifications in the crystal structure due to mutations, additions, substitutions, and/or deletions of amino acids, or other changes in any of the components that make up the crystal could also account for variations in structure coordinates. If such variations are within an acceptable standard of error such as a rmsd<1.0Å for the alpha carbons that comprise the binding pocket, as compared to the original coordinates, the resulting three-dimensional shape is considered to be the same. Thus, for example, a compound that bound to the active site binding pocket of NS5B described herein would also be expected to bind to another binding pocket whose structure coordinates defined a shape that fell within the acceptable error.

Method of Crystallizing

In another aspect of the present invention, there is provided a method for producing a crystallized HCV NS5B complex comprising an HCV NS5B polypeptide and a compound, wherein said compound forms an association within an NS5B binding pocket defined by the structural coordinates of at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 (and optionally one of: amino acid residues 37 and 496) of the HCV NS5B, or defined by a functionally equivalent analog thereof.

A preferred crystallization method comprises the steps of:

a) incubating a purified HCV NS5B polypeptide in a crystallization buffer to obtain a crystallized NS5B polypeptide;

b) solubilizing the compound; and

c) soaking the crystallized polypeptide with the solubilized compound in a soaking buffer for a suitable soaking period to yield a crystallized NS5B complex.

In an alternative method of preparing a crystallized HCV NS5B complex as defined above, the compound is added directly to a crystallization buffer containing crystallized HCV NS5B.

In another alternative method, a crystallized HCV NS5B complex as defined above is prepared by co-crystallizing the NS5B protein with the compound. This method

comprises the steps of:

a) combining purified HCV NS5B polypeptide with the compound in solubilized form to form an NS5B complex; and

b) crystallizing the complex in a crystallization buffer.

In a preferred embodiment of this aspect, the NS5B is complexed with a compound selected from the compound families identified above. More preferably, the compound is selected from one of Compounds **A**, **B** or **C**, as set out above, or a related compound thereof. As set out above, the crystallization method may involve direct addition of the compound to a crystallization buffer containing crystallized HCV NS5B or combining purified HCV NS5B polypeptide with the compound in solubilized form. When the solubilized form of the compound is used, the compound is preferably solubilized in 100% DMSO at a concentration of about 25 mM.

NS5B is admixed with the purification buffer at a concentration of about 5 mg/mL to about 15 mg/mL. Preferably, the NS5B is used at a concentration of about 7 to about 10 mg/mL in the purification buffer. The purification buffer, the nature of which is generally known in the art and which may include a salt and/or glycerol to stabilize the HCV NS5B, is used at a pH of between about 6 and about 9. The preferred pH of the purification buffer is about 7.5. A buffer such as, but not limited to, Tris-HCl, HEPES or bis-Tris can be used at a concentration of between about 0 mM and about 50 mM. Preferably, the buffer is Tris-HCl at a concentration of about 20 mM.

In order to stabilize the HCV NS5B, a salt such as NaCl, $(NH_4)_2SO_4$, or KCl can be added to the buffer at a concentration of about 200 mM to about 800 mM. Preferably, the salt is NaCl at a concentration of about 300 mM.

To further stabilize the HCV NS5B, glycerol can be added at a concentration of between about 0% and about 30%. Preferably, glycerol is at a concentration of about 10%.

More preferably, the purification buffer is at about pH 7.5 and contains Tris-HCl at a concentration of about 20 mM, glycerol at a concentration of about 10%, DTT at a concentration of about 5 mM, and NaCl at a concentration of about 300 mM.

The NS5B polypeptide can be crystallized using any one of various techniques known in the art, including for example, batch crystallization under oil, hanging drop vapor diffusion and sitting drop vapor diffusion techniques. The hanging drop vapor diffusion technique, as described in McPherson *et al.* (Preparation and Analysis of Protein Crystals, Krieger Pub. 1989) is the preferred method of crystallization for the purposes of the present invention. Briefly, this method of crystallization involves placing a droplet containing purified NS5B in a crystallization buffer over a reservoir solution. Vapor diffusion from the droplet increases protein concentration thereby promoting crystallization.

The crystallization buffer used may be selected from any one of a number of buffers known by those of skill in the art to be suitable for this purpose, including, but not limited to, MES, sodium phosphate, potassium phosphate, sodium acetate or sodium succinate at a concentration of about 50 mM to about 0.2 M. Preferably, the crystallization buffer is MES at a concentration of about 0.1 M.

The pH of the crystallization buffer is typically between about 4.5 and about 6.5, and preferably, the crystallization buffer is used at a pH of about 5.4.

The crystallization buffer may additionally contain at least one precipitating agent which facilitates crystallization of the NS5B. Examples of appropriate precipitating agents include, but are not limited to, PEG, PEG5K mme (monomethyl ether polyethylene glycol 5000), ammonium sulfate, MPD, isopropanol, ethanol, or tertiary butanol. The precipitating agent is generally used at a concentration of about 30% to about 40%. In a preferred embodiment, the precipitating agent is PEG5K mme at a concentration of about 21% and ammonium sulfate at a concentration of about 0.4 mM.

The crystallization of NS5B is conducted under standard conditions of crystallization. For example, the crystallization is carried out at a temperature of between about 0°C and about 22°C. The preferred temperature under which to conduct the crystallization is between about 4°C and about 11 °C.

In the preferred crystallization method, solubilized compound is soaked into crystallized NS5B polypeptide in the presence of a soaking buffer. The soaking

buffer can comprise any one of a number of standard buffers including, but not limited to, MES, Tris, sodium phosphate, sodium acetate and sodium succinate at a concentration of between about 50 mM to about 0.2 M. Preferably, the soaking buffer is used at a concentration of about 0.1 M. The pH of the soaking buffer is typically between about 5 and about 8, and preferably, the soaking buffer is used at a pH of about 7.0.

The protein content of the soaking buffer is supplemented to a concentration of up to about 10 mg/mL by addition of any suitable protein including, but not limited to, lysozyme, BSA or even additional NS5B.

The soaking buffer may contain additional agents which function as NS5B stabilizers. One or more salts such as NaCl, $(NH_4)_2SO_4$ or KCl can be added to the buffer at a concentration of about 100 mM to about 500 mM. Preferably, the salt is added at a concentration of between about 150 mM and about 300 nM. More preferably, NaCl and $(NH_4)_2SO_4$ are both added at concentration of about 210 mM and about 280 mM, respectively.

To further stabilize the HCV NS5B, glycerol can be added to the soaking buffer at a concentration of between about 10 and about 20%. Preferably, glycerol is added at a concentration of about 14%.

The soaking buffer may additionally contain at least one precipitating agent which facilitates crystallization. Examples of appropriate precipitating agents include, but are not limited to, PEG, PEG5K mme (monomethyl ether polyethylene glycol 5000), ammonium sulfate, MPD, isopropanol, ethanol, or tertiary butanol. The precipitating agent is generally used at a concentration of about 10% to about 18%. In a preferred embodiment, the precipitating agent is PEG5K mme at a concentration of about 14% and ammonium sulfate at a concentration of about 0.4 mM.

Crystallized NS5B is soaked with solubilized compound in a soaking buffer for a suitable soaking period of about 1 to about 12 hours, preferably about 3 to about 8 hours and most preferably for about 5 to about 6 hours. Soaking occurs at a temperature of between about 5 and about 15°C, and preferably at a temperature of about 11°C.

In one alternative crystallization method, the compound is added to crystallization buffer containing crystallized NS5B. In this method, the compound is simply sprinkled onto the buffer and allowed to solubilize and crystallize following a suitable period of incubation.

In another alternative crystallization method, NS5B and compound are co-crystallized in a crystallization buffer as described above for NS5B alone. In this method, the NS5B and solubilized compound are combined in a crystallization buffer and allowed to crystallize under crystallization conditions such as those described above.

In accordance with the present invention, it is important that the crystallized NS5B complex be amenable to X-ray crystallography. Using X-ray crystallography analysis, the crystals of the NS5B complex obtained belong to space group P2(1)2(1)2(1) with unit cell dimension of a= 105.1, b=106.6 and c=133.5 and contain two molecules per asymmetric unit. Diffraction data were measured using a MicroMax007 home source x-ray generator equipped with a R-axis IV++ image plate area detector (Rigaku, Japan). Preferably, data to a resolution of 2.8 Å were collected on a single crystal of the complex.

X-ray Coordinates

According to yet another aspect, there is provided X-ray crystal structure coordinates of an NS5B complex as defined above. Even more preferably, the set of structure coordinates for the NS5B complex are defined according to one of Figure 4, Figure 5 and Figure 6.

The three-dimensional structure of an NS5B complex of this invention is defined by a set of structure coordinates as set forth in any one of Figure 4, Figure 5 and Figure 6. The term "structure coordinates" refers to Cartesian coordinates derived from mathematical operations related to the patterns obtained on diffraction of a monochromatic beam of X-rays by the atoms (scattering centers) of the complex in crystal form. The diffraction data are used to calculate an electron density map of the repeating unit of the crystal. The electron density maps are then used to establish the positions of the individual atoms of the binding pocket known as the

structure coordinates.

Those of skill in the art will understand that a set of structure coordinates for a protein or protein-inhibitor complex or a portion thereof, is a relative set of points that define a shape in three dimensions. Thus, it is possible that an entirely different set of coordinates could define a similar or identical shape.

The variations in coordinates may be generated by mathematical manipulations of the structure coordinates. For example, the structure coordinates set forth in Figure 4, 5 or 6 could be manipulated by crystallographic permutations of the structure coordinates, fractionalization or matrix operations to sets of the structure coordinates or any combination of the above.

Various computational analyses are necessary to determine whether a molecule or molecular complex or a portion thereof is sufficiently similar to all or parts of the HCV NS5B protein or the NS5B complex described herein to be considered equivalent. Such analyses may be carried out using current software applications, such as the Molecular Similarity applications of QUANTA [Molecular Simulations Inc, San Diego, Calif.], Sybyl [Tripos Associates, St. Louis, Mo.], InsightII [Accelrys], and MOE [Chemical Computing Group Inc., Montreal, Quebec, Canada].

The Molecular Similarity application permits comparisons between different structures, different conformations of the same structure, and different parts of the same structure. The procedure used in Molecular Similarity to compare structures is divided into four steps: 1) load the structures to be compared; 2) define the atom equivalence in these structures; 3) perform a fitting (superposition) operation; and 4) analyze the results.

Computer-readable Storage Medium

Still, in another aspect of the present invention, there is provided a computer-readable storage medium having stored thereon a model of the crystal structure of an HCV NS5B complex comprising an HCV NS5B polypeptide and a compound, wherein said compound associates with an NS5B binding pocket defined by the structural coordinates of at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 (and optionally one of: amino acid

residues 37 and 496) of a native HCV NS5B as set forth in Figures 4, 5 and 6, or defined by structural coordinates of a functionally equivalent analog thereof.

Computer readable data storage media are well known to those skilled in the art and include, for example, hard-disk, CD-ROM, diskette ("floppy disks") and DVD.

In accordance with this aspect of the present invention, the structure coordinates of an NS5B complex, and portions thereof, can be stored in a machine-readable storage medium. Such data can be used for a variety of purposes, such as drug discovery and X-ray crystallographic analysis of the protein crystal.

In a preferred embodiment of this aspect, the HCV NS5B complex comprises NS5B complexed with a compound of a compound family identified above. More preferably, the NS5B complex comprises NS5B complexed with one of Compounds **A, B** or **C**.

As set out previously, the binding pocket may be additionally defined by the addition of one or more of the amino acid residues selected from 36, 426, 498 and 499, or even further defined by the amino acid clusters, 36,-37, 392-399, 424-429 and 492-503.

The coordinate data of an NS5B complex, such as that set out in Figures 4, 5 and 6, when used in conjunction with a computer programmed with software to translate those coordinates into a 3-dimensional structure, may be used for a variety of purposes, especially for purposes relating to drug discovery. Software for generating such three-dimensional graphical representations are known and commercially available. Examples include Quanta and WebLite Viewer. The ready use of the coordinate data requires that it be stored in a computer-readable format. Thus, in accordance with the present invention, data capable of being displayed as a three dimensional structure is stored in a computer-readable storage medium which is capable of displaying a graphical three-dimensional representation of an HCV NS5B complex or of an HCV NS5B binding pocket as defined herein when used by a machine programmed with instructions for using said data.

The HCV NS5B X-ray coordinate data is useful for screening compounds for

potential NS5B inhibitory activity. For example, the polypeptide NS5B binding pocket structure encoded by the data may be computationally evaluated for its ability to associate or bind with a given compound. Compounds determined to "fit" into the binding pocket defined herein via some type of association or bonding may also impede the biological activity of the HCV NS5B polymerase and, thus, represent a potential drug candidate. In addition, the data may be displayed in a graphical three-dimensional representation on a computer screen which allows visual examination of the HCV NS5B binding pocket as well as the association of compounds within the binding pocket in an NS5B complex.

Virtual Methods of Identifying Compounds that Associate/Bind HCV NS5B

In further aspects, the present invention provides virtual methods to evaluate the potential of a compound to complex with HCV NS5B. These methods represent a first screen in the search for compounds that can associate or bind to the binding pocket of the present invention, and ultimately in the search for a compound that has therapeutic effects due to an association/binding with the binding pocket that results in HCV inhibition.

Thus, in a further aspect of the present invention, a virtual screening method is provided to identify potential HCV inhibitors comprising the steps of:

a) constructing a computer model of an HCV NS5B binding pocket defined by the structural coordinates of at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 (and optionally one of: amino acid residues 37 and 496) of a native HCV NS5B, or defined by a functionally equivalent analog thereof;

b) employing computational means to perform a fitting program operation between computer models of the compound to be evaluated and the NS5B binding pocket to provide an energy-minimized configuration of the compound in the binding pocket; and

c) evaluating the results of the fitting operation to quantify the association between the compound and the binding pocket, wherein a compound that associates with the binding pocket to yield a low energy, stable complex is a potential NS5B inhibitor.

In addition, the present invention provides a method of identifying compounds that

can bind to HCV NS5B, comprising the steps of:

a) applying a 3-dimensional molecular modeling algorithm to the structural coordinates of an HCV NS5B binding pocket defined by at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 (and optionally one of: amino acid residues 37 and 496) of a native HCV NS5B to determine the spatial coordinates of the binding pocket of HCV NS5B; and

b) electronically screening stored spatial coordinates of candidate compounds against the spatial coordinates of the HCV NS5B binding pocket to identify compounds that can bind within the HCV NS5B binding pocket.

In preferred embodiments of this aspect of the invention, the binding pocket may be defined additionally by one or more of amino acid residues 36, 426, 498 and 499, or by one or more of the amino acid clusters defined by amino acid residues 36-37, 392-399, 424-429 and 492-503.

In another preferred embodiment of the invention, the structural coordinates of the binding pocket are those set out in any one of Figures 4, 5 and 6, or functionally equivalent structural coordinates as would be appreciated by one of skill in the art.

According to this aspect of the invention, any given compound may be computationally evaluated for its ability to associate with the HCV NS5B binding pocket defined herein, and thus, its potential as an NS5B inhibitor determined. As alluded to above, a computer model of a polypeptide consisting of an HCV NS5B binding pocket as defined herein is constructed using well-known software such as QUANTA [Molecular Simulations Inc, San Diego, Calif.], Sybyl [Tripos Associates, St. Louis, Mo.], InsightII [Accelrys], MOE [Chemical Computing Group Inc., Montreal, Quebec, Canada]. Selected compounds to be evaluated may then be positioned in a variety of orientations, or docked, within the binding pocket. Docking may be accomplished using software such as GRID, DOCK, AUTODOCK, FlexX, and GOLD. When a compound is docked within the binding pocket to form a "virtual" representation of an NS5B complex, computational means may be further employed to generate quantitative and qualitative maps of the complex, including for example, pharmacophore maps, surface property maps (which map Conolly, Gaussian and van der Waals surfaces) and maps of Probabilistic Receptor Potentials using software such as QUANTA, Sybyl, InsightII, and MOE.

The efficiency with which a selected compound binds to the present HCV NS5B binding pocket may be tested and optimized by computational evaluation. The quality of the fit of a given compound within the NS5B binding pocket may be evaluated, for example, by shape, size and electrostatic complementarity as determined qualitatively by visual inspection or as determined quantitatively by the use of scoring functions such as LUDI, PLP, PMF, SCORE, GOLD and FlexX. These methods of qualitative and quantitative evaluation may be employed individually or in combination, for example, as in a consensus scoring manner.

Alternatively, binding efficiency can be determined based on the interaction energy of a complex formed by the binding or association of a compound with the HCV NS5B. For example, a compound determined to form a "low energy, stable complex" with NS5B, in the manner described herein, warrants further analysis as a potential NS5B inhibitor. The term "low energy, stable complex" as used herein is defined as an NS5B complex in which the van der Waals interaction energy value, i.e. the van der Waals energy of interaction between the compound and NS5B, is less than about 8000 kcal/mol. Van der Waals interaction energy value can be determined using the software MOE, and is based on the MMFF94 force field. Accordingly, a compound determined to form a complex having a van der Waals interaction energy value of less than about 8000 kcal/mol is a potential NS5B inhibitor. Preferably, a low energy, stable complex in accordance with the present invention will have a van der Waals interaction energy value of less than about 6000 kcal/mol, and more preferably, a value of less than about 4000 kcal/mol.

Method of Using the NS5B Polypeptide Variants/Analogs of the Invention

Once a series of compounds has been screened using virtual methods such as those described above, compounds determined to be potential HCV inhibitors can be further evaluated to determine the actual propensity of each to interact with the binding pocket of the present invention and to inhibit HCV.

Thus, in still another aspect of the invention, a method of screening candidate HCV NS5B inhibitor compounds is provided comprising the steps of:

a) incubating a candidate inhibitor compound under conditions suitable for binding with a polypeptide selected from the group consisting of:

i) an isolated and purified polypeptide comprising a functional HCV NS5B binding pocket defined by at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 (and optionally one of: amino acid residues 37 and 496) of a native HCV NS5B, or defined by a functionally equivalent analog thereof, wherein said binding pocket is exposed by displacement of a finger loop chain defined by at least amino acid residues 18 to 35 and wherein said binding pocket retains its native functional configuration;

ii) an isolated and purified HCV NS5B polypeptide consisting of an HCV NS5B binding pocket defined by at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 (and optionally one of: amino acid residues 37 and 496) of a native HCV NS5B, or defined by a functionally equivalent analog thereof, wherein said binding pocket retains its native functional configuration;

iii) an isolated and purified HCV NS5B polypeptide analog comprising an HCV NS5B binding pocket defined by at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 (and optionally one of: amino acid residues 37 and 496) of a native HCV NS5B, or defined by a functionally equivalent analog thereof, wherein said binding pocket retains its native functional configuration and wherein said binding pocket is exposable;

iv) an HCV NS5B polypeptide variant comprising at least one amino acid mutation within a finger loop defined by amino acid residues 18 to 35, wherein said mutation provokes a displacement of said finger loop to expose a binding pocket essentially defined by amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 (and optionally one of: amino acid residues 37 and 496) of native HCV NS5B, or defined by a functionally equivalent analog thereof, wherein said binding pocket retains its native functional configuration;

v) an HCV NS5B polypeptide, or a functionally equivalent analog thereof, characterized by displacement of amino acid residues 18 to 35; and

vi) an HCV NS5B polypeptide, or a functionally equivalent analog thereof, in which at least amino acid residues 18 to 35 have been deleted;. and

b) determining whether or not the candidate inhibitor compound binds to the polypeptide, wherein a compound that binds to the polypeptide is a potential HCV NS5B inhibitor.

Binding of the candidate compound within the defined HCV NS5B binding pocket

can be determined using methods well-established in the art. For example, binding assays may be used in which the candidate compound is exposed to a polypeptide containing the NS5B binding pocket of the invention under conditions suitable for association or binding to occur. Binding is then assessed, for example using NMR or other known detection techniques.

Method of Designing an HCV NS5B Inhibitor

In further aspects of the present invention, there is provided methods of designing compounds which associate with the NS5B binding pocket defined herein. The present invention, thus, provides the opportunity to use molecular design techniques to identify, select or design potential inhibitors of HCV NS5B based on the structure of a novel binding pocket in NS5B. Such a predictive model is valuable in light of the high costs associated with the preparation and testing of many diverse compounds that may or may not bind to the HCV NS5B protein.

According to this invention, a potential NS5B inhibitor may be evaluated for its ability to bind an NS5B binding pocket as defined herein prior to its actual synthesis and testing. If a proposed compound is predicted to have insufficient interaction or association with the binding pocket, preparation and testing of the compound is obviated. However, if the computer modeling indicates a strong interaction, the compound may then be obtained and physically tested for its ability to bind. Testing to confirm binding may be performed using conventional assays within the purview of one of skill in the art.

In this regard, a method of designing a compound which binds to an NS5B polypeptide as defined above is provided which comprises the step of: assessing the complementarity, i.e. the "fit", between the compound and a binding pocket in the NS5B polypeptide defined by at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 (and optionally one of: amino acid residues 37 and 496) of a native HCV NS5B, or defined by a functionally equivalent analog thereof.

Similarly, a method of producing a drug which inhibits RNA replication activity of HCV NS5B is also provided which comprises identifying or designing a compound which fits into an NS5B binding pocket as defined by at least amino acid residues

392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 (and optionally one of: amino acid residues 37 and 496) of a native HCV NS5B, or defined by a functionally equivalent analog thereof, wherein said binding pocket is exposed by displacement of a finger loop chain defined by at least amino acid residues 18 to 35.

For the first time, the present invention permits the use of structure-based or rational drug design techniques to design, select, and synthesize chemical entities, including inhibitory compounds, that are capable of fitting into and/or binding with the novel NS5B binding pocket defined herein.

One particularly useful drug design technique enabled by this invention is iterative drug design. Iterative drug design is a method for optimizing associations between a protein and a compound by determining and evaluating the three-dimensional structures of successive sets of protein/compound complexes.

Those of skill in the art will realize that association of natural ligands or substrates with the binding pocket of their corresponding receptors or enzymes is the basis of many biological mechanisms of action. Similarly, many drugs exert their biological effects through association with the binding cavities of receptors and enzymes. Such associations may occur with all or any part of the binding pocket. An understanding of such associations will help lead to the design of drugs having more favorable associations with their target receptor or enzyme, and thus, improved biological effects. Therefore, this information is valuable in designing potential ligands or inhibitors of receptors or enzymes, such as inhibitors of HCV NS5B-like polypeptides, and more importantly, HCV NS5B.

In iterative drug design, crystals of a series of protein/compound complexes are obtained and then the three-dimensional structure of each complex is solved. Such an approach provides insight into the association between the proteins and compounds of each complex. This is accomplished by selecting compounds with inhibitory activity, obtaining crystals of this new protein/compound complex, solving the three-dimensional structure of the complex, and comparing the associations between the new protein/compound complex and previously solved protein/compound complexes. By observing how changes in the compound affected

the protein/compound associations, these associations may be optimized.

Embodiments of the present invention are described in the following specific examples which are not to be construed as limiting.

EXAMPLES

EXAMPLE 1 – EXPRESSION AND PURIFICATION OF HCV NS5B

The recombinant HCV genotype 1b (J4 strain) NS5B polymerase (the amino acid sequence of which is shown in **SEQ ID NO: 1**) was produced in soluble form by expression of a variant that lacks the C-terminal 21 amino acids normally found on the mature NS5B. For the purposes of the present invention, this NS5BΔ21 was expressed with a C-terminal hexa-histidine sequence. Expression of the NS5B from a pET vector in *E. coli* strain JM109 (DE3) was induced with 0.4 mM IPTG for 3 hours at 24°C. Cells were harvested (in 25 mM Tris pH 7.5, 10 % glycerol, 1 mM EDTA, 500 mM NaCl, 2 mM 2-mercaptoethanol and cocktail of protease inhibitors (standard cocktail that can be purchased from Roche Biochemicals Inc.) and lysed in a microfluidizer. The lysate, after centrifugation (30,000X), was purified according to the following sequential chromatographic steps that are known to one skilled in the art: (i) metal affinity using Ni-NTA resin (Qiagen) and elution with a buffer containing increasing concentration of imidazole (from 10 mM to 500 mM); (ii) DEAE-Sepharose chromatography wherein the NS5B flowed through the column in a buffered solution containing 300 mM NaCl; (iii) Heparin Sepharose chromatography wherein the NS5B was in a buffered solution with 200 mM NaCl, and then bound NS5B was then eluted with a buffer containing a gradient up to 1 M NaCl. Eluted fractions, enriched with the NS5B protein were concentrated on a Resource S column, and then applied to a Superdex 200 column in 20 mM Tris-HCl pH 7.5, 10 % glycerol, 5 mM DTT, 300 mM NaCl. Peak fractions containing highly pure His-tag NS5B were stored at -80°C until use.

EXAMPLE 2 – CRYSTALLIZATION OF HCV NS5B

HCV polymerase NS5B strain 1b/J4 **(SEQ. ID NO: 1)**, obtained as described in detail in Example 1, was crystallized in its apo form using monomethyl ether polyethylene glycol 5000 (PEG5Kmme) as a precipitant agent. Large single crystals were obtained by the hanging drop vapor diffusion technique (McPherson, *supra*). In

particular, 1 µL of NS5B (7.66 mg/mL in purification buffer containing 20 mM Tris, pH 7.3, 300 mM NaCl, 10% glycerol) was added to 1 µL of a solution made of 21% PEG5Kmme, 0.1 M MES, pH 5.4, 10% glycerol and 0.4 M ammonium sulfate. The resulting 2 µL drop was suspended above a 1 mL reservoir solution of 21% PEG5Kmme, 0.1 M MES, pH 5.4, 20% glycerol and 0.4 M ammonium sulfate. The crystals of the NS5B complex obtained belong to space group P2(1)2(1)2(1) with unit cell dimension of a= 105.1, b=106.6 and c=133.5 and contain two molecules per asymmetric unit. The crystals were shown to diffract to a resolution of up to 2 Å using X-rays from a regular rotating anode source.

EXAMPLE 3 – PREPARATION OF INHIBITOR COMPOUNDS A, B AND C

Inhibitor compound **A** was made as described in detail in co-pending application US 10/755,256 filed January 12, 2004.

Inhibitor compound **B** was made as described in detail in co-pending application US 10/755,256 filed January 12, 2004.

Inhibitor compound **C** was made as described in detail in WO 03/007945 published January 22, 2003.

EXAMPLE 4 – FORMATION OF NS5B-INHIBITOR COMPLEX

Inhibitor soaking

Crystals were transferred in a 5 µL drop of a soaking solution made of 14% PEG5Kmme, 14 mM Tris, pH 7.5, 70 mM MES, pH 7.0, 14% glycerol, 210 mM NaCl, 10 mg/mL lysozyme and 280 mM ammonium sulfate. The inhibitor molecule was dissolved in DMSO to a concentration of 25 mM. Inhibitor solution (0.2 µL) was added to the 5 µL NS5B crystal drop and incubated for 5-6 hours at 11°C. Following incubation, the crystals of the NS5B-inhibitor complex were transferred from the solution with a cryoloop (Hampton Research, California, USA) and cryo-cooled in liquid nitrogen. NS5B complexes were prepared using Compounds **A** and **B**.

Data collection

Diffraction data were collected on a MicroMax-007 rotating anode x-ray generator equipped with an Raxis-IV++ Image plate detector (Rigaku/MSC, Texas, USA). Data to a resolution of 2.8 Å was collected on a single crystal cryo-cooled at –180°C for

NS5B complexes prepared with Compound **A** and Compound **B**.

Phasing, Model Building and Refinement

The diffraction data were phased by Molecular Replacement (MR) using the publicly available structure of HCV NS5B (pdb code : 1C2J). Rotation and translation search were done using the program CNX (Accelrys). Model building was carried out with the software O (Alwyn Jones, Upsala University, Sweden) and model refinement was performed with the CNX software (Accelrys). The model was improved by repetition of the procedure of model building and refinement until a desirable result was obtained. The final model in each case included two molecules of NS5B, identified as NS5B A and NS5B B, (i.e. residue A1 to A149, A154 to A563, B1 to B17, B36 to B148 and B153 to B563) and one molecule of compound **A** or **B** associated with NS5B molecule B. The resulting atomic structure coordinates of the compound **A** and compound **B** NS5B complexes are shown in Fig. 4 and Fig. 5, respectively. The final crystallographic R factor the NS5B-Compound **A** complex was 21.9% and R_{free} factor was 28.0% to a resolution of 2.8Å. For the NS5B-Compound **B** complex, the final crystallographic R factor was 21.9% and R_{free} factor was 25.7% to a resolution of 2.7 Å.

EXAMPLE 5 - MODEL OF NS5B-COMPOUND C COMPLEX BASED ON AN NMR BOUND CONFORMATION.

The binding site of Compound **C** on HCV polymerase was determined by several steps which are described herein in detail. Initially, NMR spectroscopic methods (transferred NOESY) were applied to determine the structure of Compound **C** when bound to HCV polymerase. Another NMR technique, differential line-broadening (DLB), helped to identify the segments of Compound **C** which come into contact with the polymerase versus those that are solvent exposed in the bound state. The bound structure was then docked onto the X-ray-derived structure of the NS5B-Compound **A** complex. The docking procedure began by overlaying the common feature of the 5/6 aromatic rings of the indole and benzimidazole of Compound **A** and Compound **C**. All the differential line-broadening data were then accounted for and the complex was energy minimized using MOE.

The transferred NOESY and DLB data were acquired as follows. A sample tube (5 mm) containing Compound **C** was prepared by adding 15 µL of concentrated

solution in DMSO-d_6 (containing 0.31 mg of Compound **C**) to an aqueous buffer composed of 20 mM Tris-d_{11}, 2 mM DTT-d_{10}, 1 mM EDTA-d_{12}, 300 mM NaCl, and 10% (v/v) D_2O spiked with TSP-2,2,3,3-d_4. Buffer was added to a final volume of 600 µL, and the pH was adjusted to 6.0. Spectra of free Compound **C** were taken. A concentrated stock solution of HCV polymerase was added twice to the NMR tube and spectra were acquired. The concentrated stock contained 21.7 mg/mL of HCV polymerase (NS5BΔ21C-His_6) in buffer consisting of 20 mM Tris-d_{11}, 2 mM DTT-d_{10}, 1 mM EDTA-d_{12}, 300 mM NaCl, and 10% (v/v) glycerol-d_8.

NMR spectra were acquired on Bruker AVANCE 600 and 800 MHz NMR spectrometers at 22 and 27 °C. Suppression of the solvent signal was achieved by the use of pre-saturation or by inserting a 3-9-19 WATERGATE module prior to data acquisition. One-dimensional spectra were collected on both sample tubes for DLB comparisons, and 128 scans were collected for each spectrum. NOESY experiments on a sample tube containing no polymerase resulted in the observation of no significant cross-peaks, as expected. On the other hand, NOESY spectra on a sample tube containing polymerase resulted in the observation of many cross-peaks which contained the valuable inter-hydrogen distance information of the inhibitor **C** when bound to HCV polymerase. A series of NOESY mixing times including 75, 100, 150, 200 and 300 msec (at 600 and 800 MHz fields) was recorded. Two-dimensional data sets were typically acquired with 2048 points in t_2 and 512 points in t_1. 128 scans were averaged for NOESY FIDs. The data were processed and analyzed using Bruker's XWinNMR and WinNMR software (Bruker Canada, 555 Steeles Ave. East, Milton, Ontario). Data sets were zero-filled to yield 2048 by 2048 real points after transformation using a phase-shifted sinebell window function.

The structure of Compound **C** was calculated by a custom simulated annealing protocol that was implemented in the MOE molecular modeling program (CCG, Montreal, Qc, Canada). All calculations were performed using the mmff94 forcefield as implemented in MOE version 2202.03. NMR-derived flat-bottomed restraints were generated from the series of NOESY data. The relative intensity of NOESY cross-peaks was classified into three categories which were then used as restraints, strong (1.8-2.7 Å), medium (1.8-3.5 Å) or weak (1.8-5.0 Å). The force constants in the NMR-derived flat-bottomed restraints were gradually increased during the cooling stages. A single, high temperature unrestrained dynamics run was performed at

1000 K, with 100 structures collected at 10 psec intervals to generate a starting set of conformations. Each structure was cooled and minimized using the following simulated annealing protocol. During the simulations the temperature and restraint weights were changed from 1000K to 50K, and from 1/1000000 to 20, respectively. The final structures were energy minimized, the total energies were calculated, the restraint energies were calculated, and the restraint violations were determined. NMR-consistent structures were isolated, and two families of structures were identified.

One representative from each family of NMR-derived structures was then docked onto the X-ray-derived structure of the NS5B-Compound **A** complex. The docking procedure began by overlaying the common 5/6 aromatic rings of the indole and benzimidazole scaffolds of Compounds **A** and **C**. Only one structure of Compound **C** simultaneously matched the DLB data and the shape of the binding pocket. In this complex, the right-hand side of Compound **C** lies over or in the vicinity of Pro 495 and Pro 496. Using the MOE molecular modeling program, the side-chain orientations of amino acids at positions 503, 498, and 499 were slightly re-adjusted to improve the inhibitor-polymerase fit. His 428 and 502 were protonated. A combination of steepest descent, conjugate gradient and truncated Newton algorithms was used to minimize the energy of the complex which resulted in only minor changes to the structure. The protocol was set up such that the inhibitor and polymerase residues within 6 Å of the inhibitor were allowed to move during the minimization whereas all other residues were fixed, and a cutoff of long-range forces at 9.5-10 Å was used. The mmff94s forcefield as implemented in the MOE molecular modeling program version 2002.03 was applied for the calculations which included solvation. Visualization of the complex was performed using the MOE molecular modeling program.

The structural coordinates of the NS5B-Compound **C** complex were generated from the final minimization and are shown in Figure 6.

References:

- Ago *et al.* **1999**, Structure 7(11): 1417-1426
- Ausubel *et al.*, **1994**, Current Protocols in Molecular Biology, Wiley, New York.
- Bressanelli *et al.* **1999,** Proc. Natl. Acad. Sci, USA 96(23): 13034-13039.

- Bressanelli *et al.*, **2002**, J. Virol. 76: 3482-3492.
- Dayhoff, M.O., **1978**, Atlas of Protein Sequence and Structure, **5**, suppl. 3, National Biomedical Research Foundation, Washington, D.C.
- Kolykhalov, A.A. *et al.*, **2000**; J. Virol. 74: 2046-2051.
- Labonté *et al.*, **2002**, J. Biol. Chem. 277(41): 38838-38846.
- Lai, M.M.C., **1998**, Virology 244, 1-12.
- Lehninger, A. L., *et al.*, **1993**, Principles of Biochemistry (2nd edn.) Worth, New York.
- Lesburg *et al.* **1999**, Nat. Struct. Biol. 6: 937-943.
- Love *et al.*, **2003**, J. Virol. 77: 7575-7581.
- McPherson *et al.*, **1989**, Preparation and Analysis of Protein Crystals, Krieger Pub.
- O'Farrell *et al.*, **2003**, J. Mol. Biol. 326: 1025-1035.
- Pickworth-Glusker J., Trueblood K.N., Crystal Structure Analysis, 2nd Ed. Oxford University Press, 1985, New York.
- Sambrook *et al.*, **1989**, Molecular Cloning – A Laboratory Manual, Cold Spring Harbor Labs.
- Schulz G. E. and Schirmer R. H., **1985**, Principles of Protein Structure, Springer-Verlag, New York
- Strauss, J.H., and Strauss, E.G. **1999**, *Science* **283**, 802-804.
- Wang *et al.*, **2003**, J. Biol. Chem. 278(11): 9485-9495.

WO 01/047883.

WO 02/004425.

WO 03/000254.

WO 03/007945.

WO 03/010140.

WO 03/010141.

WO 03/026587.

EP 1 256 628.

US 10/755256.

US 10/755544.

US 60/546213.

CLAIMS

What is claimed is:

1. An isolated and purified polypeptide selected from the group consisting of:

i) a polypeptide comprising an HCV NS5B inhibitor binding pocket defined by at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 of a native HCV NS5B, or a functionally equivalent analog thereof, wherein said binding pocket is exposed by displacement of a finger loop chain defined by at least amino acid residues 18 to 35 and said binding pocket retains its native functional configuration;

ii) an HCV NS5B polypeptide analog comprising an HCV NS5B binding pocket defined by at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 of native HCV NS5B, or defined by a functionally equivalent analog thereof, wherein said binding pocket retains its native functional configuration and wherein said binding pocket is exposable;

iii) an HCV NS5B polypeptide consisting of an HCV NS5B binding pocket defined by at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 of native HCV NS5B, or defined by a functionally equivalent analog thereof, wherein said binding pocket retains its native functional configuration; and

iv) an HCV NS5B polypeptide variant comprising at least one amino acid mutation within a finger loop defined by amino acid residues 18 to 35, wherein said mutation provokes a displacement of said finger loop to expose a binding pocket essentially defined by amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503 of native HCV NS5B, or defined by a functionally equivalent analog thereof, wherein said binding pocket retains its native functional configuration.

2. The polypeptide as defined in claim 1, wherein the binding pocket is defined by one or more additional amino acid residues selected from: amino acid residues 37 and 496.

3. The polypeptide as defined in claim 2, wherein the binding pocket is defined by one or more additional amino acid residues selected from: amino acid residues 36, 426, 498 and 499.

4. The polypeptide as defined in claim 1, wherein the binding pocket is additionally defined by one or more amino acid clusters selected from the group consisting of amino acid residues 36-37, 392-399, 424-429 and 492-503.

5. The polypeptide as defined in claim 2, wherein the binding pocket has the amino acid sequence, V37, L392, A393, A395, A396, T399, I424, L425, H428, F429, L492, G493, V494, P495, P496, W500 and R503.

6. The polypeptide as defined in claim 3, wherein the additional amino acids have the amino acid sequence, M36, M426, R498 and T499.

7. The polypeptide as defined in claim 4, wherein the amino acid cluster 36-37 has the amino acid sequence M-V, the amino acid cluster 392-399 has the amino acid sequence L-A-R-A-A-W-E-T, the amino acid cluster 424-429 has the amino acid sequence I-L-M-T-H-F and the amino acid cluster 492-503 has the amino acid sequence L-G-V-P-P-L-R-T-W-R-H-R.

8. The polypeptide as defined in claim 1, wherein at least one of the amino acid residues at positions 30 and 31 of the HCV NS5B polypeptide variant is mutated.

9. The polypeptide as defined in claim 8, wherein the amino acid residue at position 30 is mutated.

10. The polypeptide as defined in claim 9, wherein the amino acid residue at position 30 is mutated to an amino acid selected from the group consisting of P, F, W, M, G, S, T, C, Y, N, Q, D, E, K, R and H.

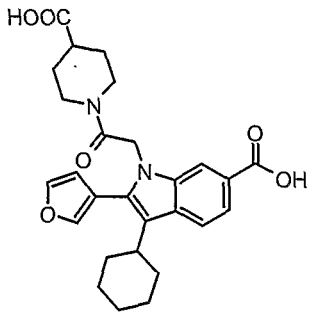
11. An HCV NS5B polypeptide, or a functionally equivalent analog thereof, characterized by displacement of amino acid residues 18 to 35.

12. An HCV NS5B polypeptide, or a functionally equivalent analog thereof, in which at least amino acid residues 18 to 35 have been deleted.

13. An HCV NS5B crystal structure comprising a binding pocket defined by the structural coordinates of at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503, and optionally one of amino acid residues 37 and 496, of a native HCV NS5B, or defined by the structural coordinates of a functionally equivalent analog thereof, wherein a native finger loop chain defined by the structural coordinates of at least amino acids 18 to 35 is displaced to expose said binding pocket.

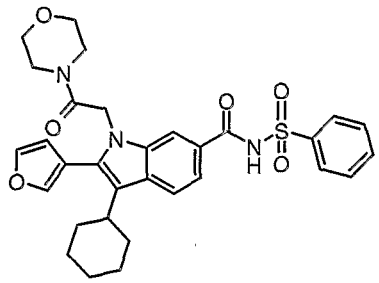
14. An NS5B complex comprising the polypeptide as defined in any one of claims 1 to 12 and a compound, wherein the compound associates with a binding pocket within the NS5B polypeptide which is defined by amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503, and optionally one of: amino acid residues 37 and 496, of native HCV NS5B, or defined by a functionally equivalent analog thereof.

15. The complex as defined in claim 14, wherein the compound has the following formula (A):



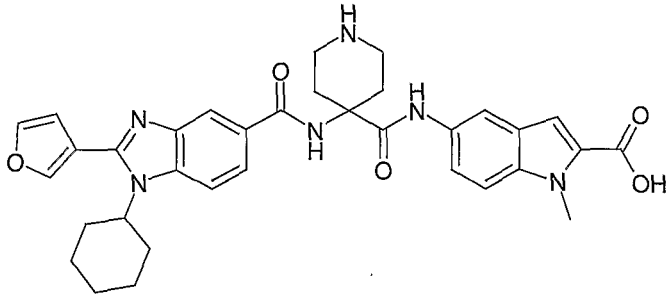
(A).

16. The complex as defined in claim 14, wherein the compound has the following formula (B):



(B).

17. The complex as defined in claim 14, wherein the compound has the following formula (C):



(C).

18. A method of producing a crystallized HCV NS5B complex comprising the polypeptide as defined in any one of claims 1 to 12 and a compound, wherein said compound associates with an NS5B binding pocket defined by at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503, and optionally one of: amino acid residues 37 and 496, of a native HCV NS5B, or defined by a functionally equivalent analog thereof, comprising the steps of:

a) incubating a purified HCV NS5B polypeptide in a crystallization buffer to obtain a crystallized NS5B polypeptide;

b) solubilizing the compound; and

c) soaking the crystallized polypeptide with the solubilized compound in a soaking buffer for a suitable soaking period to yield a crystallized NS5B complex.

19. The method as defined in claim 18, wherein the crystallization buffer is selected from the group consisting of MES, sodium phosphate, potassium phosphate, sodium acetate or sodium succinate.

20. The method as defined in claim 19, wherein the crystallization buffer is used at a concentration of between about 50 mM and 0.2 M.

21. The method as defined in claim 20, wherein the crystallization buffer is MES at a concentration of about 0.1 M.

22. The method as defined in claim 18, wherein the crystallization buffer additionally includes at least one precipitating agent.

23. The method as defined in claim 22, wherein the precipitating agent is selected from the group consisting of PEG, PEG5Kmme, ammonium sulfate, MPD, isopropanol, ethanol, and tertiary butanol.

24. The method as defined in claim 23, wherein the precipitating agent is used at a concentration of between about 30% and about 40%.

25. The method as defined in claim 24, wherein the precipitating agent is PEG5Kmme at a concentration of about 21% and ammonium sulfate at a concentration of about 0.4 mM.

26. The method as defined in claim 18, wherein the soaking buffer is selected from the group consisting of MES, Tris, sodium phosphate, sodium acetate and sodium succinate.

27. The method as defined in claim 26, wherein the soaking buffer is used at a concentration of between about 50 mM and about 0.2 M.

28. The method as defined in claim 18, wherein the soaking buffer comprises up to about 10 mg/mL of protein.

29. The method as defined in claim 18, wherein the soaking buffer additionally contains one or more salts selected from the group consisting of NaCl, $(NH_4)_2SO_4$, and KCl.

30. The method as defined in claim 29, wherein salt is present in the soaking

buffer at a concentration of about 100 mM to about 500 mM.

31. The method as defined in claim 30, wherein the soaking buffer comprises NaCl and $(NH_4)_2SO_4$.

32. The method as defined in claim 31, wherein the soaking buffer contains NaCl at a concentration of about 210 mM and $(NH_4)_2SO_4$ at a concentration of about 280 mM.

33. The method as defined in claim 18, wherein the soaking buffer additionally contains glycerol at a concentration of between about 10 and about 20%.

34. The method as defined in claim 18, wherein the soaking buffer additionally contain at least one precipitating agent.

35. The method as defined in claim 34, wherein the precipitating agent is selected from the group consisting of PEG, PEG5K mme, ammonium sulfate, MPD, isopropanol, ethanol and tertiary butanol.

36. The method as defined in claim 35, wherein the precipitating agent is present at a concentration of about 10% to about 18%.

37. The method as defined in claim 36, wherein the soaking buffer contains the precipitating agent PEG5K mme (monomethyl ether polyethylene glycol 5000) at a concentration of about 14% and ammonium sulfate at a concentration of about 0.4 mM.

38. The method as defined in claim 18, wherein the soaking period is from about 3 to about 8 hours.

39. The method as defined in claim 38, wherein the soaking period is from about 5 to about 6 hours.

40. A method of preparing a crystallized HCV NS5B complex comprising the polypeptide as defined in any one of claims 1 to 12 and a compound, wherein said

compound associates with an NS5B binding pocket defined by at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503, and optionally one of: amino acid residues 37 and 496, of a native HCV NS5B, or defined by a functionally equivalent analog thereof, wherein the method comprises adding the compound to a crystallization buffer containing crystallized HCV NS5B.

41. A method of preparing a crystallized HCV NS5B complex comprising the polypeptide as defined in any one of claims 1 to 12 and a compound, wherein said compound associates with an NS5B binding pocket defined by at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503, and optionally one of: amino acid residues 37 and 496, of a native HCV NS5B, or defined by a functionally equivalent analog thereof, comprising the steps of:

a) combining purified HCV NS5B with the compound in solubilized form to form an NS5B complex; and

b) crystallizing the complex in a crystallization buffer.

42. A set of X-ray crystal structure coordinates of an HCV NS5B complex as defined in claim 14.

43. A set of X-ray crystal structure coordinates of an HCV NS5B complex as defined in claim 15.

44. The coordinates as defined in claim 43 as set out in Figure 4.

45. A set of X-ray crystal structure coordinates of an HCV NS5B complex as defined in claim 16.

46. The coordinates as defined in claim 45 as set out in Figure 5.

47. A set of X-ray crystal structure coordinates of an HCV NS5B complex as defined in claim 17.

48. The coordinates as defined in claim 47 as set out in Figure 6.

49. A computer-readable storage medium having stored thereon structure coordinates of the crystal structure of an HCV NS5B complex as defined in claim 14.

50. A computer-readable storage medium having stored thereon structure coordinates of the crystal structure of an HCV NS5B complex as defined in claim 15.

51. The storage medium as defined in claim 50, having stored thereon the structure coordinates as set out in Figure 4.

52. A computer-readable storage medium having stored thereon structure coordinates of the crystal structure of an HCV NS5B complex as defined in claim 16.

53. The storage medium as defined in claim 52, having stored thereon the structure coordinates as set out in Figure 5.

54. A computer-readable storage medium having stored thereon structure coordinates of the crystal structure of an HCV NS5B complex as defined in claim 17.

55. The storage medium as defined in claim 54, having stored thereon the structure coordinates as set out in Figure 6.

56. A method of identifying a compound that may bind to HCV NS5B, comprising the steps of:

a) applying a 3-dimensional molecular modeling algorithm to the structural coordinates of an HCV NS5B binding pocket defined by the structural coordinates of at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503, and optionally one of: amino acid residues 37 and 496, of native HCV NS5B, or as defined by the structural coordinates of a functionally equivalent analog thereof, to determine the spatial coordinates of the binding pocket of HCV NS5B; and

b) electronically screening stored spatial coordinates of the compound against the spatial coordinates of the HCV NS5B binding pocket to determine if the compound binds within the HCV NS5B binding pocket.

57. A virtual screening method to identify potential HCV inhibitors comprising the

steps of:

a) constructing a computer model of an HCV NS5B binding pocket defined by the structural coordinates of at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503, and optionally one of: amino acid residues 37 and 496, of the HCV NS5B, or defined by the structural coordinates of a functionally equivalent analog thereof;

b) employing computational means to perform a fitting program operation between computer models of the compound to be evaluated and the NS5B binding pocket to provide an energy-minimized configuration of the compound in the binding pocket; and

c) evaluating the results of the fitting operation to quantify the association between the compound and the binding pocket, wherein a compound that associates with the binding pocket to yield a low energy, stable complex is a potential NS5B inhibitor.

58. The method as defined in claim 57, wherein the complex has a van der Waals interaction energy value of less than about 8000 kcal/mol.

59. The method as defined in claim 58, wherein the complex has a van der Waals interaction energy value of less than about 6000 kcal/mol.

60. The method as defined in claim 59, wherein the complex has a van der Waals interaction energy value of less than about 4000 kcal/mol.

61. A method of screening candidate HCV NS5B inhibitor compounds comprising the steps of:

a) incubating the polypeptide as defined in any one of claims 1 to 12 with a candidate inhibitor compound under conditions suitable for binding; and

b) determining whether or not the candidate inhibitor compound binds to the polypeptide, wherein a compound that binds to the polypeptide is a potential HCV NS5B inhibitor.

62. A method of designing a compound which binds to an NS5B polypeptide as defined in any one of claims 1 to 12, comprising assessing the complementarity between the compound and a binding pocket in the NS5B polypeptide defined by at

least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503, and optionally one of: amino acid residues 37 and 496, of a native HCV NS5B, or defined by a functionally equivalent analog thereof.

63. A method of producing a drug which inhibits RNA replication activity of HCV NS5B, comprising identifying or designing a compound which fits into an NS5B binding pocket as defined by at least amino acid residues 392, 393, 395, 396, 399, 424, 425, 428, 429, 492, 493, 494, 495, 500 and 503, and optionally one of: amino acid residues 37 and 496, of a native HCV NS5B, or defined by a functionally equivalent analog thereof, wherein said binding pocket is exposed by displacement of a finger loop chain defined by at least amino acid residues 18 to 35.