

American University in Cairo

AUC Knowledge Fountain

Theses and Dissertations

6-1-2014

Identification of drug leads against HCV and malaria using different target proteins

Reem Al Olaby

Follow this and additional works at: <https://fount.aucegypt.edu/etds>

Recommended Citation

APA Citation

Al Olaby, R. (2014). *Identification of drug leads against HCV and malaria using different target proteins* [Master's thesis, the American University in Cairo]. AUC Knowledge Fountain.

<https://fount.aucegypt.edu/etds/29>

MLA Citation

Al Olaby, Reem. *Identification of drug leads against HCV and malaria using different target proteins*. 2014. American University in Cairo, Master's thesis. *AUC Knowledge Fountain*.

<https://fount.aucegypt.edu/etds/29>

This Dissertation is brought to you for free and open access by AUC Knowledge Fountain. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of AUC Knowledge Fountain. For more information, please contact mark.muehlhaeusler@aucegypt.edu.

**IDENTIFICATION OF DRUG LEADS AGAINST HCV AND MALARIA USING
DIFFERENT TARGET PROTEINS**

REEM RAFIK AL OLABI

Spring - 2014



**Ph.D. in Applied Sciences
BIOTECHNOLOGY PROGRAM
THE AMERICAN UNIVERSITY IN CAIRO**

**IDENTIFICATION OF DRUG LEADS AGAINST HCV AND MALARIA USING
DIFFERENT TARGET PROTEINS**

A THESIS PRESENTED BY
REEM RAFIK AL OLABI

TO
THE AMERICAN UNIVERSITY IN CAIRO

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF
**DOCTOR OF PHILOSOPHY IN APPLIED SCIENCES
BIOTECHNOLOGY PROGRAM**

Spring - 2014

UNDER THE SUPERVISION OF

**Dr. Hassan ME Azzazy – Chemistry Department- The American University in
Cairo- Egypt**

Dr. Rod Balhorn – SHALs Technologies INC. – Livermore – USA

RESEARCH COMMITTEE MEMBERS

**Dr. Shoshana Levy – Department of Medicine - Stanford University- Palo Alto-
USA**

Dr. Brett Chromy – Department of Pathology – UC Davis –Davis - USA

IN COLLABORATION WITH

#	University	Professor/ Lab	Contribution
1	California Institute of Technology-USA	Dr. Jost Vielmetter Protein Expression Center	Surface Plasmon Resonance Binding Assays (Biacore t200)
2	Stanford University – USA	Dr. Shoshana Levy Department of Medicine	Provision of GST-CD81 protein / Vector Flow Cytometry Assays
3	University of California Davis	Dr. Brett Chromy Department of Pathology and Laboratory Medicine	Dual Polarization Interferometry binding assays
4	Rutgers University -USA	Dr. Joseph Marcotrigiano Department of Chemistry and Chemical Biology	Provision of Recombinant E2 protein
5		Dr. Joel Freundlich Department of Pharmacology and Physiology	Advanced docking and virtual screening runs
6	Scripps Research Institute - USA	Dr. Arthur Olson Molecular Graphics Laboratory	Preliminary Virtual Screening Runs
7	John Hopkins University – USA	Dr. Photini Sinnis Department of Molecular Microbiology and Immunology	Sporozoite Invasion Assays
8	Institute Pasteur De Lille - France	Dr. Jean Dubuisson Department of Molecular & Cellular Virology of Hepatitis C	HCV Infection Inhibition Assays
9	Lawrence Livermore National Laboratory	Dr. Adam Zemla Dr. Saphon Hok	E2 Homology Model SHAL7153 synthesis

ACKNOWLEDGEMENTS

I do believe that nothing valuable in this life is accomplished with one hand. There should be team work from different aspects. I consider the PhD as one of the major milestones in my life and I would have never been able to pursue it without the help of many precious people in my life. My parents and my sister were always of great support to me, they never ceased to pray for me on every single step. My dear lovely husband was so understanding, so encouraging, and so supportive. He is one of the main reasons that made me able to continue my postgraduate studies. My beautiful children who are the wonderful gifts that give this life a different taste, bore with me a lot, there were times I had to sacrifice playing with them and enjoying their innocence and pure laughs so as to finish working on my research and writing.

I was blessed to have Dr. Hassan Azzazy and Dr. Rod Balhorn as my great advisers and mentors throughout my M.Sc. and PhD years. They have never ceased to be of great support and I have learnt from them a lot. I will never forget how Dr. Hassan was there for me from day 1 in the Biotechnology program 5 and a half years ago. I learnt from him how to be productive, practical and always think publication wise as this is what gives value to the scientist. I also learnt from him the concept of being a bioentrepreneur, he is great by all means and I owe him a lot. I was also blessed to have Dr. Rod as my co-adviser for more than 5 years now. He taught me many things, helped me in establishing several collaborations in The US and was always there for me, supporting me from all aspects. He is a true blessing to have in my committee. I also had the optimum pleasure to have Dr. Shoshana Levy and Dr. Brett Chromy as members in the research committee.

I would like to thank you all for benefiting me with your great knowledge, experience and above all valuable deep insights. Few sentences won't express my appreciation and gratitude to you...I need tens of pages to state how lucky I was to have such magnificent, supportive, and great mentors and advisers...

I would also like to thank Dr. Jost Vielmetter (Caltech), Dr. Joseph Marcotrigiano (Rutgers), Dr. Adam Zemla (Lawrence Livermore National Laboratory), Dr. Saphon Hok

(Lawrence Livermore National Laboratory), Dr. Jean Dubuisson (Institute Pasteur De Lille), Dr. Laurence Cocquerel (Institute Pasteur De Lille), Dr. Photini Sinnis (John Hopkins University), Dr. Joel Freundlich (Rutgers University), Dr. Arthur Olson (Scripps Research Institute), Angela Eldridge (UC Davis), Felipe Vences (Stanford University), Alex Perrymen (Rutgers University), Stefano Forli (Scripps Research Institute) for their great contribution to the PhD research projects.

I will never forget the role of Mrs. Sawsan Mardini in always being there for me. And I would like to take the opportunity to thank Dr. Amr Shaarawi , Dr. Ehab Abd El Rahman and Dr. Sherif Aly for their continuous support. Additionally, I would like to extend my deepest gratitude to Dr. Rania Siam and Dr. Ahmed Moustafa for their support as well.

Last but not least, I would like to express my deepest gratitude to Nancy Aitken, who stood by me a lot throughout the time I spent in the US doing my research work. I would also like to extend my appreciation to my great Uncle Al Bannis and Gayle Bannis in addition to my lovely Tammy Abdo who also were there for me always during my stay in The US to finalize my PhD research.

To you all, I dedicate my PhD thesis, literally without you, I wouldn't have been able to move one step further.

Thank you.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS		iv
LIST OF FIGURES		viii
LIST OF TABLES		X
LIST OF ABBREVIATIONS		1
Chapter I	Abstract	2
Chapter II	Literature Review	
	<i>Hepatitis C Virus</i>	5
	History	5
	HCV Genomic Diversity	6
	HCV Life Cycle	6
	HCV Epidemiology	9
	HCV Genome and Proteins	9
	Envelope Glycoproteins	12
	E2 Homology models and resolved peptides and crystal structures	13
	Targeting HCV E2 to develop anti- HCV drugs and vaccines	14
	CD81 significance in HCV infection	15
	HCV treatment	17
	References	24
	Tables	30
	Figures	34
	<i>Malaria</i>	
	History & Life cycle	42
	Malaria Diagnosis	43
	Malaria Treatment	43
	CD81: Malaria Host Factor	44
	References	45
	Figures	47
	<i>Summary & Conclusion</i>	48
CHAPTER III	Identification of ligands that target HCV E2- binding site on CD81-large extracellular loop	50
	Introduction	50
	Materials and Methods	54
	Results and Discussion	59
	Figures	66
	Tables	68
	References	

CHAPTER IV	Generation of a Bi-dentate SHAL prototype (SH7153) using CD81 as a target	77
	Introduction	77
	Materials and Methods	78
	Results	85
	Discussion and Conclusion	88
	References	93
	Figures	96
	Tables	100
CHAPTER V	281816- A novel drug lead identified to inhibit HCV infection in a genotype-independent manner	103
	Introduction	103
	Results	106
	Discussion	113
	Materials and Methods	116
	References	122
	Tables	128
	Figures	132
CHAPTER VI	<i>In silico</i> design of a bidentate SHAL (SH2216) against HCV using HCV E2 as a target	143
	Introduction	144
	Materials and Methods	145
	Results and Discussion	157
	References	159
	Figures	
CHAPTER VII	Identification of drug leads that block the plasmodium sporozoites: CD81 interaction	155
	Introduction	155
	Materials and Methods	156
	Results	159
	Discussion	160
	References	162
	Figures	163
	Tables	171
CHAPTER VIII	Conclusions	173
CHAPTER IX	Future Perspectives	187
CHAPTER X	Novelty	181
CHAPTER XI	Originality	182

LIST OF FIGURES

CHAPTER NO.	FIGURE NO.	CAPTION	PAGE
Chapter II	1	HCV Entry Proposed mechanism	34
	2	Model for HCV Virion Assembly	35
	3	Worldwide HCV Prevalence	36
	4	HCV Genome	37
	5	The E2c crystal structure resolved by Kong <i>et al.</i>	38
	6	The crystal structure of HCV E2 resolved by Khan <i>et al.</i>	39
	7	Role of CD81 in disrupting HCV-induced immune functions	40
	8	Different classes of HCV drugs targeting different parts of the HCV genome	41
Chapter III	1	Malaria Life Cycle	47
	1	Amino acid residues that participate in HCV E2 binding to CD81-LEL	68
	2	Predicted free energy data for ligand binding sites identified on the surface of CD81-LEL by AutoLigand	69
	3	Two ligand binding sites identified by AutoLigand on the open conformation of CD81-LEL (PDB ID: 1G8Q)	70
	4	AutoLigand Analysis of the Site 1 ligand binding site	71
	5	Confirmation of ligand binding to CD81-LEL using DPI	72
	6	Binding of Ligand 1 to CD81-LEL as a function of ligand concentration	74
Chapter IV	1	The docked conformations of ligands 689002 and 93033 on CD81-LEL	96
	2	The molecular scaffold designed to link ligands 689002 and 93033 together to create SH7153	97
	3	Synthetic scheme used to synthesize SH7153	98

	4	Structures of the two SHAL by-products isolated from the SH7153 preparation by HPLC	99
	5	Binding and dissociation curves for the SHAL SH7153 and its two ligand components binding to recombinant CD81-LEL obtained by surface Plasmon resonance/	100
Chapter V	1	Comparison of structural template used for HCV E2 modelling	132
	2	Comparison of the crystal structure of E2d with the homology model (Ribbon)	133
	3	Location of five ligand binding sites to select ligands for experimental testing	134
	4	Comparison of the crystal structure of E2c with the homology model	135
	5	SPR sensogram	136
	6	281816 inhibits HCV entry in a genotype-independent manner	137
	7	281816 inhibition of HCV E2 protein binding to CD81	138 139
	8	281816 inhibits HCV entry	140
	9	281816 blocks HCV cell to cell transmission	141
	10	281816 binding sites on HCV E2	142
Chapter VI	1	The 4 binding sites selected to guide virtual screening on HCV E2	149
	2	281816 binding sites on HCV E2	150
	3	281816 analogue structure	151
	4	281816 analogue binding sites	152
	5	The <i>in silico</i> design of bidentate SH2216	154
Chapter VII	1	Stretch of 21 amino acid residues involved in plasmodium sporozoites :CD81-LEL interaction	163
	2	The 3 sites chosen on CD81-LEL to undergo virtual screening runs	164
	3	Bar chart representing the change in response units of the ligands that were found to bind to CD81-LEL using surface Plasmon resonance	165

	4	Inhibition of Plasmodium Yoelii infection of HepG2 cells	166
	5	Inhibition of Plasmodium yoleii and Plasmodium berghei infection of Hepa 1-6 cells	167
	6	Comparison of inhibition of P.y. and P.f. in HepG2 cells	168
	7	The different binding modes of 7962 to CD81	169
	8	The different binding modes of 73735 to CD81	170
Chapter IX	1	281816 analogues that were found to be FDA approved.	183

LIST OF TABLES

CHAPTER NO.	TABLE NO.	CAPTION	PAGE
Chapter I			
Chapter III	1	HCV Proteins	30
	2	HCV Protease Inhibitors	31
	3	HCV Entry Inhibitors	32
	1	Experimental analysis of ligand binding to recombinant CD81-LEL	75
	2	DPI competition experiment showing inhibition of E2 binding to CD81-LEL by Ligand 3 (689002) and comparing the inhibition to that achieved with benzyl salicylate.	76
Chapter IV	1	The change in radians of ligands 689002 and 93033 using dual polarization interferometry	101
	2	Kinetics study of 689002, 93033, and SH7153	102
Chapter V	1	<i>In silico</i> data of the 23 ligands identified by docking and SPR to bind to E2	128
	2	SPR binding data for the 23 ligands that were identified to bind to recombinant E2 protein	129
	3	The IC ₅₀ values of the 23 ligands	130
	4	Genotype independent inhibition exerted by 281816	131
Chapter VII	1	Surface Plasmon resonance experiment results showing ligands that bound to CD81-LEL and the change in their response units (RU)	171
	2	Percentage of infection inhibited by the 5 tested ligands in the 3 <i>Plasmodium sp.</i>	172

LIST OF ABBREVIATIONS

ABBREVIATION	DEFINITION
CBA	Carbohydrate Binding Agent
CD81-LEL	Cluster of Differentiation 81-Large Extracellular Loop
CLDN-1	Claudin-1
CMV	Cytomegalovirus
DAAs	Direct Acting Antivirals
DC-SIGN	Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin
DPI	Dual Polarization Interferometry
E2c	Envelope Glycoprotein Core
eIF	Eukaryotic Initiation Factor
ER	Endoplasmic Reticulum
GAGs	Glycosaminoglycans
HCV	Hepatitis C Virus
HCV E2	Hepatitis C Virus Envelope Glycoprotein 2
IFN	Interferon
LDL-R	Low Density Lipoprotein Receptor
LDs	Lipid Droplets
LVP	Lipo-Viro-Particles
MHP	Molecular Hydrophobicity Potential
MTP	Microsomal Triglyceride Transfer Protein
NANB hepatitis	Non A Non B Hepatitis
NCI-DSII	National Cancer Institute Diversity Set II
NK	Natural Killer Cells
OCLDN	Occludin
Rad	Radians
RBV	Ribavirin
RC	Replicase Complex
RU	Response Units
SHAL	Selective High Affinity Ligand
SPR	Surface Plasmon Resonance
SR-B1	Scavenger Receptor-B 1
SVR	Sustained Virologic Response
TBEV	Tick-borne encephalitis virus
UTR	Untranslated Region
VLDL	Very Low Density Lipoproteins

CHAPTER I

ABSTRACT

Hepatitis C Virus (HCV) infects 170 million individuals worldwide. Although several newly FDA approved drugs targeting the HCV serine protease and polymerase have shown promising results, there is a need for better drugs that are effective in treating all HCV genotypes and subtypes to be used in an interferon-free regimen. On the other hand, malaria is another public health burden that causes 219 million clinical episodes, and 660,000 deaths per year. In addition, 3.3 billion people live in areas at risk of malaria transmission in 106 countries. It is alarming that 86% of deaths caused by malaria globally were in children. Several challenges are faced when treating malaria, such as resistance against drugs that are used in treatment. This necessitates the development of new classes of drugs to overcome resistance.

CD81 is a target protein that plays an essential role in the internalization of HCV into hepatocytes. Thus it was also targeted to identify sets of small molecule ligands predicted to bind to several sites that were identified to be involved in HCV infection. Thirty-six ligands predicted by AutoDock to bind to these sites were tested experimentally to determine if they bound to CD81-LEL. Binding assays conducted using surface Plasmon resonance revealed that 23 out of 36 of the ligands bound *in vitro* to the recombinant CD81-LEL protein.

In an effort to create new drugs that block hepatitis C virus entry into hepatocytes, we have designed and synthesized a small molecule that targets the HCV E2 glycoprotein binding site on CD81. A selective high affinity ligand (SHAL) (11) was created by linking together two small molecules that were predicted by docking and were shown by experimental methods to bind to the same site on CD81 where E2 binds. SH7153 was found to bind to recombinant CD81-LEL with a K_d of 21 μM but wasn't found to inhibit HCV infection when tested using Raji cells (antibody neutralizing assays) and HCV infection inhibition assays. This led to the conclusion that the linkers' lengths should be optimized so as to have a SHAL that fits properly in the desired binding sites.

The HCV glycoprotein E2 has also been shown to play an essential role in hepatocyte invasion by binding to CD81 and other cell surface receptors. Recently, 2 research groups were able to resolve the core structure of HCV E2 which will largely help providing structural information that can now be used to target the E2 protein and develop drugs that disrupt the early stages of HCV infection by blocking E2's interaction with different host factors. By targeting conserved E2 residues among different genotypes and subtypes in the CD81 binding site on HCV E2, one might also be able to develop drugs that block HCV infection in a genotype-independent manner. Using the E2c structure as a template, we have used homology modeling methods to develop a structural model of the E2 protein core (residues 421-645) that includes the three amino acid segments that are not present in the E2c structure. Blind docking to this model was then performed using a library of ~4000 small molecules and a set of 40 ligands predicted to bind near conserved amino acid residues involved in the HCV E2: CD81 interaction were selected for

experimental testing. Surface Plasmon resonance was used to screen the ligands for binding to recombinant E2 protein and the best binders were subsequently tested to identify compounds that inhibit the infection of hepatocytes by HCV. One compound, 281816, inhibited infection by HCV genotypes 1a, 1b, 2a, 2b, 4a and 6a with IC₅₀'s ranging from 2.2 μ M to 4.6 μ M. Such inhibitors may represent a new paradigm for HCV treatment. In an attempt to make 281816 more promising, a SHAL prototype was designed using an analogue of 281816 (SH2216). It would be tempting to test the SHAL inhibitory effect and compare it to the 281816's inhibitory effect.

To date, human CD81 (hCD81) is the only human surface protein known to play a role in the process by which sporozoites of several *Plasmodium* species infect human hepatocytes. Blocking a human receptor that is exploited for the entry process of pathogens has been proven to be a good strategy for fighting drug-resistant mutants. Hence, we targeted the 21 amino acid stretch on CD81 large extracellular loop that was found to be involved in *Plasmodium yoelii* invasion via virtual screening runs, preliminary binding assays and sporozoite invasion assays. This led to the identification of 4 drug leads that range between moderate and strong inhibitors of infection by *Plasmodium yoelii* and *Plasmodium falciparum*. Additionally one ligand was found to potentiate the invasion of *Plasmodium yoelii*.

CHAPTER II

LITERATURE REVIEW

Parts of the literature review are adapted from the M.Sc. thesis – By: Al Olabi, Reem Rafik (2010-07-19) - The American University in Cairo “*In silico* design of selective high affinity ligands against HCV using novel computational biology tools”

A) Hepatitis C Virus

1. History

In early 1970s researchers discovered that there were more than the two known serotypes of hepatitis, hepatitis A and B. A third type was discovered and it was called none A non B (NANB) hepatitis. This discovery was a result of extensive work, which began by monitoring the participants in the study analyzing an abnormal pattern of aminotransferases bi-monthly for 6 to 12 months. This was followed by transfusion studies that were an eye opener on acute hepatitis. It was believed that the transfusion associated hepatitis would be hepatitis B but it wasn't. This was proven when the samples were found to be non-reactive to hepatitis B tests. The stored sera from the transfusion studies were then used to inoculate chimpanzees, which later on developed biochemical abnormalities (e.g. elevated aminotransferases) indicating that NANB type was a transmissible agent. After 15 years of research, hepatitis C was identified by Houghton *et al.* (1). Specific serological tests were developed and a comparison between NANB and HCV confirmed that they are the same (2).

2. HCV genomic diversity

There are six major genotypes of HCV worldwide with RNA sequences that differ by >30 %. These genotypes are further classified into subtypes, isolates and quasispecies (3). The continuous mutation of HCV is due to lack of efficient proof reading by the RNA dependent RNA polymerase in addition to the lack of the 5' to 3' exonuclease activity which leads to accumulation of mutations and viral escape (4). The host immune's response is one of the major factors that affect the persistence of HCV infection leading either to viral clearance or to chronic infection. It is believed that there are positive and negative selection processes in which the beneficial mutations are selected and favored over deleterious ones (5). The HCV genotype is one of the most important predictive viral factors for the response to therapy. Other factors affecting response to therapy include sex and age of the patient, the initial viral load and the presence or absence of liver fibrosis (4).

3. HCV Life Cycle

3.1. *Proposed model of HCV entry into hepatocytes*

HCV entry into the host cells involves several host receptors such as CD81 (6, 7, 8), the tight junction protein Claudin-1 (CLDN-1) (6, 9), scavenger receptor class BI (6, 8, 10). LDL-R (6, 8), mannose binding lectins DC-SIGN and L-SIGN (6), heparan sulfate proteoglycans (11, 6) and the asialoglycoprotein receptor (6). LDL-R and heparan sulfate play the initial role by mediating the attachment of HCV to CD81 and SRBI which upon interacting will trigger HCV movement and internalization via

CLDN-1 and occludin (11). **Figure 1** shows the proposed HCV entry mechanism (11, 12).

It was proposed that HCV particles become associated with low density lipoproteins (LDL) and very low density lipoproteins (VLDL) leading to the conclusion that LDL receptors are involved in the entry. Glycosaminoglycans (GAGs) and LDL-R function as primary receptors that mediate the attachment of HCV to CD81 and SR-BI (11, 12). CD81 and SR-BI in turn triggers HCV movement towards the tight junction proteins; claudin-1 (CLDN1) and occluding (OCLDN) which act as co-receptors (11, 12). The tight junction proteins act at a late stage of entry which is followed by the internalization of HCV through Clathrin-mediated endocytosis. It was suggested that HCV then moves inside the host cell in a low pH endosome where the acidification of the endosome helps mediate an HCV glycoprotein: endosomal membrane fusion (12, 13). It was found that the envelope proteins of other flaviviruses and alphaviruses (such as Semliki Forest virus and Dengue virus) possess an internal fusion peptide (class II fusion protein) that rearranges and trimerizes when exposed to low pH. This led to an assumption that HCV could involve class II fusion proteins in its internalization process (12).

3.2. *Polyprotein Synthesis*

This step begins with the decapsidation of the HCV nucleocapsid releasing the free positive stranded HCV genomic RNA which in turn acts as a mRNA with the newly synthesized RNA. This is followed by the translation of the HCV genome which is IRES-

dependent to produce the structural and nonstructural proteins needed for the completion of the life cycle, assembly and infecting other cells (14, 15). A cap-independent internal initiation of HCV polyprotein translation is mediated by IRES via recruiting both cellular proteins, including eukaryotic initiation factors (eIF) 2 and 3 and viral proteins (16, 17, 18). It was found that the core protein and nonstructural proteins NS4A and NS5B alters the IRES translational efficiency (19, 20).

3.3. HCV Replication

Rearrangements of intracellular membranes are induced by the Infection with the HCV positive-strand RNA virus and HCV NS4B protein was found to induce the formation of the membranous web necessary for replication (21, 22). The membranous web is derived from ER membranes (23). The membranous web consists of small vesicles embedded in a membranous matrix, forming a membrane-associated multiprotein complex that contains all of the nonstructural HCV proteins (21). HCV replication is catalyzed by the NS5B RNA-dependent RNA polymerase (RdRp). The positive-strand genome RNA serves as a template for the synthesis of a negative-strand intermediate of replication during the first step. In the second step, negative-strand RNA serves as a template to produce numerous strands of positive polarity that will subsequently be used for polyprotein translation, synthesis of new intermediates of replication or packaging into new virus particles (23).

3.4. Virus Assembly and Release

Initiation of virion assembly requires release of replicated genomes from the endoplasmic reticulum to allow their contact with core protein, which forms the HCV capsid. Assembly is thought to initiate in the cytosol before further maturation, and release occurs by transfer of nascent particles across the ER membrane to enable access to the secretory pathways in hepatocytes [24].

To sum up the assembly and release of HCV particles, it is divided into 3 stages that are interconnected: (i) an initial phase of assembly at LDs, (ii) the contribution made by viral factors that probably takes place after assembly has begun, (iii) and events on the luminal side of the ER membrane where there is engagement with the VLDL assembly pathway to facilitate virion maturation and release (figure 2).

4. HCV Epidemiology

Hepatitis C virus infects around 200 million of the world population (figure 3). In the developed countries, the chronic HCV infection accounts for 80% of the patients, and acute HCV for 20%. 40% of the infected patients' progress to end stage cirrhosis, 60% to hepatocellular carcinoma and 30% require liver transplantation (25). Chronic hepatitis C prevalence ranges from 0.1% to 14.5% in different countries (26) and Egypt has the highest HCV prevalence worldwide (15%).

5. HCV genome and proteins

HCV is a positive strand-RNA virus possessing a genome of approximately 9.5 kilobases which encodes a large polyprotein containing 3010 amino acids (figure 4) (14).

The processing and cleavage of the polyprotein is performed using the machinery of the host and enzymes produced by the virus. Ten structural and non-structural proteins are produced upon cleavage of the polyprotein. Either pathway involves peptidase cleavage signals (27). There are two untranslated regions (UTRs) one at the 5' end and one at the 3' end of the RNA (figure 4) (28, 29).

The structural proteins are the envelope glycoproteins denoted E1 and E2 and the core protein. There are two other proteins of unknown significance known as “p7” and “F” proteins (28, 30). As for the non-structural proteins; they are denoted NS2, NS3, NS4A, NS4B, NS5A, and NS5B (27, 30). The structural proteins are released by the host machinery in the endoplasmic reticulum, whereas the non-structural proteins are released following cleavage by NS2-3 and NS3/4A protease. The C terminus of the capsid protein undergoes further processing triggered by a signal peptide peptidase (29).

Table 1 provides a brief description of the HCV proteins and their functions. Since the main HCV targets in this phase are HCV E2 glycoproteins, and HCV NS3/4A complex, the next few paragraphs will provide more details on their nature and function (29).

6. HCV Envelope glycoproteins

E1 and E2 glycoproteins are type I transmembrane proteins that possess a C-terminal transmembrane domain and an N-terminal ectodomain that is N-linked to glycans. The glycans were found to play a role in virus entry and glycoprotein folding

(27).). Since HCV envelope glycoproteins are highly glycosylated, carbohydrate binding agents (CBA) such as cyanovirin-N might prove to be promising in blocking HCV entry into the cells (31). Both proteins are cleaved from the polyprotein by host signal peptidase cleavage (18). E1 and E2 form a natural non-covalent heterodimer that is retained in the endoplasmic reticulum (ER) (27, 32). Their retention in ER contributes to HCV infection conversion into the chronic form because it cannot be detected by the host immune system and also proves that the virus is released from the host cells by budding and exocytosis (27).

Several host receptors were found to be involved in HCV entry by either mediating viral attachment such as LDL-R and heparan Sulfate (23) or by interacting with HCV E2 glycoprotein such as CD81 and scavenger receptor BI (SR-BI) (24-26). These interactions trigger HCV movement to the tight junctions and its uptake by Claudin-1 and its uptake via Claudin-1 and occludin (1). Two hypervariable regions (HVR) have been identified in the HCV E2 glycoprotein sequence. HVR-1 is formed from the first 27 amino acids of the E2 N-terminus (ectodomain), and it largely contributes to the escape of the virus from the host immune response. HVR-1 has also been shown to modulate HCV entry. As for HVR-2, it has been described as an HCV entry modulator (27, 32). Viral glycoproteins such as HCV E2 are considered to be a promising target to generate drug against because it plays an important role in host cell interaction and thus if blocked will dissect the HCV life cycle in its earliest phase. HCV envelope glycoproteins are believed to belong to class II fusion proteins according to the model created by Yagnik *et al.* (33). Recently, 2 new

studies were done that showed that HCV E2 doesn't belong to the class II fusion proteins as will be shown in the next section (38,39).

7. E2 Homology models and resolved peptides and crystal structures

There is no crystal structure available for HCV E2 glycoprotein, nor are there reliable homology models except for that developed by Yagnik *et al.* (33). It is strongly believed that developing a reliable model of HCV E2 will be considered a break through because it will open another path for new category of anti-HCV drugs based on fragment based drug design and selective high affinity ligands approach.

Yagnik *et al.* (33) applied several fold recognition methods (such as TOPITS and THREADER2) to develop an HCV E2 model using the envelope protein E of Tick Borne Encephalitis virus as a template. In addition they constructed two truncated E2 proteins by PCR. These two proteins (N2 and H strains) were used in heparin and CD81 binding studies. The aim of those binding studies was to map the results obtained from experimental studies they performed onto their E2 model with the goal of identifying important residues involved in E2:CD81 interaction and a heparin binding domain. The results obtained in this study were also used to propose a rough model for the quaternary structure of the envelope glycoproteins E1 and E2 complex (34).

The C-terminal ends of the E1 and E2 glycoproteins play important roles in E1-E2 heterodimerization and retention in (ER). Charlotiaux *et al.* (35) examined the C-terminal sequences of E1 and E2 in 25 HCV strains and used these sequences to generate

several 3-D models. They found out that both domains of E1 and E2 should have a configuration of one amphipathic α helix followed by a pair of transmembrane strands. After generating 3-D models and minimizing the energy, the molecular hydrophobicity potentials (MHP) calculation and the IMPALA procedure (a Monte Carlo minimization of energy based simulation) were used to evaluate the C-terminal domains' interaction with the membranes. They concluded that the created 3-D models were consistent with the experimental studies in that both C-terminal domains are involved in ER retention and in E1-E2 heterodimerization (35).

Yu *et al.* (36) created a structural model of the HCV E2 protein using the Yagnik *et al.* (33) sequence alignment. The template used was the crystal structure of the tick-borne encephalitis virus (TBEV) soluble E protein. The ectodomain of HCV E2 in this model was similar to that of domains I and II of TBEV E protein and HCV E1 glycoprotein was similar to domain II of TBEV E protein. They proposed that E1 and E2 glycoproteins form a tetramer that forms the basic building block of the outer shell of HCV where as in case of flavivirus (TBEV) the outer shell is formed of E-protein dimers (36).

Spiga *et al.* (37) reported a 3-D structure for the HCV E2 homodimer that was developed based on a secondary structure prediction using PsiPred. This was followed by fold recognition using GenThreader v.2.1. software. They used both the TBEV E-protein as a template and the alignment of the consensus sequence of six HCV E2 genotypes (1a, 1b, 2a, 2c, 3a, 4c) in building the E2 model using DeepView v.3.7 software. The model's sequence identity with the template was 14% and the structure exhibited little secondary

structure (mainly B-sheets). In addition, molecular docking simulations were conducted to predict the HCV E2 dimeric assembly. The aim of the study conducted by Spiga *et al.* was to find surface exposed protein sequences that could be immunoreactive in the HCV E2 glycoprotein (E2 mimotopes) and also common in most genotypes. Nine fragments were selected in the developed model, synthesized on nitrocellulose membranes and tested through binding assays using phosphate conjugated anti-HCV human antibodies. This approach was considered important because it might be used to develop diagnostic kits (37).

Recently, Kong *et al.* (38) and Khan *et al.* (39) have been able to resolve the crystal structure of the core of the HCV E2 glycoprotein. Kong *et al.* (38) obtained the structure of amino acid residues 384-746 (E2c) by designing and expressing 41 soluble HCV E2 constructs and selecting 15 to screen against E2-specific Fab fragments in crystallization trials. Using a combination of x-ray crystallography and negative stain-electron microscopy, Kong *et al.* (38) discovered that the structures they obtained for E2 were globular and very different from the predicted models of E2 that were developed using class II fusion protein templates containing three β -sheet domains (figure 5).

Additionally, they were able to identify key CD81-binding residues through mutational studies. The CD81 binding sites were determined to be in the AR3C epitope, along one side of the B-sandwich (an isolated region of the CD81-binding loop) and a front layer consisting of loops, short helices and β -sheets (38). AR3C was also found to cross-neutralize HCV genotypes by blocking CD81 binding to HCV E2 (38). On the other

hand, the resolved structure of Khan et al. (39) were reported to be crystallized in complex with a FAB fragment of the mouse monoclonal antibody 2A12, and it was found to be highly similar to the previously reported structure (figure 6).

8. Significance of CD81 in HCV Infection

Several research groups have found that the CD81- large extracellular loop (CD81-LEL) is the domain involved in HCV entry via interaction with HCV E2 glycoprotein (40,41). Higginbottom *et al.* (42) identified 4 amino acid residues that are of large significance in the E2-CD81 interaction using mutational studies. They found that the D196E mutation reduced the binding to E2. In addition mutations F186L and E188K inhibited binding of CD81 to E2 whereas T163A enhanced their fusion (42). Drummer *et al.* (43) identified an E2 binding site in CD81-LEL to occupy around 806 Å². The most important contact residues they found within this site were Ile182, Phe186, Asn184, and Leu162 (43). The identification of these residues provides new targets for the design of selective high affinity ligands that could inhibit E2-CD81 interaction. Zhang *et al.* (44) also discovered a novel function for CD81 in HCV life cycle. They found that CD81 is important for efficient HCV genome replication.

This same group also discovered that the E2-CD81 interaction had several immuno-modulatory implications, among which include inducing *in vitro* a co-stimulatory signal in naive and antigen-experienced T cells leading to production of pro-inflammatory cytokine gamma interferon. Thus the CD81-E2 interaction might play a role in T-cell-mediated liver inflammation and lead to liver damage. In addition, the interaction

of these two proteins down regulates T cell receptors and suppresses the activity of natural killer (NK) cells (43). Brazzoli *et al.* found out that CD81 is a central regulator for several events within the HCV life cycle. CD81 was found to activate Rac, Rho, and Cdc42 (Rho GTPase family) mediating the contact between HCV E2/CD81 complex with CLDN-1 and occludin tight junction proteins to internalize HCV into the cell. Thus inhibiting CD81-HCV E2 interaction will block these events as stated and will massively reduce the infectivity of HCV (figure 7) (45,46). Figure 7 illustrates the role of CD81 in disrupting the host immune response against HCV. On the initiation of HCV infection, the CD81: HCV E2 interaction on the surface of different immune cells leads to activation, aggregation, increased survival of B cells and triggers VH gene hypermutation. This is thought to cause the HCV-induced mixed cryoglobulinemia and non-Hodgkin's lymphoma (45). Additionally, this interaction leads to the activation of the protein tyrosine kinase Lck which might be responsible for the presence of auto-reactive T cells in patients infected with chronic HCV (45,46). On the other hand, the HCV E2: CD81 interaction might also lead to the production of a ligand for the CCR5 receptor (RANTES) which when ligates to CCR5 at the surface of pDCs ends in priming T cells and decreasing the secretion of IFN- α and impaired maturation, proliferation and survival of these cells (45,46). On the whole, impairment of pDCs and NK cell functions might explain the establishment of chronic HCV infection (45).

To conclude the significance of CD81 in HCV infection, it is considered a key player in regulating HCV life cycle where it controls the initiation of the infection. CD81 is also involved in triggering the host response due to HCV infection, in addition to the

replication and dissemination of HCV. The significant functions that CD81 plays in HCV infection are reviewed in (46).

9. HCV treatment

Sustained virologic response (SVR) is defined as “having no detectable viral RNA after completion of antiviral therapy.” (47). The main aim of any anti-HCV treatment is to achieve SVR since it is accompanied with better clinical outcomes, lower rates of HCV-related morbidity and mortality, and stabilization of liver histology (47,48). The SVR was assessed after the end of therapy by 6 months. Recently, the assessment is also done after 12 weeks of end of therapy so as to stop the treatment after 12 weeks if the SVR is achieved (48).

The HCV Standard of Care (SOC) for many years was a combination of peginterferon (PEG-IFN) and ribavirin (RBV) owing to it being the only FDA approved combination therapy for HCV (49). This treatment was shown to have many side effects that could lead to the non-adherence to the assigned regimen in addition to having different rates of response to it depending on the HCV genotype and subtype.

Great efforts are being carried out to produce new drugs with higher efficacy, and better patient compliance being of less side effects, frequency and route of administration. The life cycle of HCV has several steps of interest that could be targeted and yield promising classes of drugs. Among the classes that are being developed are NS3 protease inhibitors, NS5B polymerase inhibitors, ribavirin analogs and cyclophilin inhibitors (figure 8) (50). On the other hand, different research groups were working on resolving

the structure of HCV E2 so as to help in creating anti-HCV vaccines and drugs by targeting HCV E2. Additionally, there are several efforts done to target the entry stage in the HCV life cycle through targeting the host factors involved in HCV entry into hepatocytes that were discussed in the HCV life cycle section.

9.1. Drugs Targeting HCV (Directly Acting Antivirals “DAAs”)

9.1.1. Protease Inhibitors

HCV serine protease has been extensively studied at the biochemical level and its structure is known (51,52). The NS3 serine protease is located in the N-terminal region of NS3 and is associated with the NS4A cofactor. Being one of the characterized HCV proteins, it became a potential anti-HCV drug target. The NS3/4A protease inhibitors can be divided into two chemical classes: macrocyclic inhibitors and linear tetrapeptide α -ketoamid derivatives.

Serine protease inhibitors range from being of low to high genetic barrier where genetic barrier is the number of aminoacid mutations needed to confer resistance to the drug used in treatment. The protease inhibitors of low genetic are more prone to HCV resistance needing only single mutation in the HCV particle whereas those of high barrier are less prone to resistant HCV strains since they need more than one mutation (53). Simeprevir is the newly FDA approved oral protease inhibitor (50). Its efficacy has been established in combination with peginterferon alfa and ribavirin in patients with HCV genotype 1 infection with compensated liver disease (including cirrhosis). In **Table 2**, the current protease inhibitors that are in different clinical phases, the manufacturing

company, the chemical structure being linear or macrocyclic, the genetic barrier, the required dose per day, their IC₅₀ values, and the genotypes they are active towards are enumerated (53).

9.1.2. *Polymerase Inhibitors*

Polymerase inhibitors bind to NS5B RNA-dependent RNA polymerase. NS5B RNA interfering with the viral replication. They are categorized into 2 different types: nucleoside inhibitors (NI) and non-nucleoside inhibitors (NNI). NI mimic the natural substrates of the polymerase thus incorporated into the RNA chain causing direct chain termination (54, 55). They require conversion to an active triphosphate form. On the other hand, NNI bind to several discrete sites outside of the HCV polymerase active center, resulting in conformational changes in the HCV polymerase before the elongation complex is formed (54,55). Sofosbuvir is the most significant newly FDA approved oral HCV drug and polymerase inhibitor. Sofosbuvir is used in combination with ribavirin or ribavirin plus interferon to treat HCV genotypes 1,2,3 and 4.

9.1.3. *NS5A Inhibitors*

The NS5A is a membrane-associated phosphoprotein in the HCV genome, which is present in phosphorylated (p56) and hyper-phosphorylated (p58) forms (55–57). It is involved in HCV virion production and has significant roles at various stages of HCV life cycle. NS5A has been shown to play a role in conferring resistance to interferon via interaction with a number of host proteins *in vivo* (55, 56). Daclastavir, an NS5A inhibitor, was found to be active at picomolar concentrations in HCV

replicons expressing a broad range of HCV genotypes (55–57). Accordingly, it might be another holy grail in the field of HCV treatment.

9.1.4. Targeting HCV E2 to develop anti-HCV drugs and vaccines

Several approaches are being used to develop anti-HCV drugs and vaccines that target the HCV E2 glycoprotein. Developing monoclonal and polyclonal antibodies is one of the approaches used. XTL Biopharmaceuticals reported that a combination of the monoclonal antibodies AB68 and AB65 (AB6865) was more effective than each one alone (58). Both monoclonal antibodies were obtained by immortalizing peripheral blood mononuclear cells obtained from the blood of HCV infected patients. It was observed that AB6865 combination can identify different conformational epitopes on E2 and immunoprecipitate different strains of HCV particles. The viral clearance takes place upon formation of immune complexes by endocytosis since that AB6865 induces phagocytosis of the immune complexes by neutrophils (58-61).

Bavituximab® is another monoclonal antibody developed by Peregrine pharmaceuticals that acts against a unique target, phosphatidyl serine. Phosphatidylserine is a lipid molecule normally found in the interior of cellular membranes and becomes exposed on the cell surface of cell membranes of viruses and host infected cells during replication. It has been reported that the presence of the phosphatidyl serine on the surface of the membranes of viruses creates a masking effect and thus helps the virus to evade the host immune response. This monoclonal antibody can be used against several viruses such as HIV, CMV, influenza and HCV. When comparing Bavituximab with other monoclonal

antibodies, it was found to be more advantageous because it has high specificity towards cells infected with viruses only. In addition it targets a molecule in the host and thus won't be affected by any viral mutations (62, 63).

Human hepatitis C immune globulin or Civacir is a polyclonal antibody obtained from human plasma enriched with HCV polyclonal antibodies collected from screened donors. It is used to prevent the re-infection of HCV infected patients who had a liver transplant (64).

Another approach is to develop a vaccine that triggers passive immunity against HCV. It was found that immunoglobulin G obtained from HCV infected patients could be a good candidate for developing such a vaccine. Vanwollegem *et al.* (64) injected 8 chimeric mice with immunoglobulin G obtained from HCV infected patients, which were then challenged with 100% infectious dose of the acute phase HCV. Five out of eight mice developed immunity against HCV. Law *et al.* (65) were able to isolate polyclonal antibodies that bind specifically to HCV E2 and neutralize several HCV strains. Full length immunoglobulins G were obtained from 3 antibody-antigen binding fragments that bound to 3 antigenic regions on E2. Monoclonal antibodies were obtained from these immunoglobulins and were injected intraperitoneally in mice with high levels of human liver chimerism. These mice were then challenged by HCV infected human serum, but the monoclonals didn't show the desired effect (65).

A third approach involves carbohydrate binding agents (CBA) against HCV E2. Since HCV envelope glycoproteins are highly glycosylated, carbohydrate binding agents

(CBA) such as cyanovirin-N might prove to be promising in blocking HCV entry into the cells (66). Several host receptors were found to be involved in HCV entry by either mediating viral attachment (such as LDL-R and heparan sulfate) (67) or by interacting with HCV E2 glycoprotein (such as CD81 and scavenger receptor BI (SR-BI)) (68-70). Based on these findings new approaches were opted such as developing small molecules that mimic the D-helix of CD81 and is known to interact with HCV E2 (71).

9.2. Drugs Targeting Host factors at the entry level

HCV entry inhibitors target the first steps of HCV life cycle. Several studies showed that they efficiently inhibit cell to cell transmission which is one of the significant contributors to the spread of the virus throughout the liver. This makes this class of drugs suitable for preventing liver graft infection which increases the rate of success of liver transplantation. On the other hand, there are other host factors that are involved in the immunomodulatory mechanisms that is controlled by HCV. They could be promising targets as well. What makes targeting host factors of potential interest is the high genetic barrier to resistance of some DAAs (72).

9.3. Targeting CD81-LEL to develop anti-HCV drugs

Several research groups have found that the CD81- large extracellular loop (CD81-LEL) plays a key role in HCV entry into cells by binding to the HCV E2 glycoprotein (10, 42, 73). Higginbottom *et al.* (42) and Drummer *et al.* (43) used mutational studies to identify residues that contribute to the E2-CD81-LEL interaction (42). This information is important because it can be used to direct the design of a series

of selective high affinity ligand-based inhibitors that block the E2-CD81-LEL interaction. (42,43). Zhang *et al.* (44) discovered a separate, additional function for CD81 in the HCV life cycle. These studies showed that CD81-LEL is important for efficient HCV genome replication where the level of CD81 expression was found to modulate the HCV RNA replication.

In addition, E2-CD81-LEL interaction induces several immuno-modulatory effects, including co-stimulatory signal in naive and antigen-experienced T cells that leads to the production of pro-inflammatory cytokine gamma interferon. Thus the CD81-E2-LEL interaction may play a role in T-cell-mediated liver inflammation and contribute to liver damage. In addition, the interaction of these two proteins down regulates T cell receptors and suppresses the activity of natural killer (NK) cells (43). Meuleman *et al.* (74) showed that using anti-CD81 as a prophylactic treatment completely protected human liver-uPA-SCID infection with different genotypes of HCV (74).

9.4. *Other host factors Entry Inhibitors*

HCV entry is considered one of the potential targets for creating an anti-HCV drug. Several host factors are involved in the entry of HCV into hepatocytes. Intervening with such mechanisms and blocking it might help in preventing both invasion of the virus, and cell to cell transmission. **Table 3** enumerates several classes of entry inhibitors that target different host factors (72).

10. References

1. Houghton, M. (2009). Discovery of the hepatitis C virus. *Liver Int.* 29 Suppl 1, 82-88.
2. Seeff LB. The history of the "natural history" of hepatitis C (1968-2009). *Liver Int.* 29 Suppl 1 89-99 (2009).
3. Le Guillou-Guillemette, H., Vallet, S., Gaudy-Graffin, C., Payan, C., Pivert, A., Goudeau, A. and Lunel-Fabiani, F. (2007). Genetic diversity of the hepatitis C virus: impact and issues in the antiviral therapy. *World J. Gastroenterol.* 13, 2416-2426.
4. Timm, J. and Roggendorf, M. (2007). Sequence diversity of hepatitis C virus: implications for immune control and therapy. *World J. Gastroenterol.* 13, 4808-4817.
5. Scott, J. D. and Gretch, D. R. (2007). Molecular diagnostics of hepatitis C virus infection: a systematic review. *JAMA.* 297, 724-732.
6. uerel, L., Voisset, C. and Dubuisson, J. (2006). Hepatitis C virus entry: potential receptors and their biological functions. *J. Gen. Virol.* 87, 1075-1084.
7. Favre, D. and Muellhaupt, B. (2005). Potential cellular receptors involved in hepatitis C virus entry into cells. *Lipids Health. Dis.* 4, 9.
8. Regeard, M., Trotard, M., Lepere, C., Gripon, P. and Le Seyec, J. (2008). Entry of pseudotyped hepatitis C virus into primary human hepatocytes depends on the scavenger class B type I receptor. *J. Viral Hepat.* 15, 865-870.
9. Zein, N. N. (2000). Clinical significance of hepatitis C virus genotypes. *Clin. Microbiol. Rev.* 13, 223-235.
10. Pileri, P., Uematsu, Y., Campagnoli, S., Galli, G., Falugi, F., Petracca, R., Weiner, A. J., Houghton, M., Rosa, D., Grandi, G. and Abrignani, S. (1998). Binding of hepatitis C virus to CD81. *Science.* 282, 938-941.
11. Budkowska, A. (2009). Mechanism of cell infection with hepatitis C virus (HCV)-a new paradigm in virus-cell interaction. *Pol. J. Microbiol.* 58, 93-98.
12. Moradpour, D., Penin, F. and Rice, C. M. (2007). Replication of hepatitis C virus. *Nat. Rev. Microbiol.* 5, 453-463.
13. Tscherne, D. M., Jones, C. T., Evans, M. J., Lindenbach, B. D., McKeating, J. A. and Rice, C. M. (2006). Time- and temperature-dependent activation of hepatitis C virus for low-pH-triggered entry. *J. Virol.* 80, 1734-1741.
14. Friebe, P., Boudet, J., Simorre, J. P. and Bartenschlager, R. (2005). Kissing-loop interaction in the 3' end of the hepatitis C virus genome essential for RNA replication. *J. Virol.* 79, 380-392.
15. Luo, G., Xin, S. and Cai, Z. (2003). Role of the 5'-proximal stem-loop structure of the 5' untranslated region in replication and translation of hepatitis C virus RNA. *J. Virol.* 77, 3312-3318.

16. Ji, H., Fraser, C. S., Yu, Y., Leary, J. and Doudna, J. A. (2004). Coordinated assembly of human translation initiation complexes by the hepatitis C virus internal ribosome entry site RNA. *Proc. Natl. Acad. Sci. U. S. A.* 101, 16990-16995.
17. Lukavsky, P. J. (2009). Structure and function of HCV IRES domains. *Virus Res.* 139, 166-171.
18. Otto, G. A. and Puglisi, J. D. (2004). The pathway of HCV IRES-mediated translation initiation. *Cell.* 119, 369-380.
19. Zhang, J., Yamada, O., Yoshida, H., Iwai, T. and Araki, H. (2002). Autogenous translational inhibition of core protein: implication for switch from translation to RNA replication in hepatitis C virus. *Virology.* 293, 141-150.
20. Kato, J., Kato, N., Yoshida, H., Ono-Nita, S. K., Shiratori, Y. and Omata, M. (2002). Hepatitis C virus NS4A and NS4B proteins suppress translation *in vivo*. *J. Med. Virol.* 66, 187-199.
21. Egger, D., Wolk, B., Gosert, R., Bianchi, L., Blum, H. E., Moradpour, D. and Bienz, K. (2002). Expression of hepatitis C virus proteins induces distinct membrane alterations including a candidate viral replication complex. *J. Virol.* 76, 5974-5984.
22. Gretton, S. N., Taylor, A. I. and McLauchlan, J. (2005). Mobility of the hepatitis C virus NS4B protein on the endoplasmic reticulum membrane and membrane-associated foci. *J. Gen. Virol.* 86, 1415-1421.
23. Bartenschlager, R., Frese, M. and Pietschmann, T. (2004). Novel insights into hepatitis C virus replication and persistence. *Adv. Virus Res.* 63, 71-180.
24. Jones, D. M. and McLauchlan, J. (2010). Hepatitis C virus: assembly and release of virus particles. *J. Biol. Chem.* 285, 22733-22739. VanCompernelle, S. E., Wiznycia, A. V., Rush, J. R., Dhanasekaran, M., Baures, P. W. and Todd, S. C. (2003). Small molecule inhibition of hepatitis C virus E2 binding to CD81. *Virology.* 314, 371-380.
25. Zeisel, M. B., Koutsoudakis, G., Schnober, E. K., Haberstroh, A., Blum, H. E., Cosset, F. L., Wakita, T., Jaeck, D., Doffoel, M., Royer, C., Soulier, E., Schvoerer, E., Schuster, C., Stoll-Keller, F., Bartenschlager, R., Pietschmann, T., Barth, H. and Baumert, T. F. (2007). Scavenger receptor class B type I is a key host factor for hepatitis C virus infection required for an entry step closely linked to CD81. *Hepatology.* 46, 1722-1731.
26. Krieger, S. E., Zeisel, M. B., Davis, C., Thumann, C., Harris, H. J., Schnober, E. K., Mee, C., Soulier, E., Royer, C., Lambotin, M., Grunert, F., Dao Thi, V. L., Dreux, M., Cosset, F. L., McKeating, J. A., Schuster, C. and Baumert, T. F. (2010). Inhibition of hepatitis C virus infection by anti-claudin-1 antibodies is mediated by neutralization of E2-CD81-claudin-1 associations. *Hepatology.* 51, 1144-1157.
27. Dubuisson, J. (2007). Hepatitis C virus proteins. *World J. Gastroenterol.* 13, 2406-2415.

28. Kwong, A. D., McNair, L., Jacobson, I. and George, S. (2008).Recent progress in the development of selected hepatitis C virus NS3.4A protease and NS5B polymerase inhibitors. *Curr. Opin. Pharmacol.* 8, 522-531.
29. Chockalingam, K., Simeon, R. L., Rice, C. M. and Chen, Z. (2010).A cell protection screen reveals potent inhibitors of multiple stages of the hepatitis C virus life cycle. *Proc. Natl. Acad. Sci. U. S. A.* 107, 3764-3769.
30. Davis, G. L., Nelson, D. R., Terrault, N., Pruet, T. L., Schiano, T. D., Fletcher, C. V., Sapan, C. V., Riser, L. N., Li, Y., Whitley, R. J., Gnann, J. W., Jr and Collaborative Antiviral Study Group (2005).A randomized, open-label study to evaluate the safety and pharmacokinetics of human hepatitis C immune globulin (Civacir) in liver transplant recipients. *Liver Transpl.* 11, 941-949.
31. Budkowska, A.(2009).Mechanism of cell infection with hepatitis C virus (HCV)-a new paradigm in virus-cell interaction. *Pol. J. Microbiol.* 58, 93-98.
32. Bartosch, B., Vitelli, A., Granier, C., Goujon, C., Dubuisson, J., Pascale, S., Scarselli, E., Cortese, R., Nicosia, A. and Cosset, F. L. (2003).Cell entry of hepatitis C virus requires a set of co-receptors that include the CD81 tetraspanin and the SR-B1 scavenger receptor. *J. Biol. Chem.* 278, 41624-41630.
33. Yagnik, A. T., Lahm, A., Meola, A., Roccasecca, R. M., Ercole, B. B., Nicosia, A. and Tramontano, A. (2000).A model for the hepatitis C virus envelope glycoprotein E2. *Proteins.* 40, 355-366.
34. Bartosch, B. and Cosset, F. L. (2006).Cell entry of hepatitis C virus. *Virology.* 348, 1-12.
35. Charloteaux, B., Lins, L., Moereels, H. and Brasseur, R. (2002).Analysis of the C-terminal membrane anchor domains of hepatitis C virus glycoproteins E1 and E2: toward a topological model. *J. Virol.* 76, 1944-1958.
36. Yu, X., Qiao, M., Atanasov, I., Hu, Z., Kato, T., Liang, T. J. and Zhou, Z. H. (2007).Cryo-electron microscopy and three-dimensional reconstructions of hepatitis C virus particles. *Virology.* 367, 126-134.
37. Spiga, O., Padula, M. G., Scarselli, M., Ciutti, A., Bernini, A., Venditti, V., Prischi, F., Falciani, C., Lozzi, L., Bracci, L., Valensin, P. E., Cudai, C. and Niccolai, N. (2006).Structurally driven selection of human hepatitis C virus mimotopes. *Antivir Ther.* 11, 917-922.
38. Kong, L., Giang, E., Nieuwma, T., Kadam, R.U., Cogburn, K.E., Hua, Y., Dai, X., Stanfield, R.L., Burton, D.R., Ward, A.B., Wilson, I.A. & Law, M. 2013, "Hepatitis C virus E2 envelope glycoprotein core structure", *Science (New York, N.Y.)*, vol. 342, no. 6162, pp. 1090-1094
39. Khan, A.G., Whidby, J., Miller, M.T., Scarborough, H., Zatorski, A.V., Cygan, A., Price, A.A., Yost, S.A., Bohannon, C.D., Jacob, J., Grakoui, A. & Marcotrigiano, J. 2014, "Structure of the core ectodomain of the hepatitis C virus envelope glycoprotein 2", *Nature* .

40. Levy, S., Todd, S. C., and Maecker, H. T. (1998). CD81 (TAPA-1): A molecule involved in signal transduction and cell adhesion in the immune system. *Annu. Rev. Immunol.* 16, 89-109
41. Zeisel, M. B., Koutsoudakis, G., Schnober, E. K., Haberstroh, A., Blum, H. E., Cosset, F. L., Wakita, T., Jaeck, D., Doffoel, M., Royer, C., Soulier, E., Schvoerer, E., Schuster, C., Stoll-Keller, F., Bartenschlager, R., Pietschmann, T., Barth, H. and Baumert, T. F. (2007). Scavenger receptor class B type I is a key host factor for hepatitis C virus infection required for an entry step closely linked to CD81. *Hepatology.* 46, 1722-1731.
42. Higginbottom, A., Quinn, E. R., Kuo, C. C., Flint, M., Wilson, L. H., Bianchi, E., Nicosia, A., Monk, P. N., McKeating, J. A. and Levy, S. (2000). Identification of amino acid residues in CD81 critical for interaction with hepatitis C virus envelope glycoprotein E2. *J. Virol.* 74, 3642-3649.
43. Drummer, H. E., Wilson, K. A. and Pombourios, P. (2002). Identification of the hepatitis C virus E2 glycoprotein binding site on the large extracellular loop of CD81. *J. Virol.* 76, 11143-11147.
44. Zhang, Y. Y., Zhang, B. H., Ishii, K. and Liang, T. J. (2010). Novel function of CD81 in controlling hepatitis C virus replication. *J. Virol.* 84, 3396-3407.
45. Brazzoli, M., Bianchi, A., Filippini, S., Weiner, A., Zhu, Q., Pizza, M. and Crotta, S. (2008). CD81 is a central regulator of cellular events required for hepatitis C virus infection of human hepatocytes. *J. Virol.* 82, 8316-8329.
46. Feneant, L., Levy, S. and Cocquerel, L. (2014). CD81 and hepatitis C virus (HCV) infection. *Viruses.* 6, 535-572.
47. Singal AG, Volk ML, Jensen D, Di Bisceglie AM, Schoenfeld PS. A sustained viral response is associated with reduced liver-related morbidity and mortality in patients with hepatitis C virus. *Clin Gastroenterol Hepatol* 2010; 8:280–288, 288.e1.
48. Camma C, Di Bona D, Schepis F, et al. Effect of peginterferon alfa-2a on liver histology in chronic hepatitis C: a meta-analysis of individual patient data. *Hepatology* 2004; 39:333–342.
49. Pol, S., Vallet-Pichard, A. and Corouge, M. (2014). Treatment of hepatitis C virus genotype 3-infection. *Liver Int.* 34 Suppl 1, 18-23.
50. Schinazi, R., Halfon, P., Marcellin, P. and Asselah, T. (2014). HCV direct-acting antiviral.
51. Kim JL, Morgenstern KA, Lin C, et al. Crystal structure of the hepatitis C virus NS3 protease domain complexed with a synthetic NS4A cofactor peptide. *Cell* 1996; 87: 343–55.
52. Yao N, Hesson T, Cable M, et al. Structure of the hepatitis C virus RNA helicase domain. *Nat Struct Biol* 1997; 4: 463–7.

53. McCown MF, Rajyaguru S, Kular S, et al. GT-1a or GT-1b subtype-specific resistance profiles for hepatitis C virus inhibitors telaprevir and HCV-796. *Antimicrob Agents Chemother* 2009; 53: 2129–32.
54. Koch U, Narjes F. Recent progress in the development of inhibitors of the hepatitis C virus RNA-dependent RNA polymerase. *Curr Top Med Chem* 2007; 7: 1302–29.
55. Gish RG, Meanwell NA. The NS5A replication complex inhibitors: difference makers? *Clin Liver Dis* 2011; 15: 627–39.
56. Gao M, Nettles RE, Belema M, et al. Chemical genetics strategy identifies an HCV NS5A inhibitor with a potent clinical effect. *Nature* 2010; 465: 96–100.
57. Asselah T. NS5A inhibitors: a new breakthrough for the treatment of chronic hepatitis C. *J Hepatol* 2011; 54: 1069–72.
58. Eren, R., Landstein, D., Terkieltaub, D., Nussbaum, O., Zauberman, A., Ben-Porath, J., Gopher, J., Buchnick, R., Kovjazin, R., Rosenthal-Galili, Z., Aviel, S., Ilan, E., Shoshany, Y., Neville, L., Waisman, T., Ben-Moshe, O., Kischitsky, A., Fong, S. K., Keck, Z. Y., Pappo, O., Eid, A., Jurim, O., Zamir, G., Galun, E. and Dagan, S. (2006). Preclinical evaluation of two neutralizing human monoclonal antibodies against hepatitis C virus (HCV): a potential treatment to prevent HCV reinfection in liver transplant patients. *J. Virol.* 80, 2654-2664.
59. Galun, E., Terrault, N. A., Eren, R., Zauberman, A., Nussbaum, O., Terkieltaub, D., Zohar, M., Buchnik, R., Ackerman, Z., Safadi, R., Ashur, Y., Misrachi, S., Liberman, Y., Rivkin, L. and Dagan, S. (2007). Clinical evaluation (Phase I) of a human monoclonal antibody against hepatitis C virus: safety and antiviral activity. *J. Hepatol.* 46, 37-44.
60. Schiano, T. D., Charlton, M., Younossi, Z., Galun, E., Pruet, T., Tur-Kaspa, R., Eren, R., Dagan, S., Graham, N., Williams, P. V. and Andrews, J. (2006). Monoclonal antibody HCV-AbXTL68 in patients undergoing liver transplantation for HCV: results of a phase 2 randomized study. *Liver Transpl.* 12, 1381-1389.
61. Borgia, G. (2004). HepeX-C (XTL Biopharmaceuticals). *Curr. Opin. Investig Drugs.* 5, 892-897.
62. Bavituximab. Peregrine Pharmaceuticals Inc. Fall Winter 2008. Available at: http://www.peregrineinc.com/images/stories/media/siteFiles/20081121_BavituximabAV.pdf. Accessed February 7, 2009.
63. Soares, M. M., King, S. W. and Thorpe, P. E. (2008). Targeting inside-out phosphatidylserine as a therapeutic strategy for viral diseases. *Nat. Med.* 14, 1357-1362.
64. Davis, G. L., Nelson, D. R., Terrault, N., Pruet, T. L., Schiano, T. D., Fletcher, C. V., Sapan, C. V., Riser, L. N., Li, Y., Whitley, R. J., Gnann, J. W., Jr and Collaborative Antiviral Study Group (2005). A randomized, open-label study to evaluate the safety and pharmacokinetics of human hepatitis C immune globulin (Civacir) in liver transplant recipients. *Liver Transpl.* 11, 941-949.

65. Law, M., Maruyama, T., Lewis, J., Giang, E., Tarr, A. W., Stamatakis, Z., Gastaminza, P., Chisari, F. V., Jones, I. M., Fox, R. I., Ball, J. K., McKeating, J. A., Kneteman, N. M. and Burton, D. R. (2008). Broadly neutralizing antibodies protect against hepatitis C virus quasispecies challenge. *Nat. Med.* 14, 25-27.
66. Helle, F., Wychowski, C., Vu-Dac, N., Gustafson, K. R., Voisset, C. and Dubuisson, J. (2006). Cyanovirin-N inhibits hepatitis C virus entry by binding to envelope protein glycans. *J. Biol. Chem.* 281, 25177-25183.
67. Budkowska, A. (2009). Mechanism of cell infection with hepatitis C virus (HCV)-a new paradigm in virus-cell interaction. *Pol. J. Microbiol.* 58, 93-98.
68. Bartosch, B., Vitelli, A., Granier, C., Goujon, C., Dubuisson, J., Pascale, S., Scarselli, E., Cortese, R., Nicosia, A. and Cosset, F. L. (2003). Cell entry of hepatitis C virus requires a set of co-receptors that include the CD81 tetraspanin and the SR-B1 scavenger receptor. *J. Biol. Chem.* 278, 41624-41630.
69. Bartosch, B. and Cosset, F. L. (2006). Cell entry of hepatitis C virus. *Virology*. 348, 1-12.
70. Cocquerel, L., Voisset, C. and Dubuisson, J. (2006). Hepatitis C virus entry: potential receptors and their biological functions. *J. Gen. Virol.* 87, 1075-1084.
71. VanCompernelle SE, Wiznycia AV, Rush JR, Dhanasekaran M, Baures PW, Todd SC (2003) Small molecule inhibition of hepatitis C virus E2 binding to CD81. *Virology* 314: 371-80.
72. Fofana, I., Jilg, N., Chung, R. T. and Baumert, T. F. (2014). Entry inhibitors and future treatment of hepatitis C. *Antiviral Res.* 104, 136-142.
73. Petracca R, Falugi F, Galli G, Norais N, Rosa D, Campagnoli S, Burgio V, Di Stasio E, Giardina B, Houghton M, Abrignani S, Grandi G (2000) Structure-function analysis of hepatitis C virus envelope-CD81 binding. *J Virol* 74: 4824-30.
74. Meuleman, P., Hesselgesser, J., Paulson, M., Vanwolleghem, T., Desombere, I., Reiser, H. and Leroux-Roels, G. (2008). Anti-CD81 antibodies can prevent a hepatitis C virus infection *in vivo*. *Hepatology*. 48, 1761-1768.
75. Seng-Lai T, Arnim P, Yuguang S, Nahum S. (2002). Hepatitis C therapeutics: current status and emerging strategies. *Nature Reviews Drug Discovery* 1, 867-881.

11. Tables

Table 1. HCV proteins. This table shows the 10 HCV proteins, their molecular mass and their function.

Protein	Molecular mass KDa	Function
Core	21	RNA binding, nucleocapsid
E1	31-35	Envelope glycoprotein, associate with E2
E2	70	Envelope glycoprotein, associate with E2, receptor binding
P7	7	Ion channel
NS2	21	Component of NS2-3 proteinase
NS3	69	N-terminal proteinase domain/ C-terminal NTPase/Helicase domain
NS4A	6	NS3/4A proteinase co-factor
NS4B	27	Induces membrane alterations
NS5A	56-58	Phosphoprotein
NS5B	68	RNA dependent RNA polymerase

Table 2. HCV protease inhibitors in different clinical phases (50).

Protease Inhibitors	Company	Clinical Phase	Chemical Structure	Required Dose per day	Genetic Barrier	IC50 (nM)	Genotypes
<i>Boceprevir</i>	VICTRELIS™ Merck	FDA/EMA approved	Linear	800 mg TID	Low	600	1,2
<i>Telaprevir</i>	INCIVO™ Janssen	FDA/EMA approved	Linear	1125 mg BID	Low	300	1,2
<i>Simeprevir</i>	OLYSIO™	FDA/EMA approved	Macrocyclic	150 mg QD	Moderate	1-10	1,2,4,5,6
<i>Faldaprevir</i>	Boehringer Ingelheim	3	Macrocyclic	120 mg or 240mg QD	Moderate	1-10	1,2
<i>BI-201 135</i>		2		240 mg QD			
<i>Danoprevir</i>	Roche	2	Macrocyclic	200/r mg BID 100/r mg BID 50/r mg BID 100/r mg BID	Moderate	1-10	1,2,4
<i>MK5172</i>	Merck	2	Macrocyclic	100 mg QD	High	1-10	1,2,4,5,6
<i>Sovaprevir</i> <i>ACH-1625</i>	Achillion	2	Macrocyclic	200-800mg QD	Moderate	1-10	1
<i>Asunaprevir</i>	BMS	2	Macrocyclic	200-600mg BID 600mg BID + Daclatasvir 60mg QD	Moderate	1-10	1,4

Table 3. HCV Entry Inhibitors. This table shows the different host targets, the corresponding examples of inhibitors and the stage of development of those inhibitors (72).

Target	Examples of compounds	Stage of development
CD81 (Cluster of Differentiation 81)	Anti-CD81 mAbs	Mouse model
	Imidazole based compounds	Cell culture
SR-BI (Scavenger Receptor – BI)	ITX5061	Phase I/IIa
	Anti-SR-BI mAbs	Mouse model
	Serum amyloid A	Cell culture
CLDN1 (Claudin – 1)	Anti-CLDN1 mAbs	Cell culture
EGFR (Epidermal Growth Factor Receptor)	Erlotinib	Phase I/IIa
NPC1L1 (Neimann-Pick C1-Like 1 gene)	Ezetimibe	Mouse model
TfR1 (Transferrin Receptor Protein 1)	Anti-TfR1 mAbs	Cell culture
Post-attachment	Flavonoids	Cell culture
Fusion/internalization	PS-ON (Phosphorothioate oligonucleotide)	Mouse model
		Cell culture
	Arbidol	Cell culture

	Chloroquine	Cell culture
	Silymarin	

12. Figures

Figure 1. HCV entry mechanism. LDL-R and Glycosaminoglycan plays the initial role by mediating the attachment of HCV associated with lipoproteins to CD81 which then interact with SR-BI triggering HCV movement to the tight junction proteins; claudin-1 (CLDN1) and occluding (OCLN) followed by clathrin mediated endocytosis (11, 12).

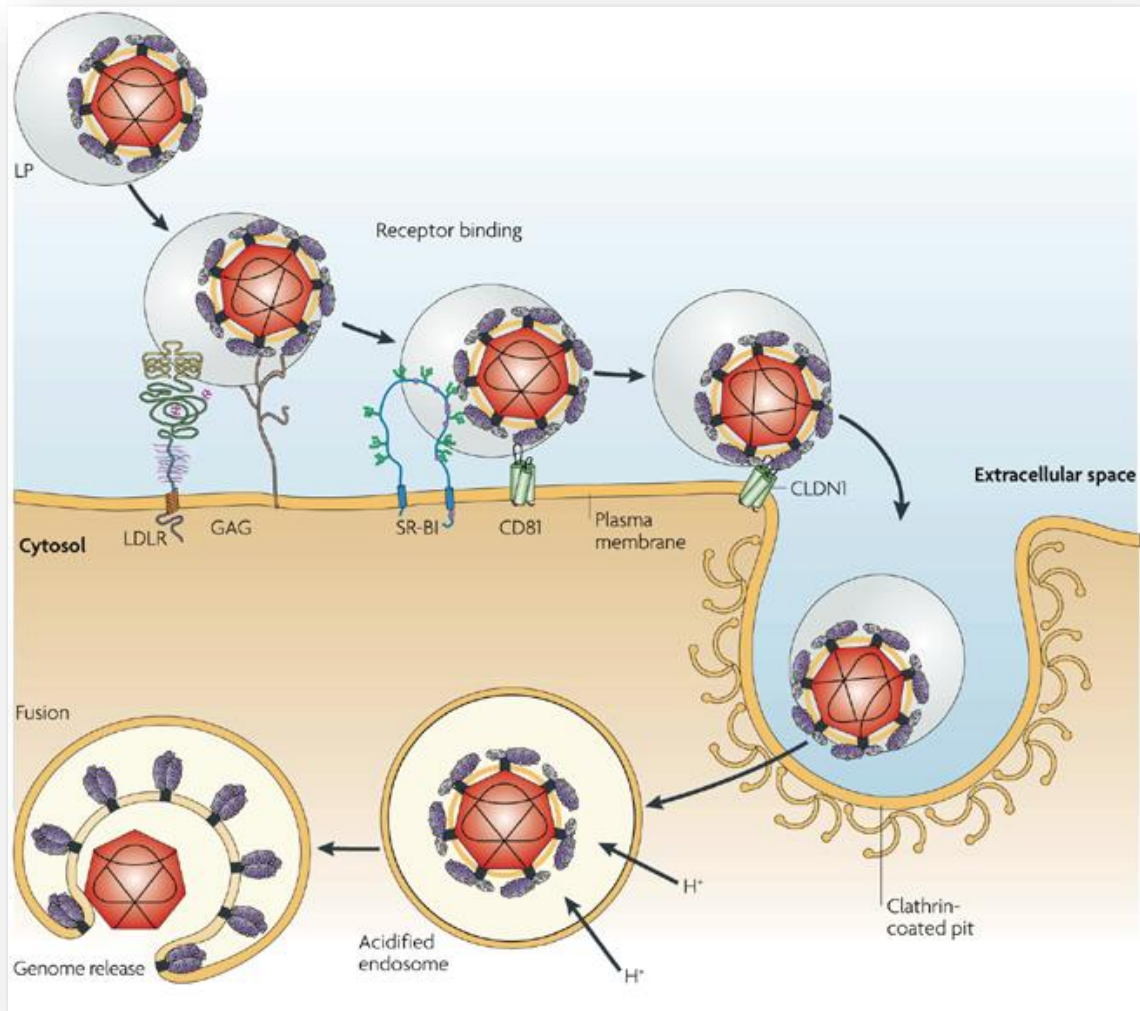


Figure 2. Model for HCV virion assembly. Based on current evidence, assembly initiates on the cytosolic side of the ER membrane (A), and complete maturation occurs in the ER lumen (B) prior to release from the cell (24). (RC: Replicase Complex, LVP: Lipo-Viro-Particles, MTP: Microsomal Triglyceride Transfer Protein)

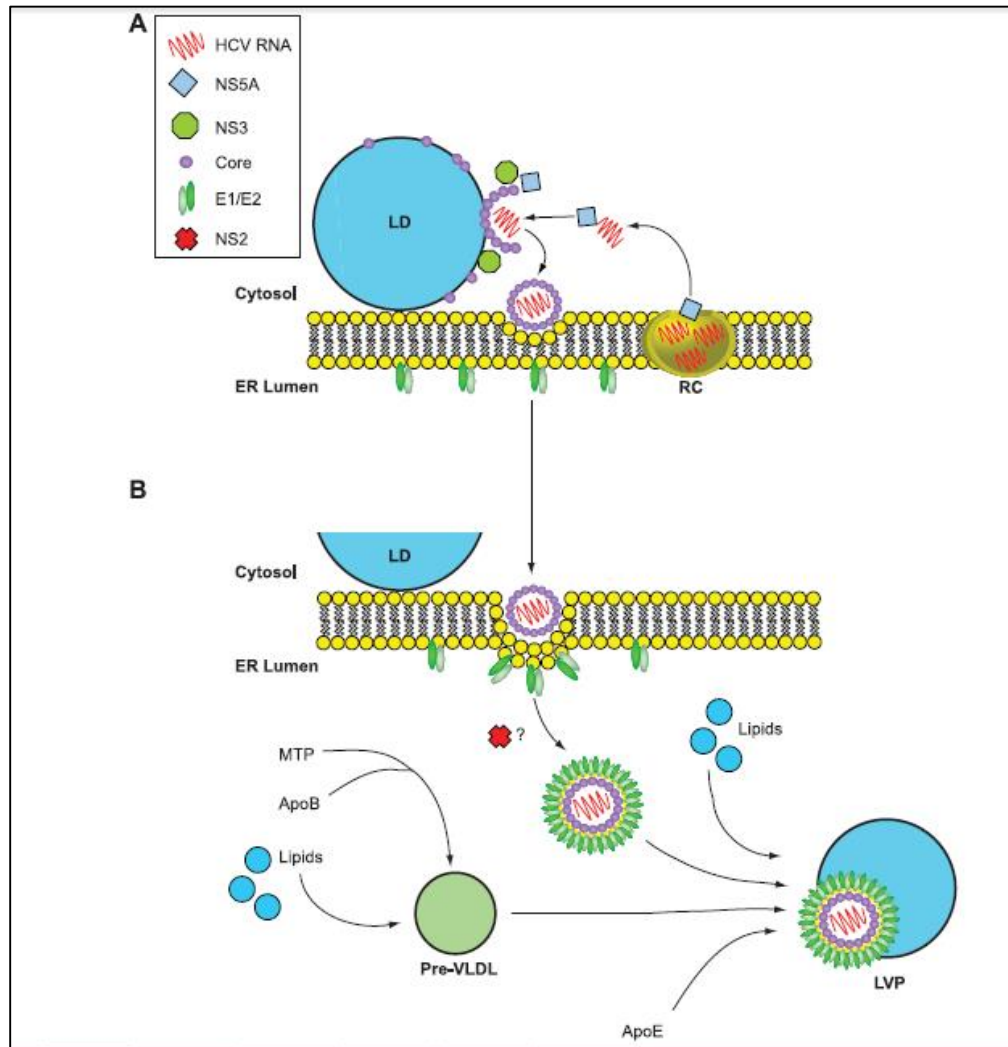


Figure 3. Worldwide HCV prevalence

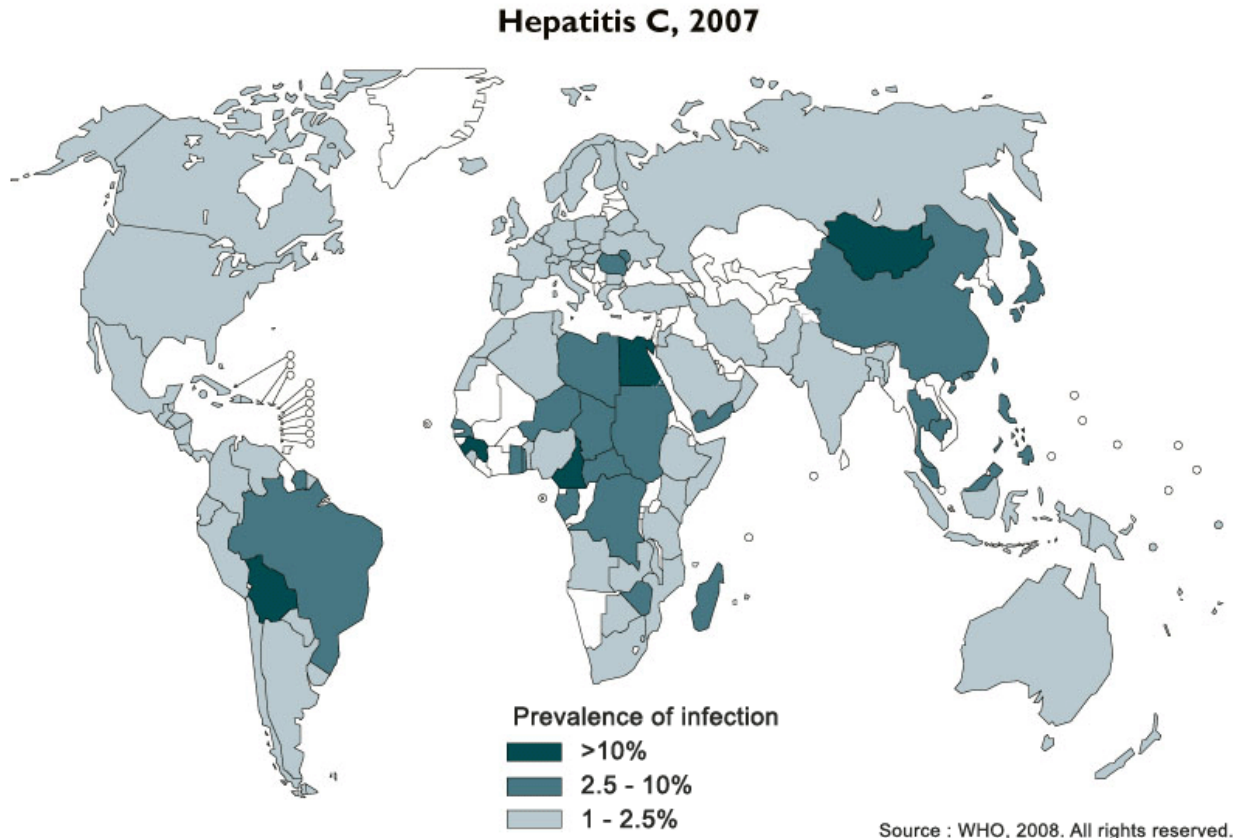


Figure 4. HCV Genome. HCV genome is a 9.3 kb polyprotein divided into 3 structural proteins and 7 non-structural proteins (75).

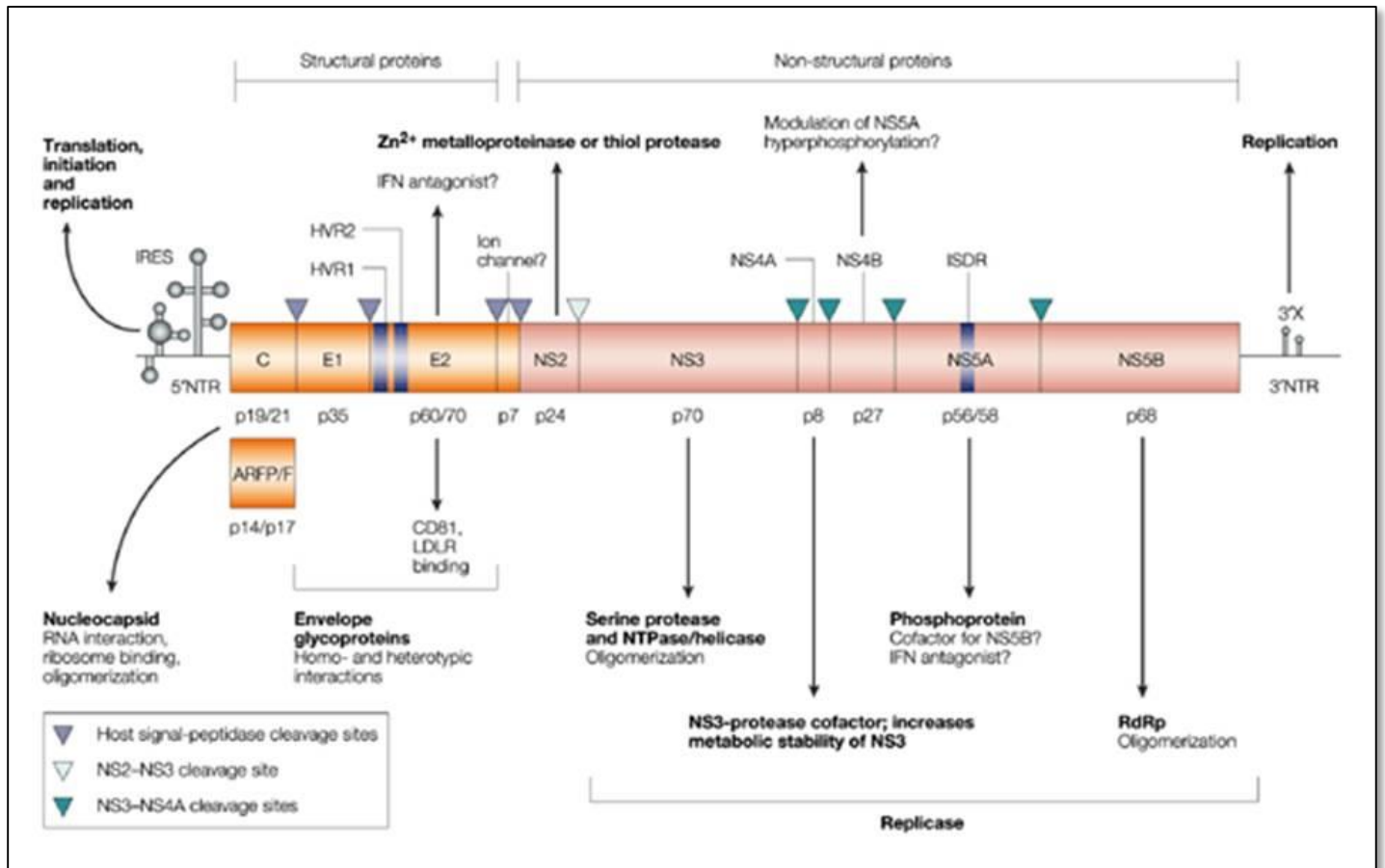


Figure 5. The E2c crystal structure resolved by Kong *et al.* This figure shows the structure of the HCV envelope glycoprotein E2 core bound to the neutralizing antibody AR3C (PDB ID: 4MWF) (38).

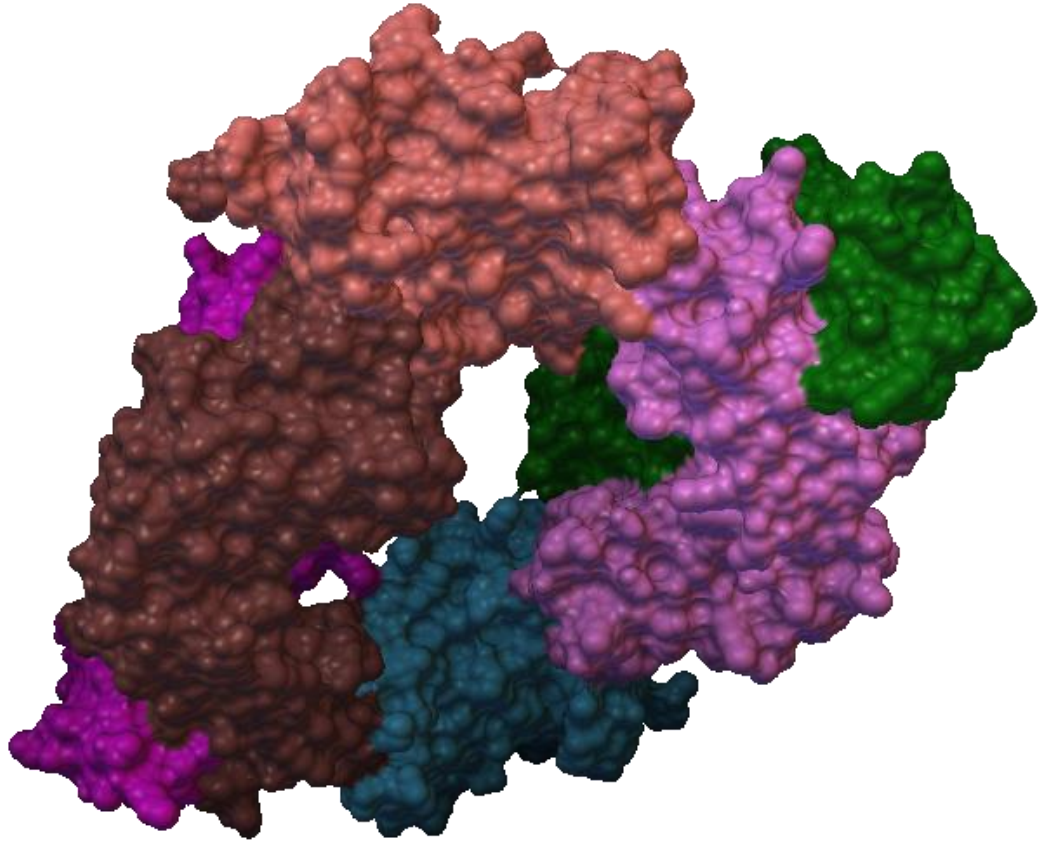


Figure 6. The crystal structure of HCV E2 resolved by Khan *et al.* This figure shows the structure of the E2 core domain in complex with an antigen-binding fragment (Fab) at 2.4 Å resolution. The E2 core has a compact, globular domain structure, consisting mostly of β -strands and random coil with two small α -helices. The strands are arranged in two, perpendicular sheets (A and B), which are held together by an extensive hydrophobic core and disulphide bonds. (PDB ID: 4NX3)

(39)

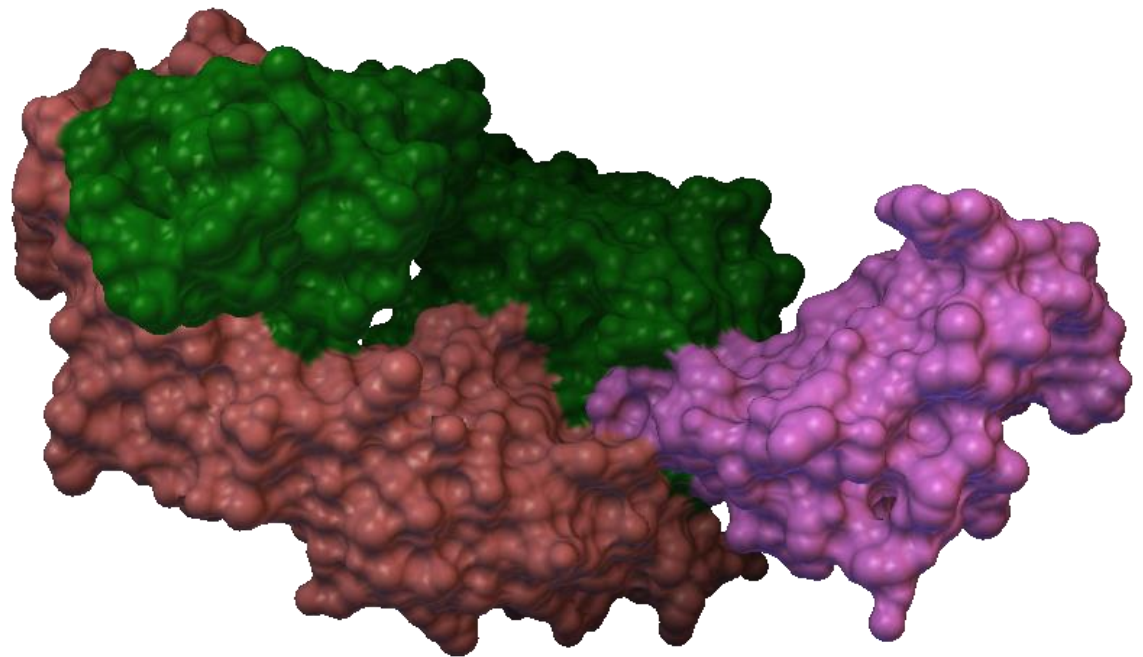


Figure 7. Role of CD81 in disrupting HCV-induced immune functions (46). During HCV infection, HCV E2 glycoprotein interacts with CD81 at the cell surface of different immune cells which in turn leads to B cell activation, aggregation, and increased survival. On the other hand, the interaction of HCV E2 with CD81 on T leads to the activation of protein tyrosine kinase Lck. This interaction leads to the production of RANTES, a ligand for the CCR5 receptor and which upon ligation to CCR5 at the surface of pDCs leads to its internalization prevents its migration from infected tissues to lymph nodes. The direct interaction of HCV E2 with CD81 at the surface of pDCs leads to a decreased secretion of IFN in addition to impaired maturation, proliferation and survival of these cells. HCV infection causes an impairment of NK cell functions. The role of CD81 in NK cell-altered function could lead to a decreased IL-12- dependent IFN- γ secretion and an increased secretion of IL-8.

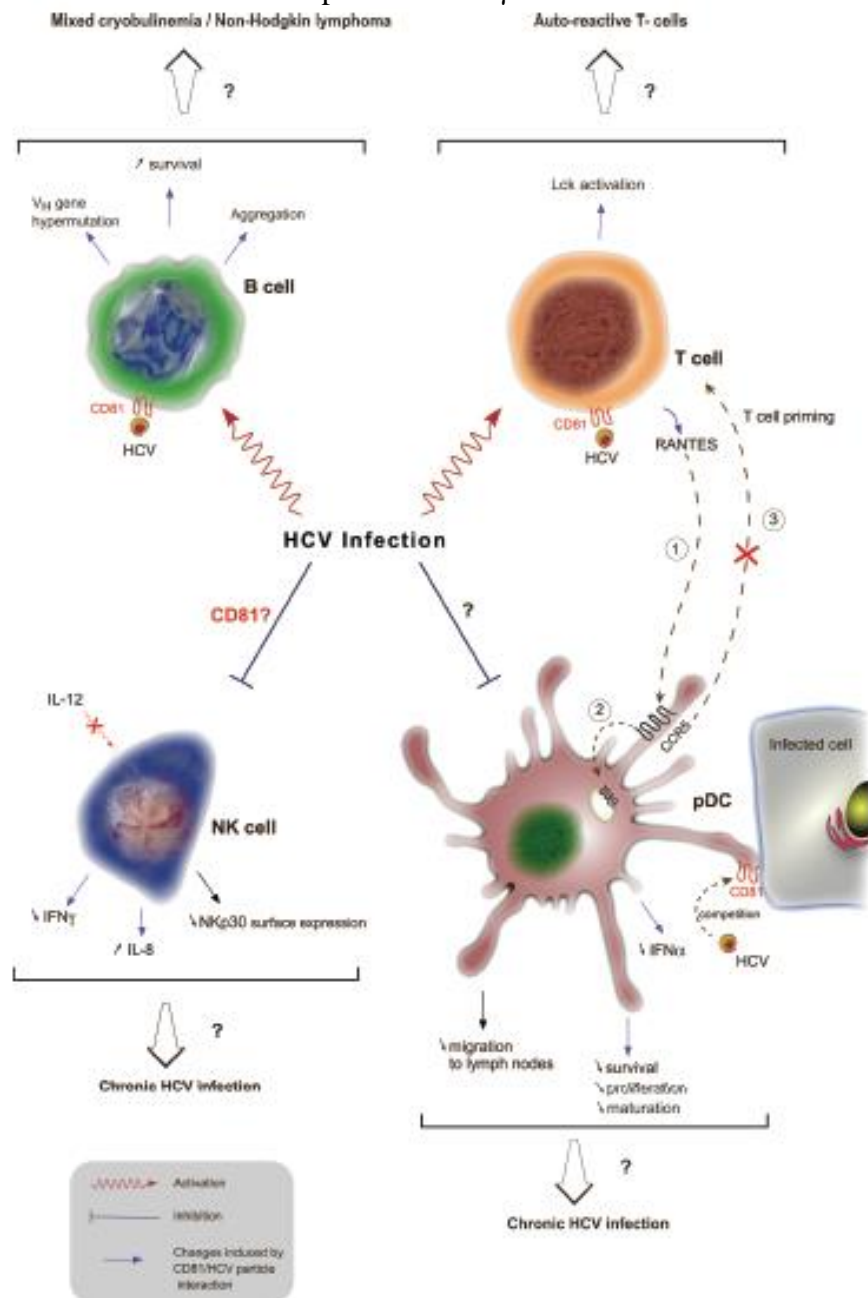
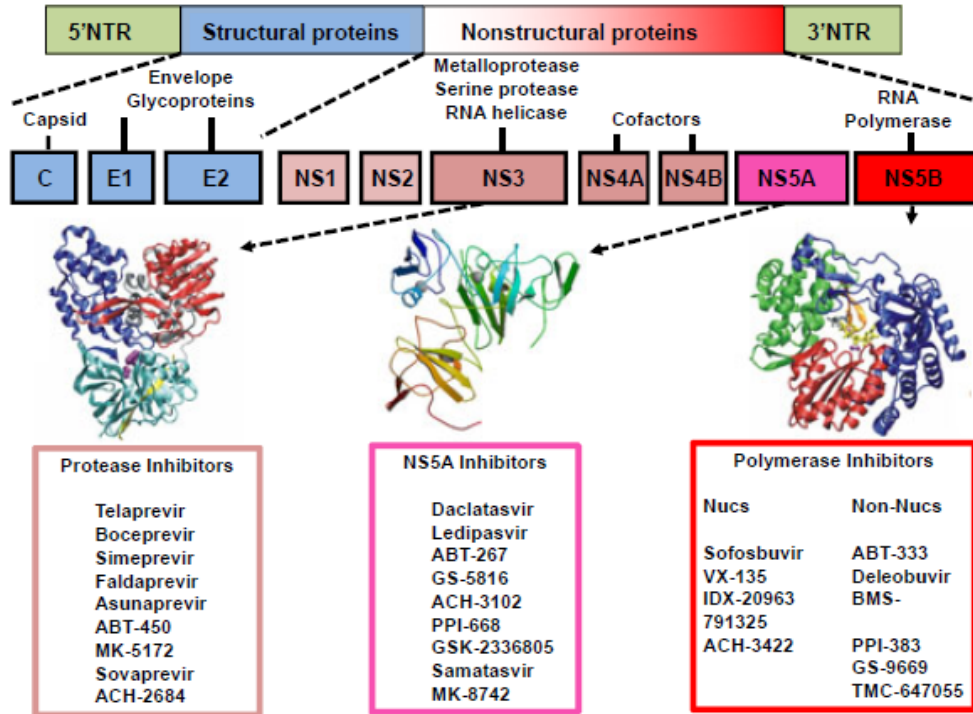


Figure 8. Different classes of HCV drugs targeting different parts of the HCV genome. The main 3 classes of drugs that are currently studied are protease inhibitors, NS5A inhibitors and polymerase inhibitors which are either nucleoside or non-nucleoside polymerase inhibitors (50).



B) Malaria

1. History

Malaria is considered one of the major parasitic diseases which is transmitted by female *Anopheles* mosquito. It is caused by protozoan parasites belonging to the genus *Plasmodium* (1). Malaria in human is caused by four types of *Plasmodium sp.* which are: *Plasmodium malariae*, *Plasmodium ovale*, *Plasmodium vivax* and *Plasmodium falciparum*. *Plasmodium falciparum* and *Plasmodium vivax* are the two major species involved in the global malaria burden (1). Malaria infection starts by the invasion of the parasite into the hepatocytes (liver stage) followed by invading the red blood cells of the host (blood stage) (figure 1).

Malaria was productively studied since the 1880. *Plasmodium falciparum* has existed for 50,000–100,000 years but the population size of the parasite increased only since 10,000 years ago (2). The causative agent of malaria was discovered by Charles Laveran (3). He discovered it through examining the blood smears of infected patients in 1884 (3, 4). In 1898, Scottish physician-Sir Ronald Ross, was able to reveal the complete life cycle of malaria and proved the vector was a mosquito. He was awarded the Nobel Prize based on his work in 1902 (5). Malaria antigen was detected in skin and lung samples of Egyptian mummies in the years 3200 and 1340 BC (6,7).

2. Diagnosis

Malaria is diagnosed mainly by the microscopic examination of the blood film to view sporozoites and merozoites. Other techniques used are antigen-based rapid diagnostic tests (RDT) (8, 9). The immune chromatographic tests (ICT) for the detection of malaria are based on the same concepts of Enzyme Linked Immunosorbent Assay (ELISA) and detect antibodies in blood or other body fluids (10, 11).

Malaria doesn't yield characteristic symptoms when infecting a person. Accordingly, when there is a suspicion of infection, recent travel history, fever, low number of platelets, enlarged spleen, and higher-than-normal levels of bilirubin in the blood combined with a normal level of white blood cells might help decide whether there is a risk of malaria infection or not (12).

3. Current treatment

The treatment of Malaria depends on the type and severity of the disease. Artemisinin-combination therapy (ACT) is considered the first-line therapy for uncomplicated *Plasmodium falciparum* malaria worldwide (13). To date, there is not enough data to help observe the emerging problem of ACT-resistant strains, hence there is a need to study the dynamics of parasite clearance in people treated with ACT in order to be able to confirm the emergence of artemisinin resistance (14).

4. CD81: Malaria host receptor

The life cycle of Malaria starts through injection of sporozoites into the host's skin by female *Anopheles* mosquitoes. The sporozoites are motile due to possessing actomyosin motor machinery which help them move in the blood stream, entering the liver in and transform to liver stages in the hepatocytes. The motile sporozoites enter the blood stream and, upon reaching the liver, transform into liver stages (14). They invade hepatic parenchymal cells and produce exo-erythrocytic forms (EEFs) through a period of asexual reproduction followed by infecting the erythrocytes causing the malaria disease (15, 16).

Silvie *et al.* (16) showed that CD81 is involved in the hepatic development of *Plasmodium falciparum* and *Plasmodium yoelii* in the liver cells and in CD81-deficient mouse hepatocytes, *Plasmodium yoelii* can't initiate the infection process. Additionally, they were able to show that antibodies against human and mouse CD81 blocked the infection with *Plasmodium falciparum* and *Plasmodium yoelii* respectively. They were able to prove that CD81 is involved in the entry stage of sporozoites into the liver though being linked to the formation of parasitophorous vacuole which one of the major steps needed for the sporozoites to transform to EEFs.

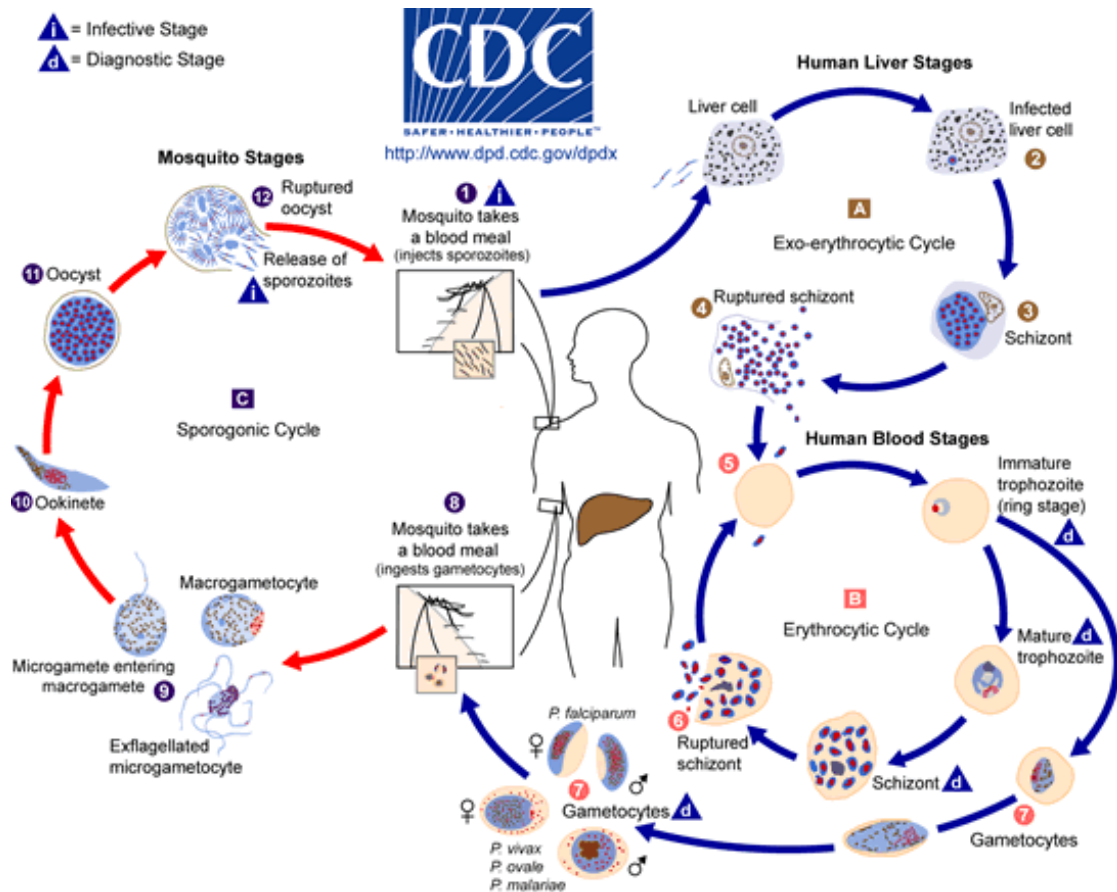
5. References

1. Druilhe P, Daubersies P, Patarapotikul J, Gentil C, Chene L, Chongsuphajaisiddhi T, *et al.* A primary malarial infection is composed of a very wide range of genetically diverse but related parasites. *J Clin Invest* 1998;101:2008–16.
2. Harper K, Armelagos G (2011). "The changing disease-scape in the third epidemiological transition". *International Journal of Environmental Research and Public Health* 7 (2): 675–97
3. Jarcho S. Malaria and murder (Joseph Jones, 1878). *Bull N Y Acad Med* 1968;44: 759–60.
4. Jarcho S. Some observations on disease in prehistoric North America. *Bull Hist Med* 1964;38:1–19.
5. "The Nobel Prize in Physiology or Medicine 1902: Ronald Ross". The Nobel Foundation. Retrieved 2012-05-14
6. Hawass Z, Gad YZ, Ismail S, Khairat R, Fathalla D, Hasan N, *et al.* Ancestry and pathology in King Tutankhamun's family. *JAMA* 2010;303:638–47.
7. Miller RL, Ikram S, Armelagos GJ, Walker R, Harer WB, Shiff CJ, *et al.* Diagnosis of *Plasmodium falciparum* infections in mummies using the rapid manual ParaSight-F test. *Trans R Soc Trop Med Hyg* 1994;88:31–2.
8. Kattenberg JH, Ochodo EA, Boer KR, Schallig HD, Mens PF, Leeflang MM (2011). "Systematic review and meta-analysis: Rapid diagnostic tests versus placental histology, microscopy and PCR for malaria in pregnant women". *Malaria Journal* 10: 321
9. Abba K, Deeks JJ, Olliaro P, Naing CM, Jackson SM, Takwoingi Y, Donegan S, Garner P (2011). "Rapid diagnostic tests for diagnosing uncomplicated *P. falciparum* malaria in endemic countries". In Abba, Katharine. *Cochrane Database of Systematic Reviews* (7): CD008122.
10. Bell D, Wongsrichanalai C, Barnwell JW. Ensuring quality and access for malaria diagnosis: how can it be achieved? *Nat Rev Microbiol* 2006;4:682–95.
11. Murray CK, Gasser Jr RA, Magill AJ, Miller RS. Update on rapid diagnostic testing for malaria. *Clin Microbiol Rev* 2008;21:97–110.
12. Nadjm B, Behrens RH (2012). "Malaria: An update for physicians". *Infectious Disease Clinics of North America* 26 (2): 243–59
13. Zani, B., Gathu, M., Donegan, S., Olliaro, P. L. and Sinclair, D. (2014). Dihydroartemisinin-piperaquine for treating uncomplicated *Plasmodium falciparum* malaria. *Cochrane Database Syst. Rev.* 1, CD010927.
14. Risco-Castillo, V., Son, O., Franetich, J. F., Rubinstein, E., Mazier, D. and Silvie, O. (2013). *Plasmodium* sporozoite entry pathways during malaria liver infection. *Biol. Aujourd'hui.* 207, 219-229.
15. Leitao, R. and Rodriguez, A. (2010). Inhibition of *Plasmodium* sporozoites infection by targeting the host cell. *Exp. Parasitol.* 126, 273-277.
16. Silvie, O., Rubinstein, E., Franetich, J. F., Prenant, M., Belnoue, E., Renia, L., Hannoun, L., Eling, W., Levy, S., Boucheix, C. and Mazier, D. (2003). Hepatocyte

CD81 is required for *Plasmodium falciparum* and *Plasmodium yoelii* sporozoite infectivity. Nat. Med. 9, 93-96.

6. Figures

Figure 1. Malaria Life cycle. The life cycle of malaria is divided into 3 significant stages; The Exo-erythrocytic stage /Liver stage (A), the erythrocytic stage/Blood stage (B) and the sporogonic stage which takes place in the gut of the female Anopheles mosquito followed by its salivary gland (Center of Disease Control and Prevention)



C) Summary and Conclusion

HCV and Malaria are two challenging infectious diseases due to the genomic diversity and not having the full picture of the mechanism of entry into the host cells in case of the former and the emerging resistant strains against the currently available medications for the later. Several research groups identified important residues in CD81-LEL involved in the infection of both HCV and malaria and their internalization into hepatocytes (16, 43, 44). Targeting the identified residues would help develop promising drugs for treating HCV and malaria.

Additionally, HCV E2 core crystal structure was recently resolved (PDB ID: 4MWF) (figure 5) (38). This would facilitate creating a reliable homology model for the whole HCV E2 protein that could be used to identify drug leads that blocks HCV E2: Host factors interaction.

Targeting HCV to develop drugs is advantageous because the side effects that might be caused when using drugs that target host receptors are unpredictable. On the other hand, targeting the host receptors to develop new classes of drugs is better when dealing with the continuous mutation of HCV and the resistance it develops against direct acting antivirals.

Different approaches could be used to generate promising drugs against HCV and malaria. Generating selective high affinity ligands is one of them where it is mainly linking to ligands that were found to bind to the target protein together to increase their

binding affinity and it might reach a nanomolar binding affinity. Another one is to identify a drug lead, optimize its chemical QSAR to be more effective with less side effects. Also, finding different analogues to the drug lead, testing them and creating a library would be a potential tool in the drug discovery process.

CHAPTER III

IDENTIFICATION OF LIGANDS THAT TARGET HCV E2-BINDING SITE ON CD81-LARGE EXTRACELLULAR LOOP

This chapter was published: Olaby, R. A., Azzazy, H. M., Harris, R., Chromy, B., Vielmetter, J. and Balhorn, R. (2013). Identification of ligands that target the HCV-E2 binding site on CD81. *J. Comput. Aided Mol. Des.* **27**, 337-346.

Objectives

The purpose of this study was to screen a library of 10,000 compounds to identify virtual screening hits that binds to significant sites on CD81-LEL. Additionally, it aimed at testing the top virtual screening hits experimentally using Lab on a chip techniques to help identify possible ligand candidates to be incorporated into a SHAL prototype. The computational biology work was done at Scripps Research Institute, Dr. Arthur Olson's Molecular Graphics Laboratory, the dual polarization interferometry binding assays were done at UC Davis at Dr. Brett Chromy's lab, and the surface plasmon resonance binding assays were done at Caltech at Dr. Jost Vielmetter's protein expression center.

1. Introduction

The World Health Organization has estimated that approximately 3% of the world population has been infected with Hepatitis C virus (HCV) and that more than 170 million of these individuals are at risk for developing liver cirrhosis or cancer (1). The lack of effective treatment or prophylactic vaccines makes HCV a serious public health problem. The virus is a blood borne pathogen that is transmitted mainly through transfusions and hemodialysis. During HCV replication, the post-translational processing and cleavage of the virus polyprotein produces ten structural and non-structural proteins. The crystal

structures that have been determined for a number of these proteins are being used to facilitate both drug and vaccine development (2-9).

Several cell surface receptors have been suggested to play a role in HCV entry into hepatocytes (10). These include LDL-R, heparan sulfate (11), scavenger receptor class BI (SR-BI) and CD81 (12,13). Pileri *et al.* was the first to identify CD81, a 26 kDa protein that belongs to the tetraspanins super family, as an important HCV receptor (14). While this protein mediates the invasion of hepatocytes by HCV, it is also widely expressed in both lymphoid and non-lymphoid tissues. CD81 contains six structural domains, four of which are trans membrane domains and two are hydrophilic extracellular domains that make up the large and small extracellular loops (15).

One reason CD81 has become such an important target for drug development is because the large extracellular loop of CD81 (CD81-LEL) has been shown to bind to the HCV E2 glycoprotein (16-19). Zhang *et al.* discovered that CD81-LEL is also important for efficient replication of the HCV genome (18). In addition, the E2:CD81-LEL interaction has been reported to induce several immuno-modulatory effects, including a co-stimulatory signal in naive and antigen-experienced T cells *in vitro* that leads to production of the pro-inflammatory cytokine α -interferon. This suggests that the E2:CD81-LEL interaction may play a role in T-cell-mediated liver inflammation and may contribute to liver damage. The interaction of these two proteins also appears to down regulate T-cell receptors and suppress the activity of natural killer cells (18).

CD81's participation in cell invasion and its contribution to liver damage make it an important target for new anti-HCV therapeutics. Some of the first inhibitors designed to block the E2:CD81-LEL interaction were CD81 mimics developed by VanCompernelle *et al.* (20). Small molecules were designed to mimic the solvent exposed hydrophobic ridge of helix D in the CD81-LEL domain and were found to bind HCV E2 reversibly and to competitively block the binding of E2 to CD81 (20). This was the first direct demonstration that CD81 is an important receptor in HCV entry (20). In addition, the mutational studies conducted by Higginbottom *et al.* (17) and Drummer *et al.* (19) identified the key amino acid residues that contribute to the E2:CD81-LEL interaction.

Kitadokoro *et al.* determined the 3D structure of CD81-LEL using X-ray crystallography, and two different crystal forms of CD81-LEL (PDB codes 1G8Q and 1IV5) were reported (21,22). In the 1G8Q structure the C and D helices form a cleft-like motif within the E2 binding site, a large cavity considered to be an excellent target site for inhibitor development. The 1IV5 conformation, in contrast, was considered to be a closed form of the CD81 structure in which this cleft is absent. Ligands binding to the closed conformation would involve interactions with 1IV5 in more shallow surface exposed sites than those present on 1G8Q (22). Molecular dynamics studies performed by Neugebauer *et al.* (23) have been used to suggest that the 1IV5 structure may be the physiologically relevant conformation. This conclusion has been attributed in part to the closure of the cleft in 1G8Q that occurred during a 50 picosecond molecular dynamic simulation. The 1G8Q conformation with the open cleft was also considered to be less stable because more

amino acid residues were found to be outside the favored energy region of the Ramachandran plot. Further analysis of the two structures suggested that the cleft observed in the open 1G8Q conformation might represent a distortion in the structure of the protein induced by crystal packing. In the closed 1IV5 structure, two of the four alpha helices (C and D) in CD81-LEL were observed to form a helix bundle with the two other helices (A and B) of an adjacent molecule in the lattice. In the 1G8Q form, a different interaction was observed between helices that appeared to distort the structure of the protein and create the cleft (22).

The discovery of these two distinct crystal forms of the CD81-LEL protein with very different surface structures in and around the E2 binding site has complicated the process of inhibitor development. The “open” form has multiple cavities surrounding the key amino acids, while the surface of the “closed” form has many fewer and shallower sites where ligands might bind. While it has been suggested that the closed form may be more stable than the open form, Neugebauer *et al.* (23) also indicated that the C and D helices exhibit a certain degree of flexibility that might make it possible to identify small molecules that fit inside the cleft between these two helices and block the interaction between CD81 and E2.

In an effort to test that possibility, we have used AutoDock and AutoLigand to screen a library of 10,000 small molecules *in silico* and identify ligands predicted to bind to two sites on the open conformation of CD81-LEL, the large cleft between the C and D helices and a smaller cavity located nearby. Both cavities are located within the E2

binding site and in close proximity to five of the amino acid residues reported to contact E2. Experimental methods have been used to test the best virtual screening hits for binding to a recombinant form of CD81-LEL, and a set of new small molecule drug candidates have been identified that bind to the protein. One of these compounds has been found to block E2 binding to CD81-LEL. Fragment-based extension methods will be used to create second-generation lead compounds from a number of these molecules. Others will be linked together to create selective high affinity ligands (SHALs) (24) that target the E2 binding site on CD81-LEL and block HCV invasion.

2. Materials and Methods

2.1. Preparation of CD81-LEL Structure and Prediction of Binding Sites

The AutoDock suite of programs developed by Dr. Arthur Olson's molecular graphics laboratory at the Scripps Research Institute was used to analyze the large extracellular domain of our target protein CD81, prepare surface grid maps, and dock a library of small molecules into cavities located in the vicinity of amino acid residues known to participate in E2 binding (25-29).

The coordinates for the crystal structure of the open conformation of CD81-LEL (PDB ID: 1G8Q) were obtained from the Protein Data Bank (PDB). AutoDock Tools (ADT) 1.5.6 (25-28) was used to delete water molecules, add polar hydrogens, assign Gasteiger charges, and create grid bounding boxes with a 1 Å spacing for use with AutoLigand and a 0.375 Å spacing for use with AutoDock 4.2. AutoGrid 4.2 was used to pre-calculate grid maps of interaction energies for various atom types and create the map

files that were used by AutoLigand to predict the CD81-LEL binding sites and by AutoDock for docking. AutoLigand was then used to rapidly scan the protein for high affinity binding pockets and identify the optimal volume, shape, and best atom types for each binding site.

The CD81-LEL protein was scanned by AutoLigand using fill sizes from 10 to 210 fill points. During this process, the structure (amino acid residues and α -carbon backbone) was kept rigid. The constructed grid box enclosed the entire protein with dimensions of 40 Å by 18 Å by 38 Å and was centered on 3.144, 34.966, and 15.812 in the protein frame of reference. Five potential ligand binding sites were identified on the open CD81-LEL structure (PDB code 1G8Q). Two sites located adjacent to amino acid residues critical for E2 binding were selected for docking.

2.2. Virtual Screening

AutoDock 4.2 (25-28) was used to perform virtual screening runs using a subset of the ZINC small molecule database containing 10,000 molecules taken from the National Cancer Institute-Diversity Set II (NCI_DSII), Sigma, and Asinex libraries. The parameters were set at 100 for the number of genetic algorithm (GA) runs, 150 as the population size, and a maximum number of generations of 25000. The Lamarckian genetic algorithm in AutoDock was used to perform the docking experiments (30). Docking results were sorted by the lowest binding energy in addition to specific ligand selection criteria that would facilitate the design and synthesis of the best SHALs. The virtual

screening runs were performed using the National Biomedical Computation Resources (NBCR) computer cluster (31). Vision (32) was used to construct the computational workflows that were used for virtual screening on the NBCR cluster. The list of small molecules predicted to bind to each site (~ 350 compounds) was ranked according to their predicted free energy of binding, and the molecules with the lowest free energies were further screened manually to identify ~120 of the best ligand candidates for experimental testing.

2.3. Ligand Evaluation

Several criteria were considered as we examined the structures of each of these ~120 small molecules and selected a subset for subsequent experimental testing and for use in the design of second-generation lead compounds and SHALs. All the molecules selected could be purchased from chemical suppliers or obtained from the Developmental Therapeutics Program at NCI. During the initial examination of the list of ligands predicted to bind to each site by AutoDock, only molecules containing a free carboxyl group or an amino group (or one of each) were selected. In the most highly ranked cases, these amino or carboxyl groups were not buried in a cavity nor did they interact with the protein surface. They were exposed to solvent and were predicted by AutoDock to bind to the protein with the functional group pointed in the general direction of the second ligand binding site. Such molecules could be easily linked together through their amino or carboxyl groups to create SHALs (24). Preference was given to ligands that were predicted to form multiple contacts with atoms or amino acid residues in or around the

perimeter of the targeted cavities. Molecules that were highly hydrophobic, highly charged, known to be toxic, exist in more than one form (such as enol-keto forms), or contained disulfide bonds were avoided. After manually filtering the ligand sets to remove the molecules that did not meet these criteria, the predicted binding energy was used to identify the top hits. Thirteen molecules predicted to bind to Site 1 were selected from this group for experimental testing and 23 molecules were selected for Site 2. Small amounts (10mg) of these 36 compounds were then obtained from the National Cancer Institute (Diversity Set II small molecule library) and tested experimentally for binding to the CD81-LEL protein.

2.4. Surface Plasmon Resonance

SPR analysis was performed using a Biacore T200 workstation (GE Healthcare, NJ, USA). A recombinant form of the CD81-LEL protein with a GST tag (generously provided by Dr. Shoshana Levy, Stanford University) was used to determine, using well established experimental technique, if the ligands bound to the protein. Briefly, 10 μ M CD81-LEL-GST diluted into 10 mM Na-Acetate buffer pH 4.5 was immobilized for 15 min at a flow speed of 5 μ l/min onto a CM5 sensor chip using amine-coupling (EDC-NHS). Approximately 20,000 RU of protein were immobilized on the chip. The ligands were prepared as 600 μ M solutions in PBS-0.05% Tween-80 (the running buffer) and they were introduced to the protein using a pre-programmed 3 minute association and 1 minute dissociation interval.

The binding affinity of selected ligands were estimated using data collected from a series of SPR binding experiments conducted at different ligand concentrations. To obtain the kinetic and affinity data needed to estimate the K_d , the original ligand sample was diluted serially with running buffer to produce seven different ligand concentrations: 1024 μM , 516 μM , 256 μM , 128 μM , 64 μM , 32 μM and 0 μM . Data were fitted using a monovalent binding model.

2.5. Dual Polarization Interferometry (DPI) Analysis

DPI analyses were performed using an AnaLight 4D workstation (Farfield Group, Manchester UK). The recombinant CD81-LEL was immobilized onto a Thiol AnaChip using Sulfo-GMBS as a cross-linker in PBS running buffer. Non-specific sites were blocked with digested casein. TRIS was used to cap the cross-linker, blocking any additional amines from covalently binding to the cross-linker on the chip surface. Ligands were prepared as 20 mM stock solutions in dimethylsulfoxide (DMSO). Each ligand was diluted to a final concentration of 500 μM in PBS just prior to injection (final DMSO concentration was 2.5%). PBS and DMSO mixed in the same ratio were used as a blank. Data collection and analysis were performed using the AnaLight Resolver.

A subset of the ligands identified to bind to CD81-LEL were also tested to determine if they might block the HCV E2 protein from binding to CD81-LEL using DPI. In these experiments, a recombinant form of the CD81-LEL protein was immobilized on the chip and unreacted cross-linker was blocked as described above. Recombinant HCV

E2 protein (Immune Technology Corp, New York, NY) was then injected to determine the magnitude of the binding response when E2 bound to CD81-LEL in the absence of the ligand. To evaluate the effect of a ligand on E2 binding to CD81-LEL, the same experiment was repeated except that the E2 protein was premixed with the ligand at a final ligand concentration of 500 μM . If the ligand inhibits E2 binding to CD81-LEL when the mix of E2 and the ligand are added to the chip, the DPI binding response in the presence of the ligand should be less than the response in the absence of the ligand. If a reduction in E2 binding is observed by DPI, the magnitude of the inhibition can be calculated using the binding responses for the ligand, the E2 protein and a mixture of the E2 protein and the ligand.

3. Results and Discussion

3.1. Target Regions on CD81-LEL

In this study we used the crystal structure of the open CD81-LEL conformation as the target for the virtual screening runs performed using AutoDock to identify small molecule ligands predicted to bind to cavities that encompass or are located near known E2 contact residues. Based on mutation studies, Higginbottom *et al.* (17) identified four residues that were considered to be essential for the HCV E2 protein to bind to CD81-LEL. The Asp196Glu mutation in CD81 was observed to reduce binding to E2. In addition mutations Phe186Leu and Glu188Lys inhibited binding of CD81 to E2, whereas the Thr163Ala mutation enhanced their interaction (17). Drummer *et al.* (19) also examined the binding site, which was estimated to cover approximately 806 \AA^2 of the CD81-LEL

surface, and identified three additional amino acid contacts, Ile182, Asn184, and Leu162 (19) (figure 1). We used these seven residues as markers to identify the best regions on the CD81-LEL protein surface to target when designing inhibitors to block the E2:CD81 interaction.

3.2. *The AutoLigand Fill Points and Energy Plot Analysis*

AutoLigand was used to analyze the surface of CD81-LEL and select the best ligand binding sites. Five binding sites were identified as potential targets by plotting the total energy per volume (Kcal/mol) for the fill points generated against the volume of the filled site and picking those sites with the lowest values. Figure 2 shows the data from each fill generated at different starting points on the surface using increasing numbers of fill points to fill larger and larger volumes. The fill volumes with less than 100 \AA^3 are small cavities within the protein structure that could be water or ion binding sites and were not considered suitable drug targets. The open diamonds are the values for the fills near amino acid Asn184, one of the five key residues shown previously to interact with E2. The best fill for the site in this region, $-0.165 \text{ Kcal/mol \AA}^3$, was obtained using 180 fill points. As more points were used and the volume of the cavity increased, the predicted free energy of binding became less favorable.

One site predicted by AutoLigand to be an excellent small molecule binding site was located in a region that contained five of the CD81 amino acid residues (Ile182, Phe186, Asn184, Glu188, Asp196) (19) that have been shown by others to interact with

E2 (figure 3a). This is a large cavity located between the C and D helices that is only present in the open conformation of CD81-LEL. A second group of fill points was generated for a neighboring cavity located on the opposite side of the protein (figure 3b). The fill points generated for these two sites were predicted to have the lowest interaction energy of all the sites identified on the open conformation of CD81-LEL. Consequently, these two sites were selected as the primary sites for use in small molecule docking.

3.3. *Docking and Analysis of Ligands Predicted to Bind to the Selected Sites*

Docking runs were performed for the sites selected on CD81-LEL using the NCI Diversity Set II, Sigma, and Asinex libraries of small molecules. The list of ligands predicted to bind to each site were ranked according to binding energy and how well the ligand's atoms mapped onto the fill points for the site. In addition to the fill points defining the rough shape of ligands that would fit best within the cavity, specific fill points were also color coded to identify particular atoms (C,H,N,O) in the ligand that would interact optimally with the surface of the protein in the regions surrounding the ligand (figure 4). The fill points predicted for the site shown in Figure 4a are colored red for hydrogen acceptors such as oxygen or nitrogen, blue for hydrogen, or grey for carbon. One of the better ligands predicted to bind to this site (figure 4b) has atoms that superimpose well with the fill point map (figure 4c). While the superimposition does not need to match perfectly, the points of contact on the protein are considered to be good if the majority of the different atom types in the molecule (75-80%) approximate the same location as the fill points. Such ligands would be expected to form multiple contacts/interactions with the

protein (such as hydrogen bonds, salt bridges, van der Waals interactions) and should bind more tightly than other ligands predicted to make only one or two contacts.

3.4. Experimental Confirmation of Ligand Binding

A total of 36 ligands were tested experimentally using surface Plasmon resonance (Biacore T200 instrument) to identify which of the molecules predicted to bind to Sites 1 and 2 on CD81 actually bind to a recombinant form of the protein (CD81-LEL). Twenty-six of the molecules provided a positive change in response units (RU) upon introduction to a chip containing the immobilized protein (Table 1), indicating the ligands bound to the protein. The measured responses for the ligands that bound varied from 2.3 to 78.4 RU. Those ligands providing the largest responses tended to be molecules that were predicted to bind more deeply inside cavities in Site 1 (ligands 30930, 98026, 7438, 5069) or Site 2 (ligands 78623, 127947, 16631, 38743). Control experiments were performed to confirm that the recombinant form of CD81-LEL we used in these experiments had the correct structure. In these experiments, the CD81-LEL protein was immobilized on a chip and then DPI was used to show the HCV E2 protein recognized and bound to the immobilized CD81-LEL (Table 2).

Six of the more interesting ligand candidates (three predicted to bind to Site 1 and three predicted to bind to Site 2) were further tested to confirm they bind to CD81-LEL using DPI. The results, shown in Figure 5, showed that all six ligands bound to the protein. The relative rank in strength of binding of the Site 1 and 2 ligands, as determined by DPI,

were also similar to the ranking obtained by SPR and the free energy of binding predicted by AutoDock for the majority of the ligands. Ligands 1 - 4 exhibited binding responses that were stronger than or equivalent to the binding observed for benzyl salicylate (0.58 radians, see Figure 5), a small molecule reported previously to block E2 binding to CD81 (33). Benzyl salicylate was identified by Holzer *et al.* (33) by performing a similar virtual screen of small molecules (using a different set of databases) to the cleft we have referred to as Site 1 in the open conformation of CD81-LEL. Thirty-seven analogs of benzyl salicylate were subsequently synthesized by Holzer *et al.* (33) in an effort to enhance the inhibitory activity of benzyl salicylate, but none of the analogs proved to be a better inhibitor than parent compound benzyl salicylate.

For some ligands, significant differences were observed in the actual binding responses obtained by SPR and DPI. As one example, Ligands 1, 2 and 5 had a very similar binding response when tested by SPR, but these ligands exhibited different responses when tested by DPI. One reason for this observed difference in the DPI response might relate to conformational changes in the protein that occur when the small molecules bind. The change in radians measured using DPI when a ligand binds to a protein is known to result from a combination of two effects: 1) the resulting increase in mass and volume when the ligand binds to the protein on the surface of the chip and 2) a conformational change in the protein induced by the binding of the ligand. Small molecules binding in deeper cavities would be expected to have more and stronger

contacts with the protein than ligands sitting exposed to solvent in shallow cavities or surface binding sites.

Those molecules predicted by AutoDock to have the lowest free energy of binding also exhibited the largest DPI radians change and SPR response. The collective data provided by the AutoDock free energy prediction, SPR, and DPI binding assays allowed us to estimate and categorize the relative strength of the ligand's binding to CD81-LEL as strong, moderate or weak. Within the set of six ligands shown in Figure 5, Ligands 1, 2 and 4 exhibit the strongest binding, followed by ligands 5 and 6, which are categorized as moderate binders. Ligand 3 appears to be the weakest binder in the group. Additional SPR analyses performed using a series of Ligand 1 concentrations (figure 6) provided an estimated K_d of $201\mu\text{M}$ for an affinity fit of Ligand 1 binding to the recombinant CD81-LEL protein.

3.5. Effect of Ligand 3 on in vitro Binding of HCV E2 Protein to CD81-LEL

Competition experiments were also performed to determine if selected Site 1 or Site 2 ligands might block E2 binding to CD81-LEL. The two strongest binders in the Site 1 group shown in Figure 5 did not block E2 binding to CD81-LEL. Ligand 3 (689002), on the other hand, was observed to reduce E2 binding to CD81-LEL by 40% (Table 2). The magnitude of the reduction in the binding response in the presence of Ligand 3 is consistent with Ligand 3 having an EC_{50} greater than $500\mu\text{M}$ and being slightly less effective than benzyl salicylate in inhibiting E2 binding to CD81-LEL. This result not

only confirms that Ligand 3 binds within the E2 binding site on CD81-LEL, but it also identifies a small molecule that could prove useful as an early stage drug lead in the development of therapeutics that block HCV invasion.

4. References

1. Ferrari C, Urbani S, Penna A, Cavalli A, Valli, A, Lamonaca V, Bertoni R, Boni C, Barbieri K, Uggeri J, Fiaccadori F (1999) Immunopathogenesis of hepatitis C virus infection. *J Hepat* 31 (Supplement 1): 3-8.
2. Bartenschlager R (1999) The NS3/4A proteinase of the hepatitis C virus: unraveling structure and function of an unusual enzyme and a prime target for antiviral therapy. *J Viral Hepat* 6: 165-81.
3. Lesburg CA, Radfar R, Weber, PC (2000) Recent advances in the analysis of HCV NS5B RNA-dependent RNA polymerase. *Curr Opin Investig Drugs*: 1: 289-96.
4. Welbourn S, Pause A (2007) The hepatitis C virus NS2/3 protease. *Curr Issues Mol Biol* 9: 63-9.
5. Venkatraman S, Njoroge FG (2009) Macrocyclic inhibitors of HCV NS3 protease. *Expert Opin Ther Pat* 19: 1277-303.
6. Enomoto M, Tamori A, Kawada N (2009) Emerging antiviral drugs for hepatitis C virus. *Rev Recent Clin Trials* 4: 179-84.
7. Chary A, Holodniy M (2010) Recent advances in hepatitis C virus treatment: review of HCV protease inhibitor clinical trials. *Rev Recent Clin Trials* 5: 158-73.
8. Sharma SD (2010) Hepatitis C virus: molecular biology & current therapeutic options. *Indian J Med Res* 131: 17-34.
9. Stoll-Keller F, Barth H, Fafi-Kremer S, Zeisel MB, Baumert TF (2009) Development of hepatitis C virus vaccines: challenges and progress. *Expert Rev Vaccines* 8: 333-45.
10. Dubuisson J (2007) Hepatitis C virus proteins. *World J Gastroenterol* 13: 2406-15.
11. Budkowska A (2009) Mechanism of cell infection with hepatitis C virus (HCV)--a new paradigm in virus-cell interaction. *Pol J Microbiol* 58: 93-8.
12. Bartosch B, Vitelli A, Granier C, Goujon C, Dubuisson J, Pascale S, Scarselli E, Cortese R, Nicosia A, Cosset FL (2003) Cell entry of hepatitis C virus requires a set of co-receptors that include the CD81 tetraspanin and the SR-B1 scavenger receptor. *J Biol Chem* 278: 41624-30.
13. Bartosch B, Cosset FL (2006) Cell entry of hepatitis C virus. *Virology* 348: 1-12.
14. Pileri P, Uematsu Y, Campagnoli S, Galli G, Falugi F, Petracca R, Weiner AJ, Houghton M, Rosa D, Grandi G, Abrignani S (1998) Binding of hepatitis C virus to CD81. *Science* 282: 938-41.
15. Levy S, Todd SC, Maecker HT (1998) CD81 (TAPA-1): a molecule involved in signal transduction and cell adhesion in the immune system. *Ann Rev Immunol* 16: 89-109.
16. Petracca R, Falugi F, Galli G, Norais N, Rosa D, Campagnoli S, Burgio V, Di Stasio E, Giardina B, Houghton M, Abrignani S, Grandi G (2000) Structure-function analysis of hepatitis C virus envelope-CD81 binding. *J Virol* 74: 4824-30.
17. Higginbottom A, Quinn ER, Kuo CC, Flint M, Wilson LH, Bianchi E, Nicosia A, Monk PN, McKeating JA, Levy S (2000) Identification of amino acid residues in

- CD81 critical for interaction with hepatitis C virus envelope glycoprotein E2. *J Virol* 74: 3642-9.
18. Zhang YY, Zhang BH, Ishii K, Liang TJ (2010) Novel function of CD81 in controlling hepatitis C virus replication. *J Virol* 84: 3396-407.
 19. Drummer HE, Wilson KA, Pountourios P (2002) Identification of the hepatitis C virus E2 glycoprotein binding site on the large extracellular loop of CD81. *J Virol* 76: 11143-7.
 20. VanCompernelle SE, Wiznycia AV, Rush JR, Dhanasekaran M, Baures PW, Todd SC (2003) Small molecule inhibition of hepatitis C virus E2 binding to CD81. *Virology* 314: 371-80.
 21. Kitadokoro K, Bordo D, Galli G, Petracca R, Falugi F, Abrignani S, Grandi G, Bolognesi M (2001) CD81 extracellular domain 3D structure: insight into the tetraspanin superfamily structural motifs. *Embo J* 20: 12-8.
 22. Kitadokoro K, Galli G, Petracca R, Falugi F, Grandi G, Bolognesi M (2001) Crystallization and preliminary crystallographic studies on the large extracellular domain of human CD81, a tetraspanin receptor for hepatitis C virus. *Acta Crystallogr D Biol Crystallogr* 57: 156-8.
 23. Neugebauer A, Klein CDP, Hartmann RW (2004) Protein-dynamics of the putative HCV receptor CD81 large extracellular loop. *Bioorg Med Chem Lett* 14: 1765-1769.
 24. Balhorn R, Hok S, Burke PA, Lightstone FC, Cosman M, Zemla A, Mirick G, Perkins J, Natarajan A, Corzett M, DeNardo SJ, Albrecht H, Gregg JP, DeNardo GL (2007) Selective high-affinity ligand antibody mimics for cancer diagnosis and therapy: initial application to lymphoma/leukemia. *Clin Cancer Res* 13: 5621s-5628s.
 25. AutoDock website: <http://autodock.scripps.edu>
 26. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AK (1998) Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Comput Chem* 19: 1639-1662.
 27. Huey R, Morris GM, Olson AJ, Goodsell DS (2007) A semi empirical free energy force field with charge-based desolvation. *J Comput Chem* 28: 1145-52.
 28. Huey R, Goodsell DS, Morris GM, Olson AJ (2004) Grid-based hydrogen bond potentials with improved directionality. *Letters in Drug Design and Discovery* 1: 178-183.
 29. Harris R, Olson AJ, Goodsell DS (2008) Automated prediction of ligand binding sites in proteins. *Proteins* 70: 1506-17.
 30. Morris GM, Huey R, Olson A (2008) Using AutoDock for Ligand-Receptor Docking. *Current Protocols in Bioinformatics Chapter 8 Unit 8.14*.
 31. NBCR website: <https://www.nbc.net/pub/wiki/index.php?title=CADD Pipeline>
 32. Sanner MF (1999) Python: a programming language for software integration and development. *J Mol Graph Model* 17: 57-61.
 33. Holzer M, Ziegler S, Neugebauer A, Kronenberger B, Klein CD, Hartmann RW (2008) Structural modifications of salicylates: inhibitors of human CD81-receptor HCV-E2 interaction. *Arch Pharm (Weinheim)* 341: 478-84.

5. Figures

Figure 1. Amino acid residues that participate in HCV E2 binding to CD81-LEL The colored residues are amino acids that have been identified by Higginbottom *et al.* (17) and Drummer *et al.* (19) to contribute to the binding of the HCV protein E2 to CD81-LEL. The structure shown is the monomer of the open conformation of CD81-LEL (PDB ID: 1G8Q). (a) Front view of the protein showing the four contact residues Leu162 (blue), Ile182 (green), Asn184 (orange), and Phe186 (red). (b) Back side of the CD81-LEL protein showing the other three contact residues Thr163 (yellow), Glu188 (cyan) and Asp196 (magenta). This figure was prepared using AutoDock Tools version 1.5.6.

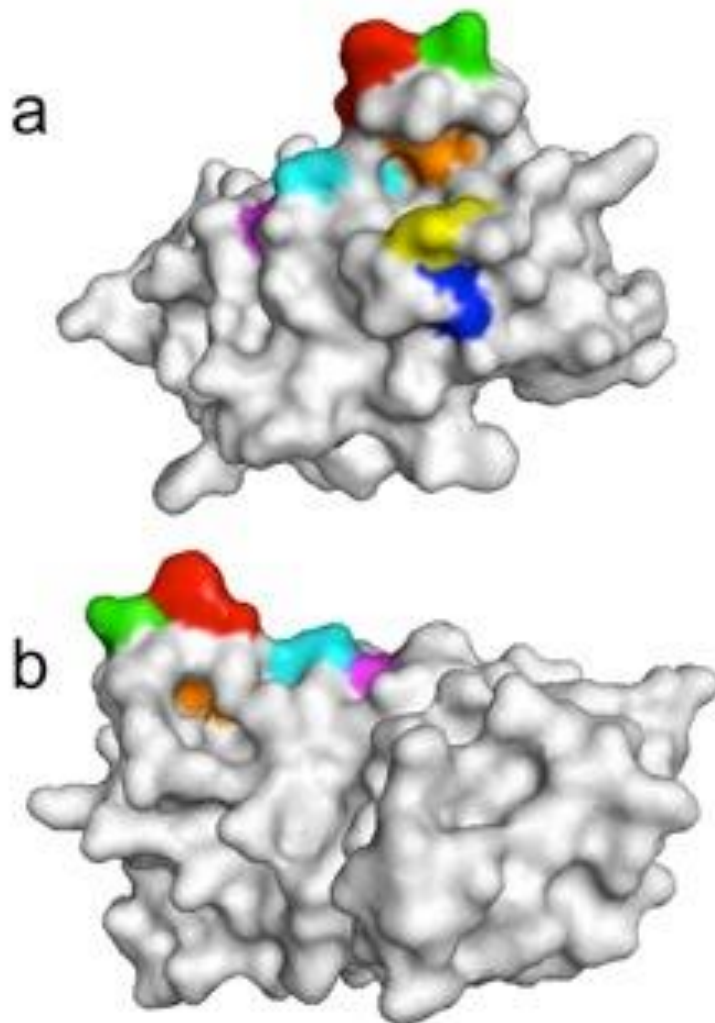


Figure 2. Predicted free energy data for ligand binding sites identified on the surface of CD81-LEL by AutoLigand. This figure was generated by plotting the total energy per volume versus the volume of each fill made from different amounts of fill points. The different symbols depict the fills that start in different locations within the five cavities/sites identified by AutoLigand. Note that there are more than five sets of symbols because some symbols represent fills starting in different locations within the same site (e.g. the large cavity called Site 1). The most efficient fills are those that have the lowest total energy per volume using the smallest volume. The fill points enclosed in the boxes labelled Site 1 and Site 2 correspond to the fills used for docking. This figure was prepared using AutoDock Tools version 1.5.6.

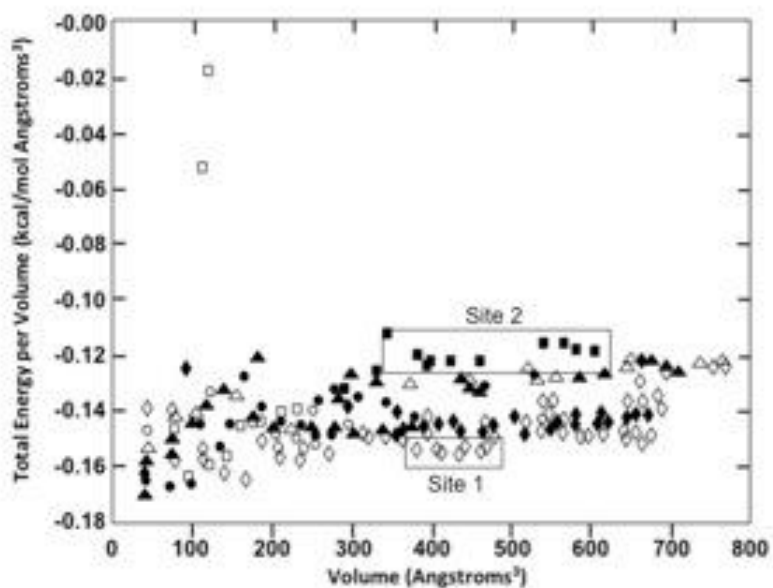


Figure 3. Two ligand binding sites identified by AutoLigand on the open conformation of CD81-LEL (PDB ID: 1G8Q) These two sites were selected as docking targets based on their proximity to the amino acid residues that contact E2 (identified on the molecular surface by blue, yellow, green, orange, cyan, red and magenta colors; see Figure 1 legend for residue numbers) and the low free energy (high affinity) predicted for ligands that would bind in this site. (a) The green spheres fill Site 1 and define its location, the large cavity located between the C and D helices predicted by AutoLigand to be the best binding site. Ligands binding to this site would bind very close to the majority of the amino acids that participate in binding to E2. The green spheres correspond to the open diamond fill points in Figure 2 located between 500 and 600 Å³. (b) The brown spheres show the location of a second binding site, Site 2, predicted by AutoLigand on the opposite side of the protein. Ligands binding to this site should also contribute to the disruption of E2 binding. The brown spheres correspond to the black square fill points shown in Figure 2 located between 550 and 650 Å³. This figure was prepared using AutoDock Tools version 1.5.6.

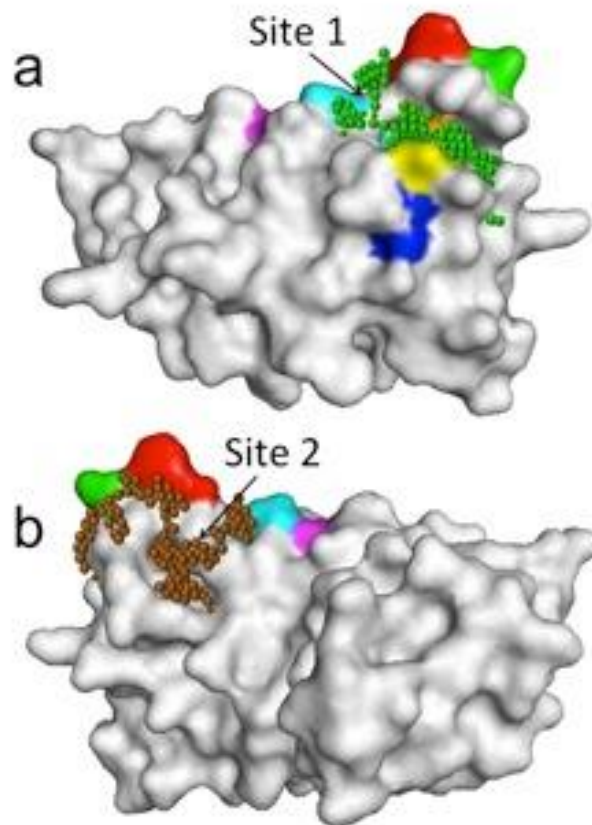


Figure 4. AutoLigand Analysis of the Site 1 ligand binding site AutoLigand fill points not only identify cavities on the surfaces of proteins, but they also predict the structural features of ligands that would bind with the best affinity and selectivity to the protein at these sites. (a) The fill points provided by AutoLigand define the rough shape of ligands that would fit best into the Site 1 cavity. Individual or groups of fill points are also color coded (gray for carbon, light blue for hydrogen, and red for hydrogen acceptors oxygen and nitrogen) to identify particular atoms in the ligand that would interact optimally with the protein's atoms or functional groups in the regions surrounding the ligand. (b) Ligand 1 is shown bound to Site 1 on CD81-LEL in the location and orientation predicted by AutoDock. (c) The superposition of fill points (small spheres) provided by AutoLigand and the actual atom types in Ligand 1 (large spheres) is high (75-80%) indicating that this ligand should bind well in this particular site. Note that the amino acid residues that contact E2 shown in Figure 1 are also shown in these figures using the same color-coding. This figure was prepared using AutoDock Tools version 1.5.6.

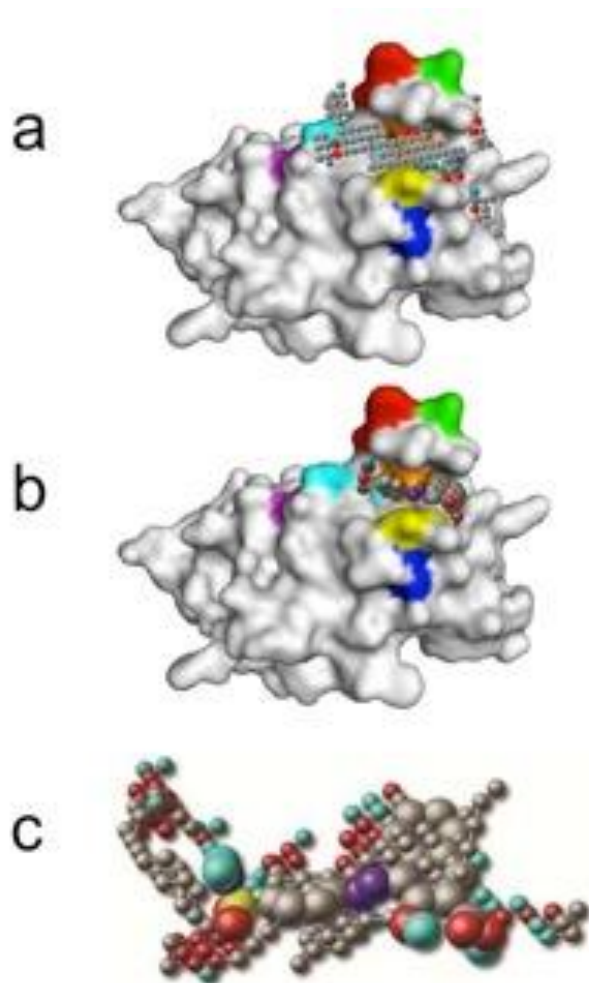


Figure 5. Confirmation of ligand binding to CD81-LEL using DPI Six of the ligands that were found to bind to CD81-LEL by SPR analysis were selected and tested by a second method, DPI, to confirm they bind to CD81-LEL. The results show that all six ligands bind to the protein. The molecules are listed according to the assessed quality of the ligand and its interaction with CD81-LEL using AutoDock's predicted free energy of binding and the DPI and SPR binding data. The relative rank in strength of binding of the Site 1 and 2 ligands, as determined by DPI, were also similar to the ranking obtained by SPR and the free energy of binding predicted by AutoDock. Ligands 1 - 4 exhibited binding responses that were stronger than or similar to the binding observed for benzyl salicylate, a small molecule reported previously to block E2 binding to CD81 (33). Criteria used to define the quality of the ligands are: Strong - makes more than 5 contacts with protein, predicted to be selective and not predicted to bind to multiple sites, not too hydrophobic in addition to having an *in silico* binding energy of > -5 , DPI binding of > 0.3 radians and SPR binding response of > 30 Response Units (RU); moderate - makes 4-5 contacts with protein, hydrophobic interactions contribute to binding in addition to having an *in silico* binding energy of > -3 , a DPI binding of > 0.15 radians and SPR binding response of > 10 ; and weak - makes 3-4 contacts with protein in addition to having an *in silico* binding energy of < -3 , a DPI binding of < 0.15 radians and SPR binding response of < 10 RU.

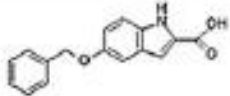
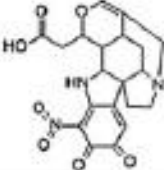
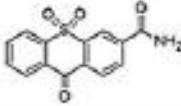
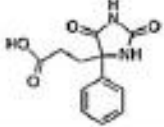
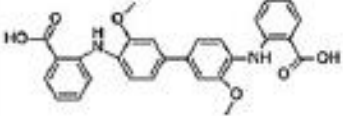
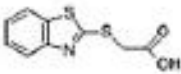
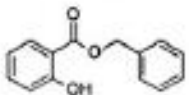
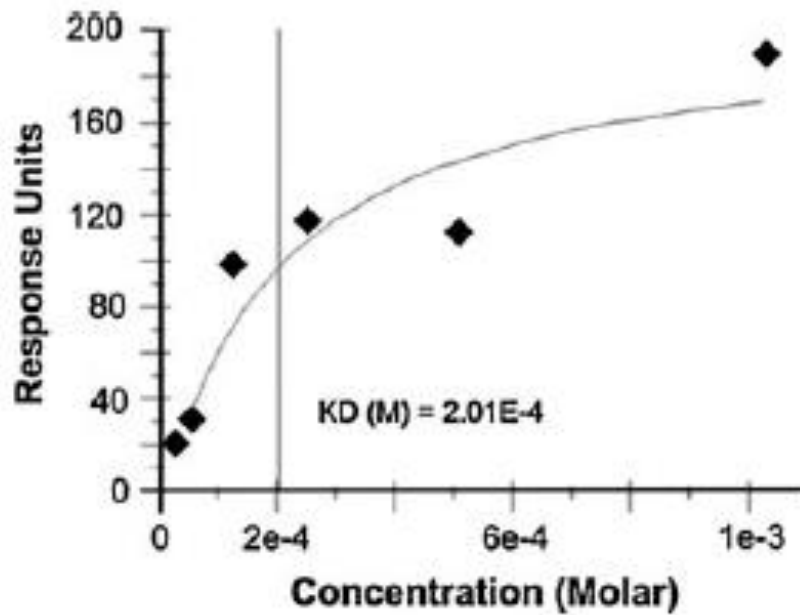
Ligand Number (NCI number)	Structure	Predicted binding energy (kcal/mole)	DPI Response (Radians)	SPR Response (RU)
Site 1				
1 (30930)		- 7.0	1.1	67.7
2 (5069)		- 5.7	0.53	67.5
3 (689002)		- 4.2	0.56	2.3
Site 2				
4 (38743)		- 5.5	0.62	30.6
5 (127947)		- 3.5	0.33	64.8
6 (11891)		- 3.9	0.16	15.2
Benzyl salicylate			0.58	n.d.

Figure 6. Binding of Ligand 1 to CD81-LEL as a function of ligand concentration

This binding experiment was performed using a Biacore T200 as described in the Materials and Methods section. Using the data shown, Ligand 1 was estimated to have a K_d of $\sim 201\mu\text{M}$ based on an affinity fit.



6. Tables

Table 1. Experimental analysis of ligand binding to recombinant CD81-LEL. Thirty-six ligands predicted by AutoDock to bind to Sites 1 and 2 on CD81-LEL were tested experimentally using surface Plasmon resonance as described in the Materials and Methods section. Ligand code numbers are those assigned by the National Cancer Institute. The data, which are the response units generated by the Biacore instrument, are shown for only the 26 ligands that were observed to bind. Because the binding experiments were performed by passing the same concentration of each ligand sequentially across the same protein coated chip, the magnitude of the response can be used to provide an approximate ranking of binding strength. Response unit values greater than 0 indicate binding.

Site 1 Ligand	SPR Response Units	Site 2 Ligand	SPR Response Units
7436	78.4	127947	64.8
30930	67.7	16631	61.8
5069	67.5	63865	58
98026	50	78623	54.5
68982	41.2	38743	30.6
123115	34.5	638134	25
21034	21.1	408734	23.9
36914	3.4	408860	18.2
689002	2.3	93033	16.9
		11891	15.2
		70980	14.1
		25368	13.3
		68971	12.1
		156957	10.6
		55573	10.6
		303800	7.8
		362639	6.5
Binding Site	% Ligands Bound (Number Tested)		
1	69.2 (13)		
2	73.9 (23)		

Table 2. DPI competition experiment showing inhibition of E2 binding to CD81-LEL by Ligand 3 (689002) and comparing the inhibition to that achieved with benzyl salicylate.

Molecule added to CD81-LEL	DPI binding response (radians)	Percent E2 Binding to CD81 (%)
Ligand 3	0.555	
Benzyl salicylate	0.582	
E2 protein	4.140	100
E2 protein + ligand 3	2.700	59.8
E2 protein + benzyl salicylate	1.691	31.2

CHAPTER IV

GENERATION OF A SHAL PROTOTYPE (SH7153) USING CD81-LEL AS A TARGET

Objectives

The aim of this study is to generate a bidentate SHAL prototype (SH7153) and test using Surface Plasmon resonance, antibody neutralizing assays and HCV infection inhibition assays to determine its inhibitory effects on HCV. The computational biology work was done at Scripps Research Institute (USA), Dr. Arthur Olson's Molecular Graphics Laboratory, the surface plasmon resonance binding assays were done at Caltech (USA) at Dr. Jost Vielmetter's protein expression center, and the HCV infections assays were done at Institute Pasteur De Lille (France) at Dr. Jean Dubuisson's lab.

1. Introduction

Several steps in the HCV life cycle are being actively targeted to create new drugs that block virus invasion, processing of the pro-protein or replication of the viral genome (1, 2). CD81 is considered a major HCV receptor that participates in the first step in the process – entry of the virus into hepatocytes (3). A number of anti-CD81 antibodies have been shown to block viral invasion and protect cells from the HCV cytopathic effect (4). Holzer *et al.* was one of the first to select for and identify small molecules that target this entry step (5,6). Two compounds identified in virtual and experimental screens performed by this group, benzyl salicylate (5) and terfenadine (6), have been shown to be moderate inhibitors of the HCV E2: CD81 interaction. we identified a second cavity within the E2 binding site adjacent to the large cleft targeted by Holzer *et al.* (5,6) using AutoLigand (7) to search the surface of CD81-LEL near the amino acid residues in CD81 that participate

in HCV envelope glycoprotein binding (8, 9). Virtual screening runs were performed using AutoDock (10) to identify additional small molecules which might bind to the large cavity targeted previously by Holzer *et al.* and to this new cavity, and 26 of those predicted to bind were confirmed experimentally to bind to CD81. One of these ligands, 689002, was also shown to block E2 binding (11). New experiments were conducted with the ligands predicted to bind to the second cavity have identified another ligand that blocks E2 binding to CD81-LEL. We have used these two ligands to design and synthesize a prototype (first generation) selective high affinity ligand (SHAL) as a steric inhibitor to block HCV E2 glycoprotein binding to CD81.

2. Materials and Methods

2.1. Recombinant Proteins, Antibodies and Ligands

The recombinant HCV E2 glycoprotein was provided by Dr. Joseph Marcotrigiano (Rutgers University-NJ/USA), His-CD81-LEL (Bioclone-CA/USA) , JS-81 antibody (BD Pharmingen, 551108) and the small molecules were provided by the NIH/NCI Developmental Therapeutics Program.

2.2. Dual Polarization Interferometry

DPI analyses were performed using an AnaLight 4D workstation (Farfield Group, Manchester UK). The recombinant CD81-LEL was immobilized onto a Thiol AnaChip using Sulfo-GMBS as a cross-linker in PBS running buffer (pH 7.4). Non-specific sites were blocked with digested casein. TRIS was used to cap the cross-linker, blocking any additional amines from covalently binding to the cross-linker on the chip surface. Ligands

were prepared as 20 mM stock solutions in dimethylsulfoxide (DMSO). Each ligand was diluted to a final concentration of 500 μ M in PBS just prior to injection (final DMSO concentration was 2.5 %). PBS and DMSO mixed in the same ratio were used as a blank. Data collection and analysis were performed using the AnaLight Resolver.

A subset of the ligands identified to bind to CD81-LEL was also tested to determine if they might block the HCV E2 glycoprotein from binding to CD81-LEL using DPI. In these experiments, a recombinant form of the CD81-LEL protein was immobilized on the chip and unreacted cross-linker was blocked as described above. Recombinant HCV E2 glycoprotein (Immune Technology Corp, New York, NY) was then injected to determine the magnitude of the binding response when E2 bound to CD81-LEL in the absence of the ligand. To evaluate the effect of a ligand on E2 binding to CD81-LEL, the same experiment was repeated except that the E2 glycoprotein was premixed with the ligand at a final ligand concentration of 500 μ M. If the ligand inhibits E2 binding to CD81-LEL when the mix of E2 and the ligand are added to the chip, the DPI binding response in the presence of the ligand should be less than the response in the absence of the ligand. If a reduction in E2 binding is observed by DPI, the magnitude of the inhibition can be calculated using the binding responses for the ligand, the E2 glycoprotein and a mixture of the E2 glycoprotein and the ligand.

2.3. *Surface Plasmon Resonance*

SPR analysis was performed using a Biacore T100 workstation (GE Healthcare, NJ, USA). A recombinant form of the CD81-LEL protein with a His-tag (Bioclone-CA/USA) was used to determine the binding affinities of ligands 689002 and 93033 in

addition to that of SH7153. Briefly, 10 μ M CD81-LEL-His diluted into 10 mM sodium acetate buffer pH 4.5 was immobilized for 15 min at a flow speed of 5 μ l/min onto a CM5 sensor chip using amine coupling (EDC-NHS). Approximately 20,000 RU of protein were immobilized on the chip. The ligands prepared as 200 μ M solutions in PBS-1% DMSO (the running buffer) and they were introduced to the protein using a pre-programmed 3 min association and 1 min dissociation interval. The binding affinities of selected ligands were estimated using data collected from a series of SPR binding experiments conducted at different ligand concentrations. To obtain the kinetic and affinity data needed to estimate the K_d , the original ligand sample was diluted serially with running buffer to produce five different ligand concentrations: 200, 66.6, 22.2, 7.41, 2.47 and 0 μ M. Data were fitted using a monovalent binding model.

2.4. *Solid phase synthesis of SH7153 and its analogues*

Unless otherwise stated, all reagents and solvents were obtained from commercial suppliers and used as received. 2-Chlorotriethylamine resin (100-200 mesh, 1% DVB, 1.2 mmol/g loading density) and all Fmoc-protected amino acids were obtained from Novabiochem Inc. Fmoc-8-amino-3,6-dioxactanonic acid (Fmoc-mini-PEG) was purchased from Peptides Int. (Louisville, Kentucky- USA). The hexafluorophosphate (PF₆) salt of 1,4,7,10-Tetraazacyclododecane-1,4,7,10-tetraacetic acid mono-N-hydroxysuccinimide (DOTA-NHS) ester was purchased from Macrocyclics (Dallas, Texas/ USA). N-(biotinyloxy)succinimide (biontynyl-OSu) was purchased from BACHEM (Torrance, California/ USA). Dabsyl-L-valine (Dv acid) was purchased from TCI Inc. (Tokyo). (3-(2-((3-Chloro-5-trifluoromethyl)-2-pyridinyl)oxy)-anilino)-3-

oxopropanionic acid (Ct acid) and 4-(4-(4chlorobenzyl)piperazino)-3-nitrobenzenecarboxylic acid (Cb acid) ligands were procured from Bionet Research (Key Organics Ltd. – Camelford, Pennsylvania/ USA)). Acetonitrile (MeCN), dichloromethane (DCM), N,N-dimethylformamide (DMF), diisopropylethylamine (DIEA), hydrazine, trifluoroacetic acid (TFA), 1,3-diisoproylcarbodiimide (DIC), 1-hydroxybenzotriazole (HOBT), 20 % pyridine/DMF and triethylsilane (Et₃SiH or TES) were purchased from Sigma-Aldrich (St. Louis, Missouri / USA).

Instrumentation. Solid support syntheses were carried out in Pierce polyethylene columns (5 mL). All SHALs were purified by semi-preparative HPLC at 10 mL/min on a Waters preparative machine with photodiode array detection (Waters Symmetryprep C18, 7 μm, 19 x 300 mm column) and lyophilized using Kinetics Flexi-Dry freeze-dryer. The purity was confirmed by analytical HPLC carried out at 1 mL/min on an Agilent 1100 machine (Waters Symmetry C18, 5 μm, 4.2 x 150 mm column, diode array detector) with a linear gradient from 95 % H₂O (1 % TFA) to 80 % MeCN (1 % TFA) over 12 min. Mass spectra were acquired on a Micromass Quattro Micro API mass spectrometer operating in positive ion mode. The samples were dissolved in MeCN / H₂O (1:1), 0.1 % formic acid for mass spectrometry analysis.

Preparation of Resin 1: 2-Chlorotrityl chloride resin (100 mg, 0.13 mmol) was suspended in 2 mL dichloromethane (DCM) for 30 min and washed 3x with anhydrous dimethylformamide (DMF). The resin was treated with Fmoc-D-Lys(Boc)-OH (20 mg, 43 μmol, 468.5 g/mol, 0.33 equiv. of the resin, the limiting reactant) and DIEA (25 μL,

144 μmol) dissolved in 1.0 mL anhydrous DMF (Scheme 1). The reaction mixture was agitated on a nutator for 3 h. Ethanol (1.0 mL) and DIEA (100 μL) was added and agitated for 15 min to quench the excess chlorotriyl chlorides on the resin. The reaction solution was removed by filtration and the resin was washed 3x with DMF and treated with 20 % piperidine/DMF (1 mL) for 20 min for fluorenylmethoxycarbonyl (Fmoc) deprotection. The Fmoc-deprotected resin-bound Boc-lysine (Resin 1) was washed with DMF and subsequently coupled to the next Fmoc-protected residue using amide-coupling chemistry employing ~4-5 molar equivalents of DIC and HOBt in DMF. The washing and filtration cycle with DMF (2 mL, 3x, 1 min each) was performed between every deprotection and coupling steps.

Preparation of Resin 2: Resin 1 was treated with a solution of Fmoc-8-amino-3,6-dioxactanonic acid (50 mg, 0.13 mmol, 385.4 g/mol, 3 equiv. of the limiting reactant), DIC (30 μL , 0.19 mmol, 126.20 g/mol, 0.815 g/mL), and HOBt (26 mg, 0.19 mmol, 135.12 g/mol) in 1.0 mL anhydrous DMF for 3-5 h (coupling cycle 2, step 1a and 1b). The resin was then treated with Fmoc-D-Lys(Dde)-OH, HOBt and DIC in 1.0 mL anhydrous DMF for 5 h (coupling cycle 3, step 2a and 2b). After Fmoc deprotection, the resultant free amine was subsequently coupled to another solution of Fmoc-8-amino-3,6-dioxactanonic acid (Fmoc-mini-PEG), HOBt and DIC in 1.0 mL anhydrous DMF for 3-5 hours (coupling cycle 4, step 3a and 3b) producing Resin 2 after 20% piperidine/DMF Fmoc deprotection and washing the resin with DMF to remove the excess reagents.

Synthesis of SH7153. Resin 2 was treated with 93033 (32 mg, 0.085 mmol, 375.81 g/mol, 1.5 equiv.), HOBt and DIC in 1.0 mL anhydrous DMF for 5-12 h (coupling cycle 8, step 1a). The reaction solution was filtered and the resin was washed and treated twice with 4 % hydrazine in DMF (40 μ L Hydrazine in 900 μ L DMF for 10 min) to remove the 1-(4,4-dimethyl-2,6-dioxacyclohexylidene)ethyl (Dde) protecting group (step 1b). After washing, the resulting resin was treated with 689002 (30 mg, 0.080 mmol, 374.71 g/mol), HOBt and DIC in 1.0 mL anhydrous DMF for 5-12 h (coupling cycle 9, step 2). In the last step, the solvent was exchanged from DMF to DCM (2 mL) and resin was washed with DCM. The washed resin was then treated with 2 mL of 20 % TFA / 1 % TES in DCM for 30 min to cleave the product from the resin as well as the Boc-protecting group on the ϵ -amine of the first lysine residue. The cleaved solution was filtered and the resin was washed with DCM. The combined DCM filtrate (red) was concentrated. The crude material was analyzed by analytical Agilent HPLC (to determine the ratio of product / impurities in the cleaved material using UV absorbance peak areas at 254 nm). Purification was performed using semi-preparative HPLC and the product fractions were lyophilized to dryness. Analytical HPLC chromatograms and electron spray ionization mass spectroscopy (ESI-MS) was used to confirm the final product.

2.5. *Antibody neutralizing assays*

For antibody neutralization assay Raji cells were used, a human B cell line that expressed high amounts of CD81 on the surface (data not shown). Cells were grown in RPMI medium (10% FCS, 1% penicillin/streptomycin, 1% L-glutamine, 1% non-essential amino acids, 1% sodium pyruvate, pH 7.4, at 37°C with 5% CO₂). 2x10⁵ cells

were incubated with or without different concentrations (50 μ M, 100 μ M, 400 μ M and 1 mM) of indicated inhibitor for 20 min at room temperature, subsequently 1 μ l (16 ng/ μ l) of FITC-labeled anti CD81 antibody (BD Pharmingen, 551108) was added to the cells and incubated for 20 min (antibody titration was performed to obtain a working dilution range, data not shown). Cells were washed and analyzed by FACS (BD FACSCalibur, software: Cell Quest Pro). Mean Fluorescence Intensity MFI was calculated using Flowjo software (TreesStar, www.flowjo.com).

2.6. HCV infection assays

Pseudotyped retroviral particles harboring HCV envelope proteins (HCVpp) from different genotypes were produced as described previously [23, 24] with plasmids kindly provided by F.L. Cosset, J. Ball, and R. Bartenschlager. A plasmid encoding the feline endogenous virus RD114 glycoprotein [25] was used for the production of RD114pp. Both HCVpp and RD114pp expressed *Firefly* luciferase.

The cell culture-produced HCV particles (HCVcc) used in this study were based on the JFH1 strain [26] and were prepared as described previously [27, 28]. They were engineered to express the A4 epitope, titer-enhancing mutations and *Gaussia* luciferase [27,28].

To identify ligands that inhibit HCV infection, Huh-7 cells were seeded in 96-well plates and treated the day after with six different concentrations of each ligand diluted in DMSO in duplicate using a Zephyr automated liquid handling workstation (Caliper BioSciences, Hopkinton, MA). The final concentration of DMSO (1%) was adjusted to be the same for all ligand concentrations. Cells treated with DMSO were used as negative controls. Cells treated with different concentrations of anti-CD81 (JS-81 from BD Pharmingen, San Jose, CA) 1 hour before infection, were also used as positive controls. The third day, RD114pp, HCVpp or HCVcc were inoculated and incubated for 30 hours

at 37°C. *Firefly* and *Gaussia* luciferase assays were performed as indicated by the manufacturer (Promega, San Luis Obispo, CA).

3. Results

Our previous analyses of two sets of small molecules predicted to bind to the large cleft on the open conformation of CD81-LEL and to an adjacent cavity nearby identified 26 small molecules that bind to the large extracellular loop of CD81. These compounds are being evaluated as potential leads for modification or optimization (e.g. by fragment based drug design or the development of SHALs) to create steric inhibitors that block the binding of HCV E2 glycoprotein to CD81 on hepatocytes. Using dual polarization interferometry, one of these compounds (689002) was shown to block the recombinant E2 glycoprotein binding to CD81-LEL immobilized on a chip. Additional binding studies have been conducted in an effort to identify other ligands that might also block the HCV E2: CD81 interaction and be used in combination with 689002 to create a bidentate SHAL with improved CD81 binding.

3.1. *Identification of a Second Ligand that Blocks E2 binding to Recombinant CD81-LEL*

Dual polarization interferometry (DPI) experiments conducted with a set of small molecules that were predicted to bind to a second cavity on CD81-LEL (11) (adjacent to key CD81 amino acids (8,9) that contribute to CD81's binding to the HCV E2 glycoprotein) identified a second small molecule that blocks recombinant E2 binding to CD81-LEL. In these experiments, recombinant HCV E2 glycoprotein was immobilized on a chip and dual polarization interferometry (DPI) was used to compare the binding of

CD81-LEL, in the absence and presence of 500 μ M ligand, to the immobilized E2. As we reported previously (11), the compound 689002 (9-Oxo-9H-thioxanthene-3-carboxamide 10,10-dioxide) was observed to block CD81 binding to E2 (Table 1). During the screening of the Site 2 ligands, a second compound, 93033 (4-(2-(2,4-dioxo-1,2,3,4-tetrahydropyrimidin-1-yl)acetamido)-2-hydroxybenzoic acid), was also found to block CD81-LEL binding to E2. In both cases, the DPI response in radians produced by CD81-LEL binding to E2 was reduced dramatically when the CD81-LEL-ligand mix was added to the E2 immobilized on the chip. In contrast, ligand 30930 (5-(benzyloxy)-1H-indole-2-carboxylic acid, which is shown as an example of a negative control) was observed to have no effect on CD81-LEL binding.

3.2. *Design and Synthesis of a Prototype Bidentate SHAL*

The lowest energy conformations of 689002 and 93033 bound to CD81-LEL, as identified by AutoDock (figure 1A), were used to design a first generation bidentate SHAL. As a first step, the distance from the amino nitrogen on 689002 and the carbon on 93033's carboxyl group was measured to define a minimum length that would be required for the linker connecting the two ligands. While the linear distance between these atoms was measured to be 8.15Å (figure 1B), the linker was designed to be longer (~12Å) to allow it to extend over the amino acid Glu188 (colored cyan) located between the two ligand binding sites. Using this length to define the distance of separation between the two ligands in the SHAL, a scaffold to which the two ligands would be conjugated was designed (figure 2). As is typical of all first generation SHALs, lysine and miniPEG

molecules were used to create the scaffold. Lysine residues were used at the terminus of the SHAL to provide a free amine and free carboxyl group to which various tags and effectors can be attached to the SHAL. Internal lysine residues were employed to provide branch points within the scaffold. MiniPEG molecules were incorporated to provide distance between linked ligands and also to make the linker flexible and contribute to overall SHAL solubility.

Solid phase synthesis and carbodiimide coupling chemistry was used to connect the components to the resin and create the molecular scaffold of the SHAL, and ligands 689002 and 93033 were subsequently conjugated to the appropriate sites on the scaffold (figure 3). Following cleavage from the resin and HPLC purification, 20 milligrams of SH7153 (molecular mass of 977.00 Da) were obtained. Two additional by-products of the reaction, SH7153A (8 mg, molecular mass of 960.04 Da) and SH7153B (7 mg, molecular mass of 993.97 Da) were also obtained (figure 4). SH7153A is a bidentate SHAL containing two 689002 ligands linked together. SH7153B is a bidentate SHAL containing two 93033 ligands linked together.

3.3. Analysis of SH7153 Binding to CD81

SH7153 was first tested using surface Plasmon resonance to determine if it would bind to recombinant CD81-LEL immobilized on a chip. As shown in Figure 5A, SH7153 not only bound to CD81-LEL, but it also dissociated slowly from the protein ($k_{\text{off}} \sim 2.17\text{E}-04$ 1/sec). This is in marked contrast to the results obtained with the two ligands. Both 689002 and 93033 also showed reasonable binding responses, but the two ligands

dissociate rapidly (figure 5B) as is characteristic of weakly bound small molecules. The off rate for 689002 was 0.002631 1/sec and the off rate for 93033 was 0.00239 1/sec. A series of binding curves were obtained for the SHAL and two ligands at different concentrations to obtain an estimate of their affinities for recombinant CD81-LEL. Both ligands bound weakly and had K_d 's that were greater than 200 μ M (Table 2). SH7153, in contrast, bound with a K_d of 21 μ M.

The relative strength of binding of SH7153 and the two by-products of the SHAL synthesis SH7153A and SH7153B to native CD81 was then tested by measuring their ability to block the JS-81 antibody binding to Raji cells. It was shown that SH7153 and the 2 by-products didn't give a significant inhibitory effect.

4. Discussion and conclusions

Virtual screens performed in an earlier study by docking a large library of diverse small molecules into two neighboring cavities on the open conformation of CD81-LEL led to the identification of two sets of compounds that were predicted to bind to CD81 (11) in the same region where the HCV glycoprotein E2 has been reported to bind (8,9). Experimental testing of these compounds revealed that 26 of the molecules bound to a recombinant form of CD81 LEL. One of the molecules predicted to bind to the largest cavity, 9-O-9H-thioxanthene-3-carboxamide 10,10-dioxide (689002), was found to block E2 binding to recombinant CD81-LEL (11). Our recent testing of the other set of molecules predicted to bind to the smaller cavity on CD81-LEL led to the identification of a second compound that blocks E2 binding to CD81-LEL, 4-(2-(2,4-dioxo-1,2,3,4-tetrahydropyrimidin-1-yl)acetamido)-2-hydroxybenzoic acid (ligand 93033).

Both ligands 689002 and 93033 were identified by docking experiments conducted using two cavities that are only present on the surface of the open conformation of CD81-LEL (11). Although molecular dynamics studies of the open conformation conducted by Neugebauer *et al.* (13) have been used to suggest that a second, closed conformation of CD81-LEL reported by Kidadokoro *et al.* (14) was more likely to be the physiologically relevant conformation, Neugebauer *et al.*'s (13) own analysis of the dynamics data suggested the two conformations might exist in dynamic equilibrium. This hypothesis has been supported by the small molecule binding studies we conducted using recombinant CD81-LEL. More than 72% of the ligands predicted by docking to bind to the large cleft/cavity present only in the open conformation of the protein were found by experiment to bind to CD81-LEL (11).

The prediction that the two molecules bind to two adjacent cavities surrounding 5 of the 7 key CD81 amino acids that bind to E2, and the docked conformations of the two molecules showing that both molecules contain carboxyl groups that do not appear to interact appreciably with CD81 (so it is available for conjugation to a linker) and are oriented toward each other all suggest 689002 and 93033 are good choices to link together to create a prototype bidentate SHAL that targets the E2 binding site on CD81. The docked conformations of the molecules bound to CD81-LEL were used to design a linker scaffold, and the first CD81 SHAL, SH7153, was then synthesized using solid phase chemistry. Upon purification by HPLC, SH7153 and two by-products (bidentate SHALs containing two of the same ligands) were obtained.

Surface Plasmon resonance analyses of SH7153 binding to CD81-LEL showed the SHAL bound to the recombinant protein with a K_d of $\sim 21 \mu\text{M}$. Estimates obtained for the K_d 's of the two ligands used to create the SHAL (both 689002 and 93033 had K_d 's $> 200 \mu\text{M}$) indicate the linking together of 689002 and 93033 resulted in the production of a SHAL with a K_d that is at least 10 fold higher than the K_d of the individual ligands. In addition, the SHAL was observed to dissociate from CD81-LEL slowly with an off-rate that is similar to or slower than the HCV E2 protein (15) and a number of HCV neutralizing antibodies (16). Flow cytometry experiments were also conducted to examine SH7153 binding to native CD81 on Raji cells using the JS-81 antibody as a competitor. While an initial set of experiments suggested SH 7153 inhibited JS-81 binding to CD81 on Raji cells in a dose dependent manner, recent experiments have not confirmed that inhibition. Subsequent HCV infection experiments also showed this first SHAL was not effective in inhibiting HCV viral infection of Huh-7 cells.

While kinetic data obtained for a specific binding event that occurs between two isolated proteins can contribute a great deal to our understanding of that particular interaction, the results may not necessarily reflect what happens between the two proteins in their native state. In the case of inhibitors designed to prevent HCV E2 binding to CD81 and block viral entry and cell infection, a number of factors can have a profound impact on how well the drug induces its effect. Recombinant protein interactions in solution, for example, may not be the same as the interactions that occur between E2 when it is associated with the virus and CD81 when it is located in the target cell's membrane. E2 is known to form a heterodimer with E1 in the membrane of HCV, and E2 in the dimer

may react differently when E2 is free in solution. Protein structural dynamics may also affect the number of available binding sites. The number of binding events required per cell to trigger viral entry and infection, the number of available CD81 molecules, and subtle structural variations in the HCV E2 protein expressed by different viral genotypes that may modulate the interaction of the protein with CD81 are also important. The impact of these factors may explain some of the variability in kinetic data others have obtained and may have led to major differences in opinion as to what the K_d of an antibody or inhibitor needs to be to effectively block HCV binding to CD81. Based on K_d 's determined for HCV E2 binding to CD81, Petracca et al (17) suggested antibodies would need to have nanomolar affinities to block HCV from binding to CD81. However, Keck et al (18) and Nakajima *et al.* (15) reported K_d 's for HCV E2 protein binding to CD81 that were only in the micromolar range. The requirement for a K_d in the micromolar range was further supported by the mathematical modeling of HCV viral kinetics conducted by Padmanabhan and Dixit (19). The K_d for E2 binding to CD81 was estimated from their mathematical treatment of HCV infections to be in the range of 10-100 μM .

The current results show that we can effectively combine rapid virtual screening of large libraries of compounds with the experimental testing of a small number of hits to identify lead small molecules that can be linked together to create a SHAL. By linking together the two weak binding ligands 689002 and 93033, we have been able to create a first generation bidentate SHAL that binds to CD81 with an affinity that is higher than the affinities of the individual ligands used to create it. While this first SHAL did not inhibit HCV infection of Huh-7 cells, the SHALs μM affinity and slow off-rate observed for the

SHAL binding to CD81-LEL is encouraging. Based on our previous experience with second generation SHALs developed for cancer therapy, improvements in binding affinity and efficacy can be accomplished through the optimization of linker lengths, replacing one ligand with another, or by adding a third ligand to create a tridentate SHAL.

5. References

- 1- Chen SL, Morgan TR. The natural history of hepatitis C virus (HCV) infection. *Int J Med Sci.* 2006;3(2):47-52.
- 2- Pawlotsky JM. The science of direct-acting antiviral and host-targeted agent therapy. *Antivir Ther.* 2012;17(6 Pt B):1109-17.
- 3- Samreen B, Khaliq S, Ashfaq UA, Khan M, Afzal N, Shahzad MA, *et al.* Hepatitis C virus entry: Role of host and viral factors. *Infect Genet Evol.* 2012 Dec;12(8):1699-709.
- 4- Fofana I, Xiao F, Thumann C, Turek M, Zona L, Tawar RG, *et al.* A novel monoclonal anti-CD81 antibody produced by genetic immunization efficiently inhibits hepatitis C virus cell-cell transmission. *PLoS One.* 2013 May 21;8(5):e64221.
- 5- Holzer M, Ziegler S, Neugebauer A, Kronenberger B, Klein CD, Hartmann RW. Structural modifications of salicylates: Inhibitors of human CD81-receptor HCV-E2 interaction. *Arch Pharm (Weinheim).* 2008 Aug;341(8):478-84.
- 6- Holzer M, Ziegler S, Albrecht B, Kronenberger B, Kaul A, Bartenschlager R, *et al.* Identification of terfenadine as an inhibitor of human CD81-receptor HCV-E2 interaction: Synthesis and structure optimization. *Molecules.* 2008 May 7;13(5):1081-110.
- 7- Harris R, Olson AJ, Goodsell DS. Automated prediction of ligand-binding sites in proteins. *Proteins.* 2008 Mar;70(4):1506-17.
- 8- Higginbottom A, Quinn ER, Kuo CC, Flint M, Wilson LH, Bianchi E, *et al.* Identification of amino acid residues in CD81 critical for interaction with hepatitis C virus envelope glycoprotein E2. *J Virol.* 2000 Apr;74(8):3642-9.
- 9- Grove J, Nielsen S, Zhong J, Bassendine MF, Drummer HE, Balfe P, *et al.* Identification of a residue in hepatitis C virus E2 glycoprotein that determines scavenger receptor BI and CD81 receptor dependency and sensitivity to neutralizing antibodies. *J Virol.* 2008 Dec;82(24):12020-9.
- 10- Morris GM, Huey R, Olson AJ. Using AutoDock for ligand-receptor docking. *Curr Protoc Bioinformatics.* 2008 Dec;Chapter 8:Unit 8.14.
- 11- Olaby RA, Azzazy HM, Harris R, Chromy B, Vielmetter J, Balhorn R. Identification of ligands that target the HCV-E2 binding site on CD81. *J Comput Aided Mol Des.* 2013 Apr;27(4):337-46.
- 12- Chockalingam K, Simeon RL, Rice CM, Chen Z. A cell protection screen reveals potent inhibitors of multiple stages of the hepatitis C virus life cycle. *Proc Natl Acad Sci U S A.* 2010 Feb 23;107(8):3764-9.
- 13- Neugebauer A, Klein CD, Hartmann RW. Protein-dynamics of the putative HCV receptor CD81 large extracellular loop. *Bioorg Med Chem Lett.* 2004 Apr 5;14(7):1765-9.
- 14- Kitadokoro K, Bordo D, Galli G, Petracca R, Falugi F, Abrignani S, *et al.* CD81 extracellular domain 3D structure: Insight into the tetraspanin superfamily structural motifs. *EMBO J.* 2001 Jan 15;20(1-2):12-8.

- 15- Nakajima H, Cocquerel L, Kiyokawa N, Fujimoto J, Levy S (2005) Kinetics of HCV envelope proteins' interaction with CD81 large extracellular loop. *Biochem Biophys Res Commun* 328: 1091–1100.
- 16- Wang Y, Keck ZY, Fong SK. Neutralizing antibody response to hepatitis C virus. *Viruses*. 2011 Nov;3(11):2127-45.
- 17- Roberto P., Fabiana F., Giuliano G., Nathalie N., Domenico R., *et al.* (2000) Structure-Function Analysis of Hepatitis C Virus Envelope-CD81 Binding. *J Virol*. 74: 4824-4830.
- 18- Keck ZY, Li SH, Xia J, von Hahn T, Balfe P, McKeating JA, *et al.* Mutations in hepatitis C virus E2 located outside the CD81 binding sites lead to escape from broadly neutralizing antibodies but compromise virus infectivity. *J Virol*. 2009 Jun;83(12):6149-60.
- 19- Padmanabhan P, Dixit NM. Mathematical model of viral kinetics *in vitro* estimates the number of E2-CD81 complexes necessary for hepatitis C virus entry. *PLoS Comput Biol*. 2011 Dec;7(12):e1002307.
- 20- Copeland, RA, Pompliano, DL, Meek, TD (2006) Drug-target residence time and its implications for lead optimization. *Nature Rev. Drug Discov*. 5: 730-9.
- 21- Tummino, PJ, Copeland, RA (2008) Residence time of receptor-ligand complexes and its effect on biological function. *Biochem*. 47: 5481-92.
- 22- Copeland, RA (2010) The dynamics of drug-target interactions: drug-target residence time and its impact on efficacy and safety. *Expert Opin. Drug Discovery* 5: 305-310.
- 23- Bartosch B, Bukh J, Meunier JC, Granier C, Engle RE, *et al.* (2003) In vitro assay for neutralizing antibody to hepatitis C virus: Evidence for broadly conserved neutralization epitopes. *Proc Natl Acad Sci U S A* 100(24): 14199-14204. 10.1073/pnas.2335981100.
- 24- Op De Beeck A, Voisset C, Bartosch B, Ciczora Y, Cocquerel L, *et al.* (2004) Characterization of functional hepatitis C virus envelope glycoproteins. *J Virol* 78(6): 2994-3002.
- 25- Sandrin V, Boson B, Salmon P, Gay W, Negre D, *et al.* (2002) Lentiviral vectors pseudotyped with a modified RD114 envelope glycoprotein show increased stability in sera and augmented transduction of primary lymphocytes and CD34+ cells derived from human and nonhuman primates. *Blood* 100(3): 823-832. 10.1182/blood-2001-11-0042.
- 26- Wakita T, Pietschmann T, Kato T, Date T, Miyamoto M, *et al.* (2005) Production of infectious hepatitis C virus in tissue culture from a cloned viral genome. *Nat Med* 11(7): 791-796. 10.1038/nm1268.
- 27- Rocha-Perugini V, Montpellier C, Delgrange D, Wychowski C, Helle F, *et al.* (2008) The CD81 partner EWI-2wint inhibits hepatitis C virus entry. *PLoS One* 3(4): e1866. 10.1371/journal.pone.0001866; 10.1371/journal.pone.0001866.
- 28- Delgrange D, Pillez A, Castelain S, Cocquerel L, Rouille Y, *et al.* (2007) Robust production of infectious viral particles in huh-7 cells by introducing mutations in

hepatitis C virus structural proteins. *J Gen Virol* 88(Pt 9): 2495-2503.
10.1099/vir.0.82872-0.

Figures. Figure 1. The binding modes of ligands 689002 and 93033 to CD81-LEL. The docked conformations of ligands 689002 (pink) and 93033 (blue) are shown binding to Sites 1 (large cleft on right) and Site 2 (smaller cavity on left) located within the E2 binding site on CD81-LEL (A). The linear distance between ligand carboxyl groups was measured to be 8.15Å, but linker was designed to be longer (~12Å) to allow it to bend over amino acid Glu188 (colored cyan) located between the two ligand binding sites (B).

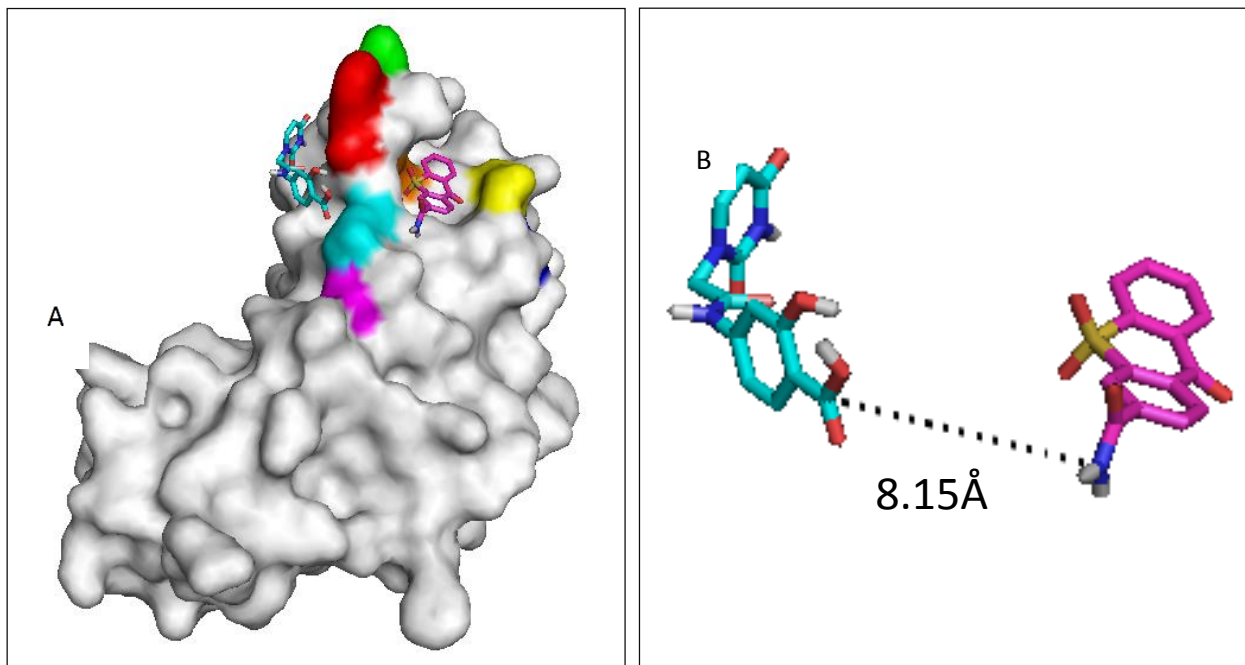


Figure 2. The Linker Molecular Scaffold of SH7153. The molecular scaffold designed to link ligands 689002 and 93033 together to create SHAL SH7153 is comprised on lysine and miniPEG residues. The terminal lysine residue provide a free amine and carboxyl group that can be used to attached tags or other effectors. Lysine residues are also incorporated to provide branch points within the linker and miniPEG molecules are used to provide distance between ligands and help optimize the SHAL's solubility.

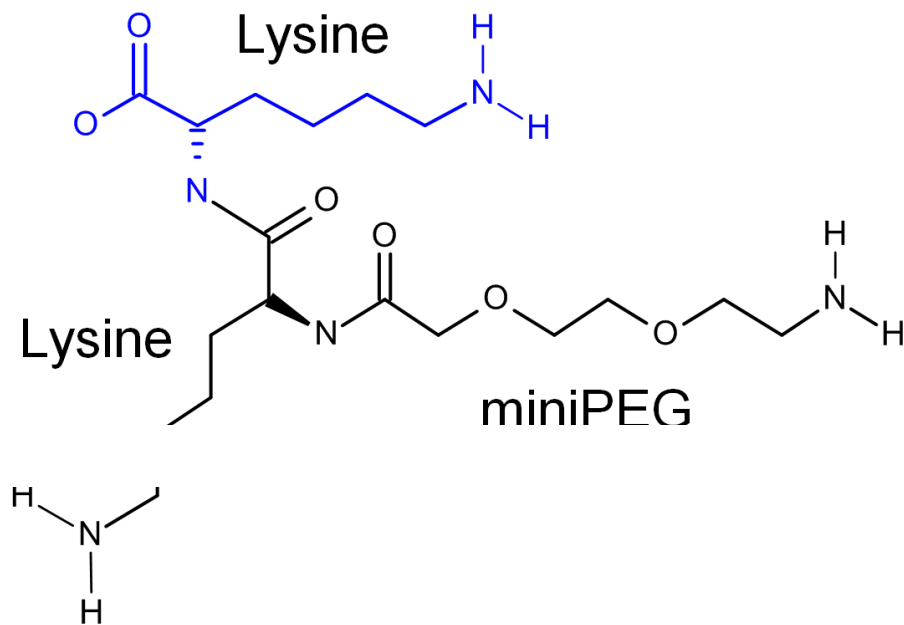


Figure 3. Synthetic scheme used to synthesize SH7153. Solid phase synthesis was performed on a chlorotriptyl resin. For a detailed description of the synthesis, please see the Experimental section

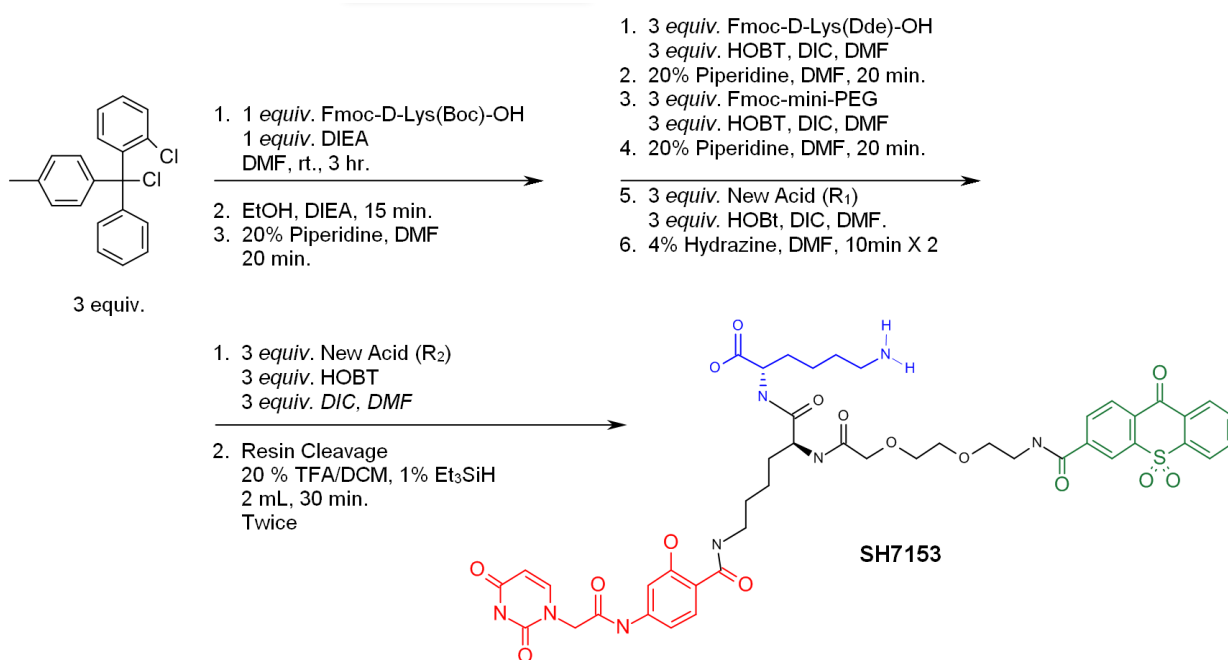


Figure 4. Structures of the two SHAL by-products isolated from the SH7153 preparation by HPLC. SH7153A was found to contain two 689002 ligands linked together. SH7153B was found to contain two 93033 ligands linked together.

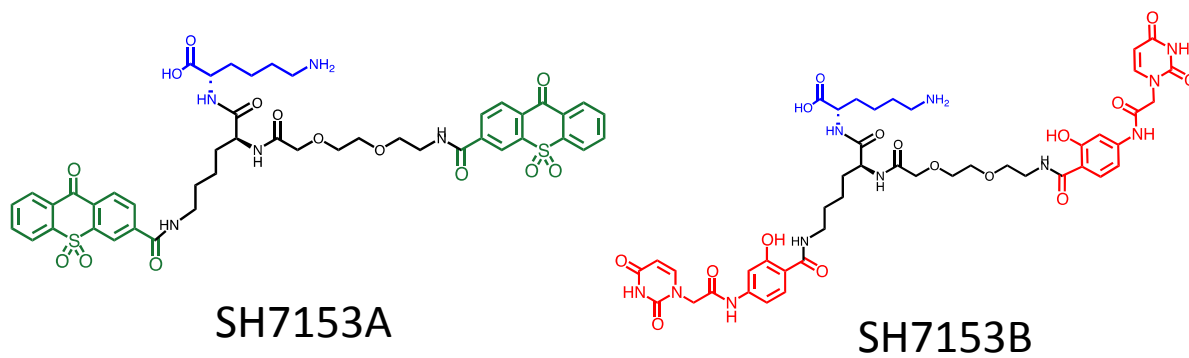
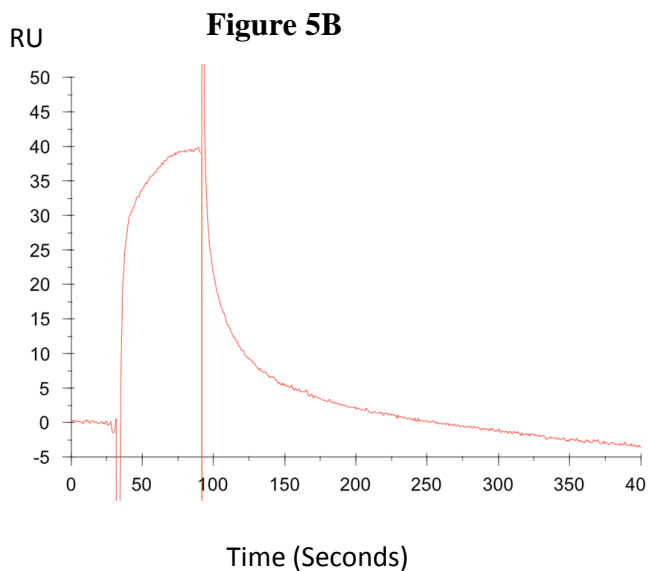
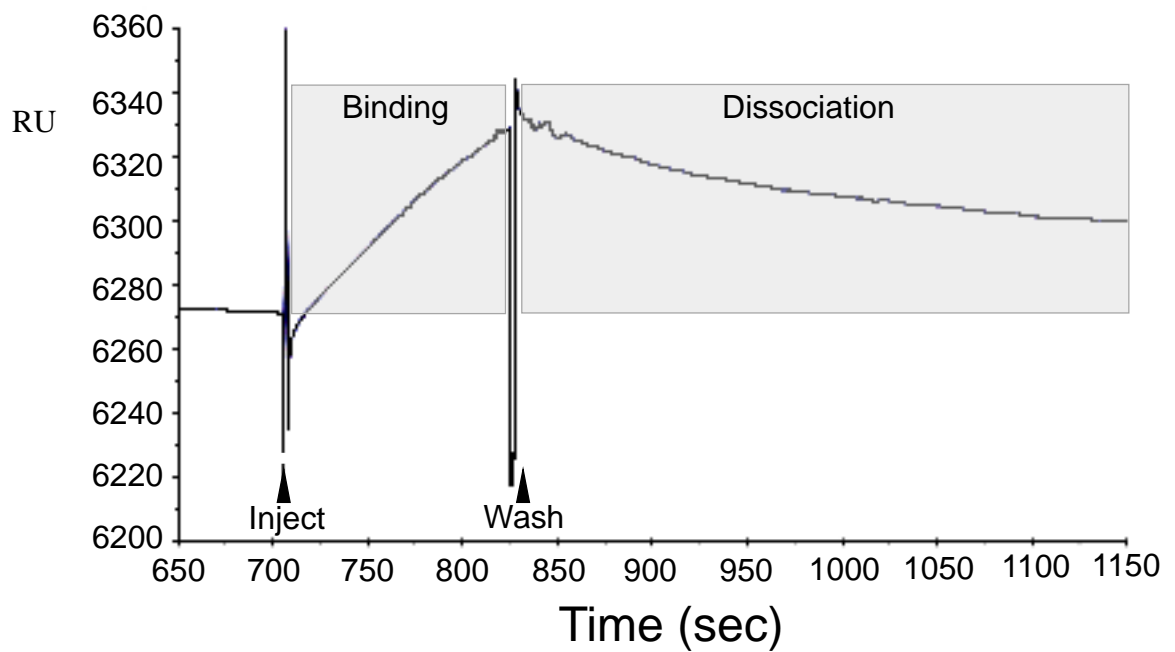


Figure 5A. Binding and dissociation curves for the SHAL SH7153 and its two ligand components binding to recombinant CD81-LEL obtained by surface Plasmon resonance. A detailed description of the experiment is described in the Experimental section. A. Binding and dissociation of SH7153. B. Binding and dissociation of the ligand component 93033.



7. Tables

Table 1. DPI data for ligands 689002 and 93033. Ligands 93033 and 689002 were found to interfere with HCV E2 binding to CD81. Ligand 30930 was used as a negative control.

Added to E2 on Chip	DPI Response in Radians (Rad)
CD81-LEL	2.52
30930	2.52
93033	0.25
689002	-0.23

Table 2. The kinetics study of SH7153, ligands 689002 and 93033.

Ligand or SHAL	Kd (μM)	Off rate (1/sec)
93033	>200	0.00239
689002	>200	0.002631
SH7153	21	2.17E-04

CHAPTER V

281816- A NOVEL DRUG LEAD IDENTIFIED TO INHIBIT HCV INFECTION IN A GENOTYPE-INDEPENDENT MANNER

This chapter is based on a paper titled “Identification of a novel drug lead that inhibits HCV infection and cell-to-cell transmission by targeting the HCV E2 glycoprotein.” that was submitted to Plos one by: Reem R. Al Olaby Laurence Coquerel, Adam Zemla, Jean Dubuisson, Laure Saas, Jost Vielmetter, Joseph Marcotrigiano, Abdul Ghafoor Khan, Felipe Vences Catalan, Alexander L. Perryman, Joel S. Freundlich, Stefano Forli, Shoshana Levy, Rod Balhorn, Hassan M.E. Azzazy.

Objectives

The aims of this study is to create a new homology model of HCV E2 based on the newly resolved crystal structures (PDB IDs: 4MWF, 4NX3), screen ~2000 ligands against the HCV E2 homology model and identifying top virtual screening hits, test the hits using lab on a chip techniques and antibody neutralizing assays, test the ligands with the best binding affinity using HCV infection inhibition assays and do further tests for the drug lead 281816 to determine its effect on cell to cell transmission of HCV and other aspects. The Homology model was created by Dr. Adam Zemla (Lawrence Livermore National Laboratory – USA), the virtual screening runs were conducted in Scripps Research Institute (USA), Dr. Arthur Olson’s molecular graphics laboratory and Rutgers University (USA) at Dr. Joel Freundlich’s lab. The HCV infection inhibition assays were done in Institute Pasteur De Lille (France) at Dr. Jean Dubuisson’s lab and the antibody neutralizing assays were done in Stanford University (USA) at Dr. Shoshana Levy’s lab.

1. Introduction

Hepatitis C Virus (HCV) is a global public health problem [1] in which nearly 85% of affected individuals have acute HCV infections and exhibit no symptoms. In addition, more than three-quarters of these cases will advance to chronic disease, which

include liver cirrhosis and liver cancer [2]. The current standard of care treatment for HCV (Peginterferon/Ribavirin, PR) can cause deleterious side effects, and a sustained virologic response (SVR) is achieved in less than 50% of genotype-1 patients [3]. The FDA approved protease inhibitors Telaprevir (TVR) and Boceprevir (BOC) have been shown to provide higher SVR rates in genotype 1 patients [3, 4] when each is combined with PR. However the poor safety profile of TVR and BOC reported in the Week 16 analysis of the French Early Access Program suggest there is still a need for better HCV drugs [5]. The two most recent FDA approvals have been for the oral drugs Simeprevir and Sofosbuvir, inhibitors that target the HCV NS3/4A protease and polymerase, respectively [6]. Simeprevir, which needs to be administered with ribavirin and peg-interferon, has a number of undesirable side effects [7]. The efficacy of Simeprevir has also been shown to be diminished significantly due to viral breakthrough (HCV RNA rebounds and becomes detectable in the patient before treatment is completed), in patients infected by HCV genotypes 4-6 containing the Q80K, R155K and D168E/V polymorphisms [7]. Recommendations for the use of Sofosbuvir indicate it should be administered with Ribavirin to HCV genotype 2 and 3 infections and that Peg-Interferon should be included in the treatment when infections involve genotypes 1 and 4. While Sofosbuvir is considered the Holy Grail in HCV treatment by some, it is recommended that treatments be limited to 12 weeks [6]. Its high cost (\$1,000 USD/pill) also puts it out of reach of many HCV infected patients. This has led many of the larger pharmaceutical companies to continue developing new drugs that target one or more steps in the HCV life cycle and block virus invasion, processing of the pro-protein or replication of the viral genome.

Several research groups have shown that the CD81-large extracellular loop (CD81-LEL) plays a key role in HCV entry into cells by binding to the HCV envelope glycoprotein 2 (E2) [8-11]. Zhang et al. [12] elucidated a separate, additional function for CD81 in the HCV life cycle. These studies revealed that CD81-LEL is important for efficient HCV genome replication. In addition, the E2-CD81-LEL interaction has been determined to induce several immuno-modulatory effects such as the production and release of pro-inflammatory cytokine gamma interferon from T-cells. In addition, this

interaction has also been shown to down regulate T-cell receptors and suppress the activity of natural killer (NK) cells [13]. Therefore, it is tempting to speculate that blocking the CD81- LEL:HCV E2 interaction might also contribute to arresting disease progression to liver cirrhosis.

Following the discovery of the E2 glycoprotein's role in HCV infection and disease progression, several approaches have been used to attempt to develop anti-HCV drugs and vaccines that target the HCV E2 glycoprotein [14-17]. These efforts have had to deal with challenges that relate to the genomic diversity and heterogeneity of HCV, limitations in animal models used to test vaccines and drugs, and the lack of a resolved crystal structure for the HCV E2 glycoprotein. Recently, two crystal structures have been reported for the core ectodomain of the HCV E2 protein [18, 19]. Kong et al. [18] obtained the structure of amino acid residues 384-746 (E2c) by designing and expressing 41 soluble HCV E2 constructs and selecting 15 to screen against E2-specific Fab fragments in crystallization trials. Using a combination of x-ray crystallography and negative stain-electron microscopy, Kong et al. [18] showed that the structures they obtained for E2 were globular and very different from the predicted models of E2 that were developed using class II fusion protein templates containing three β -sheet domains. Additionally, they were able to identify key CD81-binding residues through mutational studies. Important CD81 binding sites were determined to be in the epitope recognized by the neutralizing antibody AR3C, along one side of the α -sandwich (an isolated region of the CD81-binding loop) and a front layer consisting of loops, short helices and β -sheets [18-20]. AR3C was also found to cross-neutralize HCV genotypes by blocking CD81 binding to HCV E2 [21]. During the preparation of this manuscript, a second structure was reported for E2c (amino acid residues 492-649) by Khan et al. [19]. This new structure, which was obtained by crystallizing E2c in complex with a FAB fragment of the mouse monoclonal antibody 2A12, is highly similar to the previously reported structure. In addition to providing a second structure for the E2 core from a different HCV genotype (2a), new information was also reported on the accessibility of the E2 core amino acids within the structure using a combination of limited proteolytic degradation and deuterium exchange.

Despite the advances in the field of HCV drug development, the current drugs offer little protection against the emergence of genetic variants (escape variants) of HCV – a feature of HCV biology that complicates both drug and vaccine development. Drugs that target only one step in the HCV life cycle may be less effective in treating patients who become infected with these emerging variants. In an effort to develop a suitable drug candidate that targets the majority of the existing HCV genotypes, we developed an HCV E2 homology model based on the new HCV E2 core crystal structure reported by Kong et al. [18] and have used this model to identify small molecule drug leads that target highly conserved sites on the HCV E2 glycoprotein located within the region bound by CD81. AutoDock was used to perform virtual screening runs against 1715 small molecules and 34 of the best compounds were tested experimentally using surface Plasmon resonance (Biacore T100) to identify a set of small molecules that bind to the recombinant E2 protein. The compounds showing binding activity were then tested for their ability to block HCV infection of Huh-7 cells. One compound, 281816, was found to block infection of the cells by each of the HCV genotypes and subtypes tested (1a, 1b, 2a, 2b, 4a and 6a) in a dose-dependent manner. Experiments with Huh-7 cells have shown that both mechanisms that lead to HCV infection, cell-free and cell-to-cell transmission, are abrogated by 281816. Inhibition of cell-free infection is limited to the viral attachment step, as well as interactions occurring during viral internalization and fusion; 281816 appears to have no effect on post-entry processes.

2. Results

a. Structural model of E2

In order to maximize the likelihood that these experiments would lead to the discovery of small molecules that bind to E2 and block E2's binding to CD81, we developed a homology model of the core of the E2 protein containing the domains known to bind to CD81 to use as our docking target. This model was created using the HCV genotype 1a protein sequence NP_751921.1 and the crystal structure of E2c as the primary template (PDB entry: 4MWF) (18). Using a model, rather than the E2c crystal structure, was important because the reported structure of E2c was determined to be unsuitable for

the docking experiments we needed to perform. The atom coordinates listed in PDB chains 4mwf_C and 4mwf_D provide structural information for only 169 and 171 residues respectively out of the 363 amino acids present in the full-length E2 protein. Within each of deposited PDB chains three stretches of amino acid sequence (large loop P453-P491 containing 39 amino acids, T542-G547 or V574-N577, and F586-R596) are missing in the crystal structure (figure 1A). Unfortunately, similar regions are also missing in the crystal structure of the genotype 2a HCV E2c protein (PDB chain: 4nx3_D) reported by Khan et al. [19] which provides atom coordinates for only 119 amino acids. Structural superposition of 4mwf_C and 4nx3_D (figure 1B) shows strong conformational similarities between the experimentally solved structures of the E2 proteins with a root mean square deviation of 1.07 Ångstroms measured on 98 residues for which distances between corresponding C α atoms are under 3 Ångstroms. The most significant structural deviations are observed in the region 566-601 (numbering from 4mwf_C) which corresponds to the region that also exhibits the greatest variation in sequence (see sequence alignment in Figure 1A).

Exhaustive structure similarity searches of 90 residue structural fragments of E2 conducted using the entire PDB database (255,302 PDB chains) revealed that no additional structural homologs could be found at the level of calculated structure similarities by LGA score [22] higher than LGA_S=45%, suggesting that the HCV E2 protein represents a novel fold in the current PDB. Thus, the modeling of the structure of the insertions needed to fill in missing regions in the experimentally solved crystal structure and to complete the model was a difficult task, and it was completed with a very low degree of confidence. By applying a combination of structural modeling and analysis methods to the E2 crystal structure (see Materials and Methods section), we were able to construct a model that met all the requirements needed for docking. This model contains the peptide segments and loops that are missing in the E2c structure (figure 2), the missing fragment with amino acids known to be critical for E2 binding (W487), as well as the exact sequence for the HCV genotype 1a E2 protein. Three regions that have been identified by others to be critical for E2 binding to CD81 [24-26] are contained in the

model in their entirety (figure 2). Currently, however, only three of the twenty-one Region 1 amino acids (H421-N423) are present in the model. A comparison of our model to the two E2c structures (see bar plots in the Figure 1A and superposition of two structures in Figure 2) shows the main core regions are, as one would expect, very similar. The differences that are observed in the core region are small and appear to reflect only minor local deviations between experimentally solved structural templates. The large region that does differ corresponds to the missing peptide segments (Figure 4).

b. Ligands predicted to bind to CD81 binding sites on E2

Five ligand-binding sites on the HCV E2 homology model (Figure 3) were identified by docking of the National Cancer Institute's Diversity Set III library of ligands to the E2 model. Each of these sites is associated with or positioned next to one or more of the amino acid or peptide sequences that have been identified by others to either participate in E2 binding to CD81, E2 binding to E1 or to be important for HCV infectivity. The first sequence of importance is the peptide segment Q412-N423 that was identified to bind to the broadly neutralizing antibody AP33 [20, 27]. Alanine mutagenesis studies have shown all of the amino acids in this region appear to be important for HCV infectivity [24]. The model used in this study currently contains only three of the amino acids that correspond to this segment, H421, I422 and N423. Sequence 2 spans the second hyper-variable domain of E2, extending from amino acid Y474 to R492 [13, 24-26]. The majority of amino acids in this sequence have been shown to have no effect on E2 binding to CD81 when mutated [23], but antibodies binding to this region of the protein do inhibit HCV infectivity [25] and CD81 binding [27]. One amino acid located within this domain, W487, does however appear to be critical for E2 binding to E1. This amino acid is the first residue in one of the WHY motifs that have been reported to play a role in E1:E2 dimerization [26]. The third sequence spans amino acids S522-G551 [20,24-26] and the fourth sequence of importance is comprised of amino acids P612-P619 [25]. Mutations of residues Y527, W529, D535, Y613, R614, W616, H617 and Y618 in these two regions have all been shown to eliminate E2 binding to CD81 [23,25]. Mutating all but three of

these amino acids (D535, R614 and W616) appears to eliminate specific interactions with CD81. W616 is the first amino acid in another WHY motif that is located in a region (G600-C620) that has been shown to be involved in fusion [28]. Alanine mutagenesis of D535, R614 and W616 was found to disrupt the structure of the AR3A epitope and indirectly impact CD81 binding [25].

These five binding sites were used to guide to our selection of the top virtual screening hits to be tested experimentally for binding to recombinant E2 protein. While there is still some debate regarding the importance of the entire domains bound by neutralizing antibodies, amino acid mutagenesis studies have provided a great deal of insight into those amino acids located within the epitopes that participate in E2 binding to CD81. Based on this information, we have used the set of amino acids W420-I422, S424, G523, Y527, W529, G530, D535, P612-R614 and W616-P619, whose mutation has been shown to eliminate E2 binding to CD81, to identify locations within these sites (figure 3) where ligand binding would be expected to disrupt E2's ability to bind to CD81.

Thirty-four of the highest scoring ligands were selected from the docking run for experimental analysis. Conformers of each of the ligands bound to one or more of these five binding sites. The best ligands were considered to be those that exhibited the lowest free energy of binding and were predicted to interact with or bind nearby one or more of the E2 amino acids within the sites that were reported to be critical for E2 binding to CD81. The free energy of binding predicted for the best bound ligand conformations, shown in Table 1, ranged from -6.2 to -8.7 Kcal/mol. Additional criteria used to select among the group of ligands predicted to bind include the number of contact points/interactions (such as hydrogen bonds, salt bridges, Van Der Waals interactions) with amino acids in the model (the larger number of contacts or interactions the better) and the chemical structure of the ligands (preference is given to those that contain a free amino or carboxyl group that is exposed to solvent). Ligands with free amino or carboxyl groups can easily be linked to other ligands to create higher affinity or more selective second-generation inhibitors.

c. Experimental confirmation of ligand binding to HCV E2

Each of the 34 ligands was tested experimentally using surface Plasmon resonance detection (on a Biacore T100 instrument) to determine if it would bind to recombinant HCV E2 protein and to obtain an assessment (relative to the other ligands) of how well it binds. Twenty-three of the molecules provided a positive change in response units (RU) indicating they bound to the E2 protein immobilized on the chip (Table 2). The measured responses for the ligands that bound varied from 54 to 276 RUs. Data was also obtained on the rate of ligand dissociation by measuring the amount of ligand remaining bound at two time points, dissociation 1 (10 seconds) and dissociation 2 (50 seconds), during the rinsing of the chip with buffer (figure 5). The majority of the ligands dissociated quickly, as one might expect for small molecules that bind to the surface of a protein. A few, such as ligands 121861, 4429, 158413, 81462, and 57103, exhibited slower off rates when compared to others.

d. Inhibition of HCV Entry

The 23 compounds that were observed to bind to recombinant E2 protein were then tested to determine if they would block HCV infection of Huh-7 cells. Pseudotyped retroviral particles harboring the envelope protein of an endogenous feline retrovirus (RD114pp) were first used to determine the specificity and the safety of molecules. We excluded from a further characterization the molecules for which the half maximal inhibitory concentration (IC₅₀) against RD114pp was higher than 10 μ M or the molecules that significantly increased RD114pp infection due to exerting non-specific inhibition (Table 3). The remaining ligands were next tested against pseudotyped retroviral particles harboring genotype-2a HCV envelope proteins (HCVpp 2a), cell culture produced HCV particles (HCVcc) or RD114pp. As a positive control, an anti-CD81 antibody was included in the assays. One compound, 281816, showed an inhibitory effect on both HCVpp and HCVcc infection with IC₅₀'s of 1.02 μ M and 3.95 μ M, respectively (Table 3 and Figure 6A), indicating that this molecule inhibits the entry step of the HCV lifecycle, probably through a specific effect on the virus's interaction with CD81. While Huh-7 cell

toxicity was not observed over the range of ligand 281816 concentrations tested in the assays (the highest concentration tested was 10 μ M), a subsequent viability assay (MTS assay) showed a 50% cytotoxic concentration (CC50) for 281816 of 14 μ M (data not shown). MTS is a tetrazolium compound that can be reduced by viable cells to generate formazan products that are directly soluble in cell culture medium.

To determine if 281816 would inhibit HCV genotypes other than 2a, a series of infection assays was performed with HCVpp bearing envelope proteins from a number of different HCV genotypes. Interestingly, 281816 was found to be equally effective in inhibiting Huh-7 infection by all the HCV genotypes tested (1a, 1b, 2a, 2b, 4a and 6a, Figure 6B). The IC50 values ranged from 2.2 μ M to 4.6 μ M (Table 4).

To confirm that 281816 inhibits HCV entry with no further effect on post-entry steps, 281816 (10 μ M) was added at different time points (figure 8A) before (-2 to 0 hr, b), during (0 to 2hr, c), or after (2 to 24hr, d) inoculation of Huh-7 cells with HCVcc, as previously described [29]. Cells treated with dimethylsulfoxide (DMSO) and cells treated continuously (-2 to 24hr, a) with 281816 were used as controls. The results clearly showed that 281816 significantly inhibits HCVcc infection when present during virus infection (figure 8A, c). The decrease in HCVcc infection that was observed in condition b is likely to be due to some 281816 entering into the cells and acting on the entry step (figure 8A, b). Similarly, a slight decrease was also observed in condition d (figure 8A), which is likely related to 281816 acting on the entry of the remaining particles (those entering after 2 hr). Together, these results confirm that 281816 inhibits the entry step of HCV lifecycle.

After attachment to the cell surface and binding to entry factors, HCV virions are internalized by clathrin-mediated endocytosis [30, 31]. Following internalization, HCV is transported to early endosomes along actin stress fibers, where fusion seems to take place [31, 32]. To determine at which step HCV entry is impaired by 281816, we administered the ligand at different intervals during the early phase of infection. Virus attachment and binding were performed at 4°C (figure 8B, Steps 1 and 2). Then, cells were shifted to 37°C to allow endocytosis and fusion (figure 8B, Step 3). Cells treated with JS81 were

used as controls where JS81 antibody binds to CD81 and blocks HCV E2: CD81 interaction (data not shown). The addition of 281816 during step 2 and step 3 led to the strongest inhibition of HCV infection, as strong as the one observed when 281816 was present during all three steps. We also observed a significant inhibition of HCV infection when 281816 was added during the early attachment/binding steps (figure 8B, Step 1). Together, these results indicate that 281816 inhibits HCV infection by acting on more than the first (attachment/binding) step of viral entry. These data suggest the ligand also affects interactions during HCV internalization and fusion.

e. Blocking of E2 binding to CD81

Ligand 281816 was originally selected for testing based on the prediction by docking that it would bind to a site on the HCV E2 protein where CD81 binds. The infection assay conducted with Huh-7 cells demonstrated 281816 was effective in inhibiting the entry step in the HCV life cycle. To confirm that the binding of 281816 to E2 inhibits the HCV E2-CD81 interaction, flow cytometry was used to monitor the binding of a recombinant form of the E2 protein to native CD81 overexpressed on Raji cells as a function of 281816 concentration. The results in Figure 7 show binding of the E2 protein to Raji cells is inhibited by 281816 in a dose dependent manner.

f. 281816 abrogates HCV cell-to-cell transmission

In addition to cell-free infection, HCV can also be transmitted via cell-to-cell contact by a mechanism that is not completely understood [33, 34, 35]. Indeed, HCV is transmitted in the presence of monoclonal antibodies (mAbs) or patient-derived antibodies that are able to neutralize virus-free infectivity [33,35]. Since cell-to-cell transmission has been suggested to be a major route of transmission for HCV [34], we next analyzed the effect of 281816 on this process. For this purpose, Huh-7 cells were infected with HCVcc for 2h and then cultured with neutralizing anti-E2 antibody (3/11) in the presence of 281816 (1 μ M and 10 μ M). Cells cultured in presence of solvent (DMSO) and Epigallocatechin-3-gallate (EGCG, 50 μ M) [35] were used as negative and positive controls of inhibition, respectively. Three days post-infection, foci were

visualized by immunofluorescence (figure 9) and sizes of foci were measured by counting the number of cells per focus. The results showed that 281816 led to a significant reduction of the number of cells per focus in a dose-dependent manner. Together, these results indicate that 281816 also inhibits cell-to-cell transmission of HCV.

3. Discussion

While it has been known for some time that the E2 envelope glycoprotein plays an important role in the life cycle of HCV, we are only now beginning to learn details about the structure of the protein and how it functions. This has been attributed to the challenging intrinsic properties of the protein, such as the presence of multiple flexible loops, its tendency to form disulfide aggregates in solution and the high level of N-linked glycosylation, all of which make it difficult to determine the protein's structure. Neutralizing antibody epitope analyses and mutation studies, however, have provided a great deal of information about the regions of the E2 protein and specific amino acids that participate in CD81 binding and are important for HCV infectivity.

It became possible to use computational docking and structure-based drug design methods to begin developing anti-HCV drugs that target the conserved regions of E2 and block its interaction with host receptors due to two reasons. First, the recent determination of two HCV E2 protein core crystal structures [18,19]. Second, our use of the deposited coordinates to create a new homology model of the protein's structure containing the majority of conserved segments known to be important for viral invasion of hepatocytes. Our docking of a library of diverse small molecules to this homology model led to the identification of a set of ligands that were predicted to bind to sites near key amino acids known to participate in CD81 or E1 binding or block HCV infection, and 23 of the 34 compounds were confirmed by experiment to bind to recombinant E2 protein.

When these 23 ligands were tested for activity in blocking HCV infection of Huh-7 cells, only ligand 281816 was found to inhibit HCV infection using both HCVcc and HCVpp based assays. Upon analyzing the activity spectrum of HCV using HCVpp bearing envelope proteins from different HCV genotypes (1a, 1b, 2a, 2b, 4a and 6a),

281816 was found to inhibit the infection of all tested genotypes with IC₅₀'s ranging from 2.2 μ M to 4.6 μ M (Table 4), indicating that this small molecule inhibits HCV entry in a genotype-independent manner. Ligand 281816 was also observed to block the binding of recombinant E2 protein to Raji cells expressing CD81.

The docking experiments conducted with 281816 identified the two different binding sites on E2 shown in Figure 10. One cluster of 281816 conformers bound deep inside a cavity positioned directly above Y618 and P619, two amino acids in site 4 that are known to contribute to E2's binding to CD81 [26]. The two strongest 281816 ligand binding modes are shown bound to this site. 281816 was also predicted to bind to a shallow cavity on the side of the protein. These conformers were predicted to bind to site 1 near residues V515, G517, P515 and H421-N423. H421-N423 is part of a larger segment of E2 that has been shown to bind to CD81 [19]. As expected, the ligand positioned above Y618 and P619 in the deeper cavity was predicted to bind more strongly to the protein (free energy of binding of the best bound ligand = -8.64 Kcal/mol) than when it was bound to the shallow cavity on the side of the protein (free energy of binding = -6.39 Kcal/mol).

A subset of the 281816 conformers (not shown) in the cluster observed to bind near site 4 overlapped into site 2 and bound immediately adjacent to D481-P490, part of the epitope targeted by antibodies that block HCV infectivity and E2 binding to CD81 [27]. W487, a residue within this peptide segment whose mutation has been shown to disrupt E2:E1 dimerization [22], is also located near site 2. Other conformers in the cluster binding on the side of the protein also bound near site 5 and amino acid residues P612 and Y613. One interesting and unique feature of the 281816 ligand is that a number of its conformers are predicted to bind immediately above or next to the exposed faces of the P612-P619 domain that is known to participate in E2 binding to CD81 [26]. The limited proteolysis and deuterium exchange experiments conducted with the E2 protein core and reported by Khan et al. [19] indicate these 281816 binding sites are accessible and exposed to solvent.

To probe more deeply into the inhibition of the infection process by 281816, experiments were also run to determine if the inhibition of cell-free infection by 281816 might be limited to viral entry, which step in the entry process might be affected by the compound, and what impact, if any, 281816 might have on cell-to-cell transmission of HCV. Analyses of Huh-7 cells inoculated with HCVcc before, during or after treatment with 281816 revealed the compound only blocks HCV entry and does not inhibit post-entry processes in the HCV life cycle. A kinetic analysis of the effect of 281816, coupled with a temperature block to endocytosis and fusion, was used to examine the cell-free entry steps in more detail and showed 281816 inhibits not only the initial attachment/binding step, but it also has an effect on interactions that occur later during viral internalization and fusion. Ligand 281816 was also observed to abrogate the cell-to-cell transmission of HCV. 281816 treatment of Huh-7 cells cultured in the presence of the anti-E2 neutralizing antibody 3/11 not only led to a dose dependent reduction in the number of cells forming foci, but it was found to be more effective in blocking cell-to-cell transmission than the Epigallocatechin-3-gallate [35] used as a positive control.

In addition to identifying a promising new small molecule drug lead for treating HCV that targets the E2 glycoprotein, this study also provides a demonstration of the utility of our new E2 homology model in the discovery of small molecules that bind to important sites on E2. By targeting sites containing amino acid residues identified by others to participate in CD81 binding and CD81-dependent processes that impact HCV infectivity, a small molecule was identified that not only blocks E2 binding to CD81 and the cell-free entry process, but it is also effective in blocking the cell-to-cell transmission of HCV – the predominant mechanism of transmission that contributes to the persistence of infections [34]. This effect of 281816 corroborates other reports of a CD81-dependent cell-to-cell transmission process [61,61] that can be blocked by anti-CD81 antibodies [34,66] and soluble CD81 [32]. 281816, known as methiothepin or 1-methyl-4-(3-methylsulfanyl-5,6-dihydrobenzo[b][1]benzothiepin-5-yl)piperazine, is also interesting because it has been determined previously to block dopamine [36] and serotonin [37] receptors and has been shown to inhibit a number of other biological activities, which

include the binding or entry of two other unrelated viruses into cells (Lassa [38], Marburg [39]), *Plasmodium falciparum* proliferation [40], and *Mycoplasmodium tuberculosis* infections [41].

4. Materials and Methods

a. Creation of the homology model of E2 used for docking

A crystal structure of E2c deposited in the PDB under a code 4MWF was resolved by Kong et al. [18] at a resolution of 2.65 Angstroms. However, upon examination of the structure file prior to docking, the set of reported atom coordinates of the protein was found to be incomplete. In addition to the coordinate file containing structural information for only 171 residues out of the 363 amino acids present in the full-length protein, structural information was missing for several peptide segments or loops within the structural core of the protein. In order to prepare a more complete version of the structure for docking, we have performed several homology modeling and structure analysis tasks using the coordinates of E2c as a template. The final structural model was created using the AS2TS system [42] based on atom coordinates from the PDB chains 4mwf_C and 4mwf_D. A structural search for similar fragments in proteins in the PDB that could be used to model missing loop regions was performed using the StralSV algorithm [43], which identifies protein structures that exhibit structural similarities despite low primary amino acid sequence similarity. The side-chain prediction was accomplished using SCWRL [44] when residue-residue correspondences did not match. Residues that were identical in the template and E2 protein were copied from the template onto the model. Potential steric clashes were identified in the unrefined model using a contact-dot algorithm in the MolProbity software package [45], and the constructed model was finished with relaxation using UCSF Chimera [46].

b. Virtual screen of the NCI Diversity Set III to the HCV E2 protein model

AutoDock VINA 1.1.2 (VINA) [47] was used to perform a virtual screen of the NCI Diversity Set III against the homology model that was created using the new crystal structure developed by Kong et al. [18] (PDB ID: 4MWF) as a template. The model of

the protein was prepared for docking using the MolProbity Server (to add all of the hydrogen atoms and to flip the HIS/ASN/GLN residues if doing so significantly lowered the energy) and AutoDockTools 4.2 (which added the Gasteiger-Marsili charges and merged the non-polar hydrogens onto their respective heavy atoms) [48,49]. The NCI Diversity Set III library containing 1,715 models of compounds was obtained from the ZINC server (<http://zinc.docking.org>) [50]. The multi-molecule “mol2” files from ZINC were prepared for docking calculations using Raccoon [51], which added the Gasteiger-Marsili charges, merged the non-polar hydrogen atoms onto their respective heavy atoms, and determined which bonds should be allowed to freely rotate during the calculations, to generate the “pdbqt” docking input format.

Four different, overlapping grid boxes were used in this virtual screen to enable the docking calculations to explore almost the entire surface of this E2 model (except for the large, flexible loop that was added to the model and the relatively flat surface near it). Since large grid boxes were used in these calculations, the “exhaustiveness” setting in VINA was increased to 20. Each calculation used 8 CPUs on the Linux cluster at Rutgers University-NJMS. The first box, which included P490, was centered at 38.829, 12.968, -40.958 (x, y, z) and had the following dimensions: 24.0 x 35.0 x 30.0 (x, y, z in Angstroms). The second grid box, which included G436, was centered at 48.401, 11.791, -14.449 (x, y, z) and had a size of: 32.0 x 36.0 x 24.0 (x, y, z in Angstroms). The third grid box, which included S528, was centered at 51.644, 25.877, -27.795 (x, y, z) and encompassed 30.0 x 30.0 x 30.0 Angstroms. The fourth grid box, which was selected to include the side of E2 not covered by the previous three grid boxes, was centered at 57.777, 12.968, -34.067 (x, y, z) and enclosed 24.0 x 35.0 x 32.0 Angstroms (x, y, z).

The docking outputs generated by VINA were processed and filtered using python scripts from Raccoon2 and Fox [51]. The top-ranked VINA mode from each docking calculation was harvested, and 17 different sets of energetic and interaction-based filters were investigated to harvest the most promising docking results for visual inspection. The following parameters were explored in the filtering process: -e indicates

the minimum estimated Free Energy of Binding from the VINA score in kcal/mol, -l is the minimum ligand efficiency value in kcal/mol/heavy atom, -S is the minimum number of hydrogen bonds between the ligand and target, and -H indicates that the ligand had to form a hydrogen bond with either a backbone amino group (::N) or a backbone carbonyl oxygen (::O) of any residue in that grid box.

For the results with grid box 1, filters 12 and 13 each harvested 70 and 51 compounds, respectively. Those filtered sets were pooled together to form a set of 96 unique compounds for visual inspection. Filters 14 (which harvested 11 compounds), 15 (which harvested 21 compounds), and 1 (which harvested 34 compounds) were pooled together from the results with grid box 2, in order to identify 52 compounds for visual inspection. Similarly, for the results with grid box 3, filters 1 (which harvested 25 compounds), 14 (which identified 20 candidates), and 15 (which harvested 13 compounds) were pooled to obtain 34 compounds for visual inspection. To identify candidates in the results with grid box 4, filters 1 (which harvested 26 compounds), 14 (which harvested 19 compounds), and 15 (which harvested 14 compounds) were pooled to obtain 42 compounds. These four different pools of potentially promising compounds were then visually inspected to select the ligands to be tested experimentally for binding to recombinant E2 protein.

c. Expression and purification of the HCV E2 protein Con1eE2

A construct containing a sequence encoding amino acids 384-656 of the Con1 envelope protein 2 ectodomain (eE2) [19], a genotype 1 E2 sequence, was cloned into a lentiviral expression vector containing a carboxy-terminal Protein A tag separated by a PreScission Protease cleavage consensus sequence. eE2-ProtA was stably expressed in HEK293T cells using lentiviral infection. The protein was secreted into the media and supernatants were purified using IgG Sepharose (GE Healthcare, Piscataway, NJ). eE2-ProtA was eluted with 100 mM sodium citrate and 20 mM KCl at pH 3 directly into tubes containing 1M Tris pH 9 for immediate neutralization. PreScission Protease (GE Healthcare, Piscataway, NJ) was added to the eluted sample at a ratio of 1:50

(enzyme:eE2), and the digest was then dialyzed into 20 mM HEPES pH 7.5, 250 mM NaCl, 5% glycerol. eE2 was separated from the cleaved tag and the PreScission Protease by ion exchange chromatography [19].

d. Experimental analysis of ligand binding to recombinant E2 by surface Plasmon resonance (SPR) detection

A set of 34 of the ligands predicted by AutoDock to bind to E2 were tested experimentally to determine if they bound to recombinant E2 protein immobilized on a chip using surface Plasmon resonance detection. The SPR analyses were performed using a Biacore T100 workstation (GE Healthcare, NJ, USA) and recombinant HCV E2 protein. 1 μ M HCV E2 was diluted into 10 mM sodium acetate buffer pH 5 and immobilized for 15 min at a flow speed of 5 μ l/min onto a CM5 sensor chip using amine coupling (EDC-NHS). Approximately 10,000 response units (RU) of protein were immobilized on the chip. His-CD81-LEL (Bioclone Inc, San Diego, CA) binding to HCV E2 was tested as a positive control prior to injecting the ligands to confirm the E2 protein was functional and would bind CD81-LEL. In a typical experiment with CD81, 1 μ l of his-CD81 (50nM) in 114 μ l PBS was injected into channel 2 and 106.4 RUs of CD81 bound to the E2 on the chip. This was followed by testing the binding of the 34 virtual screening hits where the ligands were prepared as 200 μ M solutions in PBS and they were introduced to the protein using a pre-programmed 3 min association and 1 min dissociation interval. The response was measured at two time points during dissociation, 10 and 50 seconds, to obtain information on the rate of ligand dissociation from E2.

e. HCV infection assays

Pseudotyped retroviral particles harboring HCV envelope proteins (HCVpp) from different genotypes were produced as described previously [52, 53] with plasmids kindly provided by F.L. Cosset, J. Ball, and R. Bartenschlager. A plasmid encoding the feline endogenous virus RD114 glycoprotein [54] was used for the production of RD114pp. Both HCVpp and RD114pp expressed *Firefly* luciferase.

The cell culture-produced HCV particles (HCVcc) used in this study were based on the JFH1 strain [55] and were prepared as described previously [56, 57]. They were engineered to express the A4 epitope, titer-enhancing mutations and *Gaussia* luciferase [56,57].

To identify ligands that inhibit HCV infection, Huh-7 cells were seeded in 96-well plates and treated the day after with six different concentrations of each ligand diluted in DMSO in duplicate using a Zephyr automated liquid handling workstation (Caliper BioSciences, Hopkinton, MA). The final concentration of DMSO (1%) was adjusted to be the same for all ligand concentrations. Cells treated with DMSO were used as negative controls. Cells treated with different concentrations of anti-CD81 (JS-81 from BD Pharmingen, San Jose, CA) 1 hour before infection, were also used as positive controls. The third day, RD114pp, HCVpp or HCVcc were inoculated and incubated for 30 hours at 37°C. *Firefly* and *Gaussia* luciferase assays were performed as indicated by the manufacturer (Promega, San Luis Obispo, CA).

The analysis of the effect of 281816 ligand on Huh-7 infection by HCVpp bearing envelope proteins from different genotypes was performed in 24-well plates using the method described above. This ligand was also screened for toxicity to the cells using the MTS (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2Htetrazolium) assay [58].

f. Inhibition of Recombinant E2 binding to native CD81

The human B cell line Raji (ATCC, Manassas, VA), which expresses high levels of CD81 on its surface, was used to determine if ligands inhibit the binding of HCV-E2 protein to native CD81. Purified HCV-E2 protein (4 µg) was pre-incubated with 1,5,15, 50, 100 or 400 µM of the ligand 281816 for 25 min at RT. After pre-incubation the E2-ligand complex was added to the cells and incubated for 25 min. The complexes were washed from the cells and 0.5 µg of anti E2 antibody (clone H53) was added followed by secondary anti mouse-FITC (Southern Biotechnology, Birmingham, AL). The cells were washed, fixed with 3% paraformaldehyde, and analyzed by flow cytometry (BD FACS

Calibur, software: Cell Quest Pro) analysis. The mean fluorescence intensity (MFI) was calculated using Flowjo software (TreesStar, www.flowjo.com).

g. Antibodies

Mouse anti-E1 A4 [59] and rat anti-E2 3/11 [60] were produced *in vitro* using a MiniPerm apparatus (Heraeus Instruments - Germany). Alexa555-conjugated goat anti-mouse immunoglobulins were obtained from Jackson ImmunoResearch (West Grove, PA). JS-81 (PharMingen, Oxford, England) was used according to the manufacturers' protocols.

h. HCVcc cell-to-cell transmission assay

Huh-7 cells were seeded on coverslips and infected with HCVcc for 2h at 37°C. Cells were then washed and cultured for 72h at 37°C in culture medium containing neutralizing anti-E2 antibody (3/11 mAb ; 50ug/ml) in presence of 281816 at the indicated concentrations. Cells cultured in presence of DMSO and Epigallocatechin-3-gallate (EGCG, 50 µM) [35] were used as negative and positive controls of inhibition, respectively. Cells were fixed with formalin solution (formaldehyde 4%, Sigma, St Louis, MO), and foci detected by indirect immunofluorescence using the anti-E1 monoclonal antibody A4.

i. Kinetics of entry

Cells treated with 281816 at 10µM or with DMSO were infected with HCVcc for 1h at 4°C (attachment/binding period). Virus was removed, cells were washed with medium and incubated again for 1h at 4°C (post-attachment/binding period). Cells were then washed and incubated for 1h at 37°C (endocytosis/fusion period). Lastly, cells were washed and incubated in complete culture medium for 21h. Infection levels were monitored by measuring luciferase activities.

5. References

1. Anwar MI, Rahman M, Hassan MU, Iqbal M (2013) Prevalence of active hepatitis C virus infections among general public of lahore, Pakistan. *Virology* 10(1): 351-422X-10-351. 10.1186/1743-422X-10-351; 10.1186/1743-422X-10-351.
2. Blackard JT, Shata MT, Shire NJ, Sherman KE (2008) Acute hepatitis C virus infection: A chronic problem. *Hepatology* 47(1): 321-331. 10.1002/hep.21902.
3. Zeuzem S, Berg T, Moeller B, Hinrichsen H, Mauss S, et al. (2009) Expert opinion on the treatment of patients with chronic hepatitis C. *J Viral Hepat* 16(2): 75-90. 10.1111/j.1365-2893.2008.01012.x.
4. Marks KM, Jacobson IM (2012) The first wave: HCV NS3 protease inhibitors telaprevir and boceprevir. *Antivir Ther* 17(6 Pt B): 1119-1131. 10.3851/IMP2424; 10.3851/IMP2424.
5. Colombo M, Fernandez I, Abdurakhmanov D, Ferreira PA, Strasser SI, et al. (2013) Safety and on-treatment efficacy of telaprevir: The early access programme for patients with advanced hepatitis C. *Gut* . 10.1136/gutjnl-2013-305667; 10.1136/gutjnl-2013-305667.
6. Asselah T (2014) Sofosbuvir for the treatment of hepatitis C virus. *Expert Opin Pharmacother* 15(1): 121-130. 10.1517/14656566.2014.857656; 10.1517/14656566.2014.857656.
7. Lenz O, Vijgen L, Berke JM, Cummings MD, Fevery B, et al. (2013) Virologic response and characterization of HCV genotype 2-6 in patients receiving TMC435 monotherapy (study TMC435-C202). *J Hepatol* 58(3): 445-451. 10.1016/j.jhep.2012.10.028; 10.1016/j.jhep.2012.10.028.
8. Pileri P, Uematsu Y, Campagnoli S, Galli G, Falugi F, et al. (1998) Binding of hepatitis C virus to CD81, *Science* 282: 938.
9. Petracca R, Falugi F, Galli G, Norais N, Rosa D, et al. (2000) Structure-function analysis of hepatitis C virus envelope-CD81 binding. *J Virol* 74(10): 4824-4830.
10. Higginbottom A, Quinn ER, Kuo CC, Flint M, Wilson LH, et al. (2000) Identification of amino acid residues in CD81 critical for interaction with hepatitis C virus envelope glycoprotein E2. *J Virol* 74(8): 3642-3649.
11. Drummer HE, Wilson KA, Pountourios P (2002) Identification of the hepatitis C virus E2 glycoprotein binding site on the large extracellular loop of CD81. *J Virol* 76(21): 11143-11147.
12. Zhang YY, Zhang BH, Ishii K, Liang TJ (2010) Novel function of CD81 in controlling hepatitis C virus replication. *J Virol* 84(7): 3396-3407. 10.1128/JVI.02391-09.
13. Ahlenstiel G (2013). The natural killer cell response to HCV infection. *Immune Netw* 13(5): 168-176. 10.4110/in.2013.13.5.168.

14. El-Awady MK, Tabll AA, El-Abd YS, Yousif H, Hegab M, et al. (2009) Conserved peptides within the E2 region of hepatitis C virus induce humoral and cellular responses in goats. *Virology* 6: 66-422X-6-66. 10.1186/1743-422X-6-66; 10.1186/1743-422X-6-66.
15. Carlsen TH, Scheel TK, Ramirez S, Fountis SK, Bukh J (2013) Characterization of hepatitis C virus recombinants with chimeric E1/E2 envelope proteins and identification of single amino acids in the E2 stem region important for entry. *J Virol* 87(3): 1385-1399. 10.1128/JVI.00684-12; 10.1128/JVI.00684-12.
16. Li YP, Kang HN, Babiuk LA, Liu Q (2006) Elicitation of strong immune responses by a DNA vaccine expressing a secreted form of hepatitis C virus envelope protein E2 in murine and porcine animal models. *World J Gastroenterol* 12(44): 7126-7135.
17. Ray R, Meyer K, Banerjee A, Basu A, Coates S, et al. (2010) Characterization of antibodies induced by vaccination with hepatitis C virus envelope glycoproteins. *J Infect Dis* 202(6): 862-866. 10.1086/655902; 10.1086/655902.
18. Kong L, Giang E, Nieuwma T, Kadam RU, Cogburn KE, et al. (2013) Hepatitis C virus E2 envelope glycoprotein core structure. *Science* 342(6162): 1090-1094. 10.1126/science.1243876.
19. Khan AG, Whidby J, Miller MT, Scarborough H, Zatorski AV, et al. (2014) Structure of the core ectodomain of the hepatitis C virus envelope glycoprotein 2. *Nature*. 10.1038/nature13117.
20. Law M, Maruyama T, Lewis J, Giang E, Tarr AW, et al. (2008) Broadly neutralizing antibodies protect against hepatitis C virus quasispecies challenge. *Nat Med* 14(1): 25-27. 10.1038/nm1698.
21. Zemla A (2003) LGA: A method for finding 3D similarities in protein structures. *Nucleic Acids Res* 31(13): 3370-3374.
22. Yi M, Nakamoto Y, Kaneko S, Yamashita T, Murakami S (1997) Delineation of regions important for heteromeric association of hepatitis C virus E1 and E2. *Virology* 231:119.
23. Owsianka AM, Timms JM, Tarr AW, Brown RJ, Hickling TP, et al. (2006) Identification of conserved residues in the E2 envelope glycoprotein of the hepatitis C virus that are critical for CD81 binding. *J Virol* 80(17): 8695-8704. 10.1128/JVI.00271-06.
24. Roccasecca R, Ansuini H, Vitelli A, Meola A, Scarselli E, et al. (2003) Binding of the hepatitis C virus E2 glycoprotein to CD81 is strain specific and is modulated by a complex interplay between hypervariable regions 1 and 2, *J Virology* 77:1856.
25. Rothwangl KB, Manicassamy B, Uprichard SL, Rong L (2008) Dissecting the role of putative CD81 binding regions of E2 in mediating HCV entry: Putative CD81

- binding region 1 is not involved in CD81 binding. *Virology* 5: 46-422X-5-46. 10.1186/1743-422X-5-46; 10.1186/1743-422X-5-46.
26. Tarr AW, Owsianka AM, Timms JM, McClure CP, Brown RJ, et al. (2006) Characterization of the hepatitis C virus E2 epitope defined by the broadly neutralizing monoclonal antibody AP33, *Hepatology* 43: 592.
 27. Lavillette D, Pecheur EI, Donot P, Fresquet J, Molle J, et al (2007) Characterization of fusion determinants points to the involvement of three discrete regions of both E1 and E2 glycoproteins in the membrane fusion process of hepatitis C virus, *J Virology* 81:8752.
 28. Goueslain L, Alsaleh K, Horellou P, Roingeard P, Descamps V, et al (2010) Identification of GBF1 as a cellular factor required for hepatitis C virus RNA replication, *J Virology* 84: 773.
 29. Blanchard E, Belouzard S, Goueslain L, Wakita T, Dubuisson J, et al. (2006) Hepatitis C virus entry depends on clathrin-mediated endocytosis. *J Virol* 80(14): 6964-6972. 10.1128/JVI.00024-06.
 30. Meertens L, Bertaux C, Dragic T (2006) Hepatitis C virus entry requires a critical postinternalization step and delivery to early endosomes via clathrin-coated vesicles. *J Virol* 80(23): 11571-11578. 10.1128/JVI.01717-06.
 31. Coller KE, Berger KL, Heaton NS, Cooper JD, Yoon R, et al. (2009) RNA interference and single particle tracking analysis of hepatitis C virus endocytosis. *PLoS Pathog* 5(12): e1000702. 10.1371/journal.ppat.1000702; 10.1371/journal.ppat.1000702.
 32. Timpe JM, Stamataki Z, Jennings A, Hu K, Farquhar MJ, et al. (2008) Hepatitis C virus cell-cell transmission in hepatoma cells in the presence of neutralizing antibodies. *Hepatology* 47(1): 17-24. 10.1002/hep.21959.
 33. Witteveldt J, Evans MJ, Bitzegeio J, Koutsoudakis G, Owsianka AM, et al. (2009) CD81 is dispensable for hepatitis C virus cell-to-cell transmission in hepatoma cells. *J Gen Virol* 90(Pt 1): 48-58. 10.1099/vir.0.006700-0; 10.1099/vir.0.006700-0.
 34. Brimacombe CL, Grove J, Meredith LW, Hu K, Syder AJ, et al. (2011) Neutralizing antibody-resistant hepatitis C virus cell-to-cell transmission. *J Virol* 85(1): 596-605. 10.1128/JVI.01592-10; 10.1128/JVI.01592-10.
 35. Calland N, Albecka A, Belouzard S, Wychowski C, Duverlie G, et al. (2012) (-)-Epigallocatechin-3-gallate is a new inhibitor of hepatitis C virus entry. *Hepatology* 55(3): 720-729. 10.1002/hep.24803; 10.1002/hep.24803.
 36. HTS assay for allosteric agonists of the human D1 dopamine receptor: Primary screen for antagonists. NIH Molecular Libraries Probe Production Network. BioAssay AID 488983. Accessed on March 21, 2014. Available at: <https://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=488983>
 37. Antagonists at human 5-hydroxytryptamine receptor 5-HT_{1E}. Extracted from literature and IUPHAR database. BioAssay 624232. Accessed on March 21,

2014. Available at:

<https://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=624232>

38. qHTS for inhibitors of binding or entry into cells for Lassa virus. NIH Molecular Libraries Probe Production Network. BioAssay 540256. Accessed on March 21, 2014. Available at:
<https://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=540256>
39. qHTS for inhibitors of binding or entry into cells for Marburg virus. NIH Molecular Libraries Probe Production Network. BioAssay 540256. Accessed on March 21, 2014. Available at:
<https://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=720532>
40. qHTS for inhibitors of Plasmodium falciparum proliferation. NIH National Institute of Allergy and Infectious Diseases, Xinzhuan Su. BioAssay 504749. Accessed on March 21, 2014. Available at:
https://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?mid=504749_53
41. High throughput screen to identify inhibitors of Mycobacterium tuberculosis H37Rv. Southern Research Institute, Birmingham, AL. E. Lucile White. BioAssay AID 1332. Accessed on March 21, 2014. Available at:
<https://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=1332>
42. Zemla A, Zhou CE, Slezak T, Kuczmarski T, Rama D, et al. (2005) AS2TS system for protein structure modeling and analysis. *Nucleic Acids Res* 33(Web Server issue): W111-5. 10.1093/nar/gki457.
43. Zemla AT, Lang DM, Kostova T, Andino R, Ecale Zhou CL (2011) StralSV: assessment of sequence variability within similar 3D structures and application to polio RNA-dependent RNA polymerase, *BMC Bioinformatics* 12:226.
44. Krivov GG, Shapovalov MV, Dunbrack Jr RL (2009) Improved prediction of protein side-chain conformations with scwrl4. *Proteins* 77: 778.
45. Chen VB, Arendall 3rd WB, Headd JJ, Keedy DA, Immormino RM, et al. (2010) MolProbity: all-atom structure validation for macromolecular crystallography, *Acta Crystallogr D Biol Crystallogr* 66:12.
46. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, et al. (2004) UCSF chimera-a visualization system for exploratory research and analysis, *J Comput Chem* 25: 1605.
47. Trott O, Olson AJ (2010) AutoDock vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* 31(2): 455-461. 10.1002/jcc.21334.
48. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino, RM et al. (2010) MolProbity: All-atom structure validation for macromolecular crystallography, *Acta Cryst D*66: 12.
49. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, et al. (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem* 30(16): 2785-2791. 10.1002/jcc.21256.

50. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG (2012) ZINC: A free tool to discover chemistry for biology. *J Chem Inf Model* 52: 1757.
51. Forli S Raccoon <<http://autodock.scripps.edu/resources/raccoon>>. Accessed 2013. Molecular Graphics Laboratory, The Scripps Research Institute, La Jolla, CA, 2010.
52. Bartosch B, Bukh J, Meunier JC, Granier C, Engle RE, et al. (2003) *In vitro* assay for neutralizing antibody to hepatitis C virus: Evidence for broadly conserved neutralization epitopes. *Proc Natl Acad Sci U S A* 100(24): 14199-14204. 10.1073/pnas.2335981100.
53. Op De Beeck A, Voisset C, Bartosch B, Ciczora Y, Cocquerel L, et al. (2004) Characterization of functional hepatitis C virus envelope glycoproteins. *J Virol* 78(6): 2994-3002.
54. Sandrin V, Boson B, Salmon P, Gay W, Negre D, et al. (2002) Lentiviral vectors pseudotyped with a modified RD114 envelope glycoprotein show increased stability in sera and augmented transduction of primary lymphocytes and CD34+ cells derived from human and nonhuman primates. *Blood* 100(3): 823-832. 10.1182/blood-2001-11-0042.
55. Wakita T, Pietschmann T, Kato T, Date T, Miyamoto M, et al. (2005) Production of infectious hepatitis C virus in tissue culture from a cloned viral genome. *Nat Med* 11(7): 791-796. 10.1038/nm1268.
56. Rocha-Perugini V, Montpellier C, Delgrange D, Wychowski C, Helle F, et al. (2008) The CD81 partner EWI-2wint inhibits hepatitis C virus entry. *PLoS One* 3(4): e1866. 10.1371/journal.pone.0001866; 10.1371/journal.pone.0001866.
57. Delgrange D, Pillez A, Castelain S, Cocquerel L, Rouille Y, et al. (2007) Robust production of infectious viral particles in huh-7 cells by introducing mutations in hepatitis C virus structural proteins. *J Gen Virol* 88(Pt 9): 2495-2503. 10.1099/vir.0.82872-0.
58. Malich G, Markovic B, Winder, C (1997) The sensitivity and specificity of the MTS tetrazolium assay for detecting the *in vitro* cytotoxicity of 20 chemicals using human cell lines, *Toxicology* 124: 179.
59. Dubuisson J, Hsu HH, Cheung RC, Greenberg HB, Russell DG, et al. (1994) Formation and intracellular localization of hepatitis C virus envelope glycoprotein complexes expressed by recombinant vaccinia and sindbis viruses. *J Virol* 68(10): 6147-6160.
60. Flint M, von Hahn T, Zhang J, Farquhar M, Jones CT, et al. (2006) Diverse CD81 proteins support hepatitis C virus infection. *J Virol* 80(22): 11331-11342. 10.1128/JVI.00104-06.
61. Catanese MT, Loureiro J, Jones CT, Dorner M, von Hahn T, et al. (2013) Different requirements for scavenger receptor class B Type 1 in hepatitis C virus cell-free versus cell-to-cell transmission. *J Virol* 87 (15): 8282-8293.

62. Fofana I, Xiao F, Thumann C, Turek M, Zona L, et al. (2013) A novel monoclonal anti-CD81 antibody produced by genetic immunization efficiently inhibits hepatitis C virus cell-to-cell transmission. *PLOS One* 8(5): e64221.

6. Tables

Table 1. Ligands predicted to bind to the HCV E2 protein by blind docking of the NCI Diversity set III small molecule library to the HCV E2 structural model.

Ligand ID	Free Energy of Binding (Kcal/mol)	Ligand ID	Free Energy of Binding (Kcal/mol)
670283	-7.69	211490	-8.7
86467	-7.47	113486	-6.26
639174	-7.81	144694	-7.27
81462	-6.81	4429	-7.3
403379	-7.58	133071	-7.5
213700	-7.89	163910	-7.4
359472	-7.91	54709	-7.3
146554	-7.67	135618	-8.7
204232	-8.54	281254	-6.5
281816	-8.64	319990	-7.4
308835	-8.4	369070	-6.3
60785	-7.48	59620	-7.3
84100	-6.99	38968	-3.9
158413	-7.9	171303	-5.8
57103	-6.36	228155	-8.7
121861	-8.16	13316	-6.8
3076	-7.71	117268	-7.6

Table 2. Magnitude of surface Plasmon resonance binding response obtained for the 23 ligands that were identified to bind to recombinant E2 protein immobilized on a CM5 sensor chip. The rate of ligand dissociation is assessed by measuring the response units at two time points (10 sec and 50 sec) after the chip with bound ligand is rinsed with buffer.

Ligand ID	Binding (RU)	Dissociation 1 (RU)	Dissociation 2 (RU)
670283	54.3	4	1.4
86467	54.9	1.9	0.8
639174	55.4	2.3	0.6
81462	57.2	9.2	6.5
403379	58	2.8	1.1
213700	62	3.1	0.8
359472	62	2.5	0.8
146554	63.4	3.1	0.8
204232	63.4	2.5	0.4
281816	64.5	3.7	0.9
308835	64.8	7.1	5.2
60785	70.4	2.8	0.6
84100	71.2	4.2	2.2
158413	71.2	10.3	8.5
57103	81.6	11.4	2.5
121861	88.4	26.1	20.4
117268	88.5	4.1	1.2
3076	92.2	3.2	1.6
211490	102.9	6.1	2.1
113486	104.7	7	2.6
144694	118.8	6	2.3
4429	155.3	28.9	14.2
133071	276.3	1.8	-2

Table 3. The IC₅₀ values obtained for the 23 ligands screened for their ability to inhibit HCVcc, HCVpp and RD114pp infection of Huh-7 cells. ND refers to molecules for which the half maximal inhibitory concentration (IC₅₀) against RD114pp was higher than 10 μ M or the molecules that significantly increased RD114pp infection.

Ligand ID	IC ₅₀ (μ M)		
	RD114 pp	HCVpp	HCVcc
670283	3	ND	ND
86467	>10	>10	>10
639174	0.03	ND	ND
81462	>10	>10	>10
403379	>10	>10	>10
213700	>10	>10	>10
359472	>10	>10	>10
146554	>10	ND	ND
204232	>10	>10	>10
281816	>10	1.02	3.95
308835	>10	>10	>10
60785	3.5	ND	ND
84100	>10	>10	>10
158413	>10	>10	>10
57103	0.3	ND	ND
121861	>10	>10	>10
117268	0.1	>10	>10
3076	0.25	ND	ND
211490	0.5	ND	ND
113486	>10	>10	>10
144694	>10	>10	>10
4429	>10	>10	>10
133071	0.10	ND	ND
Anti-CD81	>10	0.17	0.36

Table 4. Genotype independent inhibition of HCVpp infection of Huh-7 cells by ligand 281816.

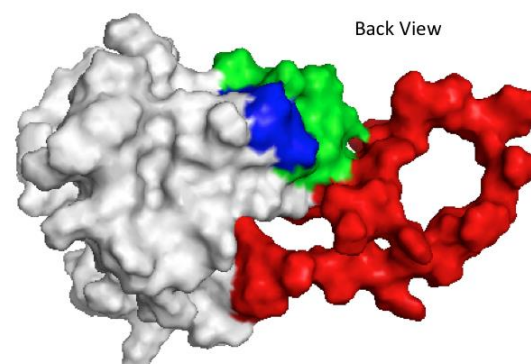
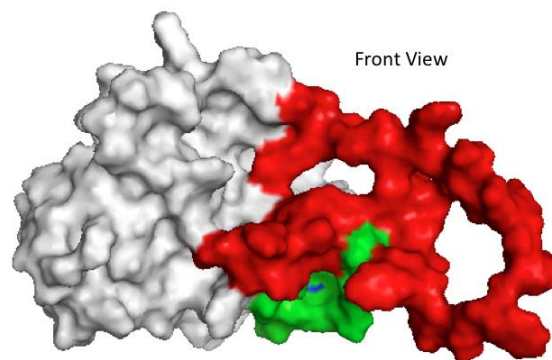
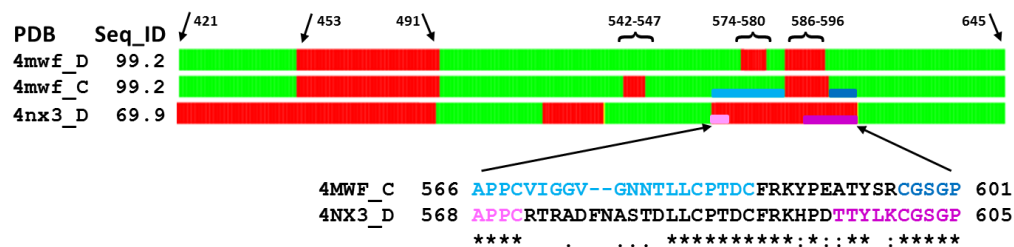
Subtypes	IC50 (μM)
HCVpp 1a	2.95
HCVpp 1b	4.66
HCVpp 2a	2.22
HCVpp 2b	2.93
HCVpp 4a	3.44
HCVpp 6a	3.30

Figures

Figure 1. Comparison of structural templates used for modeling of HCV E2c protein.

(A) Bar representation of structural similarities between crystal structures 4MWF chains C and D, and 4NX3 chain D. Regions reported in coordinates span amino acid residues from H421 to N645. In the column Seq_ID are provided sequence identities between amino acid sequences taken from coordinates and corresponding sequence fragments from HCV E2 protein of genotype 1a. In GREEN are colored regions where structural deviations are below 3 Ångstroms measured as C α -C α distances between corresponding residues from the superimposed structures. In RED are regions where structural data is missing or deviations are greater than 3 Ångstroms. (B) Structural superposition of 4mwf_C and 4nx3_D shows strong conformational similarities between experimentally solved structures of E2 proteins for which the level of sequence identity is 69%. In blue and purple are colored structural fragments where two structures 4mwf_C (566-601; BLUE: light-dark) and 4nx3_D (568-605; PURPLE: light-dark) significantly differ.

(A)



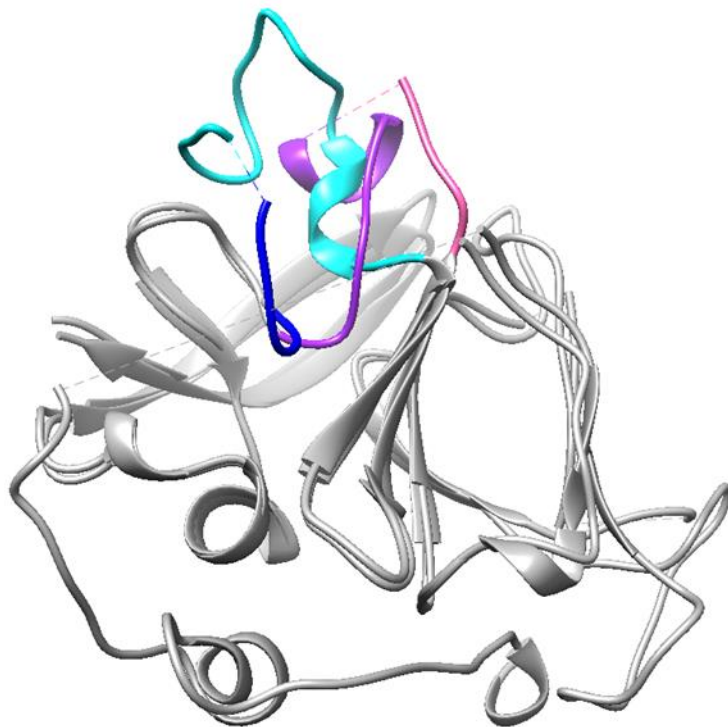
(B)

Figure 2. Comparison of the crystal structure of E2c with the homology model. Structural superposition between E2c crystal structure from the PDB chain 4mwf_D (red) and the homology model (black) is illustrated using ribbons representation. It shows overlap in structure conformations in most of the regions, except the fragments where coordinates in the experimental structure are missing (red dashed lines).

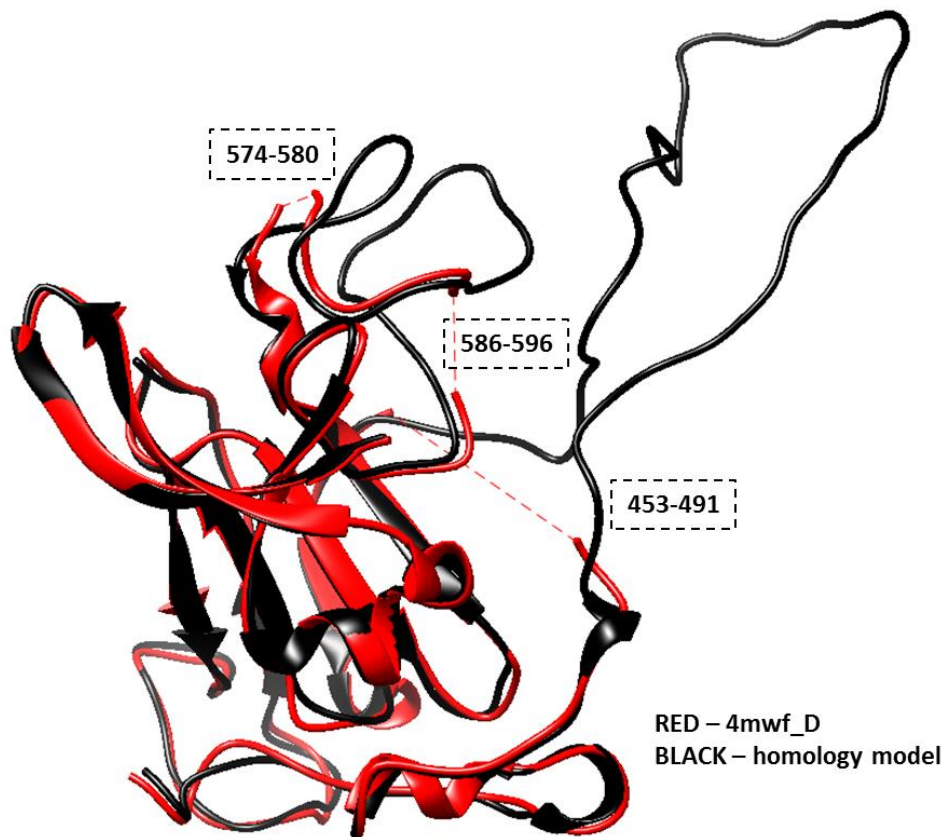


Figure 3. . Location of five ligand binding sites used to select ligands for experimental testing. Each of these sites either covers or is located immediately adjacent to amino acids or peptide segments of the E2 protein known to be important for HCV infectivity. H421-N423 (yellow): each amino acid in this region important for infectivity. Amino acids Y474-R492 (light cyan) have been shown to prevent infectivity, but this region of the protein has no effect on E2 binding CD81. W487 (dark cyan) is a key amino acid that is involved in E2 binding to E1. S522-G551 (light green) and Y527 and W529 (dark green) are critical for E2 binding to CD81. Site 4: P612, Y613, and H617-P619 (red) are critical for E2 binding to CD81; mutations to R614-W616 (pink) disrupts the structure of the region.

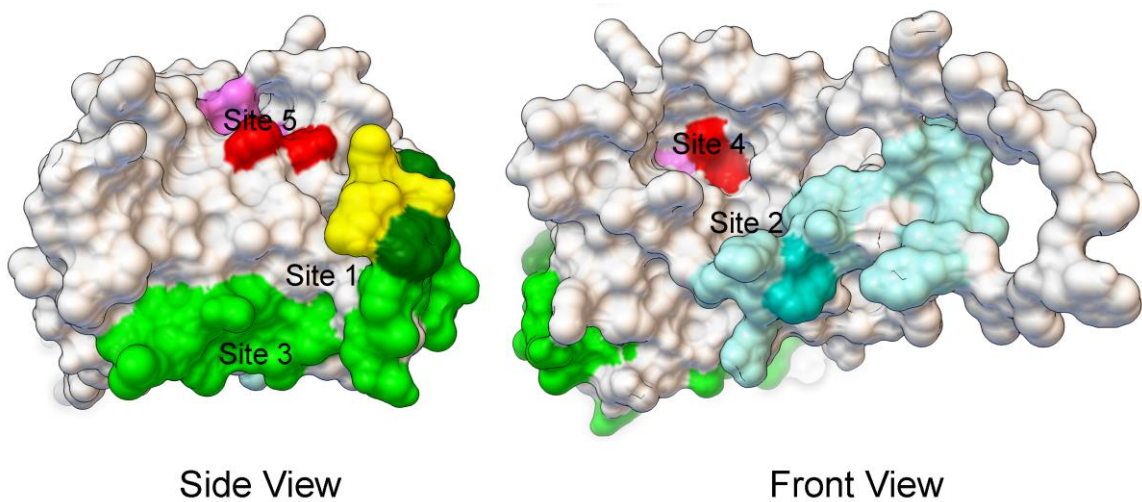


Figure 4. Comparison of the crystal structure of E2c with the homology model. A. E2c crystal structure (red). B. Homology model (black). C. Superposition of the two structures. The residues colored red remain unchanged from the crystal structure. The residues colored black show the structure is either different or not present in the crystal structure.

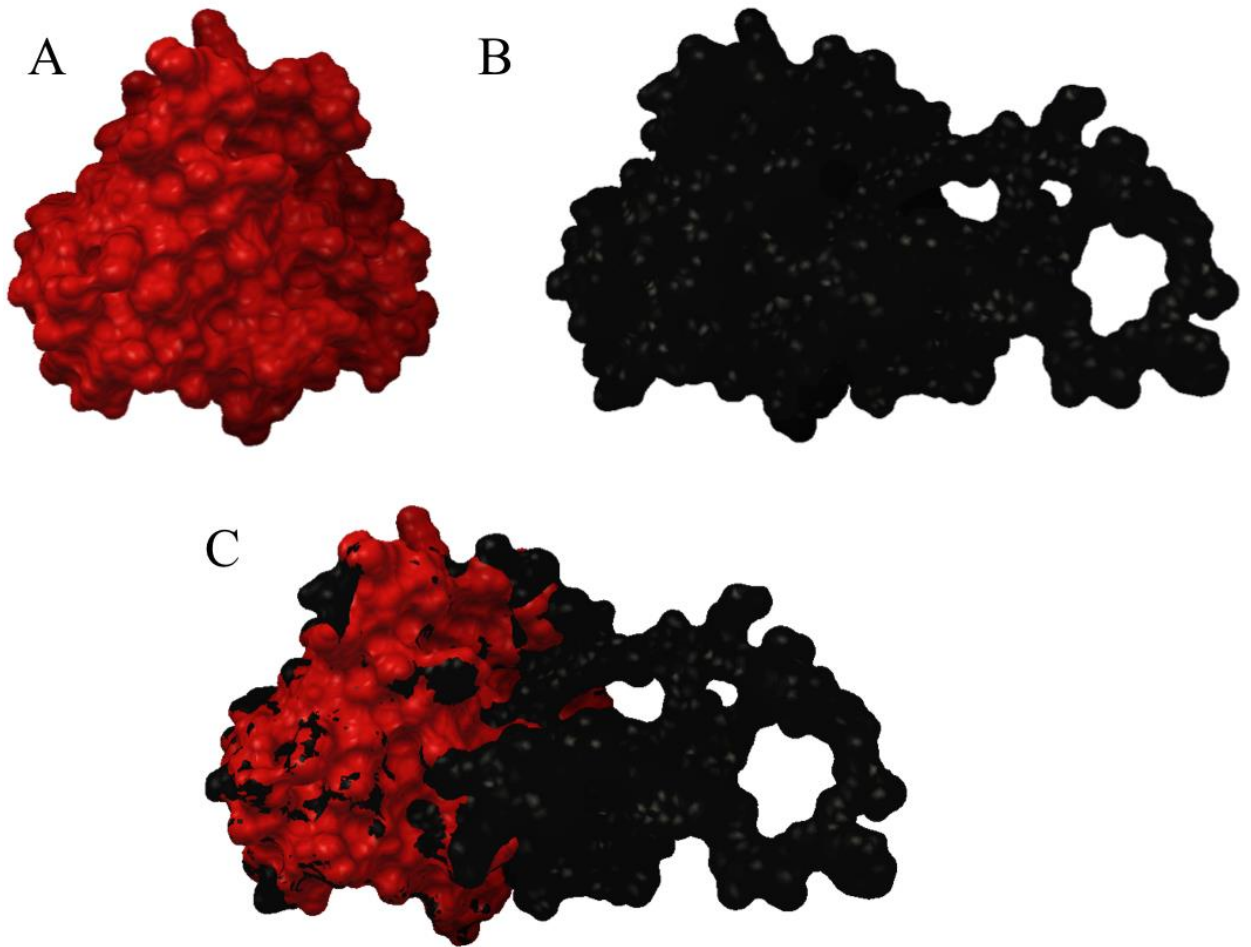


Figure 5. Surface Plasmon resonance sensogram (Biacore T100). This figure shows sensorgrams (binding and dissociation plots) for two of the ligands that bound to the recombinant E2 protein immobilized on a CM5 chip, 86467 (green) and 121861 (red), and the 3 reference points that are used to measure the binding and dissociation (dissociation 1 and dissociation 2) of the compound expressed in response units (RU).

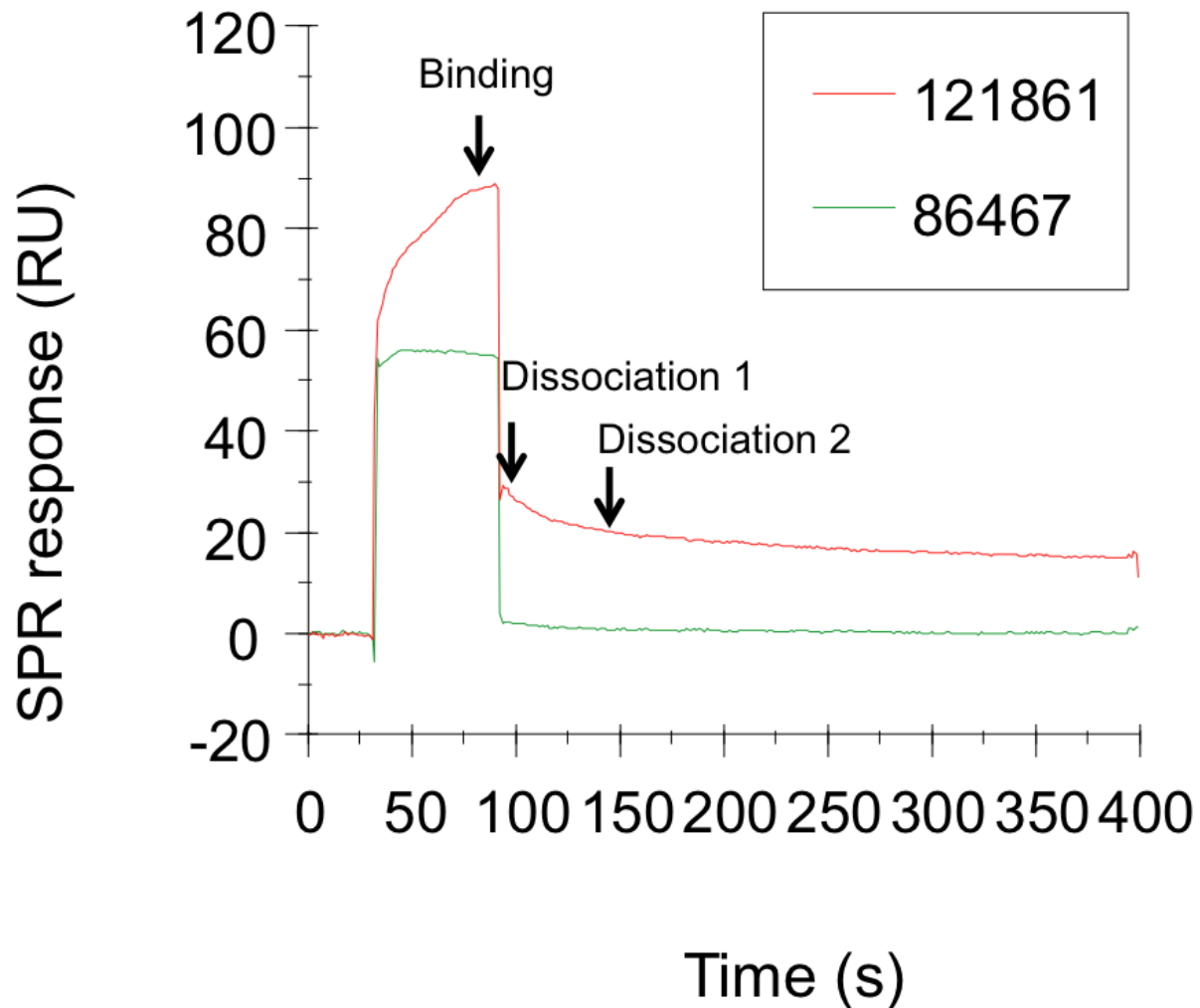


Figure 6. 281816 inhibits HCV entry in a genotype-independent manner. (A) Huh-7 cells in 96-well plates were pre-treated with 281816 (left and middle panels) or anti-CD81 antibody (right panel) at the indicated concentrations and then infected with HCVpp 2a or HCVcc. (B) Huh-7 cells in 24-well plates were pre-treated with 281816 at the indicated concentrations and infected with HCVpp expressing envelope proteins from the indicated genotype. After 30 hours of infection, cells were lysed and luciferase activity quantified. HCVpp infections were normalized to RD114pp infections.

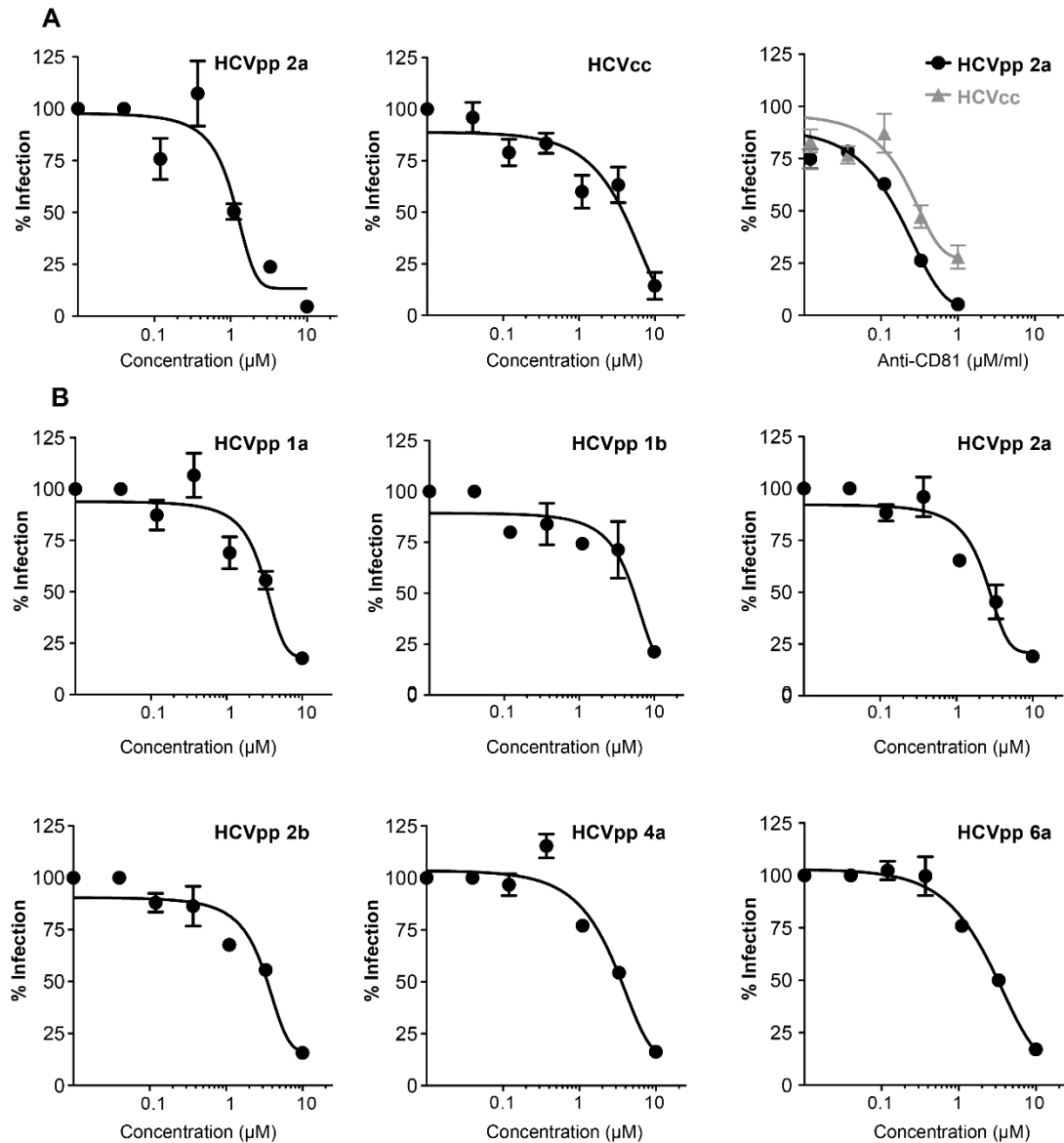


Figure 7. 281816 inhibition of HCV E2 protein binding to CD81 on Raji cells. Flow cytometry was used to quantify recombinant HCV E2 protein binding to native CD81 over-expressed on Raji cells. Binding of the recombinant E2 protein to native CD81 on the surface of Raji cells was detected using the mouse monoclonal E2 antibody clone H53 followed by staining with a secondary anti mouse-FITC antibody as described in the Materials and Methods section. E2 binding is inhibited by 281816 in a dose-dependent manner up to 100 μ M.

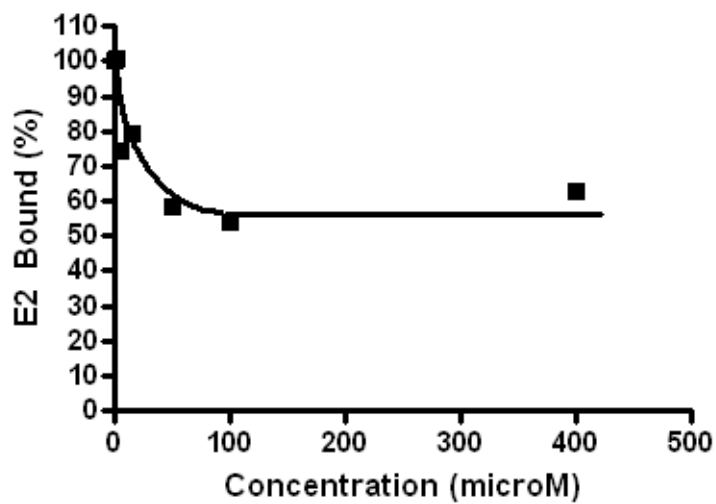


Figure 8. 281816 inhibits HCV entry. A, Huh-7 cells in 24-well plates were treated at different time points with 281816 at 10 μ M and infected with HCVcc for 2h at 37°C. 281816 was added full-time during the experiment (a), 2h before virus inoculation (b), 2h during virus inoculation (c), or full-time after virus inoculation (d). B, Huh-7 cells were infected with HCVcc for 1h at 4°C (Step 1: attachment/binding), then virus was removed and cells incubated again at 4°C for 1h (Step 2: post-attachment/binding). Finally, cells were shifted at 37°C for 1h (Step 3: endocytosis/fusion) and left at 37°C for 21h. 281816 was added at 10 μ M either during the Step 1, Step 2, Step 3 or Steps 1-2-3. * and *** mean p values below 0.05 and 0.0001, respectively.

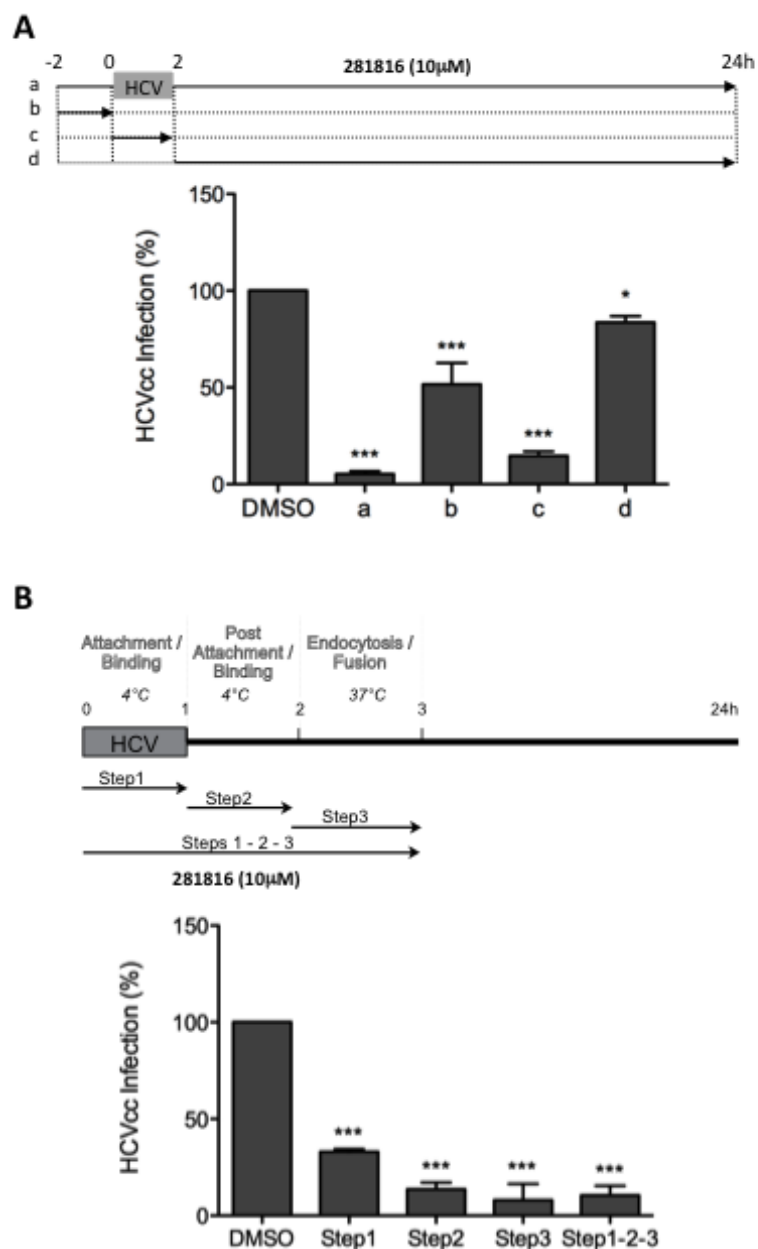


Figure 9. 281816 blocks HCV cell-to-cell transmission. Huh-7 cells were seeded on coverslips and infected with HCVcc for 2h at 37°C. Cells were then washed and cultured for 72h at 37°C in culture medium containing the 3/11 neutralizing mAb (50 µg/ml) in presence or in absence of 281816 at indicated concentrations. Cells cultured in presence of DMSO or EGCG at 50 µM were used as controls. Number of infected cells per focus was determined by A4 indirect immunofluorescence. ** and *** mean p values below 0.001 and 0.0001.

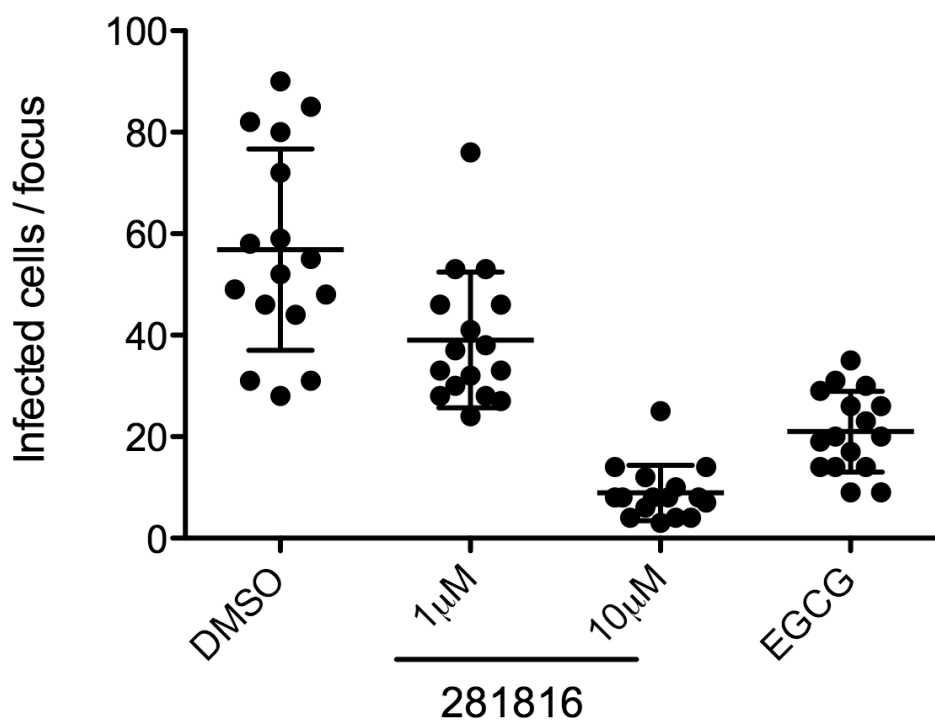
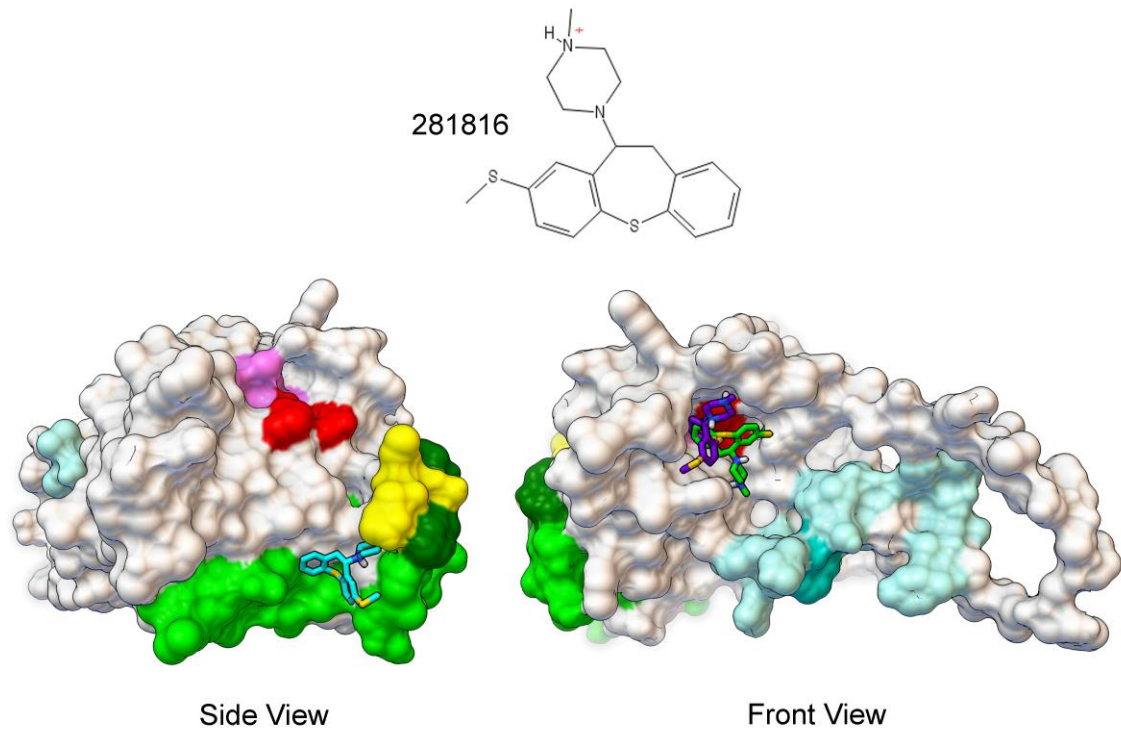


Figure 10. 281816 binding sites on HCV E2. 281816 (structure, top) is predicted to bind to two sites on the E2 protein. Different conformers of 281816 are shown bound in the two sites. The E2 amino acids are color coded as described in Figure 2.



CHAPTER VI

***IN SILICO* DESIGN OF SHAL PROTOTYPE (SH2216) AGAINST HCV USING HCV E2 AS A TARGET PROTEIN**

Objective

Designing a bidentate SHAL prototype by selecting 2 potential ligands to be incorporated into the SHAL and measuring the distance between their free carboxyl and aminogroups.

1. Introduction

HCV E2 is considered one of the significant viral proteins that are involved in the internalization of HCV into hepatocytes through its interactions with several host proteins (1-4). The newly resolved crystal structure of the core of HCV E2 made it easier to develop a homology model for the whole or most of the E2 protein (figure 1) (5,6). This in turn helps in targeting the model by doing virtual screening runs, analyzing the results and determining top virtual screening hits. Upon validating true binders using surface Plasmon resonance, those are further tested using infection inhibition assays so as to see which has an inhibitory effect on HCV.

The previous chapter discussed how 281816 was identified to be a genotype-independent inhibitor in a dose dependent manner. It was also found that 281816 binds to several potential binding sites on the E2 model *in silico*. These results have suggested we try to design an *in silico* bidentate SHAL (SH2216) which is formed of a pair of 281816 ligand analogue.

2. Materials and methods

2.1. Creation of the homology model of E2 used for docking

Two crystal structures of E2c have recently been deposited in the PDB under codes 4MWF and 4NX3 were resolved by Kong *et al.* (5) at a resolution of 2.65 Angstroms and Khan *et al.* (6) at a resolution of 2.4Å. However, upon examination of the structure prior to docking, both sets of reported atom coordinates of the protein were found to be incomplete. In addition to the coordinate file containing structural information for only 171 residues out of the 363 amino acids present in the full-length protein structure obtained by Kong *et al.* (5), structural information was missing for several peptide segments or loops within the structure. Several peptide segments were also missing from the Khan *et al.* (6) structure file. In order to prepare a more complete version of the structure for docking, we have performed several homology modeling and structure analysis tasks using the coordinates of Kong *et al.*'s (5) E2c as a template. The final structural model was created using the AS2TS system (7) based on atom coordinates from the PDB chains 4mwf_C and 4mwf_D. A structural search for similar fragments in proteins in the PDB that could be used to model missing loop regions was performed using the StralSV algorithm (8), which identifies protein structures that exhibit structural similarities despite low primary amino acid sequence similarity. The side-chain prediction was accomplished using SCWRL (9) when residue-residue correspondences did not match. Residues that were identical in the template and E2 protein were copied from the template onto the model. Potential steric clashes were identified in the unrefined model using a contact-dot algorithm in the MolProbity software package (10), and the constructed model was finished with relaxation using UCSF Chimera (11).

2.2. 281816 Docking to HCV E2 homology model

AutoDock 4.2 (12,13) was used to perform 281816 docking to HCV E2 model. The parameters were set at 100 for the number of genetic algorithm (GA) runs, 150 as the population size, and a maximum number of generations of 25000. The Lamarckian genetic algorithm in AutoDock was used to perform the docking experiment. Docking results were sorted to identify the binding sites that 281816 binds to and the different between free binding energies.

2.3. Determining an analogue for 281816 and docking it to HCV E2 homology model

One of 281816 Analogues (Ligand 22594527) was chosen to design an *in silico* prototype of SHAL. It was docked to the HCV E2 homology model, and the binding mode was analyzed. The binding sites were determined and 2 were chosen. The distance between the 2 sites were measured and the number of mini-PEG moieties to be used was determined. Avogadro was used to link the 2 281816 ligands with mini-PEG and Lysine to create a bi-dentate SHAL (SH2216).

3. Results and Discussion

Ligand 281816 was predicted to bind to two different sites on E2 shown in **Figure 2**. One site is located deep inside a cavity positioned directly above Y618 and P619, two amino acids known to contribute to E2's binding to CD81. The second site was located in a cavity on the opposite of the protein. In this second site, the ligand is positioned directly above residues R614-W616 and immediately adjacent to P612 and Y613 and H421-N423. These amino acid residues comprise Sites 1 and 4 and the residues in both sites are amino

acids that have been shown to be critical for E2 binding to CD81. As expected, the ligand positioned above Y618 and P619 in the deeper cavity was predicted to bind more strongly to the protein (free energy of the best bound ligand = -8,64 Kcal/mol) than when it was bound to the more shallow cavity on the other side of the protein (free energy = -6.39 Kcal/mol). One interesting and unique feature of the 281816 ligand is that it is predicted to bind to both of the exposed faces of the 612-619 domain that is known to participate in E2 binding to CD81.

Owing to the fact that 281816 doesn't have neither a free amino group nor a free carboxyl group, a search for an analogue that has either of these groups was necessary. Ligand 22594527 was one of the analogues that were found to be suitable for a SHAL synthesis (figure 3). This ligand was docked to HCV E2 model and was found to bind in the first site to PRO619 and TRP616 (figure 4a), and in the second site to PRO612 and TYR613 (figure 4b) which makes it having similar binding mode to 281816. These 2 binding sites were chosen depending on the binding modes of Ligand 22594527 and the significant residues that were previously determined to be involved in HCV E2: CD81-LEL interaction. This was followed by measuring the distance between the 2 sites (figure 4c) which was found to be in the range between 26-27.5 Å and determining the number of mini-PEG and lysine residues that are needed to link the two molecules together in such a manner that one ligand can bind to each site simultaneously.

4. References

1. El-Awady, M.K., Tabll, A.A., El-Abd, Y.S., Yousif, H., Hegab, M., Reda, M., El Shenawy, R., Moustafa, R.I., Degheidy, N. & El Din, N.G. 2009, "Conserved peptides within the E2 region of Hepatitis C virus induce humoral and cellular responses in goats", *Virology journal*, vol. 6, pp. 66-422X-6-66.
2. Carlsen, T.H., Scheel, T.K., Ramirez, S., Foung, S.K. & Bukh, J. 2013, "Characterization of hepatitis C virus recombinants with chimeric E1/E2 envelope proteins and identification of single amino acids in the E2 stem region important for entry", *Journal of virology*, vol. 87, no. 3, pp. 1385-1399.
3. Li, Y.P., Kang, H.N., Babiuk, L.A. & Liu, Q. 2006, "Elicitation of strong immune responses by a DNA vaccine expressing a secreted form of hepatitis C virus envelope protein E2 in murine and porcine animal models", *World journal of gastroenterology : WJG*, vol. 12, no. 44, pp. 7126-7135.
4. Ray, R., Meyer, K., Banerjee, A., Basu, A., Coates, S., Abrignani, S., Houghton, M., Frey, S.E. & Belshe, R.B. 2010, "Characterization of antibodies induced by vaccination with hepatitis C virus envelope glycoproteins", *The Journal of infectious diseases*, vol. 202, no. 6, pp. 862-866.
5. Kong, L., Giang, E., Nieuwma, T., Kadam, R.U., Cogburn, K.E., Hua, Y., Dai, X., Stanfield, R.L., Burton, D.R., Ward, A.B., Wilson, I.A. & Law, M. 2013, "Hepatitis C virus E2 envelope glycoprotein core structure", *Science (New York, N.Y.)*, vol. 342, no. 6162, pp. 1090-1094.
6. Kong, L., Giang, E., Nieuwma, T., Robbins, J.B., Deller, M.C., Stanfield, R.L., Wilson, I.A. & Law, M. 2012, "Structure of hepatitis C virus envelope glycoprotein E2 antigenic site 412 to 423 in complex with antibody AP33", *Journal of virology*, vol. 86, no. 23, pp. 13085-13088.
7. Yagnik, A. T., Lahm, A., Meola, A., Roccasecca, R. M., Ercole, B. B., Nicosia, A. and Tramontano, A. (2000).A model for the hepatitis C virus envelope glycoprotein E2. *Proteins*. 40, 355-366.
8. Zemla A, Zhou CE, Slezak T, Kuczmarski T, Rama D, Torres C, Sawicka D, Barsky D: AS2TS system for protein structure modeling and analysis. *Nucleic Acids Res* 2005, 33:W111-115.
9. Zemla AT, Lang DM, Kostova T, Andino R, Ecale Zhou CL: StralSV: assessment of sequence variability within similar 3D structures and application to polio RNA-dependent RNA polymerase. *BMC Bioinformatics* 2011, 12:226.
10. Krivov, G.G., Shapovalov, M.V., and Dunbrack, Jr, R.L.: Improved prediction of protein side-chain conformations with scwrl4. *Proteins* 77, 4 (Dec 2009), 778-95.
11. Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC: MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 2010, 66:12-21.
12. Morris, G. M., Huey, R. and Olson, A. J. (2008).Using AutoDock for ligand-receptor docking. *Curr. Protoc. Bioinformatics*. Chapter 8, Unit 8.14.

13. Morris, G. M. , Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E. , Belew, R. K., Olson, A. J. (1998). Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function, *J. Comput. Chem.* 19, 1639–1662.

5. Figures

Figure 1. The four binding sites selected to guide the virtual screening runs on HCV E2 homology model. Site 1: Yellow (412-423) each amino acid in this region important for infectivity. Site 2 Light cyan (474-492) prevents infectivity, no effect on E2 binding CD81; Dark cyan 487 (involved in E2 binding to E1). Site 3 Light green (522-551); Dark green 527, 529 (critical for E2 binding to CD81). Site 4: Red (612, 613, 617-619, critical for E2 binding to CD81); Pink (614-616) disrupts structure of region.

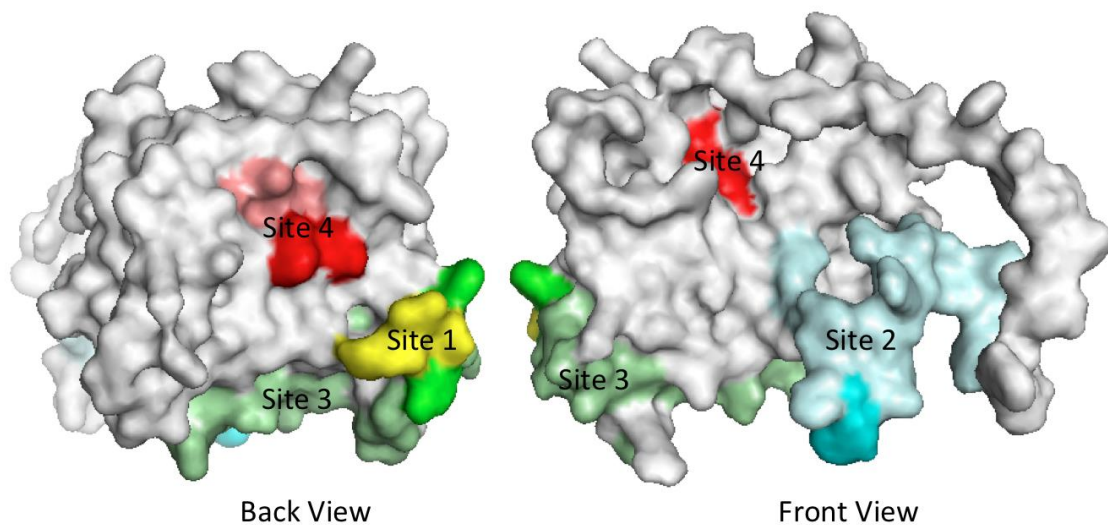


Figure 2. 281816 binding sites on HCV E2. 281816 binds three sites. Clusters of different conformations of bound 281816 molecules are shown bound in the three sites. Sites and amino acids are color coded as described in Figure 2.

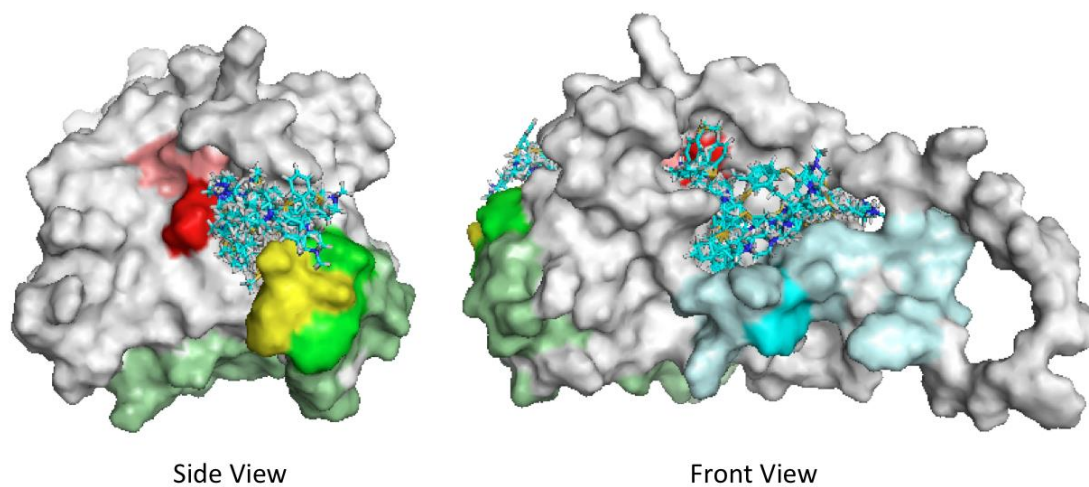


Figure 3. The structure of 281816 analogue. 281816 analogue (22594527)- *Z*-but-2-enedioic acid;N'-(3-methylsulfanyl-5,6-dihydrobenzo(b)(1)benzothiepin-5-yl)ethane-1,2-diamine.

CSC1=CC2=C(C=C1)SC3=CC=CC=C3CC2NCCN.C(=CC(=O)O)C(=O)O

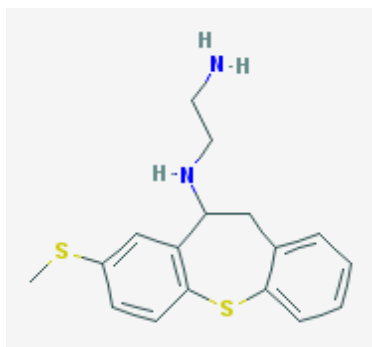
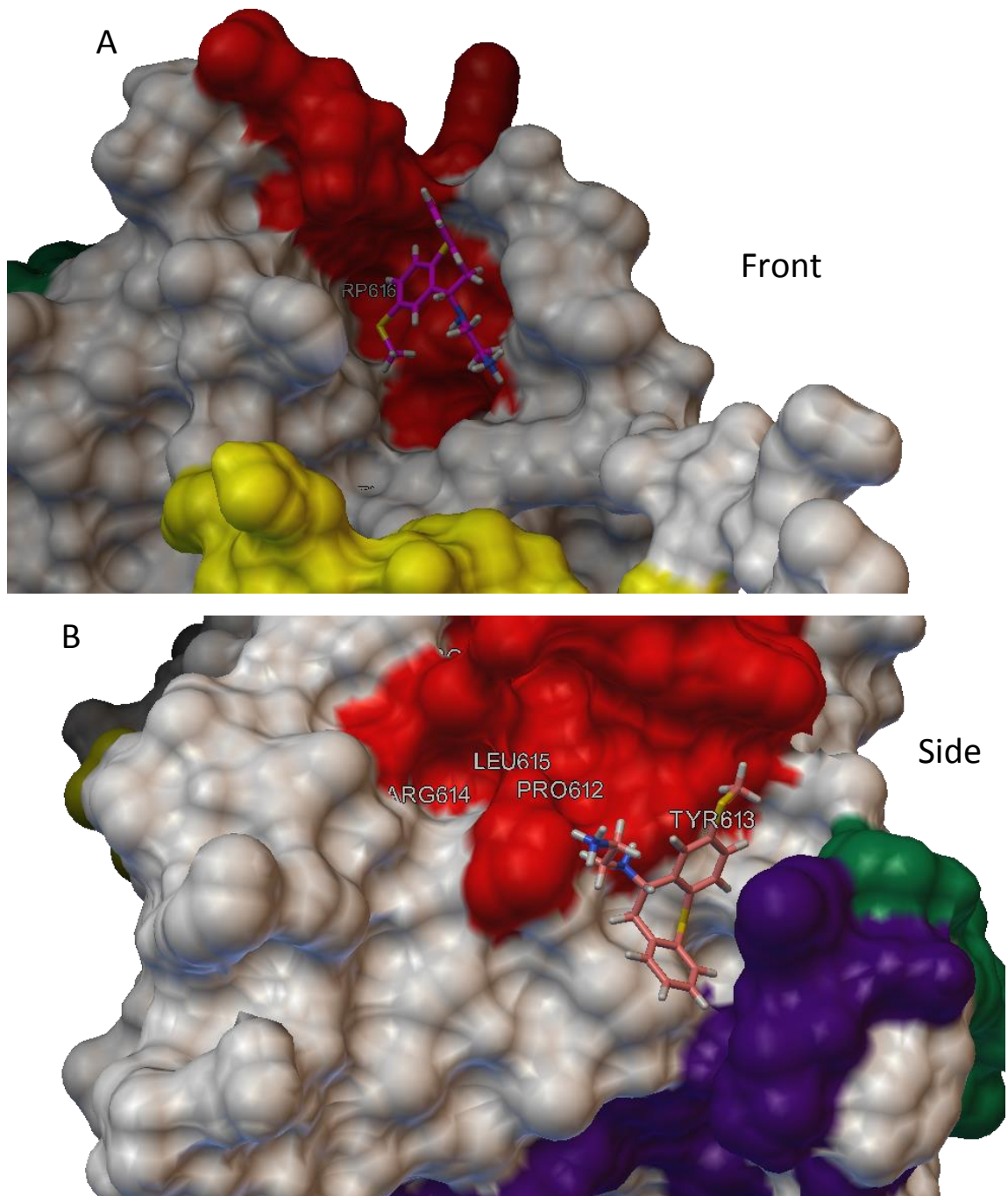


Figure 4. The 2 chosen binding sites that ligand 22594527 binds to. The first site comprises residues PRO619 and TRP616 (a) and the second site compromises PRO612 and TYR613 (b). The distance between the free aminogroup of the first ligand binding to site (a) and that binding to site (b) was found to be in the range 26-27.5 Å (c).



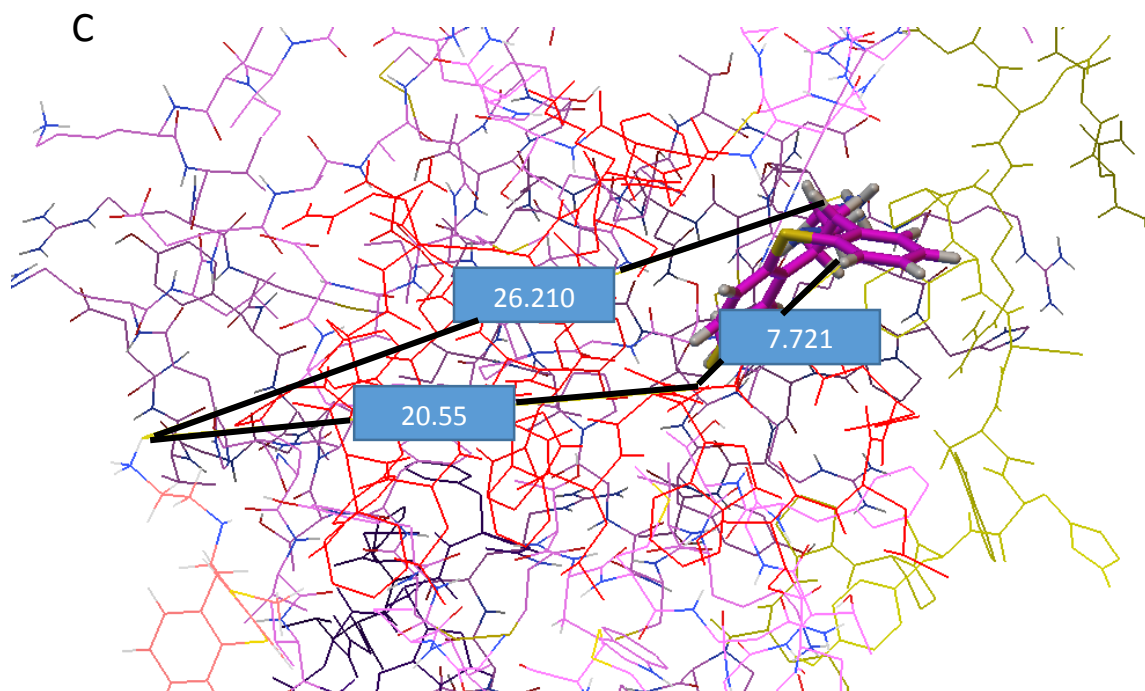
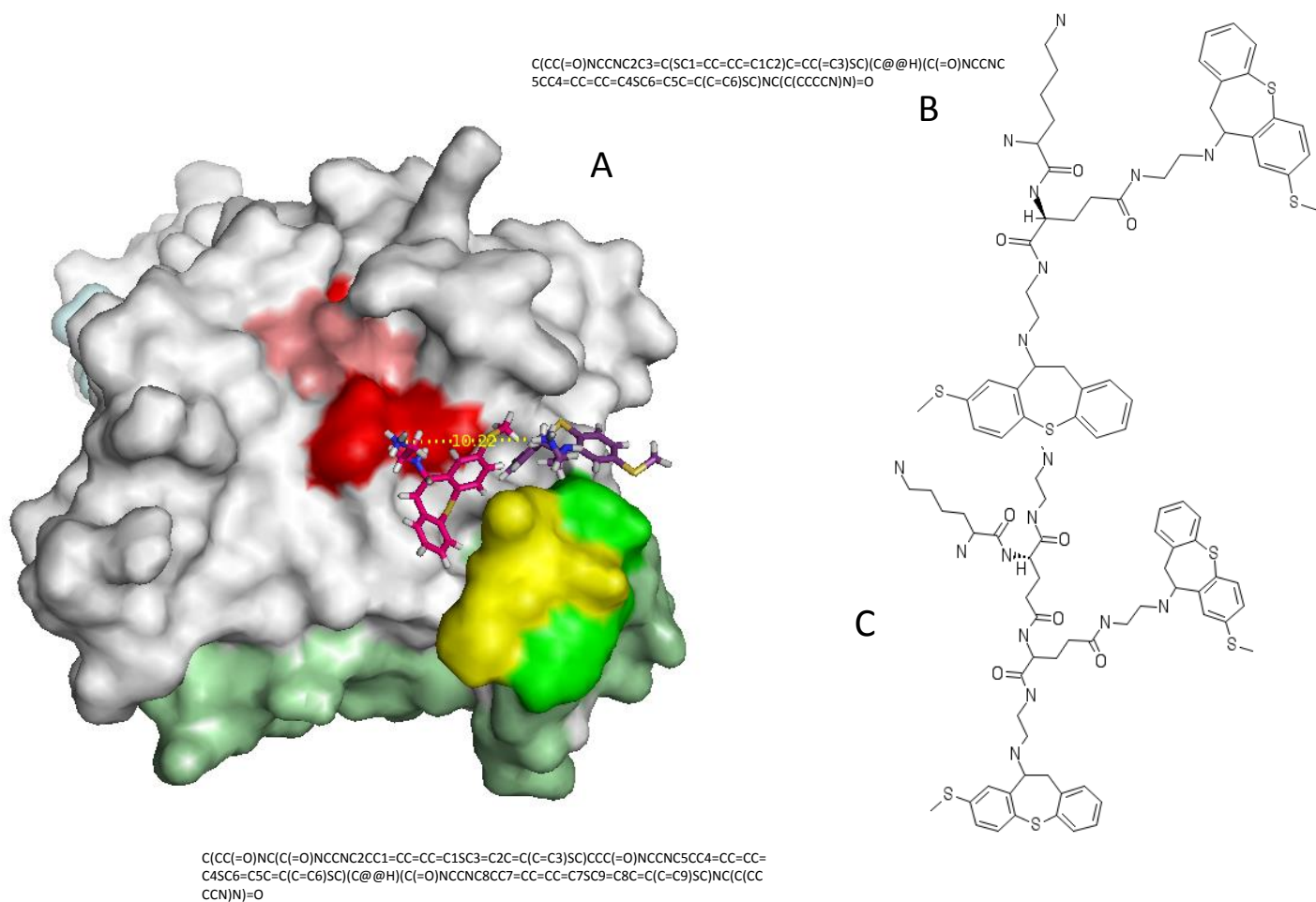


Figure 5. Bidentate SHAL from 2 ligands 22594527. The distance between the 2 ligands binding on the same site is around 10.22 Å. Which enables generating a bidentate (b) and tridentate (c) SHAL without the use of linkers.



CHAPTER VII

IDENTIFICATION OF DRUG LEADS THAT BLOCK PLASMODIUM SPOROZOITES: CD81-LARGE EXTRACELLULAR LOOP INTERACTION

Objectives

The aim of this study is to screen ~2000 ligands against the residues in CD81-LEL that are involved in CD81: Plasmodium sporozoites interaction, test the hits using lab on a chip techniques and antibody neutralizing assays followed by testing the ligands with the best binding affinity using Sporozoite Invasion assays. The Sporozoite invasion assays were conducted in John Hopkins University (USA) at Dr. Photini Sinnis's lab.

1. Introduction

Malaria is caused by protozoan parasites belonging to the genus Plasmodium. Malaria in human is caused by four types which are: *Plasmodium malariae*, *Plasmodium ovale*, *Plasmodium vivax* and *Plasmodium falciparum*. *P. falciparum* and *P. vivax* are the two major species involved in global malaria (1). Malaria infection starts by the invasion of the parasite into the hepatocytes (Liver stage) followed by blood stage.

In order for Plasmodium species to propagate in the host and establish an infection, they should be internalized into the hepatocytes (2). The mechanism of interaction of plasmodium sporozoites with hepatocytes is still not clear however, hCD81 was found to be one of the significant surface proteins on the hepatocytes required by the sporozoites to invade the liver cells (3-5).

CD81 belongs to tetraspanin family. It possesses ubiquitous functions such as cell adhesion, migration, cell fusion, co-stimulation, signal transduction, and differentiation (Reviewed in 6). It was found that antibodies to CD81 or CD81 silencing largely inhibits the infection of hepatocytes by *Plasmodium falciparum* sporozoites. Additionally, it was found that *Plasmodium yoelii* sporozoites do not have the ability to infect CD81-deficient mouse hepatocytes both *in vitro* and *in vivo* (3-5).

Yalaoui *et al.* (7) identified specific amino acid residues in CD81 that are important for infection through conducting mutational studies. They found a stretch of 21 amino acid residues that are required for invasion (135-VVDDDANNAKAVVKTFHETLD-155) (figure 1) (7).

In our study we targeted those amino acid residues by creating 3 grid parameter files (figure 2), used them to guide the virtual screening runs against ~ 4000 ligands, analyzed 500 of the resulting docked ligands, determined the virtual screening hits and tested the hits *in vitro* using surface Plasmon resonance and antibody neutralizing assays to validate binding to CD81 and inhibition of CD81: antibody interaction.

2. Materials and Methods

2.1. In silico Experiments

In this study we used three different grid boxes (3 regions) to encapsulate the region of CD81 that has the amino acid residues involved in malaria infection of liver cells. A total of ~4000 ligands from the NCI Library were docked against each of these grid boxes in the crystal structure of hCD81 (PDB ID: 1G8Q). This was followed by

analyzing the results (filtering, visually inspecting and determining virtual screening hits. Thirty one hits were chosen to be tested *in vitro*.

2.2. *Surface Plasmon Resonance*

SPR analysis was performed using a Biacore T100 workstation (GE Healthcare, NJ, USA). A recombinant form of the CD81-LEL protein with a GST-tag (Shoshana Levy Lab - Stanford) was used to determine the binding affinities of the 31 ligands. 10 μM CD81-LEL-His diluted into 10 mM sodium acetate buffer pH 4.5 was immobilized for 15 min at a flow speed of 5 $\mu\text{l}/\text{min}$ onto a CM5 sensor chip using amine coupling (EDC-NHS). Approximately 20,000 RU of protein were immobilized on the chip. The ligands were prepared as 200 μM solutions in PBS-1% DMSO (the running buffer) and they were introduced to the protein using a pre-programmed 3 min association and 1 min dissociation interval.

2.3. *Antibody Neutralizing assays*

For antibody neutralization assay Raji cells were used, a human B cell line that expressed high amounts of CD81 on the surface (data not shown). Cells were grown in RPMI medium (10% FCS, 1% penicillin/streptomycin, 1% L-glutamine, 1% non-essential amino acids, 1% sodium pyruvate, pH 7.4, at 37°C with 5% CO₂). 2×10^5 cells were incubated with or without different concentrations (50 μM , 100 μM , 400 μM and 1 mM) of indicated inhibitor for 20 min at room temperature, subsequently 1 μl (16 ng/ μl) of FITC-labeled anti CD81 antibody (BD Pharmingen, 551108) was added to the cells

and incubated for 20 min (antibody titration was performed to obtain a working dilution range, data not shown). Cells were washed and analyzed by FACS (BD FACSCalibur, software: Cell Quest Pro). Mean Fluorescence Intensity MFI was calculated using Flowjo software (TreesStar, www.flowjo.com).

2.4. Mosquito Infection

Anopheles stephensi mosquitoes were fed on mice infected with *Plasmodium yoelii* 17XNL parasites. On day 14 post bloodmeal, salivary glands were harvested, homogenized, sporozoite number was determined using a hemocytometer and the indicated number of sporozoites was added to cell monolayers. *Plasmodium falciparum* sporozoites were generated by membrane feeding *An. stephensi* mosquitoes using gametocyte cultures generated *in vitro*.

2.5. Cells and antibodies.

HepG2-CD81 cells (8), which are derived from human liver carcinoma and express CD81, were maintained in DMEM supplemented with 10% fetal calf serum and 1 mM glutamine (DMEM/FCS). Monoclonal antibody (mAb) 2E6 was used for development assays and is directed against *Plasmodium* Hsp 70 (9) and mAb 2A10 is directed against the repeat region of *P. falciparum* CSP and was used for invasion assays.

2.6. Invasion and development assays.

5×10^5 HepG2-CD81 cells were plated on collagen coated Lab-Tek wells and allowed to adhere overnight. The following day, cells were preincubated with 400 μ M of

compound for 30 minutes at 37°C. Controls were pre-incubated with DMSO at the same concentration in which it was found in the compound solutions. 10^5 sporozoites were added to each well and in the case of *P. falciparum*, incubated for 1.5 hours and scored for invasion whereas in the case of *P. yoelii*, incubated for 44 hours and scored for exoerythrocytic stages.

3. Results

Eight out of 31 ligands were found to bind to CD81 using surface Plasmon resonance based assays (Table 1 – Figure 3). Ligand 87504 was found to be the one with the most desirable binding-dissociation behavior having the highest binding RUs. Ligand 90444 and 93033 are considered of comparable binding and dissociation.

In addition, the eight ligands had inhibitory effect on CD81: antibody interaction (Table 2). As shown, they were found to have a better inhibitory effect at concentration of 400uM when compared to that of 100uM.

When sporozite infection assays were conducted with the ligands, four ligands were found to inhibit *P. yoelii* (7962, 73735, 90444 and 75866) whereas ligand 40614 exhibited an interesting action where it enhanced the infection with *P. yoelii* (figure 4). Inhibition of *P. berghei* by both ligands 7962 and 73735 was detected when using Hep1-6 hepatoma cells whereas 90444 and 75866 didn't inhibit *P. yoelii* or *P. berghei* in Hepa1-6 cells. Ligand 40614 was found to be consistent in enhancing development in both *P. yoelii* and *P. berghei* in both Hepa1-6 and Hep2G cells (figure 5). As for the infection by

P. falciparum, 7962 and 73735 were found to inhibit the infection with this species. On the other hand, 40614 wasn't found to enhance infection with *Pl. falciparum* (figure 6).

During our *in silico* analysis of the binding modes of 7962 (the most promising drug lead among the five tested compounds), we found that there are a number of 7962 conformations that interacted with the 21 amino acid sequence and Asp137 that were known to be important for malaria infection (figure 7). As for 73735, some conformations interacted with the peptide sequence and others bound near Asp137 (figure 8). The majority of 40614 ligand conformers did not interact with the peptide sequence nor near Asp 137. Only two 90444 ligand conformers were found to bind (weakly) nearby the peptide but not close to Asp137. As for 75866, only few 75866 ligand conformers were found to bind nearby the peptide sequence but none were found to bind near to Asp137.

Upon calculating the percentage of infection inhibition of the 3 *Pl.* species by the five ligands, 7962 was found to exhibit a 100% inhibition of all 3 species whereas 73735 was found to possess an inhibitory percentage of 96.8%, 77%, and 82.3% of infection with *P. falciparum*, *P. yoleii* and *P. bergerii* respectively.

4. Discussion

Upon conducting virtual screening runs, 31 ligands were selected to be the first set for conducting experimental validation using surface Plasmon resonance. Eight of the virtual screening hits were found to bind to CD81 using surface Plasmon resonance. Ligand 87504 was one of the ligands that showed best binding and dissociation thus when going back and analyzing it's *in silico* binding behavior it was found to bind to several

residues in the stretch of 21 amino acid residues identified by Yalouli *et al.* (7) to be involved in Plasmodium : CD81-LEL interaction (figure 1).

Those eight ligands were further tested using antibody neutralizing assays and were found to inhibit anti-CD81 binding to CD81-LEL in a dose dependent manner where the 400uM concentration gave better inhibitory effect compared to the 100uM of the ligands ranging from % of inhibition of 17% to ~30%.

There are three features of 7962 and 73735 (ligands that block sporozite infection) that docking experiments suggest may distinguish them from 40614, 90444 and 75866 (ligands that do not block infection): (i) 7962 and 73735 conformers, in general, bind more strongly to CD81 (1-3kcal/mole). (ii) The strongest binding conformers of 7962 and 73735 were found to bind to the red site A (figure 7, Figure 8). (iii) The higher number of 7962 and 73735 conformers are predicted to bind to or nearby Asp137; where none of the 40614, 90444 or 75866 conformers were predicted to bind near or to Asp137 which was shown to be essential for infection. Collectively, these data suggest that ligand 7962 would help generate a promising drug to treat several species of malaria.

5. References

1. Druilhe P, Daubersies P, Patarapotikul J, Gentil C, Chene L, Chongsuphajaisiddhi T, *et al.* A primary malarial infection is composed of a very wide range of genetically diverse but related parasites. *J Clin Invest* 1998;101:2008–16.
2. Prudencio M, Rodriguez A, Mota MM (2006) The silent path to thousands of merozoites: the Plasmodium liver stage. *Nat Rev Microbiol* 4: 849–856.
3. Silvie O, Rubinstein E, Franetich JF, Prenant M, Belnoue E, *et al.* (2003). Hepatocyte CD81 is required for Plasmodium falciparum and Plasmodium yoelii sporozoite infectivity. *Nat Med* 9: 93–96.
4. Silvie O, Greco C, Franetich JF, Dubart-Kupperschmitt A, Hannoun L, *et al.* (2006) Expression of human CD81 differently affects host cell susceptibility to malaria sporozoites depending on the Plasmodium species. *Cell Microbiol* 8:1134–1146.
5. Silvie O, Charrin S, Billard M, Franetich JF, Clark KL, *et al.* (2006) Cholesterol contributes to the organization of tetraspanin-enriched microdomains and to CD81-dependent infection by malaria sporozoites. *J Cell Sci* 119: 1992–2002.
6. Levy S, Shoham T (2005) The tetraspanin web modulates immune-signalling complexes. *Nat Rev Immunol* 5: 136–148.
7. Yalaoui S, Zougbedé S, Charrin S, Silvie O, Arduise C, *et al.* (2008) Hepatocyte Permissiveness to Plasmodium Infection Is Conveyed by a Short and Structurally Conserved Region of the CD81 Large Extracellular Domain. *PLoS Pathog* 4(2).
8. Olivier Silvie, Céline Greco, Jean-François Franetich, Anne Dubart-Kupperschmitt, Laurent Hannoun, Geert-Jan van Gemert, Robert W. Sauerwein, Shoshana Levy, Claude Boucheix, Eric Rubinstein and Dominique Mazier. Expression of human CD81 differently affects host cell susceptibility to malaria sporozoites depending on the Plasmodium species. *Cellular Microbiology* 8:1134-46; 2006.
9. Tsuji M, Mattei D, Nussenzweig RS, Eichinger D and Zavala F. Demonstration of heat-shock protein 70 in the sporozoite stage of malaria parasites. *Parasitol Res* 80:16-21; 1994.

6. Figures

Figure 1. Stretch of 21 amino acid residues involved in plasmodium sporozoites: CD81-LEL interaction. Red, Green, Blue: amino acids surrounding three main ligand binding sites. Light orange: V135-D155 - Dark orange: D137 (7).

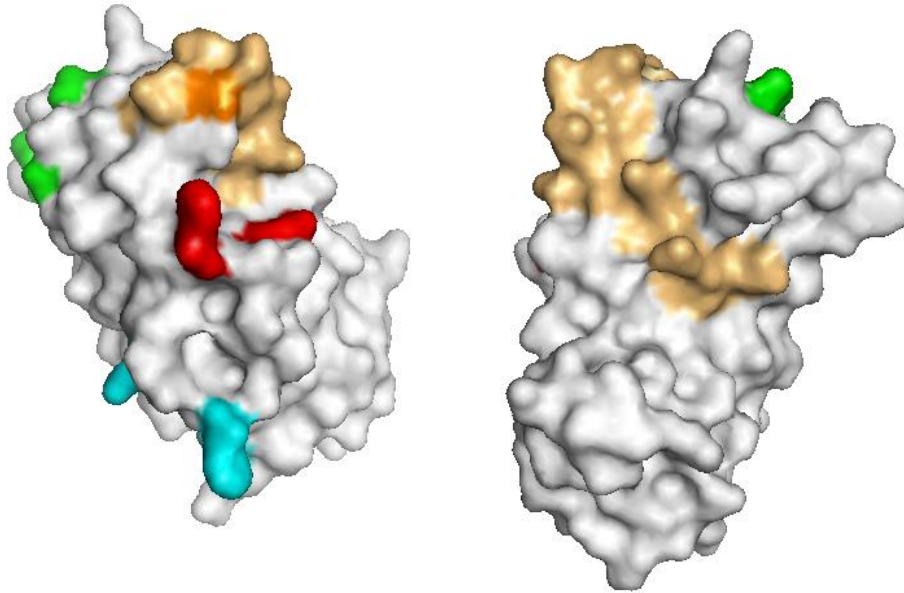


Figure 2. The 3 sites chosen on CD81-LEL to undergo virtual screening runs. Those sites were based on the stretch of 21 amino acids where the site containing those residues was fragmented into 3 subsites and each of them included other residues.

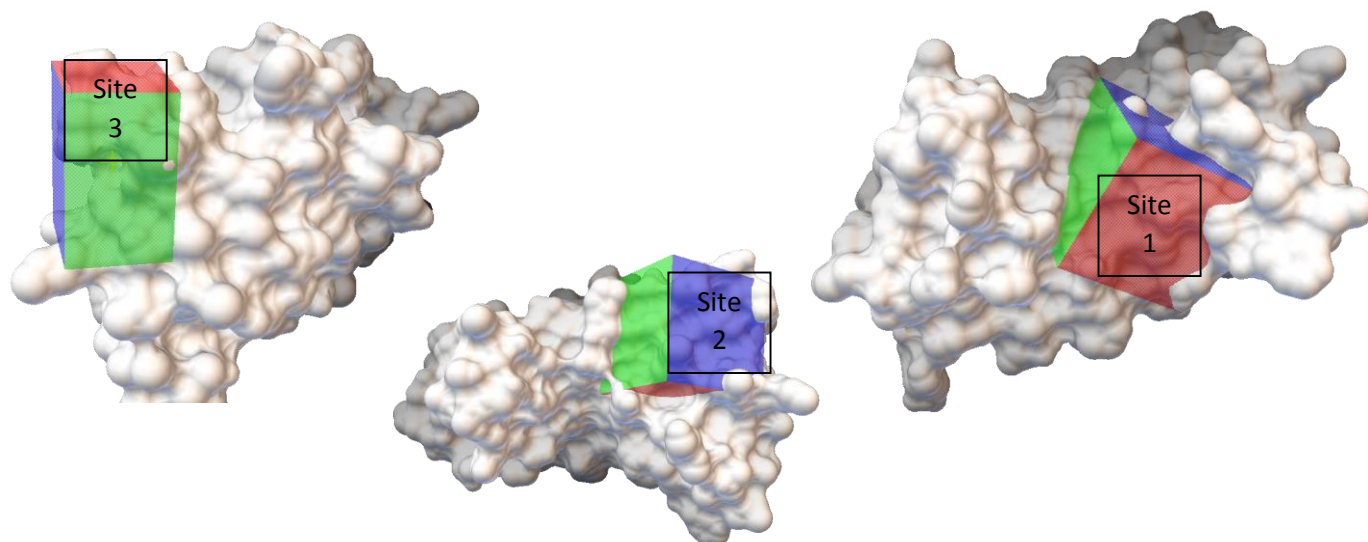
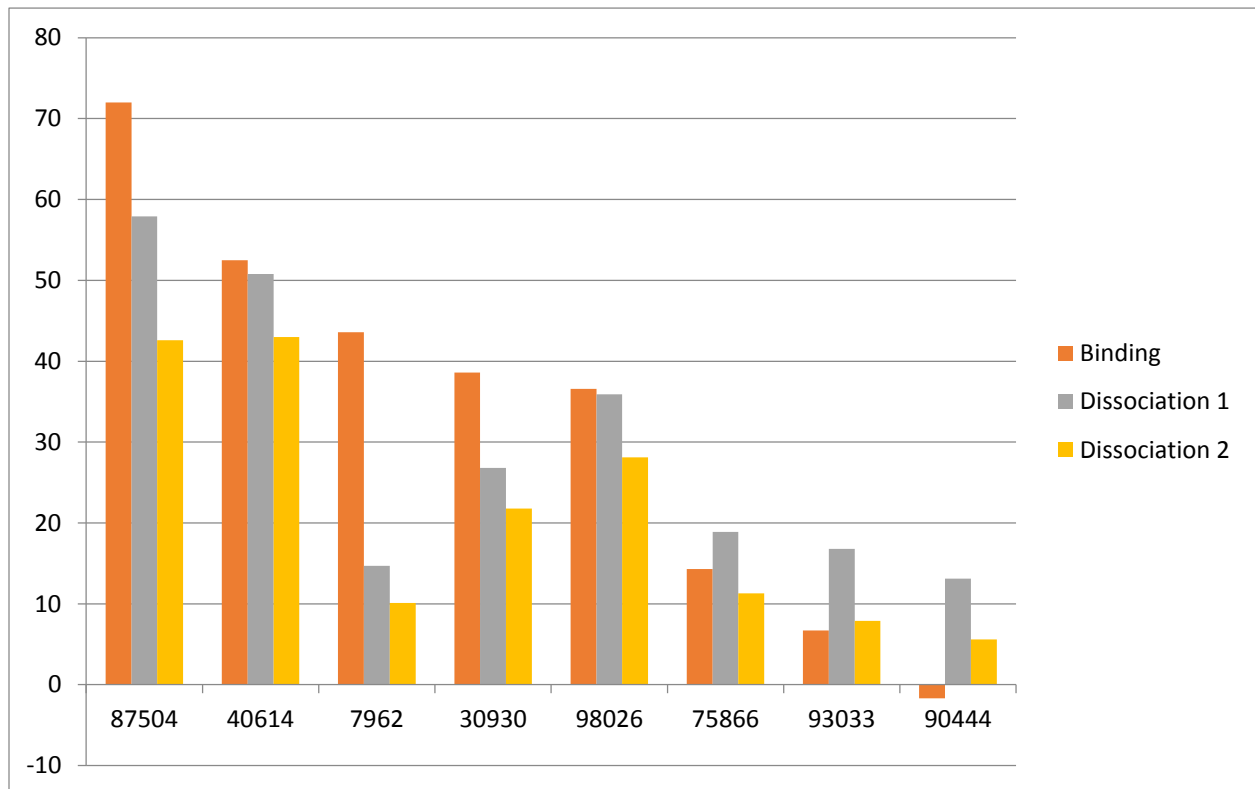
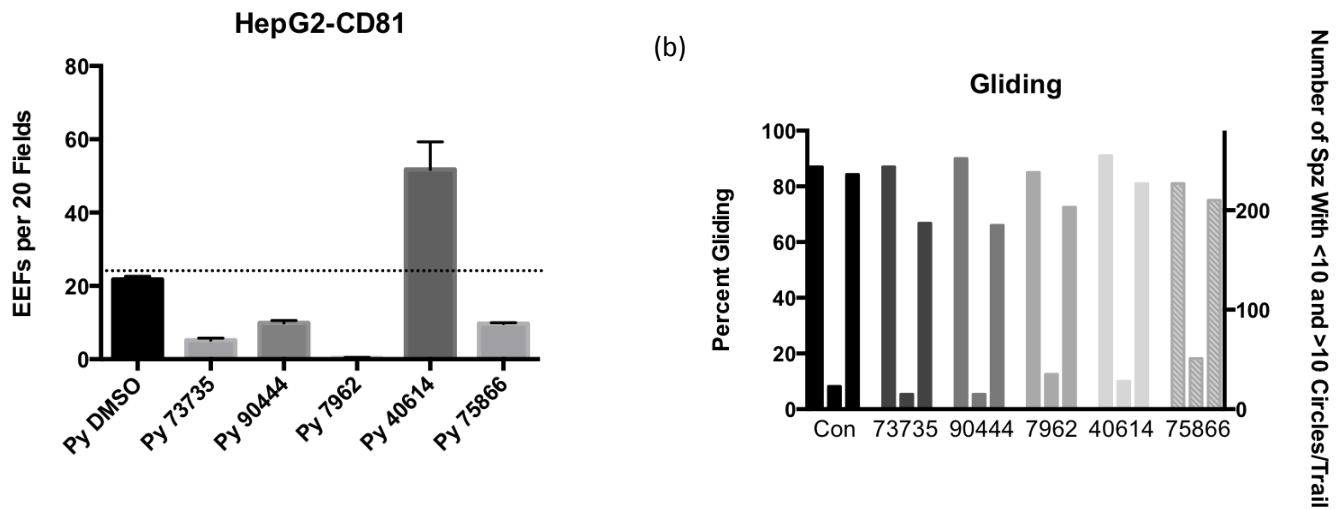


Figure 3. Bar chart representing the change in response units of the ligands that were found to bind to CD81-LEL using surface Plasmon resonance. The chart shows that 87504 had the best binding – dissociation behavior with CD81-LEL whereas 90444 had the least desirable binding-dissociation behavior.



(a) **Figure 4. Inhibition of *P. yoelli* Infection of HepG2 Cells.** Ligand 7962 was found to have the strongest inhibitory effect followed by ligand 73735, 75866 and 90444. Ligand 40614 was found to enhance the infection with *Pl. Yoleii* in HepG2-CD81(a). The gliding Assay(b).



- Figure 5. Inhibition of *P. yoelli* & *P. berghei* Infection of Hepa1-6 Cells.** Ligands were tested using Hepa 1-6 cells which express less CD81 so as to see the difference in behavior compared to HepG2 cells that express higher no. of CD81...It was found that both Ligands 7962 and 73735 exhibited an inhibitory effect yet, ligands 90444 and 75866 didn't show any inhibitory effect in case of *Plamodium yoleii* (a) and *Plasmodium berghei* (b).

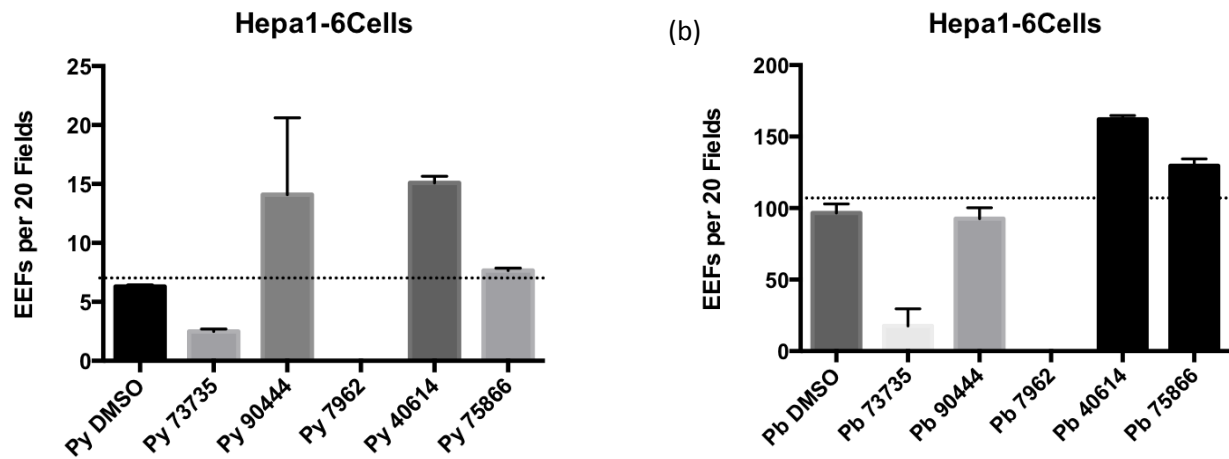


Figure 6. Comparison of Inhibition of *P.y.* and *P.f.* Infection of HepG2 Cells. The results of ligands tested using HepG2 cell line (a) were compared to its effect on infectivity of *P.y* and *P.f.* It was shown that ligands 7962 and 73735 affect the invasion step and not post invasion whereas ligands 90444 and 75866 affects steps other than invasion where in case of *P.f.* invasion is the only step observed (parasites inside versus outside of the cell). 40614 does not enhance infection in Pf (may affect post invasion event) (b).

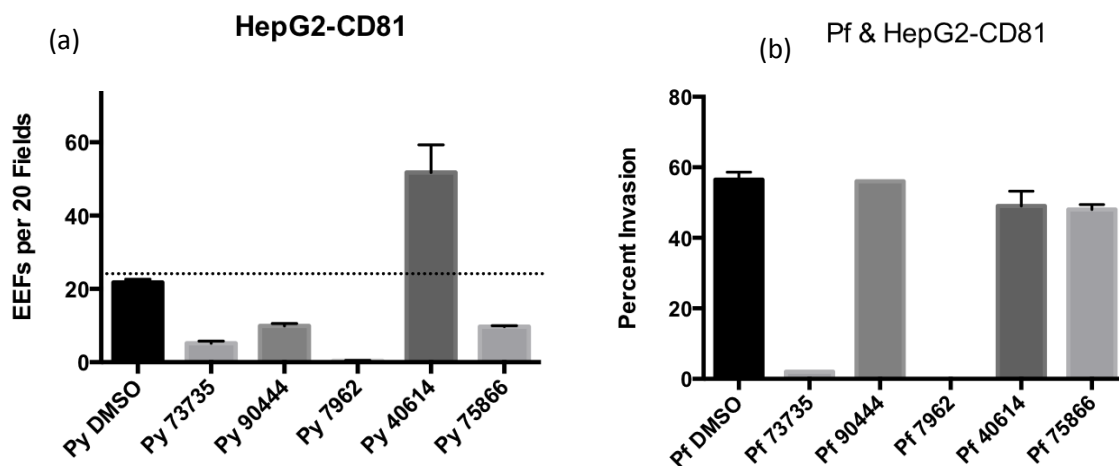


Figure 7. The different binding modes of 7962 to CD81. A number of 7962 conformations interacted with the 21 amino acid sequence and Asp137 that were known to be important for malaria infection.

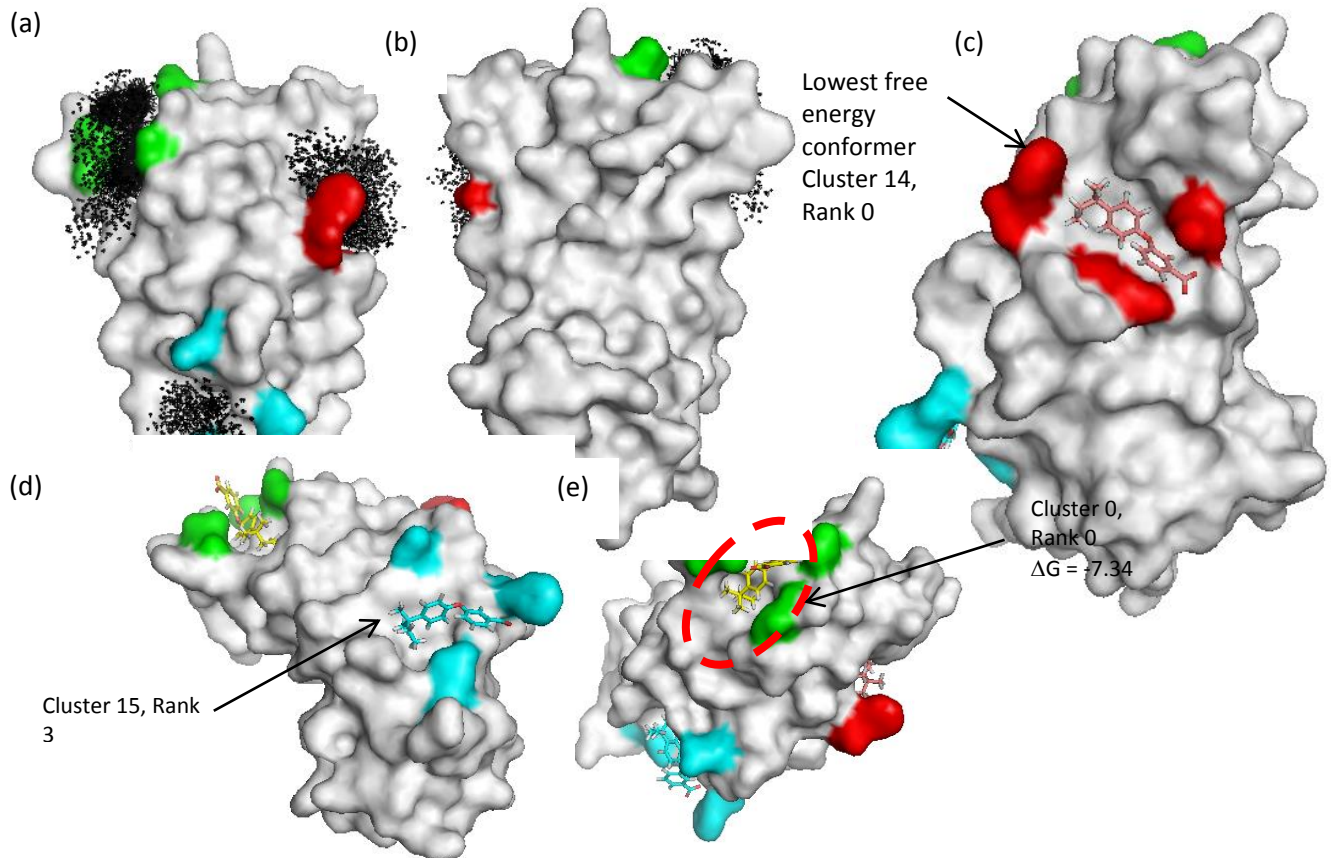
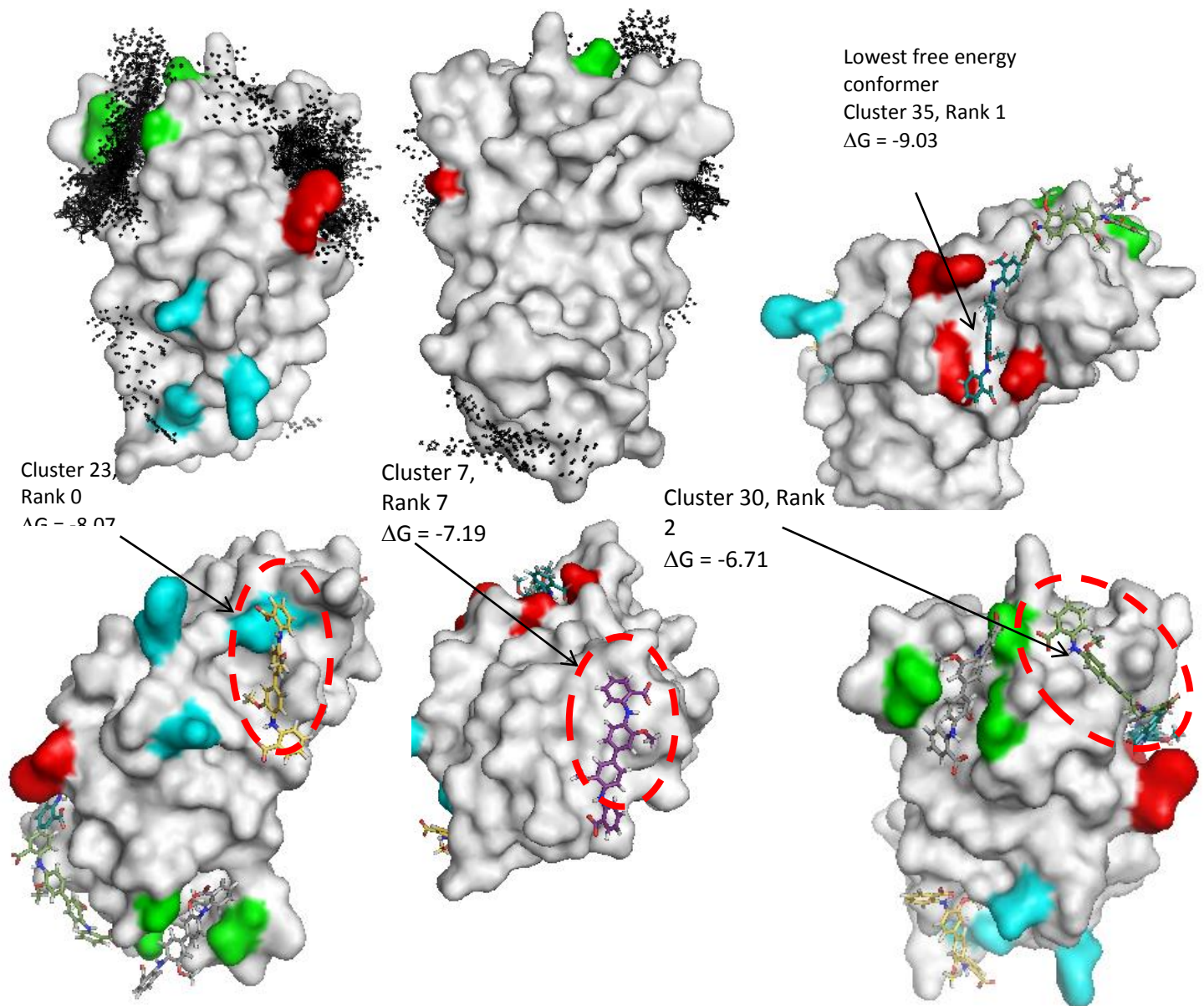


Figure 8. The different binding modes of 73735 to CD81. A number of 7962 conformations interacted with the 21 amino acid sequence and Asp137 that were known to be important for malaria infection.



7. Tables

Table 1. Surface Plasmon Resonance Experiment. This table shown the change in response units in relation to binding and dissociation of the 8 ligands that were tested.

Ligand ID	Binding (RU)	Dissociation 1 (RU)	Dissociation 2 (RU)
87504	72	57.9	42.6
40614	52.5	50.8	43
7962	43.6	14.7	10.1
30930	38.6	26.8	21.8
98026	36.6	35.9	28.1
75866	14.3	18.9	11.3
93033	6.7	16.8	7.9
90444	-1.7	13.1	5.6

Table 2. Percentage of infection inhibited by the 5 tested ligands in the 3 *Plasmodium* sp. Two cell lines were used to show the reduction in infectivity of *Pl. sporozoites* (HepG2 cells and Hepa 1-6)

Ligand	Percent Infection			
	HepG2-CD81		Hepa1-6 Cells	
	P.f.	P.y.	P.y.	P.b.
73735	3.2	23	38.3	17.7
90444	98.7	44.3	222	95.6
7962	0	0	0	0
40614	86.2	234	238	167
75866	84.2	42.6	122	135

CHAPTER VIII

CONCLUSIONS

AutoDock and its tool AutoLigand proved to be very helpful in identifying potential ligand binding sites on the surface of CD81-LEL. In addition to generating fill points for each cavity and using the collective points to provide information about the volume and depth of the cavity, a feature common to most docking programs, properties were identified for specific point groupings (features equivalent to atoms or functional groups) that would optimize the ligand's interaction with specific atoms lining the inner surface of the cavity. Using AutoLigand, we also increased our efficiency of identifying new molecules that bound to the protein. Previous studies using earlier versions of AutoDock that did not contain AutoLigand yielded results in which 25-55% of predicted binders actually bound to the target protein when tested experimentally. The virtual ligand screens (docking runs) performed in this study led to the identification of a diverse group of new small molecules that bind to CD81-LEL. A set of 36 ligands were tested by SPR and DPI to obtain a set of 26 new molecules we can use to develop inhibitors that block HCV invasion. Four of these ligands were observed to exhibit stronger or similar binding to CD81-LEL as benzyl salicylate, a small molecule reported to be a moderate inhibitor blocking the binding of HCV E2 to CD81. One of these ligands, 689002, has been found to inhibit the binding of HCV E2 to CD81-LEL. 689002 and the other ligands identified in this study will be used to develop second generation leads that bind more tightly to CD81-LEL using fragment-based drug design methods, and different combinations of Site

1 and Site 2 ligands will be linked together to create selective high affinity ligands called SHALs.

Two ligands, 689002 and 93033 were linked together to form a bidentate SHAL (SH7153). Using SPR analysis, SH7153 was found to bind to CD81-LEL when tested using SPR. But flow cytometry assays and HCV infection inhibition assays showed that SH7153 didn't have an inhibitory effect. This might be either due to the length of linkers which might need optimization or exchanging one of the two ligands with a new ligand that shows better binding mode *in silico*.

HCV E2 glycoprotein has a major role in the viral entry and its pathogenic effects being involved in different interactions with different host factors. Several research groups targeted different HCV proteins like serine protease and polymerase. The availability of the core structure of HCV E (E2c) (12) makes it feasible to generate a drug that targets the early stage of HCV via blocking HCV E2 interaction with different host factors. Although E2 has several hypervariable regions where HVR1 is highly mutative, targeting conserved residues in HCV E2 might help generate a drug that treats several HCV genotypes and subtypes.

In this study a structural model of the HCV E2 protein has been created using the E2c crystal structure as a template. Docking studies conducted with this model have led to the identification of 23 small molecule drug candidates that bind to the E2 protein. We were able to identify a new drug lead, 281816, which was found to inhibit HCV genotypes

and subtypes (1a, 1b, 2a, 2b, 3, 4a and 6a). This new drug lead might offer a new paradigm in HCV treatment. It might also be interesting to develop a cocktail of drugs that target E2 binding site on CD81 and CD81 binding site on E2, hence completely blocking the HCV E2: CD81 interaction and downstream pathogenic effects.

In several studies, SHAL were shown to increase the binding affinity when compared to single ligands. Hence, an *in silico* bidentate SHAL (SH2216) based on the studies on HCV E2 model was designed *in silico*. 281816 was shown to bind to several potential binding sites on the created homology model that comprises several significant residues such as PRO619 and TRP613. Owing to the fact that 281816 doesn't have neither free amino group nor free carboxyl group, a search for an analogue that do have one of these groups was conducted. Ligand 22594527 was found to have a free amino group. It was docked to the model to compare its binding modes to that of 281816 and it was found to have the same desirable binding to HCV E2. Two binding sites were chosen and the distance between the aminogroup of the first ligand and that of the second between the 2 sites was measured and an *in silico* designed SHAL was designed accordingly linking the 2 ligands with mini-PEG and lysine moieties.

Knowing that CD81 is a common target for HCV and malaria, we identified the significant residues involved in plasmodium sporozoite: CD81 interaction that are responsible for sporozoite invasion into the hepatocytes based on previous research done. We targeted those residues using AutoDock and did virtual screening runs against a library of ~4000 ligands. Eight out of thirty one ligands tested using surface Plasmon

resonance were found to bind to CD81. They were further tested using antibody neutralizing assays to see which inhibit binding CD81 to JS-81 antibody. They were also found to inhibit the JS-81: CD81 binding. Upon testing 5 of them using sporozoite invasion assay, Ligands 7962 and 73735 were found to inhibit *P. falciparum*, *P. yoleii* and *P. berghei* at 400uM conc.

Collectively, the approach adopted throughout the PhD project depended on determining virtual screening hits based on *in silico* analysis of the virtual screening results, testing them using surface Plasmon resonance to confirm binding to the target protein, followed by antibody neutralizing assays to validate that they may have an inhibitory effect of the infection and ending with infection inhibition assays and sporozoite invasion assays.

CHAPTER IX

FUTURE PERSPECTIVES

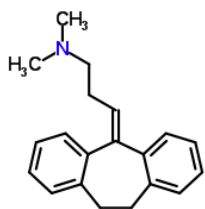
A library of promising compounds resulted from this study and this in turn necessitates the continuation of analyzing other subsets of these compounds for binding either to CD81 or HCV E2. Additionally, the drug leads that were already identified should be further tested in cell culture and animal models to be able to conduct pharmacokinetics and pharmacodynamics studies.

Since generating SHALs is considered a beneficial approach when it comes to increasing the binding affinity to the target protein. Selecting different hits against CD81 and generating a cocktail of bi-dentate and tridentate SHALs from them would help come up with a better drug candidate; since as the number of ligands increase in a SHAL, the binding affinity increase (Bi-dentate → Tri-dentate). It would also be helpful to try different lengths of the linkers in SH7153 so as to see whether the non-inhibitory effect was attributed to the linkers' length being not suitable or to the included ligands themselves (689002 and 93033).

281816, one of the promising drug leads was found to inhibit HCV in a genotype-independent manner. Testing it using a suitable HCV animal model would help take a step further towards confirming if this is a suitable drug candidate for the treatment of HCV. Additionally, it might be useful to synthesize SH2216 based on the *in silico* designed one, test it and compare its inhibitory effect to that of 281816 alone. It would also be interesting to test other 281816 analogues (figure 1) to observe their inhibitory effect as well.

The identified drug leads that are targeting the plasmodium sporozoite: CD81 interaction should be further tested using sporozoite invasion assay. This will help identify the best inhibitors and optimize them further to be suitable drug candidates.

Figure 1. 281816 analogues that were found to be FDA approved for treating different diseases.



Amitriptyline

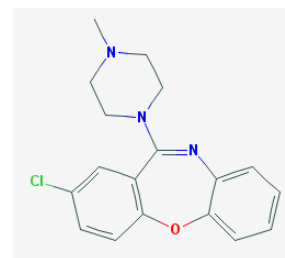
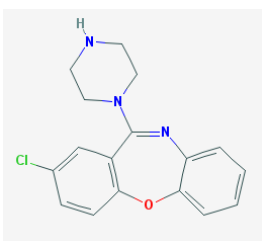
CN(C)CCC=C1C2=CC=CC=C2CCC3=CC=CC=C31

zinc968257

zinc93

C1CN(CCN1)C2=NC3=CC=CC=C3OC4=C2C=C(C=C4)Cl

Amoxapine (211µM)



Loxapine

CN1CCN(CC1)C2=NC3=CC=CC=C3OC4=C2C=C(C=C4)Cl

zinc1639

Quetiapine

C1CN(CCN1CCOCCO)C2=NC3=CC=CC=C3SC4=CC=CC=C42

zinc19632628

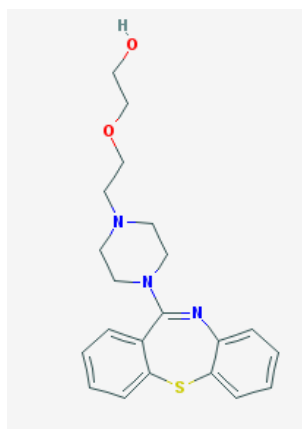


Table 1. The 281816 drug analogs and their corresponding indications

Analog/Drug	Indications
Amoxapine	Antidepressant
Amytryptiline	Tricyclic antidepressant
Loxapine	Typical Antipsychotic
Quetiapine	Atypical Antipsychotic

CHAPTER X

NOVELTY

The novelty of the work done in this thesis project lies in targeting the entry phase of HCV instead of the replication and translation process which is targeted by polymerase and protease inhibitors. The aim of this project was to identify drug leads that have the ability to block the entry phase of HCV via blocking HCV E2: CD81-LEL interaction and thus altering any upstream processes. The two main protein targets that were used were CD81-LEL and HCV E2. In case of HCV E2, a homology model was created based on the newly resolved crystal structure of the core of HCV E2 (E2c) (PDB ID: 4MWF), and this was further used in virtual screening to identify potential hits that could block the entry phase. This helped identify a promising drug lead (281816) that was found to block HCV infection and cell to cell transmission. On the other hand, specific residues in CD81-LEL that were known to interact with HCV E2 from previously published literature were targeted so as to identify additional set of hits that could block the entry phase.

Beside targeting HCV, CD81-LEL was known to be involved in malaria infection through interacting/facilitating the internalization of the *Plasmodium sp.* Sporozoites invasion into the hepatocytes. Accordingly, we targeted the specific residues in CD81-LEL that were known to be involved in internalization of malaria sporozoites. This in turn led to the identification of 2 drug leads that were found to exert inhibitory effects on *Pl. falciparum* which is one of the most aggressive human plasmodium parasites.

CHAPTER XI
ORIGINALITY STATEMENT

‘I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at AUC or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at AUC or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.’

Signed: [Reem R. Al Olabi](#).

Date: 5/18/2014