

RESEARCH ARTICLE

# Identification of Entry Factors Involved in Hepatitis C Virus Infection Based on Host-Mimicking Short Linear Motifs

Austin W. T. Chiang<sup>1a,2</sup>, Walt Y. L. Wu<sup>3</sup>, Ting Wang<sup>1b,2c</sup>, Ming-Jing Hwang<sup>\*</sup>

Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan

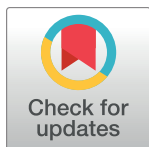
<sup>1</sup> These authors contributed equally to this work.

<sup>1a</sup> Current address: Department of Pediatrics, University of California, San Diego, CA, United States of America

<sup>1b</sup> Current address: Chemical Biology and Molecular Biophysics Program, Taiwan International Graduate Program (TIGP), Academia Sinica, Taipei, Taiwan

<sup>1c</sup> Current address: Institute of Biochemical Sciences, College of Life Science, National Taiwan University, Taipei, Taiwan

\* [mjhwang@ibms.sinica.edu.tw](mailto:mjhwang@ibms.sinica.edu.tw)



**OPEN ACCESS**

**Citation:** Chiang AWT, Wu WYL, Wang T, Hwang M-J (2017) Identification of Entry Factors Involved in Hepatitis C Virus Infection Based on Host-Mimicking Short Linear Motifs. *PLoS Comput Biol* 13(1): e1005368. doi:10.1371/journal.pcbi.1005368

**Editor:** Claus O. Wilke, University of Texas at Austin, UNITED STATES

**Received:** September 8, 2016

**Accepted:** January 17, 2017

**Published:** January 27, 2017

**Copyright:** © 2017 Chiang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This work was funded by the Ministry of Science and Technology, Taiwan (<https://www.most.gov.tw/>), under grant number MOST 104-2811-B-001-004. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

## Abstract

Host factors that facilitate viral entry into cells can, in principle, be identified from a virus-host protein interaction network, but for most viruses information for such a network is limited. To help fill this void, we developed a bioinformatics approach and applied it to hepatitis C virus (HCV) infection, which is a current concern for global health. Using this approach, we identified short linear sequence motifs, conserved in the envelope proteins of HCV (E1/E2), that potentially can bind human proteins present on the surface of hepatocytes so as to construct an HCV (envelope)-host protein interaction network. Gene Ontology functional and KEGG pathway analyses showed that the identified host proteins are enriched in cell entry and carcinogenesis functionalities. The validity of our results is supported by much published experimental data. Our general approach should be useful when developing antiviral agents, particularly those that target virus-host interactions.

## Author Summary

Viruses recruit host proteins, called entry factors, to help gain entry to host cells. Identification of entry factors can provide targets for developing antiviral drugs. By exploring the concept that short linear peptide motifs involved in human protein-protein interactions may be mimicked by viruses to hijack certain host cellular processes and thereby assist viral infection/survival, we developed a bioinformatics strategy to computationally identify entry factors of hepatitis C virus (HCV) infection, which is a worldwide health problem. Analysis of cellular functions and biochemical pathways indicated that the human proteins we identified usually play a role in cell entry and/or carcinogenesis, and results of the analysis are generally supported by experimental studies on HCV infection, including the ~80% (15 of 19) prediction rate of known HCV hepatocyte entry factors. Because

molecular mimicry is a general concept, our bioinformatics strategy is a timely approach to identify new targets for antiviral research, not only for HCV but also for other viruses.

## Introduction

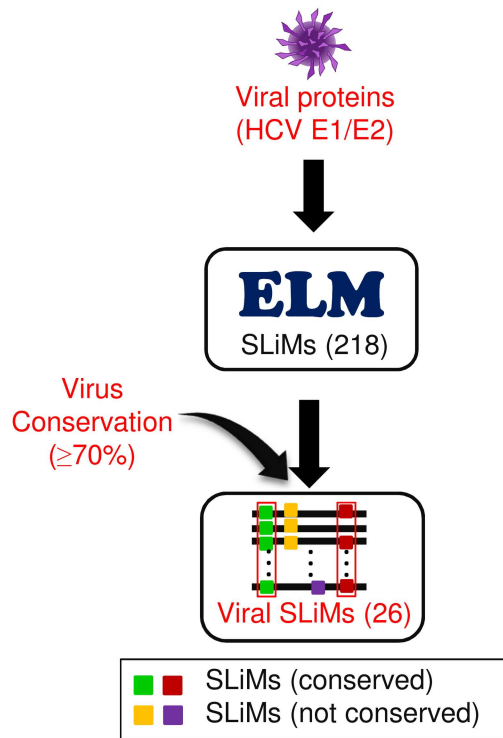
The conventional approach to countering viral infections has been to develop drugs that target viral genetic material or proteins. However, two major roadblocks to this strategy exist: 1) the limited number of druggable viral proteins owing to small viral genomes, and 2) drug resistance that occurs on a relatively short time scale owing to substantial viral genomic mutation rates. To circumvent these problems, over the past decade antiviral drug development has shifted from targeting viral proteins to host proteins that interact with components of the virus [1]. For example, compounds that inhibit interactions between viral and human proteins have been identified [2], including the compound LEDGIN, which targets the interaction between HIV integrase and human transcriptional coactivator p75 [3]. Cell-based genomic and proteomic assays that screen for host targets that interact with viral proteins have also been reported [4–6]. Nevertheless, given the large amount of biological data that has been accumulated from high-throughput omics-type experiments, development of a bioinformatics-sleuthing strategy that identifies potential antiviral host targets to complement experimental screens should be of considerable merit.

Herein, we describe the development of and evaluate such a bioinformatics strategy, the premise of which is based on viral “molecular mimicry,” an ability that viruses have developed over millions of years of evolution to antagonize their hosts [7]. Specifically, regions in viral proteins apparently can mimic short amino acid sequences found in human proteins involved in normal host protein-protein interactions (PPIs), so that a virus can hijack the PPI for its own purposes, such as hijacking a cellular process(es) to create the cell context needed for infection [8]. Consistent with this viral strategy, their proteins often contain host-like SLiMs (Short Linear Motifs) allowing them to interact with complementary host proteins [9, 10]. Viral SLiMs can be identified by sequence comparison with those with the ability to bind eukaryotic protein domains as catalogued in the database ELM (Eukaryotic Linear Motif) [11]. The viral SLiMs, the host proteins that contain a matched SLiM-binding domain, and these proteins’ interacting partners in the human PPI network then form a putative virus-host interaction network, which can be integrated with known functional and network properties of cellular pathways, including those involved in disease states, thereby allowing identification of host factors whose native functions may be altered or hijacked by the virus to facilitate its infection and/or another of its life cycle stages.

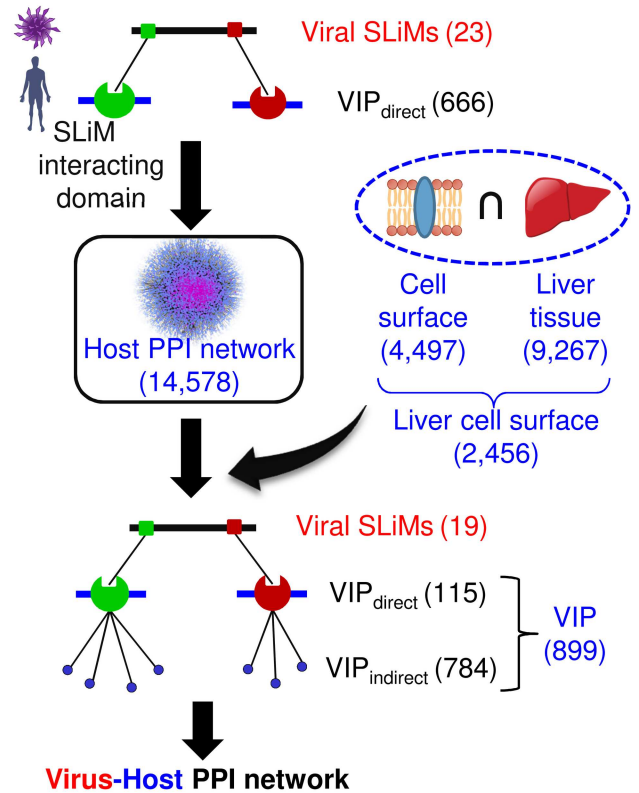
To examine this molecular mimicry strategy and the feasibility of using human PPI network data to complement experimental studies, we focused on the hepatitis C virus (HCV) envelope proteins, E1 and E2, for the following reasons: First, HCV infection is a major health problem worldwide [12], and HCV E1 and E2 are known to play essential roles in HCV entry into human hepatocytes [13]; investigating E1 and E2 might therefore lead us to identify novel HCV entry factors as targets for drug design—an important step toward developing more effective anti-HCV drugs. Second, the complexity of the network and functional analysis required was significantly reduced because only liver cell surface proteins of the human proteome need to be considered. Third, many HCV entry-facilitating human proteins have been identified, which allowed us to compare *in silico* predictions with published experimental data.

Using the HCV E1 and E2 sequences as examples, Fig 1 schematically depicts the four main components of our approach, which are detailed in Methods. First, conserved E1/E2 sequences

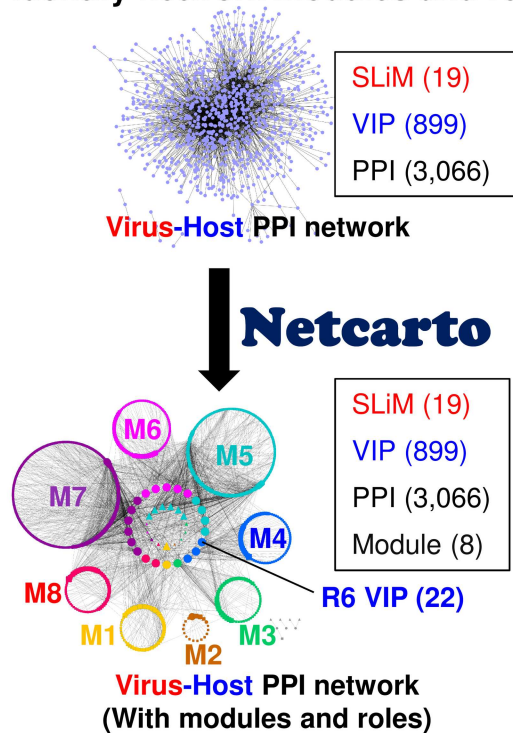
**A. Identify viral SLiMs**



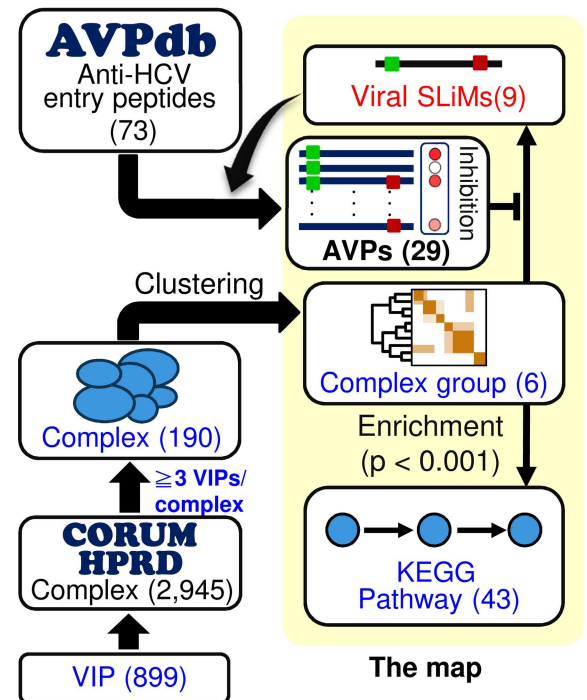
**B. Construct Virus-Host PPI network**



**C. Identify network modules and roles**



**D. Build AVP/SLiM-Complex map**



**Fig 1. The four components of our bioinformatics strategy.** (A) Identify human-type short linear motifs (SLiMs) found in a viral protein(s). (B) Construct a virus-host-PPI network. (C) Identify network modules and roles of the network nodes. (D) Build a map of AVP/SLiM-protein complexes. In parentheses are the numbers of motifs, proteins, modules, pathways, etc., identified in this study (see [Methods](#) for details).

doi:10.1371/journal.pcbi.1005368.g001

from various HCV strains are identified that correspond to SLiMs found in the eukaryotic linear motif (ELM) database ([Fig 1A](#)). Next, proteins on the surface of human hepatocytes known to bind such SLiMs are identified (they are called  $VIP_{s_{direct}}$  for Virus-Interacting host Proteins), as are host proteins ( $VIP_{s_{indirect}}$ ) that bind  $VIP_{s_{direct}}$  ([Fig 1B](#)). Taking the experimentally determined interactions between  $VIP_{s_{direct}}$  and  $VIP_{s_{indirect}}$  from the original human PPI network ([14]; see [Methods](#)), a virus-host PPI network is then extracted that is constructed of the viral SLiMs and human VIPs ([Fig 1B and 1C](#)). This network contains modules (communities) of functionally related host proteins (nodes) and links (PPIs) within and between the modules connecting interacting nodes ([Fig 1C](#)). Finally, a map connecting SLiMs to known antiviral peptides (AVPs) and complexes containing multiple ( $\geq 3$ ) SLiM-interacting proteins is produced ([Fig 1D](#)), for which statistical analyses are carried out to find enriched functionalities and pathways that correlate with published experimental data.

As shown below, the results show that the proteins we identified as possible hepatitis C virus (HCV) entry factors: 1) have a statistically significant propensity to be found in the PHISTO and EHCO lists, which contain experimentally identified HCV-interacting proteins and genes differentially expressed in HCV-induced hepatocellular carcinoma, respectively [15, 16]; 2) have greater coverage of known HCV entry factors than a functional genomics screening experiment [5]; and 3) contain domains that can bind short linear motifs that are also present in many antiviral peptides with experimentally demonstrated activities against HCV infection. These results suggest that, to eliminate viral infection, more attention should be paid to sequence motifs involved in host protein-protein interactions because these motifs may be subject to molecular mimicry by viruses.

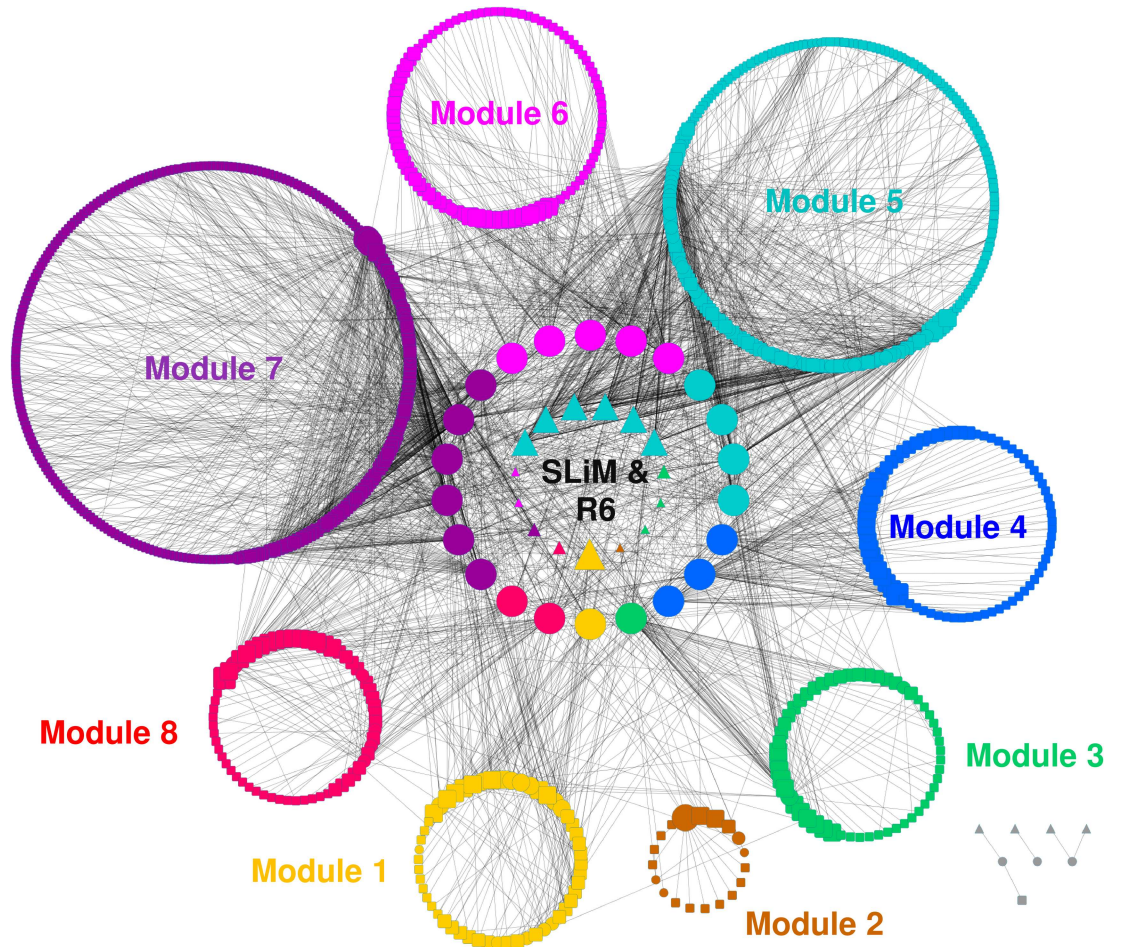
## Results

### A SLiM-derived HCV-human PPI network

We identified 19 SLiMs on HCV E1 and E2 that might bind various human protein domains ([Fig 1B](#)). Screening for human hepatocyte surface proteins in conjunction with the available human PPI network yielded 115  $VIP_{s_{direct}}$  containing at least one SLiM-binding domain. These proteins and their  $VIP_{s_{indirect}}$ , which interact with them, constitute a subset of the experimentally derived human PPI network [14] that might be directly or indirectly influenced by the mimicking HCV SLiMs. It follows, according to the premise of our molecular mimicry strategy, that the host  $VIP_{s_{direct}}$  and  $VIP_{s_{indirect}}$  potentially facilitate or inhibit HCV entry and, along with the viral SLiMs, they formed a viral-host PPI network ([Fig 1B and 1C](#)).

### Functional modules and network roles

Given a network, algorithms are available to extract network properties [17]. Using NetCarto, a tool for network module discovery [18], we found that the resulting viral-host PPI network for HCV infection is organized into eight modules with 23 R6 (global connector) hubs ([Fig 2](#)). A global connector hub (R6) is defined as a node with many links to most of the other network modules (see [Methods](#) for definitions on roles of network nodes), and as such it is thought to play an important role in connecting different functional modules. Consistent with the definition, whereas most of the 115  $VIP_{s_{direct}}$  interact with only a few other host proteins, these 22 R6 hub proteins (the twenty-third R6 hub is a viral SLiM) have many interaction partners.



<u>Node type/role</u>	R7	R6	R5	R4	R3	R2	R1
	(0)	(23)	(9)	(2)	(83)	(304)	(497)
SLiMs (19)	△ (0)	△ (1)	△ (6)	△ (0)	△ (0)	△ (3)	△ (9)
VIP <sub>direct</sub> (115)	○ (0)	○ (22)	○ (3)	○ (0)	○ (11)	○ (63)	○ (16)
VIP <sub>indirect</sub> (784)	□ (0)	□ (0)	□ (0)	□ (2)	□ (72)	□ (238)	□ (472)

**Network Module**

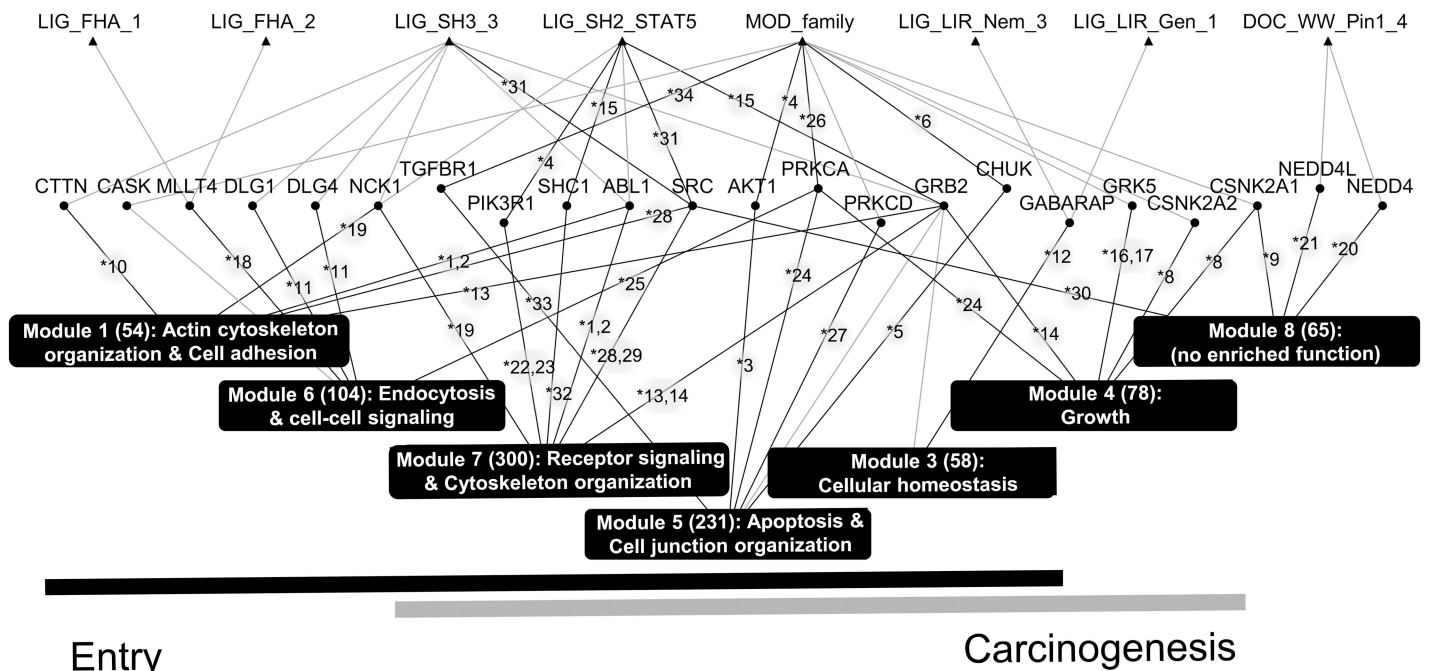
- Module 1 (54): Actin cytoskeleton organization & Cell adhesion**
- Module 2 (20): Peptide hormone processing**
- Module 3 (58): Cellular homeostasis**
- Module 4 (78): Growth**
- Module 5 (231): Apoptosis & Cell junction organization**
- Module 6 (104): Endocytosis & cell-cell signaling**
- Module 7 (300): Receptor signaling & Cytoskeleton organization**
- Module 8 (65): (no enriched function)**

**Fig 2. Modules and roles of the nodes of the HCV-Human PPI network.** The network was produced using the procedures described in Fig 1A–1C, with the modules and the roles of network nodes determined by NetCarto (see Methods). The three types of network nodes are represented as triangles for SLIMs; circles for VIPs<sub>direct</sub>; and squares for VIPs<sub>indirect</sub>. Their network roles (R7 to R1) are depicted as symbols of decreasing size. All nodes are color-coded according to their module. A representative function(s) of each module was derived from an enrichment analysis of GO terms associated with its nodes followed by a summary of Revigo representatives [21] (see Methods and S1 Table). Colored in gray in the lower right corner are eight isolated nodes.

doi:10.1371/journal.pcbi.1005368.g002

This may imply that these R6 hubs could be important host factors for HCV infection, and could serve as targets for designing anti-HCV drugs (see AVPs analysis below). Further analysis showed that 15 out of the 22 R6 hubs ( $P < 2.2 \times 10^{-16}$ ; S1 Fig) were also hub proteins in the experimentally derived PPI network of human liver cell surface proteins, suggesting that most of these VIPs<sub>direct</sub> R6 hubs have important functions for the host, irrespective of HCV infection. This is in line with the finding that viruses tend to target host hub proteins for perturbing key pathways (or biological processes) to benefit viral infections [19]. Interestingly, Gene Ontology (GO) enrichment analysis [20] and Revigo summarization [21] of the enriched GO annotations (S1 Table) revealed that the representative functions of seven of the eight modules belong to one, or both, of the two main functionalities: *entry* and *carcinogenesis* (Fig 3 and S2 Table).

As described below, much experimental data is available to support our *in silico* observations. For example, “Cytoskeleton organization” (modules 1 and 7; Figs 2 and 3) is an essential cellular process that allows HCV to migrate to the tight junction where internalization and



**Fig 3. Relationships between SLIMs, R6 VIPs, and network modules.** Of the 19 SLIMs identified for HCV E1 and E2 in Fig 1C, 13 (including six grouped in the MOD\_family; top row) are directly connected to one or more of the 22 R6 VIPs<sub>direct</sub> (middle row) in the virus-host PPI network (Fig 2). The MOD\_family contains MOD\_CK1\_1, MOD\_CK2\_1, MOD\_GSK3\_1, MOD\_NEK2\_1, MOD\_NEK2\_2, and MOD\_ProDKin\_1; all are targets of a kinase. An R6 VIP<sub>direct</sub> and a module(s) (bottom, boxed in black) are considered to be connected if more than 10% of the interacting partners of the VIP<sub>direct</sub> belong to the module. Based on this criterion, module 2 is not connected to an R6 VIP<sub>direct</sub> and, therefore, is not included in the figure. The validity of the connections displayed as solid dark lines is supported by published experimental evidence. The corresponding reference number(s) (indicated by an asterisk) are provided in S3 Table. The dark and light horizontal bars at the bottom of the figure identify modules with the functionality of *entry* and/or *carcinogenesis*, respectively.

doi:10.1371/journal.pcbi.1005368.g003

endocytosis of the virion occur [22]. Notably, some of the proteins of the R6 hubs are involved in this cellular process. According to our hypothesis, cellular processes involved in “*Cytoskeleton organization*” might be hijacked by HCV if one or more of its E1/E2 SLiMs can bind at least one of the following six R6 proteins (in this work we use official gene symbols to represent proteins encoded by the corresponding genes): PIK3R1, which enhances actin reorganization by activating PI3K-AKT signaling [23]; SRC, which induces changes in the cytoskeleton by binding and activating FAK [24]; and ABL1 [25], GRB2 [26], NCK1 [27], and CTTN [28], proteins which regulate cytoskeleton rearrangement.

A function found for module 5 is “*apoptosis*” (Figs 2 and 3), which is an essential cellular process as it prevents HCV from spreading in the host by inducing the death of HCV-infected cells [29]. However, E2 can suppress cellular apoptosis resulting in HCV survival [30]. In our viral-host PPI network, there are five module 5-associated R6 proteins, four of them, AKT1 [23, 31], CHUK [32, 33], PRKCA [34, 35], and TGFBR1 [36, 37] have a role in apoptosis and their activities and/or expressions are known to be affected by HCV infection, although the specific effects of E1 and/or E2 on CHUK, PRKCA, and TGFBR1 activities have yet to be determined. The apoptosis-regulating role of the fifth R6 protein PRKCD [38] during HCV infection has also not been examined.

Another representative function of module 7, “*receptor signaling*,” contributes to HCV entry [39] and carcinogenesis [40]. Specifically, HCV infection triggers EGFR signaling and stimulates its downstream signaling, including those of HRAS and PI3K-AKT [39]. Activation of these pathways enhances HCV entry [23, 39] and increases the proliferation of hepatocytes, which may contribute to hepatocellular carcinogenesis [40]. Three R6 proteins are associated with receptor signaling: PIK3R1, a PI3K subunit and a crucial participant in PI3K-AKT signaling [41], and GRB2 [42] and SHC1 [43], two key adaptor proteins of EGFR signaling, which when silenced substantially impair HCV entry [39].

Additional published experimental data that support the relationships between the R6 proteins and the functions of the viral-host PPI network modules are summarized in S3 Table.

## VIPs also found in PHISTO and EHCO

We identified 899 VIPs using our scheme. To evaluate the validity of our findings, we compared our list of VIPs with those in the PHISTO (Pathogen-Host Interaction Search Tool) dataset, which contains a list of experimentally verified HCV-interacting human proteins [16]. There are a total of 698 HCV-interacting human proteins in PHISTO, of which 160 are in the set of 2,456 liver cell surface proteins (Fig 1B). Of the 160 HCV-interacting hepatocyte surface proteins, 158 are annotated specifically as interacting with the polyprotein of HCV (S4 Table), which contains E1 and E2. As shown in S2A Fig, the 899 VIPs tended to contain members from the 158 subset of the PHISTO list, with 92 proteins overlapped between the two ( $P = 9.2 \times 10^{-9}$ ; S2A Fig). Similarly, the predicted interactions between HCV(SLiMs) and VIPs, i.e. the edges of the viral-host PPI network, were enriched in PHISTO ( $P = 4.3 \times 10^{-4}$  and  $9.6 \times 10^{-5}$  for direct and indirect interactions, respectively; see S3 Fig and S4 Table). These results indicate that our bioinformatics approach preferentially identified HCV-interacting human proteins.

Four of the identified virus-host PPI network modules are associated with carcinogenic processes (Fig 3), which is a somewhat unexpected result as E1 and E2 are usually only considered to be entry proteins [13]; however, this association is consistent with known oncogenic effects of E1 and E2 [44, 45]. To further evaluate the involvement of the VIPs in hepatocellular carcinoma (HCC), we assessed if a significant number of those VIPs are found in the EHCO (Encyclopedia of Hepatocellular Carcinoma genes Online) [15] dataset. Of the 614 genes that

are differentially expressed in HCV-caused HCC in EHCO, 194 are expressed on liver cell surface, and 91 of their encoded proteins were identified as VIPs ( $P = 1.4 \times 10^{-3}$ ; S2B Fig), further supporting the notion that E1/E2 interact with host proteins that have a role in carcinogenesis.

### Comparison with a siRNA-screening experiment list

Recently, a set of host factors for HCV entry was identified in a large-scale siRNA screening experiment reported by Li and coworkers [5]. However, the authors of that study identified only four of the 15 human proteins found on hepatocyte surface and known to be associated with HCV entry (Table 1). By comparison, we identified 11 of these entry factors known before Li and colleagues performed their study, and we also identified the four proteins found by them (Table 1). Three of the known entry factors that our study did not find, CLDN1, SCARB1, and CD209, are connected to at least one VIP<sub>indirect</sub> in the human PPI network, but are not VIPs themselves (S4A–S4C Fig). The fourth, NPC1L1, which we did not identify, lacks information of interaction with any of the VIPs in the network (S4D Fig).

CD81, OCLN, CLDN1 and SCARB1 are arguably the four best known HCV entry-related factors [48]. Of them, we identified CD81 and OCLN as VIPs<sub>indirect</sub> but failed to find CLDN1 and SCARB1 as noted above. Interestingly, transgenic expression of human CD81 and OCLN in mouse enabled HCV infection of mouse hepatocytes, whereas transgenic expression of human CLDN1 or SCARB1 was not necessary for mouse cell to be infected by HCV [49]. As a

**Table 1. Comparison with known entry factors of HCV infection.**

Gene symbol	Protein name	Identified <sup>a</sup>		
		This study (15)	Li 2014 <sup>b</sup> (8)	Literature <sup>c</sup> (15)
CD81	CD81 antigen	Y	Y	Y
CDC42	CDC42	Y	Y	Y
RAC1	Ras related C3 botulinum toxin substrate 1	Y	Y	Y
APOE	Apolipoprotein E	Y	N	Y
EGFR	EGF receptor	Y	N	Y
EPHA2	EphA2	Y	N	Y
HRAS	H-Ras	Y	N	Y
LDLR	Low density lipoprotein receptor	Y	N	Y
OCLN	Occludin	Y	N	Y
PCSK9	Proprotein convertase PC9	Y	N	Y
TFRC	Transferrin receptor	Y	N	Y
CLDN1	Claudin 1	N	Y	Y
CD209	DC-SIGN	N	N	Y
NPC1L1	NPC1 like 1	N	N	Y
SCARB1	Scavenger receptor class B member 1	N	N	Y
ARHGEF7	Rho guanine nucleotide exchange factor 7	Y	Y	N
CDH1	Cadherin-1 precursor	Y	Y	N
RAB34	Ras-related protein Rab-34, isoform NARR	Y	Y	N
RBP4	Retinol-binding protein 4 precursor	Y	Y	N

<sup>a</sup> Y, yes; N, no.

<sup>b</sup>Eight proteins (ARRDC2, CHKA, CYBA, DDX3X, FASN, MAP4, PIK4CA, and ROCK2) identified in [5] as HCV entry factors are not expressed on the liver cell surface according to the HPRD [46] and the Human Proteinpedia [47]; thus, these factors were excluded from our analysis.

<sup>c</sup>Proteins annotated with “Entry” and “Attachment” in the column of “HCV Life Cycle Stages Affected” in S9 Table in [5] were compared, however CLEC4M, FASN, IFITM1, PIK4CA, ROCK2, and SDC1 in S9 Table of the report were not included because they were not expressed on the liver cell surface according to the HPRD [46] and the Human Proteinpedia [47].

doi:10.1371/journal.pcbi.1005368.t001



VIP<sub>indirect</sub> CD81 is not predicted to possess a domain that can directly bind an E1/E2 SLiM, which might be considered to contradict a report suggesting a physical interaction between HCV E2 and CD81 [50] can occur. Nonetheless, viruses employ other strategies to interact with host proteins [7], and indeed, substitution mutation experiments suggested that HCV E2 uses a non-sequential motif to bind CD81 [51], which we would not have uncovered using the SLiM-based approach. Altogether, our approach found more known HCV hepatocyte entry host factors (15 of 19 (79%); Table 1) than did the siRNA functional genomics-screening assay of Li and colleagues. Furthermore, an ROC (Receiver Operating Characteristic [52]) curve analysis accounting for not only sensitivity but also specificity showed that the *in silico* predictions yielded an AUC (area under curve) of 0.68, which is almost as good as that (0.70) of Li et al.'s experimental screening (S5 Fig). Note that including protein complexes in the analysis (see below) significantly improved on specificity for the *in silico* predictions while maintaining the overall performance; similarly, in Li et al.'s experiment, a large number of genes (19,277) had been removed by a genome-wide genetic screen [53] prior to the siRNA functional analysis [5], and these were not included in the ROC curve analysis, giving rise to a much higher specificity for Li et al.'s experiment (S5 Fig). In addition, because all of those we predicted as novel ones were treated as false positives in these calculations, the actual performance of our predictions could possibly be better.

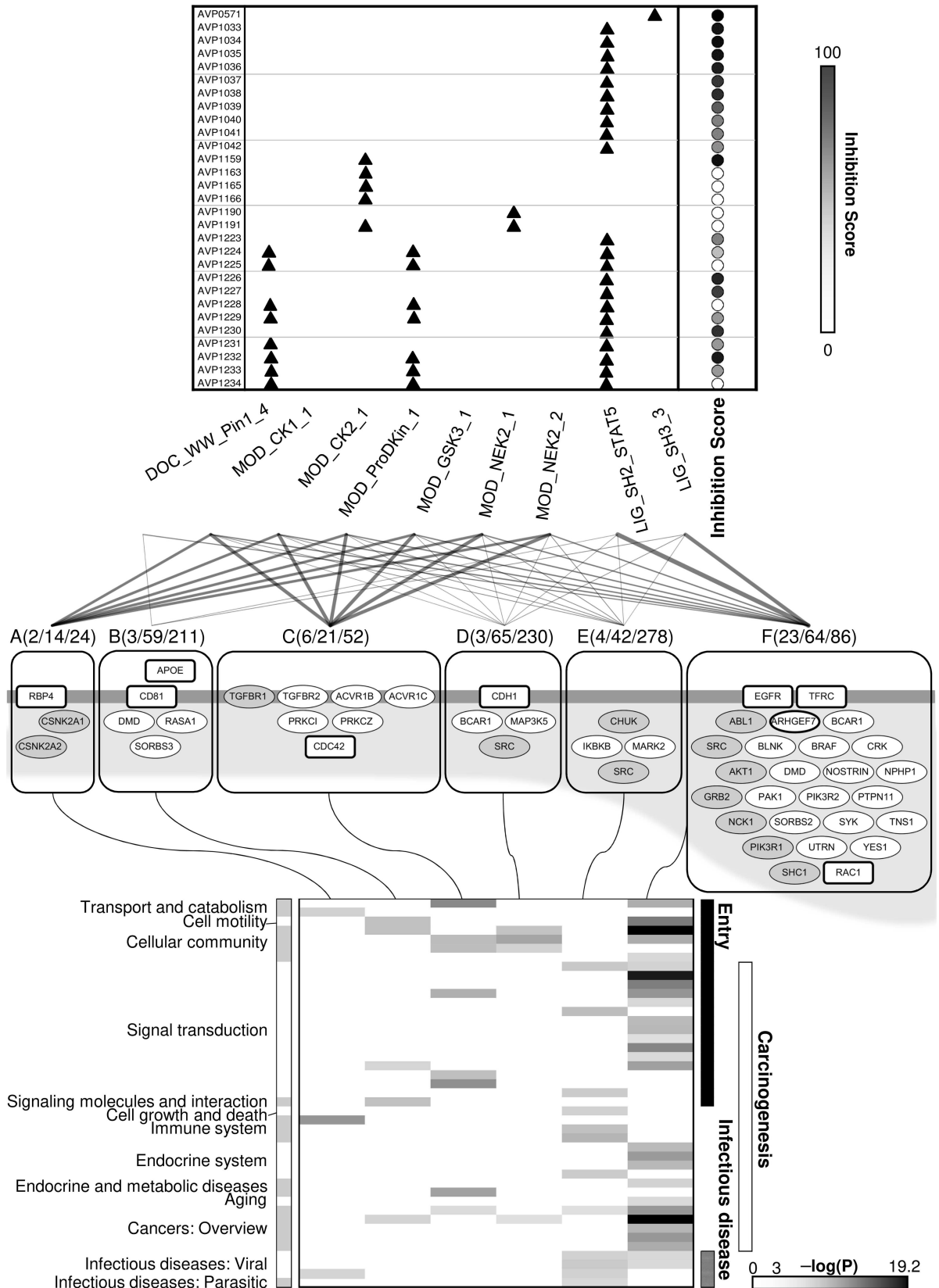
## Protein complexes and KEGG pathways

Viruses are known to target protein complexes so as to heavily perturb host functions and induce disease states, e.g., carcinogenesis [6, 54]. It is, therefore, of interest to identify human protein complexes that might be perturbed by HCV infection, including those involving VIP<sub>direct</sub> and VIP<sub>indirect</sub>. Knowledge of such protein complexes would also greatly reduce the number of VIPs needed to be considered for pathway analysis and experimentation.

By mapping the VIPs to protein complexes in the HPRD [46] and CORUM [55] databases (these databases categorize human and mammalian protein complexes, respectively) and clustering according to shared subunits, we identified six groups of protein complexes (Fig 4 and S6 Fig and S5 Table) that HCV may target by interaction with the VIPs of the complexes. These six groups contain 231 of the 899 VIPs and nine of the 19 viral SLiMs, with six of the nine SLiMs capable of binding the same protein domains (Fig 4). A comparison of our results with those obtained using the same mapping procedure and the same number of randomly selected proteins showed that the tendency of the HCV SLiMs-derived VIPs to be present in these protein complexes is unlikely to occur by chance ( $P < 1.0 \times 10^{-5}$ ; S7 Fig). Furthermore, 11 of the 22 R6 hub proteins are present in these HCV-targeted complexes, ( $P = 1.1 \times 10^{-2}$ ; S2C Fig). Moreover, these 231 VIPs are significantly enriched by members of the PHISTO, EHCO, and known HCV entry factor lists ( $P = 1.5 \times 10^{-11}$ , S2D Fig;  $P = 1.8 \times 10^{-7}$ , S2E Fig; and  $P = 4.8 \times 10^{-3}$ , S2F Fig; respectively).

Two known HCV entry factors, APOE and HRAS, and a novel entry factor ITGB identified by Zona and colleagues [39] are components of a single complex with CD81 [39, 57]. Of this CD81-complex, we identified several subunits, ADAM10, APOE, CD59, CD9, HRAS, ITGB1 and SCAM, as VIPs, this complex hence meets our criterion of  $\geq 3$  VIPs for an HCV-targeted host protein complex even though information of this complex was not used in our analysis because the latest compilations of CORUM and HPRD, in 2012 and 2009, respectively, predate the work of Zona and colleagues. This example therefore illustrates the merit of our approach in general, and the inclusion of protein complexes in the analysis in particular.

As detailed in the S1 Appendix, there are three main functionalities associated with KEGG pathways that are enriched in the HCV-targeted protein complexes: *entry*, *carcinogenesis*, and



**Fig 4. The AVP/SLiM-protein complex map.** Top panel: (left column) a list of AVPs with their AVPdb identification numbers that have an amino acid sequence containing one or more HCV E1/E2 SLiMs (indicated by the triangles and named at the bottom of the panel) that can bind to a subunit of a protein complex belonging to one of the six main complex groups. The relative efficacies of these AVPs in inhibiting HCV entry according to data provided in the AVPdb [56] are indicated by the circles column in the panel with the shading score shown to the right of the panel. Middle panel: the network connecting the nine SLiMs to their group A-F complex target(s). Within each group are VIPs of all the complexes: ovals represent VIPs<sub>direct</sub> (gray ovals are R6 VIPs<sub>direct</sub>), and rectangles represent VIPs<sub>indirect</sub> (not all VIPs<sub>indirect</sub> are shown). The ovals and rectangles in bold outline are known HCV entry factors. The horizontal bar represents the plasma membrane, the VIPs below the bar and within the shaded area are 'peripheral membrane proteins', and those spanning the bar are 'integral membrane proteins'. APOE is a peripheral membrane protein located at the extracellular side of the cell membrane. A connection between a viral SLiM and a complex group indicates that at least one protein in the group is targeted by HCV E1 and/or E2 via the viral SLiM. The thickness of the connection roughly scales to the number of proteins targeted by the SLiM in the group. The three numbers in the parenthesis are the number of VIPs<sub>direct</sub>, total VIPs (i.e., VIPs<sub>direct</sub> plus VIPs<sub>indirect</sub>), and unique subunits in the complex group. Bottom panel: the KEGG pathways enriched in the complex group (Benjamini-Hochberg adjusted  $P < 0.001$ ) (see S6 Table for the pathway names). The gray scale at bottom right indicate the significance of the P-values. The functions of the individual KEGG pathways are shown at the left of the panel and the main functionality at the right of the panel.

doi:10.1371/journal.pcbi.1005368.g004

*infectious disease*; the first two overlap substantially, as is shown by the analysis of the GO enrichment terms associated with the network modules (Fig 3). Examination of the KEGG pathways enriched in complex-forming VIPs helps us to understand the roles the complexes might play during HCV infection. For example, *TGF-beta signaling* ( $P = 1.2 \times 10^{-6}$ ), *Endocytosis* ( $P = 4.8 \times 10^{-11}$ ), and *Adherens junction* ( $P = 7.6 \times 10^{-7}$ ) are among KEGG pathways enriched in group C complexes (S6 Table) in which the VIPs<sub>direct</sub> TGFBR1, TGFBR2, ACVR1B, and ACVR1C are signaling receptors [58] and PRKCZ and PRKCI are protein kinases involved in endocytosis and adherens junction remodeling [59]. Our analysis revealed that several of the enriched pathways are involved in more than one step of HCV entry, and that the envelope proteins may regulate immune responses to HCV infection, affect hormone-related signaling pathways, and modulate HCC progression. These suggestions are in line with many experimental studies (S1 Appendix), including the report that HCV E1 and E2 can alter *RIG-I-like receptor signaling* and *Toll-like receptor signaling* [60]. Finally, the presence of enriched pathways associated with *infectious disease* suggests that some of the VIP-containing protein complexes may also be targets of other viruses, which is consistent with reports of coinfection of two or more different viruses (see, e.g. [61]).

However, when compared to all the SLiM-binding hepatocyte surface proteins and their binding partners in the human PPI network, the identified HCV-targeted protein complexes were shown to be significantly enriched in KEGG pathways belonging to the functionality of *cell entry* ( $P = 3.7 \times 10^{-2}$ ) and *carcinogenesis* ( $P = 2.7 \times 10^{-2}$ ), but not *infectious disease* ( $P = 3.8 \times 10^{-1}$ ) (see S8 Fig). This may suggest that *infectious disease* is more likely than the other two functionalities to be influenced by many of these proteins with a domain that can bind other SLiMs of the ELM database.

## AVPs and SLiMs

Peptides that can interfere with virus-host interactions are potential antiviral drugs [62, 63]. The viral SLiMs that we identified may, therefore, be useful scaffolds upon which to build AVPs. A search of the AVPdb [56] returned 73 AVPs that have been examined for entry-related, anti-HCV infection (Fig 1D), with more than one-third (29) harboring at least one of the nine identified SLiMs that might target a VIP residing in at least one of the six main protein complex groups. A statistical test indicated that these nine SLiMs were as likely to be also present in other AVPs ( $P = 8.8 \times 10^{-1}$ , S9A Fig), suggesting that, besides SLiMs, other parts of the AVP sequences are required to determine entry-related anti-HCV activities. Notably, however, many of the 29 AVPs have been shown to actively suppress HCV entry in cell-based assays (Fig 4, top panel). Although the molecular mechanisms associated with the anti-HCV

activities of these AVPs have not been fully elucidated, it is tempting to speculate that, because they contain an HCV SLiM, they may interfere with an HCV-host protein interaction by preferentially binding the host protein and, thereby, inhibiting HCV entry.

## Case examples

We present two examples to demonstrate how our bioinformatics procedures can be used to identify protein complexes known to play a role in HCV infection and/or pathology.

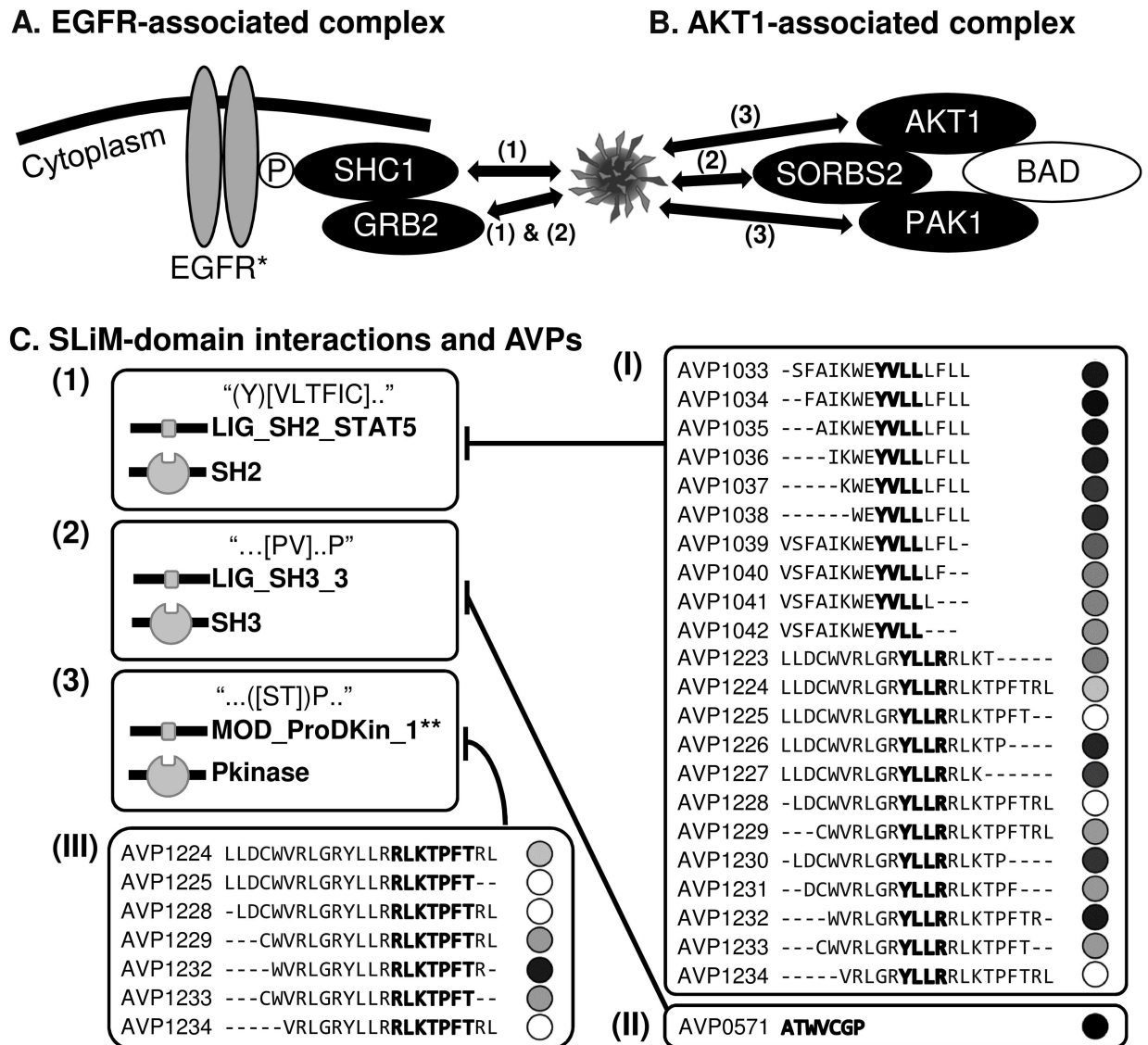
The first example (Fig 5A), involves the complex containing EGFR, SHC1, and GRB2 (group F complex, Fig 4). This complex (S5 Table, complex ID: 176) mediates HRAS signaling critical for HCV entry [39]. According to our results, HCV E1/E2 might interact directly with SHC1 and GRB2 (both are R6 VIPs<sub>direct</sub>) and indirectly with EGFR (a VIP<sub>indirect</sub>) via the SLiMs of LIG\_SH2\_STAT5 and/or LIG\_SH3\_3 (Fig 5A and 5C). Furthermore, of the 23 AVPs that contain the same SLiM (22 with LIG\_SH2\_STAT5 and one with LIG\_SH3\_3), 20 (86%) are shown to inhibit HCV entry (Fig 5C, I and II), and this percentage is shown to be statistically significant ( $P < 1.0 \times 10^{-4}$ , S9B Fig).

The second example, the AKT1-associated complex (also a group F complex, Fig 4), is presented in Fig 5B. When being a part of this complex (S5 Table, complex ID: 180), AKT1 can inhibit apoptosis induced by BAD overexpression [64], and inhibition of BAD-mediated apoptosis contributes to HCC progression [65]. Our analysis suggests that HCV E1 and/or E2 may interact with the AKT1-associated complex by targeting one or more of its members, i.e., AKT1 and PAK1 through SLiMs of the MOD\_family (MOD\_NEK2\_2 of E1; MOD\_CK1\_1, MOD\_CK2\_1, MOD\_GSK3\_1, MOD\_NEK2\_1 of E2, and MOD\_ProDKin\_1 of E1 and E2), and SORBS2, an SH3 domain-containing protein, through the SLiM of LIG\_SH3\_3. Together with the report that HCV E2 can induce AKT phosphorylation to facilitate HCV entry [23], these results suggest that a viral SLiM and protein domain association may be involved in viral entry and virus-induced carcinogenesis. Furthermore, as with the first example, the majority (four out of seven;  $P < 1.0 \times 10^{-4}$ , S9C Fig) of the AVPs containing a MOD\_family SLiM inhibits HCV entry (Fig 5C (III)).

## Discussion

Despite recent, rapid advances in high-throughput experiments, all characterized networks of virus-host interactions remain vastly incomplete. To fill this void, several studies have incorporated bioinformatics information [66] such as that obtained by text mining experimental reports from the literature [67]. In this work, we described a strategy, identifying eukaryote-interacting SLiMs found in viruses to find human proteins that may be “hijacked” by HCV for its entry via “molecular mimicry” of the SLiMs. With the identification of these human proteins, i.e., VIPs, we then built a virus-host PPI network by exploring the human PPI network for functions that may be modulated and/or productively used by the virus.

Per our “molecular mimicry” hypothesis, more than half of the hepatocyte surface proteins could be VIPs for a SLiM from the ELM database to interact with directly or indirectly (1,320/2,456, see S8 Fig legend), and more than one third (899/2,456, Fig 1B) for a SLiM harbored by HCV E1/E2 alone (Fig 1B). This suggests that SLiMs by themselves are of low binding specificity to protein domains, and thus most of the predicted VIPs are likely false positives. However, as demonstrated by the results presented above, by integrating with a variety of experimental data and information, especially with network and functional analysis, the number of VIPs (hence false positives) predicted can be greatly reduced and a manageable list of viable candidate proteins can be extracted to complement and guide further experimental investigations.



**Fig 5. Case examples.** (A) The EGFR-associated complex (S5 Table, Complex ID: 176). (B) The AKT1-associated complex (S5 Table, Complex ID: 180). (C) Types of SLiM-domain interactions (indicated by (1), (2), and (3)) that mediate the targeting of protein complexes, and three sets of AVPs (indicated by (I), (II) and (III)) containing the corresponding SLiM that exhibits inhibition activity (indicated by bar-headed arrows) are shown. Proteins represented by black ovals are VIPs<sub>direct</sub>; by gray ovals are VIPs<sub>indirect</sub>; and by white ovals are not VIPs (i.e. not in the virus-host PPI network). Double-headed arrows in panel A and B indicate predicted HCV-host PPIs in our analysis. Viral SLiM(s) are highlighted within the amino acid sequence of the AVP and are accompanied by its relative efficacy (in shaded circles) in inhibiting HCV entry (see the vertical bar for the normalized inhibition score in Fig 4 on the right of its top panel). The regular expression of each SLiM sequence as annotated in the ELM database [11] is shown within quotation marks in the SLiM-domain interactions (1), (2), and (3). \*EGFR is a known HCV entry factor. \*\*MOD\_ProDKin\_1 is representative of the MOD\_family SLiMs (see Fig 3).

doi:10.1371/journal.pcbi.1005368.g005

Our functional analysis shows that the SLiMs-derived VIPs of HCV infection and the related host protein complexes are involved in two major types of cellular functions, one associated with viral entry and the other with carcinogenesis and/or infectious disease (Figs 3 and 4). Although the inclusion of the second category was somewhat unexpected because HCV E1 and E2 were used to find SLiMs, a role for E1/E2 in carcinogenesis has been demonstrated [44, 45] and multi-functionality of other viral envelope proteins has also been documented [68–

[70]: For example, hemagglutinin, an envelope protein of the influenza virus, is involved in viral entry but also activates NF- $\kappa$ B when expressed in 293T and HeLa cells [69]. Many of the predicted interactions between the SLiMs found in HCV E1/E2 and the human VIPs that occupy a prominent role (R6, global connector hub) in the virus-host PPI network, and the relationships between these R6 VIPs and their deduced functional modules, are supported by published experimental evidence (Fig 3). Together with the results from a similar approach used to predict HIV-interacting human proteins [71] and the report that virus-host interactions may be assisted by host-like SLiMs [10], evidence is mounting to support the suggestion that, via SLiMs, host proteins can be “hijacked” and host functions rewired by pathogens, a phenomenon that has been extensively reviewed at the pathway level [8]. Further supporting this premise, some HCV E1/E2 SLiMs, particularly those that might interact with major protein complexes, are found in many AVPs that inhibit HCV infection (Fig 4).

Although recently approved direct-acting antiviral (DAA) treatments have improved the virologic response rate to >90% for most types of chronic hepatitis C infections, new, hard-to-treat HCV strains including genotype 3 and DAA-resistant variants generated from DAA-treated patients, have appeared [72]. Because DAA-resistant strains are disseminated mainly through cell-to-cell transmission rather than cell-free transmission [73], the former route will be key to eradicating HCV infection. Several known HCV entry factors for cell-free transmission, e.g., EGFR, CLDN1, OCLN, and SCARB1 are also involved in cell-to-cell transmission [74–76]. Additionally, cell-to-cell transmission independent of CD81, the most well studied binding receptor for HCV E2 [13], has been reported [77–79]. Taken together and given that HCV E1/E2 are indispensable for cell-to-cell transmission of the virus [79], other host factors that can interact with these viral proteins need to be identified. Although the mechanism(s) of HCV cell-to-cell transmission is not yet understood, this type of transmission has been associated with cell adhesion molecules [80], many of which belong to the KEGG pathway category of Cellular community. Interestingly, pathways in this category are significantly enriched in four of the six protein complex groups (all except group A and E complexes; Fig 4), suggesting that some of the complex-associated VIPs may be a good starting point to study cell-to-cell transmission of HCV. Because host-targeting antivirals are generating enthusiastic interest for the development of treatments for hard-to-treat hepatitis C infections [81], our bioinformatics strategy is a timely approach to identify new targets for antiviral research, not only for HCV but also for other viruses as the concept of SLiM involvement in molecular mimicry is a general one.

## Methods

The four main components of our *in silico* approach are shown in Fig 1.

### A) Identify viral SLiMs

The HCV E1/E2 sequences were scanned against sequences in the ELM database (<http://elm.eu.org/>) [11] to find short, matching linear sequences in mammalian proteins known to interact with other mammalian proteins, which might, therefore, be mimicked by HCV E1/E2 sequences (Fig 1A). E1 and E2 sequences (from 41 and 35 HCV strains, respectively) were extracted from the UniProtKB/SwissProt database (<http://www.uniprot.org/>) [82] for use in our study. We required that the sequences of the matched viral SLiM needed to be conserved at a rate of at least 70%, a cutoff used for a study of HIV SLiMs by others [71] and also yielded the best performance to balance sensitivity and specificity in predicting known HCV entry factors (S5 Fig). The search retrieved 26 distinct and conserved SLiMs (Fig 1A), which were then used to find the human proteins (VIPs<sub>direct</sub>) that they might interact with (Fig 1B).

As shown in [S10 Fig](#), the number of SLiMs that can be found in HCV protein sequences is roughly proportional to protein size, and all these proteins can confer the “molecular mimicry” mechanism hypothesized and be targets of our investigation; however, as explained in Introduction, we focused on E1/E2 sequences because we were primarily concerned with the viral entry process, in addition to other considerations.

## B) Construct of a virus-host PPI network

Accompanying each SLiM in the ELM database is an annotation of the protein domain(s) to which the SLiM can bind. Of the 26 identified SLiMs, 23 have information for human proteins, and of the 23, 19 were mapped to hepatocyte surface proteins that are present in the Human Integrated Protein-Protein Interaction Reference database [14] (HIPPIE release v1.7; <http://cbdm-01.zdv.uni-mainz.de/~mschaefer/hippie/>), which is a constantly updated human PPI database that integrate multiple experimental PPI datasets, to derive a PPI network. The list of hepatocyte surface proteins used to develop our virus-host PPI network was collected from Human Protein Reference Database [46] (HPRD; <http://www.hprd.org>) and Human Protein-pedia Database (<http://www.humanproteinpedia.org>) [47] by using the keywords “Liver” or “Hepatocyte” for the tissue type and “Extracellular region,” “Plasma membrane,” “Cell surface,” or “Cell junction” for the type of cellular component to search for potential host proteins expressed on the hepatocyte surface that might facilitate the entry of HCV. A set of 2,456 human proteins matched these keywords, and we termed this set “liver cell surface proteins”. Of these proteins, 115 (VIPs<sub>direct</sub>) contained at least one protein domain to which one of the 19 viral SLiMs could bind which would mimic the corresponding human SLiM. A set of 784 proteins (denoted VIPs<sub>indirect</sub>) were identified as binding partners to the VIPs<sub>direct</sub>. The interactions between the VIPs<sub>direct</sub> and VIPs<sub>indirect</sub>, and those between the viral SLiMs and VIPs<sub>direct</sub> were combined to form the virus-host network, which was then subjected to a network module/functional analysis, as described in the next step and [Fig 1C](#).


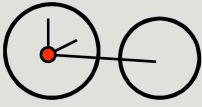
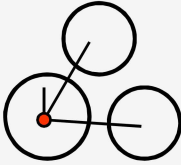
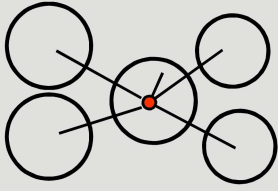

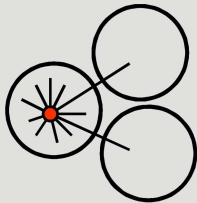
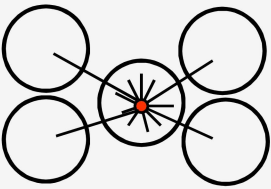
## C) Identify network modules and their functional roles

Given a network, it is useful to determine whether it is formed of modules (i.e., communities), which are often indicative of distinct functions. For this task, we applied the network analysis tool NetCarto [18] using its large-network default settings to identify modules. The roles of importance for each network node, including every VIP, could also be assigned, which ranged from the least significant, R1 (ultra-peripheral) and R2 (peripheral), to the most significant, R6 (global connector hub) and R7 (kinless hub) (a hub is a node connected to many nodes). Following NetCarto [18], the definitions of the seven classes of network nodes and their roles in the network are summarized and schematically depicted in [Fig 6](#).

The biological function of each network module was inferred using DAVID (<https://david.ncifcrf.gov/>) [83] to search for enriched GO functions [20] under the category of “biological process.” For each module, Revigo [21] was used to obtain representatives for enriched GO terms (Benjamini–Hochberg-adjusted  $P < 0.05$ ), and a main functionality was deduced to cover these representatives.

## D) Build the AVP/SLiM-protein complex map

Because host VIPs<sub>direct</sub> and VIPs<sub>indirect</sub> may be subunits of the same protein complex targeted by the virus, we mapped the VIPs to the human protein complexes in HPRD [46] and in the mammalian protein complexes database (CORUM) [55]. The 190 complexes containing three or more VIPs were then clustered by GAP [84], a tool for matrix visualization and clustering, based on similarity related to the number of common subunits. The choice of three VIPs was

Role	Term	Schematic view	Description
R1	Ultra-peripheral		Nodes with all the links in a module.
R2	Peripheral		Nodes with most ( $\geq 60\%$ ) links within their module.
R3	Satellite connector		Nodes with many ( $< 50\%$ ) links to other modules.
R4	Non-hub kinless		Nodes with links ( $\geq 65\%$ ) uniformly distributed among all modules.
R5	Provincial hub		Hub nodes with most ( $\geq 80\%$ ) of links within their module.
R6	Global connector hub		Hubs with many ( $< 50\%$ ) links to most of the other modules.
R7	Kinless hub		Hubs with links ( $\geq 50\%$ ) uniformly distributed among all modules.

**Fig 6. Definitions and schematic depictions of network roles.** According to NetCarto [18], nodes (small red circles) with z-score of within-module degree  $\geq 2.5$  are defined as module hubs (nodes with many links, i.e. spikes in the schematic view), and those with z-score  $< 2.5$  are non-hubs. Large circles represent modules. See [18] for further details.

doi:10.1371/journal.pcbi.1005368.g006



made to reduce the total number of VIPs for enrichment analysis while maintaining their coverage of known HCV entry factors as much as possible. This procedure yielded six major protein-complex clusters, or groups (small groups containing fewer than five complexes were excluded), which altogether contained 177 complexes and 231 VIPs that were linked to nine viral SLiMs. The VIPs of each of the six complex groups were then subjected to KEGG pathway enrichment analysis [85] using the clusterProfiler R package [86]. A total of 43 significantly enriched hepatocyte-expressed pathways were identified (Benjamini-Hochberg-adjusted  $P < 0.001$ ). In addition, a search of AVPdb [56] identified 73 peptides annotated with “Hepatitis C virus” and “Virus entry” whose inhibiting activities against HCV entry have been examined by experimental screening. Of the 73 sequences, 29 matched at least one of the nine viral SLiMs. Matching of the AVP sequences and viral SLiMs was performed at the ELM database website.

## Supporting Information

**S1 Fig. Statistical significance of VIPs<sub>direct</sub> being hub proteins (R6 or R7) in the host PPI network.** The P-value was computed using a binomial proportion test on the difference between the group of 22 R6 VIPs<sub>direct</sub> and that of the remaining 93 VIPs<sub>direct</sub>. On top of the bar is the number of VIPs<sub>direct</sub> that are also hub proteins (R6 or R7) in the host PPI network of liver cell surface proteins.

(PDF)

**S2 Fig. Statistical significance of overlaps between protein/gene sets.** (A) The overlap between all VIPs and proteins in PHISTO. (B) The overlap between all VIPs and proteins in EHCO. (C) The overlap between the VIPs in the six main groups of HCV-targeted complexes and the R6 proteins. (D) The overlap between VIPs in the six main groups of HCV-targeted complexes and proteins in PHISTO. (E) The overlap between VIPs in the six main groups of HCV-targeted complexes and proteins in EHCO. (F) The overlap between VIPs in the six main groups of HCV-targeted complexes and known HCV entry factors. The background for panels A, B, D, and E is for all hepatocyte surface proteins and the background for panels C and F is for all VIPs.

(PDF)

**S3 Fig. Statistical significance of overlaps between sets of PPIs (network edges).** (A) The overlap between HCV-VIPs<sub>direct</sub> PPIs and PPIs in PHISTO that were determined as direct based on their experimental methods (see S4 Table). (B) The overlap between HCV-VIPs<sub>indirect</sub> PPIs and PPIs in PHISTO that could not be determined as direct. Note that in PHISTO, other than core, NS3-4A and NS5A, the identity of the individual HCV protein(s) involving in the interaction with host proteins is not known; consequently, HCV was considered as a single node in the PPI network from PHISTO, and, therefore, the number of virus-host PPIs (i.e. network edges) is the same as that of the host proteins in this enrichment test.

(PDF)

**S4 Fig. Four known entry factors (white nodes) of HCV infection not identified in this work.** (A) CLDN1; (B) SCARB1; (C) CD209; (D) NPC1L1. Blue nodes are VIPs<sub>indirect</sub>; gray nodes are not VIPs. The connections between the nodes represent physical interactions extracted from HIPPIE [14].

(PDF)

**S5 Fig. The ROC performance against known HCV entry factors of the *in silico* predictions and Li et al.’s siRNA experiment.** As described in the main text, using a cutoff of at least 70%

sequence conservation to find SLiMs (Fig 1A), we identified 15 of the 19 known HCV entry factors before complex analysis and 9 after, and Li et al. [5] identified 8 (see Table 1). In all, the *in silico* method identified 899 (231 after complex analysis), and Li et al.'s experiment 45, hepatocyte surface proteins as potential HCV entry factors. In this figure, the cutoff of HCV E1/E2 sequence conservation was varied from 0% to 100%, at which the same procedure as described in Fig 1 was carried out, and sensitivity and specificity for the resulting predictions were calculated to generate the ROC curves, on which the performance obtained at 70% sequence conservation cutoff is indicated. AUC: Area Under Curve.

(PDF)

**S6 Fig. Grouping of HCV-targeted protein complexes.** Using GAP [84], 190 HCV-targeted protein complexes (containing 258 VIPs) were hierarchically clustered based on the number of shared subunits. The threshold (Cut, bottom right) for the clustering is indicated. The six main clusters (groups) are boxed and labeled A-F. Within these six groups there are 231 VIPs. \*Complex ID refers to notations in S5 Table, where detailed information of the HCV-targeted complexes is provided.

(PDF)

**S7 Fig. VIPs versus randomly selected proteins found in protein complexes.** The distribution plot illustrates 100,000 randomly sampled sets, each consisting of 899 proteins (the same number as VIPs) sampled from the set of 2,456 human hepatocyte surface proteins (Fig 1B). Each set was searched for protein complexes, which were required to contain at least three sampled proteins as indicated in Fig 1D. The number of proteins retained in their targeted complexes in each sampling was counted as “number of sampled proteins in complexes”. The observed number (i.e., 258, indicated by dotted line) is the number of the VIPs found in the 190 HCV-targeted protein complexes.

(PDF)

**S8 Fig. Statistical significance of the three main enriched functionalities of HCV E1/E2-derived VIPs compared to those containing a binding domain for any SLiM in the ELM database.** The distribution plots for the functionality of (A) *entry*, (B) *carcinogenesis*, and (C) *infectious disease* were derived from results of 10,000 randomly sampled sets of proteins. In each set, 899 proteins (the same number as VIPs) were randomly sampled from a set of 1,320 proteins, which is the number of proteins containing a binding domain for any SLiM in the ELM database (348 proteins) and their first PPI neighbors (972 proteins) in the human PPI network of liver cell surface proteins. The same procedure as described in Fig 1D (see Methods) for protein complex and KEGG pathway analyses was carried out for each sample set. The number of enriched KEGG pathways in functionality of *entry*, *carcinogenesis*, and *infectious disease*, was counted respectively. The observed number (indicated by dotted line) is the number of KEGG pathways belonging to the given functionality enriched in the set of VIPs found in the six main groups of HCV-targeted protein complexes (Fig 4).

(PDF)

**S9 Fig. Statistical significance of SLiM-containing AVPs.** (A) Of the 73 AVPs that were tested for entry-related anti-HCV activities in AVpdb, 29 harbor an HCV E1/E2 SLiM (one of nine) that can bind to VIPs found in the six main groups of HCV-targeted protein complexes (Fig 4); here, the number of AVPs containing at least one of the nine SLiMs for all sample sets, each consisting of 73 AVPs randomly sampled from a pool of 2059 AVPs (the total number of AVPs in AVpdb), was counted. (B) In this test, the sampling pool was the set of the 431 AVPs found to contain LIG\_SH2\_STAT5 and/or LIG\_SH3\_3 in AVpdb; the sample size was 23, of which those shown to be effective in inhibiting HCV entry in AVpdb were counted. (C) In

this test, the sampling pool was the set of the 154 AVPs found to contain MOD\_ProDKin\_1 in AVPdb; the sample size was 7, of which those shown to be effective in inhibiting HCV entry in AVPdb were counted. For (A), (B) and (C), the statistics was calculated based on 10,000 sample sets.

(PDF)

**S10 Fig. The number of conserved (in  $\geq 70\%$  sequences) SLiMs found in HCV component proteins and protein length.**

(PDF)

**S1 Table. GO terms enriched in the virus-host PPI network modules and representative functions according to Revigo.**

(XLSX)

**S2 Table. Representative function(s) and supporting published experimental data for HCV-human PPI network modules 1–7.**

(PDF)

**S3 Table. Published experimental evidence for relations between SLiMs and R6 proteins and between R6 proteins and module functions in Fig 3.**

(PDF)

**S4 Table. Types of PPIs between HCV polyprotein and human protein in PHISTO and their detection methods.**

(PDF)

**S5 Table. HCV-targeted protein complexes.**

(XLSX)

**S6 Table. KEGG pathways enriched in the six main protein complex groups.**

(PDF)

**S1 Appendix. Main functionalities associated with KEGG pathways enriched in the HCV-targeted protein complexes.**

(PDF)

## Acknowledgments

We thank Drs Yi-Ling Lin, Mi-Hua Tao, and Wei-Chung Liu for careful reading of the manuscript and helpful suggestions.

## Author Contributions

**Conceptualization:** TW MJH.

**Data curation:** TW AWTC WYLW.

**Formal analysis:** TW AWTC WYLW MJH.

**Funding acquisition:** MJH.

**Investigation:** TW AWTC WYLW.

**Methodology:** AWTC TW WYLW.

**Project administration:** MJH.

**Resources:** MJH.

**Supervision:** MJH.

**Validation:** AWTC TW WYLW.

**Visualization:** AWTC WYLW.

**Writing – original draft:** AWTC WYLW TW.

**Writing – review & editing:** AWTC WYLW MJH.

## References

1. de Chassey B, Meyniel-Schicklin L, Aublin-Gex A, André P, Lotteau V. New horizons for antiviral drug discovery from virus-host protein interaction networks. *Current opinion in virology*. 2012; 2(5):606–613. doi: [10.1016/j.coviro.2012.09.001](https://doi.org/10.1016/j.coviro.2012.09.001) PMID: [23025912](https://pubmed.ncbi.nlm.nih.gov/23025912/)
2. de Chassey B, Meyniel-Schicklin L, Vonderscher J, André P, Lotteau V. Virus-host interactomics: new insights and opportunities for antiviral drug discovery. *Genome Medicine*. 2014; 6:115. doi: [10.1186/s13073-014-0115-1](https://doi.org/10.1186/s13073-014-0115-1) PMID: [25593595](https://pubmed.ncbi.nlm.nih.gov/25593595/)
3. Christ F, Voet A, Marchand A, Nicolet S, Desimmie BA, Marchand D, et al. Rational design of small-molecule inhibitors of the LEDGF/p75-integrase interaction and HIV replication. *Nat Chem Biol*. 2010; 6(6):442–448. doi: [10.1038/nchembio.370](https://doi.org/10.1038/nchembio.370) PMID: [20473303](https://pubmed.ncbi.nlm.nih.gov/20473303/)
4. Zhou H, Xu M, Huang Q, Gates AT, Zhang XD, Castle JC, et al. Genome-scale RNAi screen for host factors required for HIV replication. *Cell Host Microbe*. 2008; 4(5):495–504. doi: [10.1016/j.chom.2008.10.004](https://doi.org/10.1016/j.chom.2008.10.004) PMID: [18976975](https://pubmed.ncbi.nlm.nih.gov/18976975/)
5. Li Q, Zhang Y-Y, Chiu S, Hu Z, Lan K-H, Cha H, et al. Integrative Functional Genomics of Hepatitis C Virus Infection Identifies Host Dependencies in Complete Viral Replication Cycle. *PLoS Pathog*. 2014; 10(5):e1004163. PubMed Central PMCID: PMC4095987. doi: [10.1371/journal.ppat.1004163](https://doi.org/10.1371/journal.ppat.1004163) PMID: [24852294](https://pubmed.ncbi.nlm.nih.gov/24852294/)
6. Rozenblatt-Rosen O, Deo RC, Padi M, Adelmant G, Calderwood MA, Rolland T, et al. Interpreting cancer genomes using systematic host network perturbations by tumour virus proteins. *Nature*. 2012; 487(7408):491–495. doi: [10.1038/nature11288](https://doi.org/10.1038/nature11288) PMID: [22810586](https://pubmed.ncbi.nlm.nih.gov/22810586/)
7. Alcami A. Viral mimicry of cytokines, chemokines and their receptors. *Nat Rev Immunol*. 2003; 3(1):36–50. doi: [10.1038/nri980](https://doi.org/10.1038/nri980) PMID: [12511874](https://pubmed.ncbi.nlm.nih.gov/12511874/)
8. Davey NE, Travé G, Gibson TJ. How viruses hijack cell regulation. *Trends Biochem Sci*. 2011; 36(3):159–169. doi: [10.1016/j.tibs.2010.10.002](https://doi.org/10.1016/j.tibs.2010.10.002) PMID: [21146412](https://pubmed.ncbi.nlm.nih.gov/21146412/)
9. Garamszegi S, Franzosa EA, Xia Y. Signatures of Pleiotropy, Economy and Convergent Evolution in a Domain-Resolved Map of Human–Virus Protein–Protein Interaction Networks. *PLoS Pathog*. 2013; 9(12):e1003778. doi: [10.1371/journal.ppat.1003778](https://doi.org/10.1371/journal.ppat.1003778) PMID: [24339775](https://pubmed.ncbi.nlm.nih.gov/24339775/)
10. Hagai T, Azia A, Babu MM, Andino R. Use of Host-like Peptide Motifs in Viral Proteins Is a Prevalent Strategy in Host-Virus Interactions. *Cell Reports*. 2014; 7(5):1729–1739. doi: [10.1016/j.celrep.2014.04.052](https://doi.org/10.1016/j.celrep.2014.04.052) PMID: [24882001](https://pubmed.ncbi.nlm.nih.gov/24882001/)
11. Dinkel H, Michael S, Weatheritt RJ, Davey NE, Van Roey K, Altenberg B, et al. ELM—the database of eukaryotic linear motifs. *Nucleic Acids Res*. 2011; 40(D1):D242–251. PubMed Central PMCID: PMC3245074.
12. Lazarus JV, Harmon K, Sperle I. Global policy report on the prevention and control of viral hepatitis. Geneva, Switzerland: World Health Organization; 2013.
13. Lindenbach BD, Rice CM. The ins and outs of hepatitis C virus entry and assembly. *Nat Rev Microbiol*. 2013; 11(10):688–700. doi: [10.1038/nrmicro3098](https://doi.org/10.1038/nrmicro3098) PMID: [24018384](https://pubmed.ncbi.nlm.nih.gov/24018384/)
14. Schaefer MH, Fontaine J-F, Vinayagam A, Porras P, Wanker EE, Andrade-Navarro MA. HIPPIE: Integrating protein interaction networks with experiment based quality scores. *PLoS one*. 2012; 7(2):e31826. PubMed Central PMCID: PMC3279424. doi: [10.1371/journal.pone.0031826](https://doi.org/10.1371/journal.pone.0031826) PMID: [22348130](https://pubmed.ncbi.nlm.nih.gov/22348130/)
15. Hsu C-N, Lai J-M, Liu C-H, Tseng H-H, Lin C-Y, Lin K-T, et al. Detection of the inferred interaction network in hepatocellular carcinoma from EHCO (Encyclopedia of Hepatocellular Carcinoma genes Online). *BMC Bioinformatics*. 2007; 8(1):66.
16. Tekir SD, Çakır T, Ardiç E, Sayılırbaş AS, Konuk G, Konuk M, et al. PHISTO: pathogen-host interaction search tool. *Bioinformatics*. 2013; 29(10):1357–1358. doi: [10.1093/bioinformatics/btt137](https://doi.org/10.1093/bioinformatics/btt137) PMID: [23515528](https://pubmed.ncbi.nlm.nih.gov/23515528/)

17. Barabási A-L, Oltvai ZN. Network biology: understanding the cell's functional organization. *Nature Reviews Genetics*. 2004; 5(2):101–113. doi: [10.1038/nrg1272](https://doi.org/10.1038/nrg1272) PMID: [14735121](https://pubmed.ncbi.nlm.nih.gov/14735121/)
18. Guimerà R, Nunes Amaral LA. Functional cartography of complex metabolic networks. *Nature*. 2005; 433(7028):895–900. doi: [10.1038/nature03288](https://doi.org/10.1038/nature03288) PMID: [15729348](https://pubmed.ncbi.nlm.nih.gov/15729348/)
19. Dyer MD, Murali TM, Sobral BW. The Landscape of Human Proteins Interacting with Viruses and Other Pathogens. *PLoS Pathog*. 2008; 4(2):e32. doi: [10.1371/journal.ppat.0040032](https://doi.org/10.1371/journal.ppat.0040032) PMID: [18282095](https://pubmed.ncbi.nlm.nih.gov/18282095/)
20. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: tool for the unification of biology. *Nat Genet*. 2000; 25(1):25–29. doi: [10.1038/75556](https://doi.org/10.1038/75556) PMID: [10802651](https://pubmed.ncbi.nlm.nih.gov/10802651/)
21. Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. *PloS one*. 2011; 6(7):e21800. doi: [10.1371/journal.pone.0021800](https://doi.org/10.1371/journal.pone.0021800) PMID: [21789182](https://pubmed.ncbi.nlm.nih.gov/21789182/)
22. Brazzoli M, Bianchi A, Filippini S, Weiner A, Zhu Q, Pizza M, et al. CD81 is a central regulator of cellular events required for hepatitis C virus infection of human hepatocytes. *J Virol*. 2008; 82(17):8316–8329. doi: [10.1128/JVI.00665-08](https://doi.org/10.1128/JVI.00665-08) PMID: [18579606](https://pubmed.ncbi.nlm.nih.gov/18579606/)
23. Liu Z, Tian Y, Machida K, Lai MMC, Luo G, Fong SKH, et al. Transient activation of the PI3K-AKT pathway by hepatitis C virus to enhance viral entry. *J Biol Chem*. 2012; 287(50):41922–41930. PubMed Central PMCID: [PMC3516739](https://pubmed.ncbi.nlm.nih.gov/PMC3516739/). doi: [10.1074/jbc.M112.414789](https://doi.org/10.1074/jbc.M112.414789) PMID: [23095753](https://pubmed.ncbi.nlm.nih.gov/23095753/)
24. Yeatman TJ. A renaissance for SRC. *Nat Rev Cancer*. 2004; 4(6):470–480. doi: [10.1038/nrc1366](https://doi.org/10.1038/nrc1366) PMID: [15170449](https://pubmed.ncbi.nlm.nih.gov/15170449/)
25. Panjarian S, Iacob RE, Chen S, Engen JR, Smithgall TE. Structure and Dynamic Regulation of Abl Kinases. *J Biol Chem*. 2013; 288(8):5443–5450. doi: [10.1074/jbc.R112.438382](https://doi.org/10.1074/jbc.R112.438382) PMID: [23316053](https://pubmed.ncbi.nlm.nih.gov/23316053/)
26. Carlier M-F, Nioche P, Broutin L-Hermite I, Boujemaa R, Clainche CL, Egile C, et al. GRB2 links signaling to actin assembly by enhancing interaction of neural wiskott-aldrich syndrome protein (N-WASP) with actin-related protein (ARP2/3) complex. *J Biol Chem*. 2000; 275(29):21946–21952. doi: [10.1074/jbc.M000687200](https://doi.org/10.1074/jbc.M000687200) PMID: [10781580](https://pubmed.ncbi.nlm.nih.gov/10781580/)
27. Chaki SP, Rivera GM. Integration of signaling and cytoskeletal remodeling by Nck in directional cell migration. *BioArchitecture*. 2013; 3(3):57–63. doi: [10.4161/bioa.25744](https://doi.org/10.4161/bioa.25744) PMID: [23887203](https://pubmed.ncbi.nlm.nih.gov/23887203/)
28. Cosen-Binker LI, Kapus A. Cortactin: The Gray Eminence of the Cytoskeleton. *Physiology*. 2006; 21(5):352–361.
29. Ke P-Y, Chen SS-L. Hepatitis C Virus and Cellular Stress Response: Implications to Molecular Pathogenesis of Liver Diseases. *Viruses*. 2012; 4(10):2251–2290. doi: [10.3390/v4102251](https://doi.org/10.3390/v4102251) PMID: [23202463](https://pubmed.ncbi.nlm.nih.gov/23202463/)
30. Lee SH, Kim YK, Kim CS, Seol SK, Kim J, Cho S, et al. E2 of hepatitis C virus inhibits apoptosis. *J Immunol*. 2005; 175(12):8226–8235. PMID: [16339562](https://pubmed.ncbi.nlm.nih.gov/16339562/)
31. Song G, Ouyang G, Bao S. The activation of Akt/PKB signaling pathway and cell survival. *Journal of Cellular and Molecular Medicine*. 2005; 9(1):59–71. PMID: [15784165](https://pubmed.ncbi.nlm.nih.gov/15784165/)
32. Luo J-L, Kamata H, Karin M. IKK/NF-kappaB signaling: balancing life and death—a new approach to cancer therapy. *The Journal of clinical investigation*. 2005; 115(10):2625–2632. PubMed Central PMCID: [PMC1236696](https://pubmed.ncbi.nlm.nih.gov/PMC1236696/). doi: [10.1172/JCI26322](https://doi.org/10.1172/JCI26322) PMID: [16200195](https://pubmed.ncbi.nlm.nih.gov/16200195/)
33. Li Q, Pène V, Krishnamurthy S, Cha H, Liang TJ. Hepatitis C virus infection activates an innate pathway involving IKK- $\alpha$  in lipogenesis and viral assembly. *Nat Med*. 2013; 19(6):722–729. doi: [10.1038/nm.3190](https://doi.org/10.1038/nm.3190) PMID: [23708292](https://pubmed.ncbi.nlm.nih.gov/23708292/)
34. Michie AM, Nakagawa R. The link between PKC alpha regulation and cellular transformation. *Immunol Lett*. 2005; 96(2):155–162. doi: [10.1016/j.imlet.2004.08.013](https://doi.org/10.1016/j.imlet.2004.08.013) PMID: [15585319](https://pubmed.ncbi.nlm.nih.gov/15585319/)
35. Fimia GM, Evangelisti C, Alonzi T, Romani M, Fratini F, Paonessa G, et al. Conventional Protein Kinase C Inhibition Prevents Alpha Interferon-Mediated Hepatitis C Virus Replicon Clearance by Impairing STAT Activation. *J Virol*. 2004; 78(23):12809–12816. doi: [10.1128/JVI.78.23.12809-12816.2004](https://doi.org/10.1128/JVI.78.23.12809-12816.2004) PMID: [15542633](https://pubmed.ncbi.nlm.nih.gov/15542633/)
36. Di Guglielmo GM, Le Roy C, Goodfellow AF, Wrana JL. Distinct endocytic pathways regulate TGF- $\beta$  receptor signalling and turnover. *Nature Cell Biology*. 2003; 5(5):410–421. doi: [10.1038/ncb975](https://doi.org/10.1038/ncb975) PMID: [12717440](https://pubmed.ncbi.nlm.nih.gov/12717440/)
37. Choi S-H, Hwang SB. Modulation of the Transforming Growth Factor- $\beta$  Signal Transduction Pathway by Hepatitis C Virus Nonstructural 5A Protein. *J Biol Chem*. 2006; 281(11):7468–7478. doi: [10.1074/jbc.M512438200](https://doi.org/10.1074/jbc.M512438200) PMID: [16407286](https://pubmed.ncbi.nlm.nih.gov/16407286/)
38. Reyland ME. Protein kinase C $\delta$  and apoptosis. *Biochemical Society Transactions*. 2007; 35(5):1001–1004.
39. Zona L, Lupberger J, Sidahmed-Adrar N, Thumann C, Harris HJ, Barnes A, et al. HRas signal transduction promotes hepatitis C virus cell entry by triggering assembly of the host tetraspanin receptor complex. *Cell Host Microbe*. 2013; 13(3):302–313. doi: [10.1016/j.chom.2013.02.006](https://doi.org/10.1016/j.chom.2013.02.006) PMID: [23498955](https://pubmed.ncbi.nlm.nih.gov/23498955/)

40. Whittaker S, Marais R, Zhu AX. The role of signaling pathways in the development and treatment of hepatocellular carcinoma. *Oncogene*. 2010; 29(36):4989–5005. doi: [10.1038/onc.2010.236](https://doi.org/10.1038/onc.2010.236) PMID: [20639898](https://pubmed.ncbi.nlm.nih.gov/20639898/)
41. Miled N, Yan Y, Hon W-C, Perisic O, Zvelebil M, Inbar Y, et al. Mechanism of two classes of cancer mutations in the phosphoinositide 3-kinase catalytic subunit. *Science*. 2007; 317(5835):239–242. doi: [10.1126/science.1135394](https://doi.org/10.1126/science.1135394) PMID: [17626883](https://pubmed.ncbi.nlm.nih.gov/17626883/)
42. Downward J. The GRB2/Sem-5 adaptor protein. *FEBS Letters*. 1994; 338(2):113–117. PMID: [8307166](https://pubmed.ncbi.nlm.nih.gov/8307166/)
43. Ravichandran KS. Signaling via Shc family adapter proteins. *Oncogene*. 2001; 20(44):6322–6330. doi: [10.1038/sj.onc.1204776](https://doi.org/10.1038/sj.onc.1204776) PMID: [11607835](https://pubmed.ncbi.nlm.nih.gov/11607835/)
44. Zhao L-J, Wang L, Ren H, Cao J, Li L, Ke J-S, et al. Hepatitis C virus E2 protein promotes human hepatoma cell proliferation through the MAPK/ERK signaling pathway via cellular receptors. *Experimental Cell Research*. 2005; 305(1):23–32. doi: [10.1016/j.yexcr.2004.12.024](https://doi.org/10.1016/j.yexcr.2004.12.024) PMID: [15777784](https://pubmed.ncbi.nlm.nih.gov/15777784/)
45. Naas T, Ghorbani M, Alvarez-Maya I, Lapner M, Kothary R, De Repentigny Y, et al. Characterization of liver histopathology in a transgenic mouse model expressing genotype 1a hepatitis C virus core and envelope proteins 1 and 2. *J Gen Virol*. 2005; 86(8):2185–2196.
46. Keshava Prasad TS, Goel R, Kandasamy K, Keerthikumar S, Kumar S, Mathivanan S, et al. Human Protein Reference Database—2009 update. *Nucleic Acids Res*. 2009; 37(suppl 1):D767–D772. PubMed Central PMCID: [PMC2686490](https://pubmed.ncbi.nlm.nih.gov/PMC2686490/).
47. Kandasamy K, Keerthikumar S, Goel R, Mathivanan S, Patankar N, Shafreen B, et al. Human Proteomepedia: a unified discovery resource for proteomics research. *Nucleic Acids Res*. 2009; 37(suppl 1): D773–D781. PubMed Central PMCID: [PMC2686511](https://pubmed.ncbi.nlm.nih.gov/PMC2686511/).
48. Ploss A, Evans MJ, Gaysinskaya VA, Panis M, You H, de Jong YP, et al. Human occludin is a hepatitis C virus entry factor required for infection of mouse cells. *Nature*. 2009; 457(7231):882–886. doi: [10.1038/nature07684](https://doi.org/10.1038/nature07684) PMID: [19182773](https://pubmed.ncbi.nlm.nih.gov/19182773/)
49. Dorner M, Horwitz JA, Donovan BM, Labitt RN, Budell WC, Friling T, et al. Completion of the entire hepatitis C virus life cycle in genetically humanized mice. *Nature*. 2013; 501(7466):237–241. doi: [10.1038/nature12427](https://doi.org/10.1038/nature12427) PMID: [23903655](https://pubmed.ncbi.nlm.nih.gov/23903655/)
50. Pileri P, Uematsu Y, Campagnoli S, Galli G, Falugi F, Petracca R, et al. Binding of Hepatitis C Virus to CD81. *Science*. 1998; 282(5390):938–941. PMID: [9794763](https://pubmed.ncbi.nlm.nih.gov/9794763/)
51. Kong L, Giang E, Nieuwma T, Kadam RU, Cogburn KE, Hua Y, et al. Hepatitis C Virus E2 Envelope Glycoprotein Core Structure. *Science*. 2013; 342(6162):1090–1094. doi: [10.1126/science.1243876](https://doi.org/10.1126/science.1243876) PMID: [24288331](https://pubmed.ncbi.nlm.nih.gov/24288331/)
52. Fawcett T. An introduction to ROC analysis. *Pattern recognition letters*. 2006; 27(8):861–874.
53. Li Q, Brass AL, Ng A, Hu Z, Xavier RJ, Liang TJ, et al. A genome-wide genetic screen for host factors required for hepatitis C virus propagation. *PNAS*. 2009; 106(38):16410–16415. doi: [10.1073/pnas.0907439106](https://doi.org/10.1073/pnas.0907439106) PMID: [19717417](https://pubmed.ncbi.nlm.nih.gov/19717417/)
54. Davis Zoe H, Verschueren E, Jang Gwendolyn M, Kleffman K, Johnson Jeffrey R, Park J, et al. Global Mapping of Herpesvirus-Host Protein Complexes Reveals a Transcription Strategy for Late Genes. *Molecular Cell*. 2015; 57(2):349–360. doi: [10.1016/j.molcel.2014.11.026](https://doi.org/10.1016/j.molcel.2014.11.026) PMID: [25544563](https://pubmed.ncbi.nlm.nih.gov/25544563/)
55. Ruepp A, Waegel B, Lechner M, Brauner B, Dunger-Kaltenbach I, Fobo G, et al. CORUM: the comprehensive resource of mammalian protein complexes—2009. *Nucleic Acids Res*. 2010; 38(suppl 1): D497–D501.
56. Qureshi A, Thakur N, Tandon H, Kumar M. AVPdb: a database of experimentally validated antiviral peptides targeting medically important viruses. *Nucleic Acids Res*. 2014; 42(D1):D1147–D1153.
57. Zona L, Tawar RG, Zeisel MB, Xiao F, Schuster C, Lupberger J, et al. CD81-Receptor Associations—Impact for Hepatitis C Virus Entry and Antiviral Therapies. *Viruses*. 2014; 6(2):875–892. doi: [10.3390/v6020875](https://doi.org/10.3390/v6020875) PMID: [24553110](https://pubmed.ncbi.nlm.nih.gov/24553110/)
58. Massagué J. TGFβ signalling in context. *Nature Reviews Molecular Cell Biology*. 2012; 13(10):616–630. doi: [10.1038/nrm3434](https://doi.org/10.1038/nrm3434) PMID: [22992590](https://pubmed.ncbi.nlm.nih.gov/22992590/)
59. Chen J, Zhang M. The Par3/Par6/aPKC complex and epithelial cell polarity. *Experimental Cell Research*. 2013; 319(10):1357–1364. doi: [10.1016/j.yexcr.2013.03.021](https://doi.org/10.1016/j.yexcr.2013.03.021) PMID: [23535009](https://pubmed.ncbi.nlm.nih.gov/23535009/)
60. Eksioglu EA, Zhu H, Bayouth L, Bess J, Liu H-y, Nelson DR, et al. Characterization of HCV Interactions with Toll-Like Receptors and RIG-I in Liver Cells. *PloS one*. 2011; 6(6):e21186. doi: [10.1371/journal.pone.0021186](https://doi.org/10.1371/journal.pone.0021186) PMID: [21695051](https://pubmed.ncbi.nlm.nih.gov/21695051/)
61. Jamma S, Hussain G, Lau DTY. Current Concepts of HBV/HCV Coinfection: Coexistence, but Not Necessarily in Harmony. *Curr Hepat Rep*. 2010; 9(4):260–269. doi: [10.1007/s11901-010-0060-4](https://doi.org/10.1007/s11901-010-0060-4) PMID: [21258658](https://pubmed.ncbi.nlm.nih.gov/21258658/)

62. Hernandez B, Tarrago T, Giralt E, Escribano JM, Alonso C. Small Peptide Inhibitors Disrupt a High-Affinity Interaction between Cytoplasmic Dynein and a Viral Cargo Protein. *J Virol*. 2010; 84(20):10792–10801. doi: [10.1128/JVI.01168-10](https://doi.org/10.1128/JVI.01168-10) PMID: [20686048](https://pubmed.ncbi.nlm.nih.gov/20686048/)
63. Chertov O, Zhang N, Chen X, Oppenheim JJ, Lubkowski J, McGrath C, et al. Novel Peptides Based on HIV-1 gp120 Sequence with Homology to Chemokines Inhibit HIV Infection in Cell Culture. *PloS one*. 2011; 6(1):e14474. doi: [10.1371/journal.pone.0014474](https://doi.org/10.1371/journal.pone.0014474) PMID: [21264298](https://pubmed.ncbi.nlm.nih.gov/21264298/)
64. Yuan Z-q, Kim D, Kaneko S, Sussman M, Bokoch GM, Kruh GD, et al. ArgBP2y Interacts with Akt and p21-activated Kinase-1 and Promotes Cell Survival. *J Biol Chem*. 2005; 280(22):21483–21490. doi: [10.1074/jbc.M500097200](https://doi.org/10.1074/jbc.M500097200) PMID: [15784622](https://pubmed.ncbi.nlm.nih.gov/15784622/)
65. Fabregat I, Roncero C, Fernandez M. Survival and apoptosis: a dysregulated balance in liver cancer. *Liver International*. 2007; 27(2):155–162. doi: [10.1111/j.1478-3231.2006.01409.x](https://doi.org/10.1111/j.1478-3231.2006.01409.x) PMID: [17311609](https://pubmed.ncbi.nlm.nih.gov/17311609/)
66. Durmuş S, akır T, zgur A, Guthke R. A review on computational systems biology of pathogen–host interactions. *Front Microbiol*. 2015: 235.
67. Thieu T, Joshi S, Warren S, Korkin D. Literature mining of host–pathogen interactions: comparing feature-based supervised learning and language-based approaches. *Bioinformatics*. 2012; 28(6):867–875. doi: [10.1093/bioinformatics/bts042](https://doi.org/10.1093/bioinformatics/bts042) PMID: [22285561](https://pubmed.ncbi.nlm.nih.gov/22285561/)
68. Eckardt-Michel J, Lorek M, Baxmann D, Grunwald T, Keil GM, Zimmer G. The Fusion Protein of Respiratory Syncytial Virus Triggers p53-Dependent Apoptosis. *J Virol*. 2008; 82(7):3236–3249. doi: [10.1128/JVI.01887-07](https://doi.org/10.1128/JVI.01887-07) PMID: [18216092](https://pubmed.ncbi.nlm.nih.gov/18216092/)
69. Pahl HL, Baeuerle PA. Expression of influenza virus hemagglutinin activates transcription factor NF-kappa B. *J Virol*. 1995; 69(3):1480–1484. PMID: [7853480](https://pubmed.ncbi.nlm.nih.gov/7853480/)
70. Ray RB, Basu A, Steele R, Beyene A, McHowat J, Meyer K, et al. Ebola virus glycoprotein-mediated anoikis of primary human cardiac microvascular endothelial cells. *Virology*. 2004; 321(2):181–188. doi: [10.1016/j.virol.2003.12.014](https://doi.org/10.1016/j.virol.2003.12.014) PMID: [15051379](https://pubmed.ncbi.nlm.nih.gov/15051379/)
71. Evans P, Dampier W, Ungar L, Tozeren A. Prediction of HIV-1 virus-host protein interactions using virus and host sequence motifs. *BMC Medical Genomics*. 2009; 2:27. doi: [10.1186/1755-8794-2-27](https://doi.org/10.1186/1755-8794-2-27) PMID: [19450270](https://pubmed.ncbi.nlm.nih.gov/19450270/)
72. Ferenci P. Treatment of hepatitis C in difficult-to-treat patients. *Nat Rev Gastroenterol Hepatol*. 2015; 12(5):284–292. doi: [10.1038/nrgastro.2015.53](https://doi.org/10.1038/nrgastro.2015.53) PMID: [25895822](https://pubmed.ncbi.nlm.nih.gov/25895822/)
73. Xiao F, Fofana I, Heydmann L, Barth H, Soulier E, Habersetzer F, et al. Hepatitis C Virus Cell-Cell Transmission and Resistance to Direct-Acting Antiviral Agents. *PLoS Pathog*. 2014; 10(5):e1004128. doi: [10.1371/journal.ppat.1004128](https://doi.org/10.1371/journal.ppat.1004128) PMID: [24830295](https://pubmed.ncbi.nlm.nih.gov/24830295/)
74. Zahid MN, Turek M, Xiao F, Thi VLD, Guerin M, Fofana I, et al. The postbinding activity of scavenger receptor class B type I mediates initiation of hepatitis C virus infection and viral dissemination. *Hepatology (Baltimore, Md)*. 2013; 57(2):492–504.
75. 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; 8(5):e64221. doi: [10.1371/journal.pone.0064221](https://doi.org/10.1371/journal.pone.0064221) PMID: [23704981](https://pubmed.ncbi.nlm.nih.gov/23704981/)
76. Lupberger J, Zeisel MB, Xiao F, Thumann C, Fofana I, Zona L, et al. EGFR and EphA2 are host factors for hepatitis C virus entry and possible targets for antiviral therapy. *Nat Med*. 2011; 17(5):589–595. doi: [10.1038/nm.2341](https://doi.org/10.1038/nm.2341) PMID: [21516087](https://pubmed.ncbi.nlm.nih.gov/21516087/)
77. Jones CT, Catanese MT, Law LMJ, Khetani SR, Syder AJ, Ploss A, et al. Real-time imaging of hepatitis C virus infection using a fluorescent cell-based reporter system. *Nature biotechnology*. 2010; 28(2):167–171. doi: [10.1038/nbt.1604](https://doi.org/10.1038/nbt.1604) PMID: [20118917](https://pubmed.ncbi.nlm.nih.gov/20118917/)
78. Timpe JM, Stamataki Z, Jennings A, Hu K, Farquhar MJ, Harris HJ, et al. Hepatitis C virus cell-cell transmission in hepatoma cells in the presence of neutralizing antibodies. *Hepatology*. 2008; 47(1):17–24. doi: [10.1002/hep.21959](https://doi.org/10.1002/hep.21959) PMID: [17941058](https://pubmed.ncbi.nlm.nih.gov/17941058/)
79. Witteveldt J, Evans MJ, Bitzegeio J, Koutsoudakis G, Owsianka AM, Angus AGN, et al. CD81 is dispensable for hepatitis C virus cell-to-cell transmission in hepatoma cells. *J Gen Virol*. 2009; 90(1):48–58.
80. Mothes W, Sherer NM, Jin J, Zhong P. Virus Cell-to-Cell Transmission. *J Virol*. 2010; 84(17):8360–8368. doi: [10.1128/JVI.00443-10](https://doi.org/10.1128/JVI.00443-10) PMID: [20375157](https://pubmed.ncbi.nlm.nih.gov/20375157/)
81. Baugh JM, Garcia-Rivera JA, Gallay PA. Host-targeting agents in the treatment of hepatitis C: A beginning and an end? *Antiviral Research*. 2013; 100(2):555–561. doi: [10.1016/j.antiviral.2013.09.020](https://doi.org/10.1016/j.antiviral.2013.09.020) PMID: [24091203](https://pubmed.ncbi.nlm.nih.gov/24091203/)
82. The UniProt Consortium. UniProt: a hub for protein information. *Nucleic Acids Res*. 2015; 43(D1):D204–D212.
83. Dennis G, Sherman B, Hosack D, Yang J, Gao W, Lane HC, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biology*. 2003; 4(5):P3. PMID: [12734009](https://pubmed.ncbi.nlm.nih.gov/12734009/)

84. Wu H-M, Tien Y-J, Chen C-h. GAP: A graphical environment for matrix visualization and cluster analysis. *Computational Statistics & Data Analysis*. 2010; 54(3):767–778.
85. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res*. 2000; 28(1):27–30. PMID: [10592173](#)
86. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A Journal of Integrative Biology*. 2012; 16(5):284–287. doi: [10.1089/omi.2011.0118](#) PMID: [22455463](#)