J Comput Aided Mol Des (2013) 27:337–346
DOI 10.1007/s10822-013-9649-3

Identification of ligands that target the HCV-E2 binding site on CD81

Reem Al Olaby · Hassan M. Azzazy · Rodney Harris · Brett Chromy · Jost Vielmetter · Rod Balhorn

Received: 27 December 2012 / Accepted: 15 April 2013 / Published online: 24 April 2013

Abstract Hepatitis C is a global health problem. While many drug companies have active R&D efforts to develop new drugs for treating Hepatitis C virus (HCV), most target the viral enzymes. The HCV glycoprotein E2 has been shown to play an essential role in hepatocyte invasion by binding to CD81 and other cell surface receptors. This paper describes the use of AutoDock to identify ligand binding sites on the large extracellular loop of the open conformation of CD81 and to perform virtual screening runs to identify sets of small molecule ligands predicted to bind to two of these sites. The best sites selected by AutoLigand were located in regions identified by mutational studies to be the site of E2 binding. Thirty-six ligands predicted by AutoDock to bind to these sites were subsequently tested experimentally to determine if they bound to CD81-LEL. Binding assays conducted using surface Plasmon resonance revealed that 26 out of 36 (72 %) of the ligands bound in vitro to the recombinant CD81-LEL protein. Competition experiments performed using dual polarization interferometry showed that one of the ligands predicted to bind to the large cleft between the C and D helices was also effective in blocking E2 binding to CD81-LEL.

Keywords Hepatitis C virus · HCV glycoprotein E2 · AutoDock · CD81 receptor · Viral entry inhibitors · Dual polarization interferometry · Surface plasmon resonance

R. A. Olaby · H. M. Azzazy
Department of Chemistry, The American University in Cairo, Cairo, Egypt

R. Harris
The Scripps Research Institute, La Jolla, San Diego, CA, USA

B. Chromy
Department of Pathology, UC Davis Medical Center, Sacramento, CA, USA

J. Vielmetter
Protein Expression Center, California Institute of Technology, Pasadena, CA, USA

R. Balhorn (✉)
Department of Applied Science, University of California Davis, Davis, CA 95616, USA
e-mail: rodbalhorn@hughes.net

Present Address:
R. Balhorn
SHAL Technologies Inc, 15986 Mines Road, Livermore, CA, USA

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 sulphate [11], scavenger receptor class BI (SR-BI) and CD81 [12, 13]. Pileri et al. [14] was

the first to identify CD81, a 26 kDa protein that belongs to the tetraspanins super family, as an important HCV receptor. 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. [18] discovered that CD81-LEL is also important for efficient replication of the HCV genome. 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 γ-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 VanCompernolle 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. [21, 22] 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. 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 favoured 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.

Materials and methods

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.

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 25,000. 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 small molecules predicted to bind to each site (~350 compounds) were 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.

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 (10 mg) 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.

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 a well established experimental technique, if the ligands bound to the protein. Briefly, 10 μM CD81-LEL-GST diluted into 10 mM sodium acetate buffer pH 4.5 was immobilized for 15 min at a flow speed of 5 μl/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 μM solutions in PBS-0.05 % Tween-80 (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 Kd, the original ligand sample was diluted serially with running buffer to produce seven different ligand concentrations: 1024, 516, 256, 128, 64, 32 and 0 μM. Data were fitted using a monovalent binding model.

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 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 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 glycoprotein and a mixture of the E2 glycoprotein and the ligand.

Results and discussion

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 glycoprotein 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 $Å^2$ of the CD81-LEL surface, and identified three additional amino acid contacts, Ile182, Asn184, and Leu162 [19] (Fig. 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.

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 $Å^3$) 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 $Å^3$ 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 kcal/mol $Å^3$, 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 favourable.

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 (Fig. 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

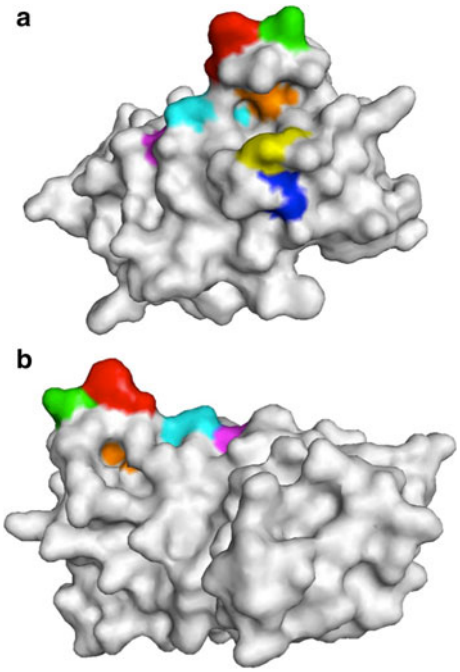


Fig. 1 Amino acid residues that participate in HCV E2 binding to CD81-LEL The colored residues are amino acids that have been identified by Higginbottom et al. [17] and Drummer et al. [19] to contribute to the binding of the HCV protein E2 to CD81-LEL. The structure shown is the monomer of the open conformation 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

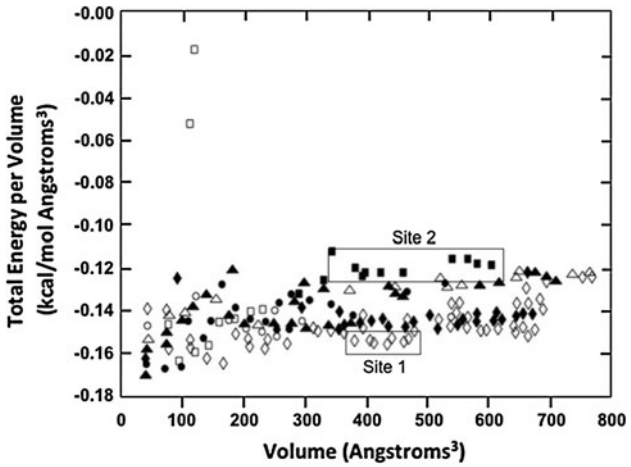


Fig. 2 Predicted free energy data for ligand binding sites identified on the surface of CD81-LEL by AutoLigand This figure was generated by plotting the total energy per volume versus the volume of each fill made from different amounts of fill points. The different symbols depict the fills that start 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

second group of fill points was generated for a neighboring cavity located on the opposite side of the protein (Fig. 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.

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 (carbon, hydrogen, nitrogen or oxygen) in the ligand that would interact optimally with the surface of the protein in the regions surrounding the ligand (Fig. 4). The fill points predicted for the site shown in Fig. 4a are colored red for hydrogen acceptors such as oxygen or nitrogen, blue for hydrogen, or gray for carbon. One of the better ligands predicted to bind to this site (Fig. 4b) has atoms that superimpose well with the fill point map (Fig. 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.

Experimental confirmation of ligand binding

A total of 36 ligands were tested experimentally using surface Plasmon resonance (on a 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

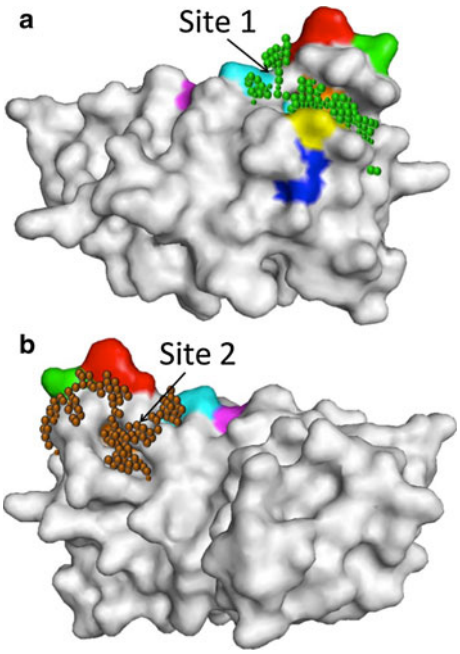


Fig. 3 Two ligand binding sites identified by AutoLigand on the open conformation of CD81-LEL (PDB ID: 1G8Q) These two sites were selected as docking targets based on their proximity to the amino acid residues that contact E2 (identified on the molecular surface by *blue*, *yellow*, *green*, *orange*, *cyan*, *red* and *magenta* colors; see Fig. 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 Fig. 2 located between 500 and 600 Å^3. **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 Fig. 2 located between 550 and 650 Å^3. This figure was prepared using AutoDock Tools version 1.5.6

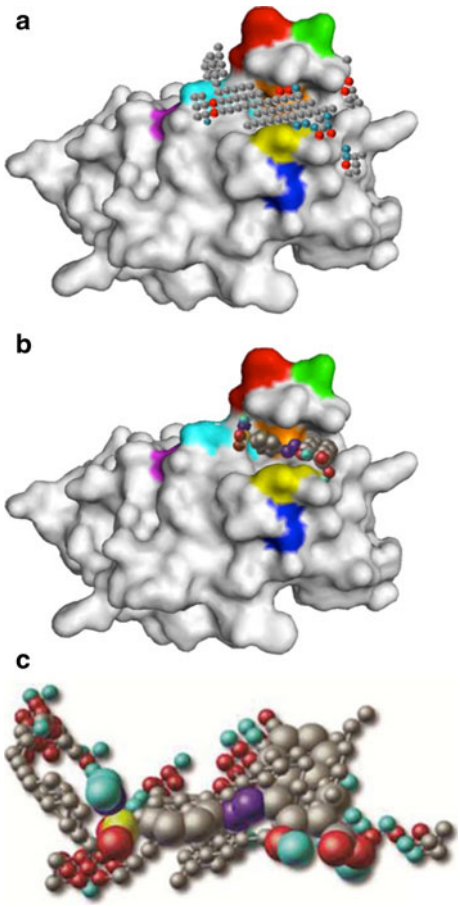


Fig. 4 AutoLigand analysis of the Site 1 ligand binding site AutoLigand fill points not only identify cavities on the surfaces of proteins, but they also predict the structural features of ligands that would bind with the best affinity and selectivity to the protein at these sites. **a** The fill points 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 Fig. 1 are also shown in these figures using the same color-coding. This figure was prepared using AutoDock Tools version 1.5.6

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 glycoprotein 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 Fig. 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 Fig. 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

Table 1 Experimental analysis of ligand binding to recombinant CD81-LEL

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
Binding Site	% Ligands Bound (Number Tested)	25368	13.3
		68971	12.1
		156957	10.6
1	69.2 (13)	55573	10.6
2	73.9 (23)	303800	7.8
		362639	6.5

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 >0 indicate binding

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

Binding of the HCV protein E2 to CD81-LEL immobilized on a chip in the absence and presence of two small molecules, Ligand 3 and benzyl salicylate, was determined by DPI as described in the Materials and methods section. Both benzyl salicylate and Ligand 3 were observed to reduce E2 binding to CD81-LEL when mixed with the protein prior to its addition to4 the chip containing CD81-LEL

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 Fig. 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 (Fig. 6) provided an estimated Kd of 201 μM for an affinity fit of Ligand 1 binding to the recombinant CD81-LEL protein.

Effect of ligand 3 on in vitro binding of HCV E2 glycoprotein 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 Fig. 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 EC50 greater than 500 μM 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.

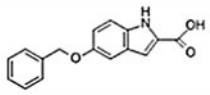
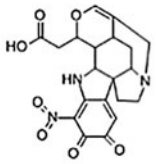
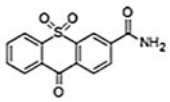
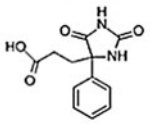
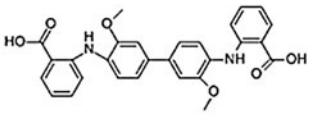
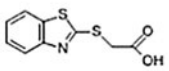
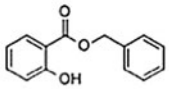
Ligand Number (NCI number)	Structure	Predicted binding energy (kcal/mole)	DPI Response (Radians)	SPR Response (RU)
Site 1				
1 (30930)		- 7.0	1.1	67.7
2 (5069)		- 5.7	0.53	67.5
3 (689002)		- 4.2	0.56	2.3
Site 2				
4 (38743)		- 5.5	0.62	30.6
5 (127947)		- 3.5	0.33	64.8
6 (11891)		- 3.9	0.16	15.2
Benzyl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

Fig. 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

Conclusion

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

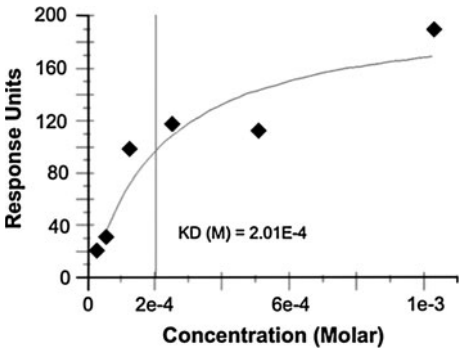


Fig. 6 Binding of Ligand 1 to CD81-LEL as a function of ligand concentration This binding experiment was performed using a Biacore T200 as described in the Materials and methods section. Using the data shown, Ligand 1 was estimated to have a Kd of ~201 μM based on an affinity fit

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. Because such a high percentage of small molecules predicted by AutoDock to bind CD81-LEL were found to bind to the protein experimentally (72 %), only a small number of ligands (36) had to be 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 by Holzer et al. [33] 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 [24]. These new molecules will be synthesized and tested in a series of HCV cell culture assays and customized mouse models to assess their ability to target the E2 binding site on CD81 and to block HCV infectivity.

Acknowledgments This work was supported by The Nadhmi Auchi Fellowship, The American University in Cairo, awarded to R. Al Olaby. Computer time was provided by the National Biomedical Computational Resource at the University of California, San Diego. We would like to thank Dr. Arthur Olson's molecular graphics laboratory at The Scripps Research Institute for supporting the AutoDock tutorial at Scripps Research Institute and providing assistance with AutoDock 4.2 and AutoDock Tools 1.5.6. We would also like to thank Dr. Shoshana Levy (Stanford University) for generously providing us with the human GST-CD81-LEL protein. Biacore support and instrument use was provided by the Protein Expression Center at Caltech. The ligands tested in this study were provided by the National Cancer Institute through its Developmental Therapeutics Program. Special thanks goes to Dr. David Tirrell (Tirrell lab-Caltech) and Nancy Aitken (Social Entrepreneur) for their great support.

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 6:165–181
3. Lesburg CA, Radfar R, Weber PC (2000) Recent advances in the analysis of HCV NS5B RNA-dependent RNA polymerase. Curr Opin Investig Drugs 1:289–296
4. Welbourn S, Pause A (2007) The hepatitis C virus NS2/3 protease. Curr Issues Mol Biol 9:63–69
5. Venkatraman S, Njoroge FG (2009) Macrocyclic inhibitors of HCV NS3 protease. Expert Opin Ther Pat 19:1277–1303
6. Enomoto M, Tamori A, Kawada N (2009) Emerging antiviral drugs for hepatitis C virus. Rev Recent Clin Trials 4:179–184
7. Chary A, Holodniy M (2010) Recent advances in hepatitis C virus treatment: review of HCV protease inhibitor clinical trials. Rev Recent Clin Trials 5:158–173
8. Sharma SD (2010) Hepatitis C virus: molecular biology and 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 8:333–345
10. Dubuisson J (2007) Hepatitis C virus proteins. World J Gastroenterol 13:2406–2415
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. 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 278:41624–41630
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: 938–941
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 74:4824–4830
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 74:3642–3649
18. Zhang YY, Zhang BH, Ishii K, Liang TJ (2010) Novel function of CD81 in controlling hepatitis C virus replication. J Virol 84:3396–3407
19. Drummer HE, Wilson KA, Poumbourios P (2002) Identification of the hepatitis C virus E2 glycoprotein binding site on the large extracellular loop of CD81. J Virol 76:11143–11147
20. VanCompernolle SE, Wiznycia AV, Rush JR, Dhanasekaran M, Baures PW, Todd SC (2003) Small molecule inhibition of hepatitis C virus E2 binding to CD81. Virology 314:371–380
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 20:12–18
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 57:156–158
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 28:1145–1152
28. Huey R, Goodsell DS, Morris GM, Olson AJ (2004) Grid-based hydrogen bond potentials with improved directionality. Lett Drug Des Discov 1:178–183
29. Harris R, Olson AJ, Goodsell DS (2008) Automated prediction of ligand binding sites in proteins. Proteins 70:1506–1517
30. Morris GM, Huey R, Olson A (2008) Using autodock for ligand-receptor docking. Curr Protoc Bioinforma 8–14
31. NBCR website: https://www.nbcr.net/pub/wiki/index.php?title=Main_Page
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) 341:478–484