

Identification of Novel Hepatitis C Virus NS3-4A Protease Inhibitors by Virtual Screening Approach

Anish Kumar¹, Rashmi Gupta¹, Kanika Verma¹, Kshitija Iyer² Shanthi V¹ and K Ramanathan^{1*}

¹Industrial Biotechnology Division, School of Bio Sciences and Technology, VIT University, Vellore, Tamil Nadu, India
²M.Sc Biotechnology, VIT University, Vellore, Tamil Nadu, India

Abstract

Hepatitis C Virus (HCV) is a leading cause of cirrhosis and liver cancer worldwide. The replication and viral polyprotein maturation of HCV crucially depends on the cleavage of the polyprotein precursor into 10 viral proteins. The NS3-4A serine protease cleaves the nonstructural region of the polyprotein at four out of five junctions, thus is a promising target for the development of antiviral inhibitors. There are many inhibitors of HCV NS3/4A protease in the clinical trial and improvement indicating significant reduction in the viral infection rate. However, most PIs develop resistance associated variants while treatment and are restricted to one or two HCV genotypes. The second generation PI, MK-5172, is the only exception, which potently inhibits most variants associated with resistance to first generation PIs and is pan-genotypic. In this study, we investigated the potent lead compound(s) based on similarity search using the most potent proved protease inhibitor, MK-5172. We have performed virtual screening techniques using PubChem database available in NCBI to identify lead like molecules. The database has yielded 32 hits for 95% similarity search and the pharmacokinetic analysis (ADME) was performed for screened compounds. This structure based drug design identified three lead compounds that can work better against NS3/4A protease.

Keywords: NS3-4A Protease; Ligand; Virtual screening; Docking; Protease inhibitors

Abbreviations: HCV; Hepatitis C virus, NCBI; National Centre for Biotechnological Information, PDB; Protein Data Bank, ADME; absorption, distribution, metabolism and excretion, SVR; Sustained Viral Response, PI; protease inhibitor, DAA; Direct acting antiviral, SAR; Structure Activity Relationship, VP; Virtual Screening

Introduction

Hepatitis C virus (HCV) infects an estimated 200 million people worldwide. About 3% of the world's population is chronically infected with HCV and 3-4 million people are newly infected each year [1], often leading to cirrhosis, hepatic failure, and hepatocellular carcinoma. Hepatitis C virus (HCV) is an enveloped RNA virus which belongs to the Flaviviridae family and is the unique member of the Hepaci virus genus with at least 6 genotypes and numerous subtypes [2,3]. Efficient vaccine against HCV does not exist and the only available standard therapy is a combination of pegylated interferon- α (IFN-Peg) and ribavirin, efficient in only 50% of patients chronically infected [4]. Though, IFN α /Ribavirin helps in controlling the viral outbreak inside hosts, these are indirect antivirals and do not target a specific HCV protein or RNA element. Moreover, patients undergoing interferon-based therapies experience significant adverse effects, including flu-like symptoms, anemia, and depression. It is therefore of major importance to develop anti-virals acting directly on the viral protein. Most importantly, direct-acting antiviral agents (DAA) have the potential to improve Sustained virologic response (SVR) rates and minimize treatment duration.

HCV RNA genome encodes a polyprotein precursor which contains structural proteins (C, E1, E2 and p7) and non-structural proteins (NS1, NS2, NS3, NS4A, NS4B, NS5A and NS5B). Non-structural proteins are involved in the replication of HCV genome and the building up of virions. Among them, NS2 and NS3 protease are essential enzymes for polyprotein processing. Therefore, they are the potential targets for screening anti HCV compounds [5]. NS3 protease activated by cofactor NS4A causes the cleavage of polyprotein at four

junctions, producing the nonstructural proteins 4A, 4B, 5A and 5B and is therefore complementary for the replication of virus [6,7]. Hence, the direct role of NS3/4A protease in viral replicate machinery is proved an important therapeutic target for the cure of hepatitis C [8]. The active site configuration of NS3 protease comprises the residues His-57, Asp-81, and Ser-139 [9]. Many inhibitors of HCV NS3/4A protease are in the clinical trials and improvement; indicating significant reduction in the viral load of patients [10]. However, the active site for protease inhibitors is a long shallow groove and even a single-point mutation is sufficient to hinder the binding of inhibitors. Therefore, pipelines of compounds certainly essential to meet the drug resistance situations. In silico screening has proven to be useful to meet the challenges of antiviral drug discovery. Large virtual compound libraries are filtered by different computational screening methods such as docking, ligand-based similarity searches or pharmacophore based screening. This approach is helpful in reducing the number of candidate molecules to a smaller set of promising candidates that are then tested biologically. Since years, computational techniques like virtual screening have proven to be of great use to make the drug development process faster and less expensive [11].

In the past decades, a lot of work in the field of drug discovery and optimization for antiviral has been done using computational techniques [12-14]. Further development in computational techniques will certainly aid in the identification of potential DAA. Hence, in the present study we identified potent lead compound (s) as DAA/PI

***Corresponding author:** K Ramanathan, Industrial Biotechnology Division, School of Bio Sciences and Technology, VIT University, Vellore- 632014, Tamil Nadu, India, Tel: +91 4162202545; Fax: +91 4162243092; E-mail: kramanathan@vit.ac.in

Received May 23, 2013; Accepted June 10, 2013; Published June 17, 2013

Citation: Kumar A, Gupta R, Verma K, Iyer K, Shanthi V, et al. (2014) Identification of Novel Hepatitis C Virus NS3-4A Protease Inhibitors by Virtual Screening Approach. J Microb Biochem Technol R1: 005. doi:10.4172/1948-5948.R1-005

Copyright: © 2014 Kumar A, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

(protease inhibitors) which should have unique mechanisms of action, higher potency, pharmacokinetics, novel binding properties, improved mutant potency profile and pan-genotype activity (effective against all HCV genotypes (1 to 7)); with the help of virtual screening techniques considering the existing most efficient drug in the pipeline as the reference compound.

Methodology

Preparation of data set

HCV NS3/4A protease structures were obtained from Protein Data Bank [15]. Currently, RCSB Protein Data Bank (PDB) is presented with about 108 structures of NS3-4A protease of HCV both in complex and free forms which provide a valuable source for the development of novel and potential drug against HCV. The native and mutant (A156T) structures of NS3-4A protease were taken from RCSB protein data bank (PDB) [16]. The corresponding PDB codes are 2P59 and 3SUG respectively. MK-5172 was used as the protease inhibitor for the ligand-based virtual screening of lead compounds in our investigation. The SMILES (Simplified Molecular Input Line Entry System) strings of the drug and lead compounds such as MK-5172, CID 58428446, CID 71276250 and CID 71276290 were collected from PubChem database and submitted to CORINA in order to generate 3D structures [17,18]. The three-dimensional structure of target proteins (2P59 and 3SUG) and lead compounds were energy-minimized using GROMACS package 4.5.3 [19] adopting the GROMOS43a1 force field parameters before performing the computational analysis.

Virtual Screening

Virtual Screening is an important tool in computer-assisted drug discovery and requires prior biological information to predict active compounds. It is the computational cognate of biological screening and is popularly used for lead identification in pharmaceutical research [20]. VP reduces the massive virtual chemical space of small organic molecules, to screen against a specific target protein, to a manageable set of promising compounds that exhibit the highest chance to be a lead compound [21]. The PubChem database (<http://www.ncbi.nlm.nih.gov/pccompound>) was used for searching similarity based lead compounds by employing the MK-5172 as query [22]. MK-5172 is a P2-P4 macrocyclic competitive inhibitor of NS3/4A protease with a broad HCV genotypic activity high mutant potency profile. It exhibits excellent selectivity over other serine proteases and shows improved inhibitory potency. Its molecular formula is C₂₉H₃₈N₄O₇ with a molecular weight of 554.63 and a density of 1.33. Structural likeness increases the chance to share a common bioactive profile so ligand-based similarity methods are preferred. Selecting compounds similar to known drugs increases the possibility of choosing a potential lead [21]. Based on a structural similarity search among small molecules, it is possible to retrieve compounds containing identical substructures that share affinity to the same receptors [23]. The number of molecules found in the database after 95% similarity search is around 32 compounds. The candidate compounds were further screened using molecular docking studies, bioavailability and ADME analysis.

ADME and Toxicity

Molecular properties viz, membrane permeability, polar surface area and bioavailability are always identified with some basic molecular descriptors such as logP (partition coefficient), molecular weight (MW), TPSA or counts of hydrogen bond acceptors and donors in a molecule [24]. These molecular properties were used in devising "rule of five" [25]. The rule states that most molecules with

drug likeliness and good membrane permeability have MW ≤ 500, calculated Octanol–water partition coefficient, log P ≤ 5 and hydrogen bond donors ≤ 5, acceptors ≤ 10 [26]. Therefore, Lipinski's Rule of Five was used to test the bioavailability characteristics such as adsorption, distribution, metabolism, elimination (ADME) of the lead compounds. In this study, these molecular properties and bioactivity for all the lead compounds were estimated using MOLINSPIRATION program (<http://www.molinspiration.com/cgi-bin/properties>) [27]. Successful drug discovery requires high-quality lead structures which may need to be more drug-like than commonly accepted [28]. Toxicity and poor pharmacokinetics should be eliminated in the early stages of drug discovery. Hence, the hits were further screened using drug-likeness, toxicity characteristics, and drug score. These physico-chemical properties were thus calculated for the filtered set of hits using the program OSIRIS Property Explorer (<http://www.organic-chemistry.org/prog/peo/>) [29].

The OSIRIS program calculates the drug-likeness based on a list of about 5,300 distinct substructure fragments created by 3,300 traded drugs as well as 15,000 commercially available chemicals yielding a complete list of all available fragments with associated drug-likeness. The drug score combines drug-likeness, cLogP, logS, MW, and toxicity risks as a total value which may be used to compute the compound's drug-score and its overall potential to qualify for a drug [29].

Molecular docking

The process of docking is involved with specification of ligand binding site in a receptor molecule and then docking the candidate ligands into the specified site. The lead compounds obtained from the ligand-based VS analysis were used in docking calculation and binding energy estimation. The SMILES strings were used for constructing the 3D-structures of all lead compounds using CORINA software. After docking, the compounds were ranked based on the geometric matching score with target proteins. The geometric matching score of MK-5172 with the target proteins (native and mutant structures) were used as reference for filtering the lead compounds. The energy minimized structures of NS3-4A protein were used as a template molecule to dock the screened inhibitors. In this study, rigid docking analysis was performed by means of Patchdock program (<http://bioinfo3d.cs.tau.ac.il/PatchDock/>) [30].

It is geometry-based molecular docking algorithm defined on shape-complementary principles. The PatchDock algorithm divides the Connolly dot surface representation [31] of the molecules into concave, convex, and flat patches. Then, complementary patches are matched to generate candidate transformations. Each candidate transformation is further evaluated by a scoring function that considers both geometric fit and atomic desolvation energy [32].

Finally, root mean square deviation (RMSD) clustering is applied to the candidate solutions to discard redundant solutions. The input parameters for the docking are the PDB coordinate file of the protein and ligand molecule. This algorithm has three major stages: (i) molecular shape representation, (ii) surface patch matching (iii) filtering and scoring. The refinement and rescore of the docking solutions from Patch Dock was performed using Fire Dock. It consists of two main steps: (i) rearrangement of the interface side-chains and (ii) adjustment of the relative orientation of the molecules. The method accounts for the observation that most interface residues that are important in recognition and binding do not change their conformation significantly after complex process. It restricts side-chain movements and thus manages to reduce the false-positive rate

noticeably. It is worth to mention that Fire Dock prediction results are comparable to current state-of-the-art refinement methods [33]. The service is available at <http://bioinfo3d.cs.tau.ac.il/FireDock/>.

Molecular dynamics simulation

GROMACS package 4.5.3 was used to perform simulations for the native and mutant NS3-4A protease-MK-5172 as well for CID 71276250 molecule [34]. Gromos 43a1 force field was implemented with GROMACS package 4.5.3. With the help of periodic boundary conditions and the SPC water model, the protein was solvated in cubic 0.9 nm [35]. PRODRG server was used to generate topology of the ligand. This server uses GROMOS force field for generating topology file and assigning atom types [36]. Seven Chlorine (7 CL⁻) counter ions were added to neutralize the total charge of the system. 1000 steps of steepest descent energy minimization were carried out for the proteins. After energy minimization step, the system was equilibrated at constant temperature and pressure. The system was equilibrated at constant temperature of 300 K and at the constant pressure of 1 atm. Using an atom-based cutoff of 8 Å, the non-bonded list was generated. The long range electrostatic interactions were handled by particle-mesh Ewald algorithm [37] whereas constrains bond lengths at their equilibrium values were handled by SHAKE algorithm handled long range electrostatic interactions and constrain bond lengths at their equilibrium values respectively [38]. The total simulation time was set to 1000 ps with integration time step of 2 fs. Trajectories were stored in traj.trr file and structural analysis was done at every picoseconds. Root mean square deviation (RMSD) was analyzed through Gromacs utilities g_rms.

Results and Discussion

We obtained the crystal structure of NS3-4A protease (2P59) and its mutant (3SUG) from PDB database. It is believed that similar structure compounds may have similar function.

Most importantly, the importance of virtual screening is highlighted number of times in the recent literatures [39-42]. Therefore, 95% similarity was applied to filter the candidate compounds. In silico ligand-based Virtual Screening result indicates that 32 compounds were identified for NS3-4A protease inhibition, similar to the currently active drug molecule, MK-5172. Many drug candidates fail in the clinical trials; reasons are unrelated in the potency against the intended drug target. Bioavailability, pharmacokinetics and toxicity issues are blamed for more than half of all failure in the clinical trials of HCV drugs. Therefore, it is essential to evaluate the oral bioavailability, pharmacokinetics and toxicity of small molecules.

Bioavailability analysis

The molecular properties and bioactivity for the candidate compounds was anticipated using Molinspiration program. The LogKow program [43] estimates the log octanol/water partition coefficient (logP) of organic chemicals and drugs by an atom/fragment contribution method [44]. Molecular polar surface area (TPSA) is calculated based on the methodology as a sum of fragment contributions. PSA has been shown to be a very good descriptor characterizing drug absorption, including intestinal absorption, blood-brain barrier penetration and bioavailability. The two important predictors of oral bioavailability of drug molecules are Log P value and PSA values [45]. Thus, we calculated log P and PSA values along with other physicochemical properties such as molecular mass, the number of hydrogen bond acceptors, and the number of hydrogen bond donors for the all the 32 candidate compounds obtained from

ligand-based VS. It is believed that lesser the nviolations (number of violations of Lipinski's Rule of Five) better is the drug molecule. The results showed that only 3 molecules (CID 58428446, CID 71276250 and CID71276290) had minimum 1violations, though MK-5172 shows 2 violations; which confirms that these molecules act as better drug molecules against NS3-4A protease (Table1).

Toxicity and physicochemical properties

The toxicity and drug properties of molecules were predicted using OSIRIS Property Explorer. It analyzes the drug likeness and thus the pharmacokinetic property by considering parameters like toxicity risks, PSA, clogP and solubility of the candidate compounds. clogP is a well-established measure of the compound's hydrophilic and therefore low logP/high hydrophilic values may indicate good absorption or permeation properties. It has been shown for compounds exhibiting reasonable probability of being well absorbed; their logP value must not be greater than 5.0. Drug solubility is an important factor that affects the movement of a drug from the site of administration into the blood. It is known that insufficient solubility of drug can lead to poor absorption [25]. The estimated log S value is a unit stripped logarithm (base ten) of a compound's solubility measured in mol/liter. Table 2 shows properties of all the candidate compounds. It is clear from the table that three compounds such as (CID 58428446, CID 71276250 and CID71276290) may fulfill the pharmacokinetics and could be considered as lead molecules for HCV NS3 protease inhibition. The toxicity risk predictor locates fragments within a molecule, which indicate a potential toxicity risk. Toxicity risk alerts are an indication that the drawn structure may be harmful concerning the risk category specified. From the data evaluated in Table 2, indicates that all the compounds including the standard MK-5172 possess some irritant property excluding only the compound 71276250.

Drug-Likeliness and Drug-Score

Drug-likeness is an important parameter because drug-like molecules exhibit favorable absorption, distribution, metabolism, excretion, toxicological (ADMET) parameters. Currently, there are many approaches to assess a compound drug-likeness based on topological descriptors, fingerprints of molecular drug-likeness structure keys or other properties such as clog P and MW. In this study, Osiris program [29] was used for calculating the fragment based drug-likeness of the lead compounds and comparing them with MK-5172. The drug score combines drug-likeness, miLogP, log S, MW, and toxicity risks in one convenient value than may be used to judge the compound's overall potential to qualify for a drug. It is interesting to note that CID 71276250 have significantly higher drug-score value (0.27), compared to MK-5172 (0.13) and other lead compounds. Furthermore, almost all the compounds including MK-5172 shows irritant property with the only exception of this compound showing no toxicity risks at all. This result indicates that CID 71276250 has distinct property in comparison with other lead molecules considered in our study. The toxicity, drug-likeness, and drug-score results for the compound CID 71276250 are illustrated in Figure 1.

Molecular docking analysis

Protein-ligand docking was performed to gain an insight into the binding affinity of lead compounds to the target protein. In the docking study, we employed Patchdock algorithm along with Fire Dock refinement. It is believed that this combined approach of docking analysis is certainly helpful for the elimination of false-positive in our prediction. The docking result is shown in Figure 2. It is clear from the

CID	miLogP	TPSA	natoms	MW	nON	nOHNH	nviolations	nrotb	volume
MK-5172 (46930991)	4.747	140.192	40	554.644	11	2	2	3	505.606
59178018	4.747	140.192	40	554.644	11	2	2	3	505.606
71278306	4.747	140.192	40	554.644	11	2	2	3	505.606
71307589	4.747	140.192	40	554.644	11	2	2	3	505.606
67304193	5.356	140.192	41	568.671	11	2	3	3	522.193
67304412	5.456	140.192	41	566.655	11	2	3	3	516.006
71276289	6.614	140.192	41	570.687	11	2	3	14	532.983
68491802	4.747	140.192	40	552.628	11	2	2	3	499.42
68491803	4.747	140.192	40	552.628	11	2	2	3	499.42
59178019	6.82	129.198	41	568.671	11	1	3	4	523.134
67304190	6.82	129.198	41	568.671	11	1	3	4	523.134
71276290	1.259	140.192	36	498.536	11	2	1	2	439.394
71278250	7.23	129.198	42	584.714	11	1	3	15	550.511
72535811	6.82	129.198	41	568.671	11	1	3	4	523.134
71278254	5.452	175.731	47	658.749	14	2	3	16	602.545
71278255	6.462	175.731	49	686.803	14	2	3	18	636.149
71278248	5.182	158.215	42	586.686	12	3	3	16	541.612
68236439	5.044	140.192	40	556.66	11	2	3	13	516.181
59178016	6.92	129.198	41	566.655	11	1	3	4	516.947
67304410	6.92	129.198	41	566.655	11	1	3	4	516.947
71276250	4.672	116.889	38	526.678	9	2	1	13	504.868
72535809	6.92	129.198	41	566.655	11	1	3	4	516.947
59178020	7.828	129.198	43	594.709	11	1	3	15	555.835
67304070	7.828	129.198	43	594.709	11	1	3	15	555.835
72535812	7.828	129.198	43	594.709	11	1	3	15	555.835
71276269	6.129	140.192	41	566.655	11	2	3	11	521.523
68236192	7.117	129.198	41	570.687	11	1	3	14	533.709
71529660	5.978	175.731	49	682.771	14	2	3	15	624.689
71276247	4.698	158.215	42	582.654	12	3	2	13	530.153
71532366	4.698	158.215	42	582.654	12	3	2	13	530.153
68236191	7.217	129.198	41	568.671	11	1	3	13	527.522
58428446	3.521	128.165	40	553.656	10	1	1	3	510.006

Table 1: Bioavailability of lead compounds-molinspiration data.

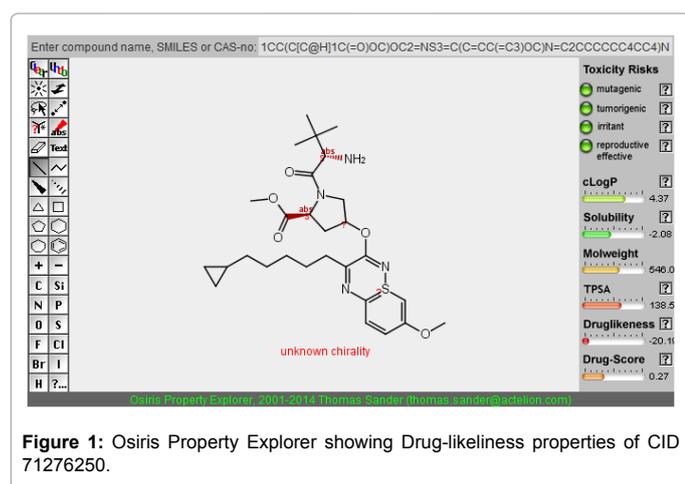


figure that the compound CID 71276250 possesses the binding energy of -41.36 and -42.37 kcal/mol for native and mutant NS3/4A protease respectively. In particular, the binding energy of this compound is considerable to the binding energy of MK-5172, indicating the effectiveness of this compound. The docked complex of native and mutant NS3-4A protease with each: MK-5172, CID 58428446, CID 71276250 and CID 71276290 and their intermolecular interactions are shown in Figures 3-6.

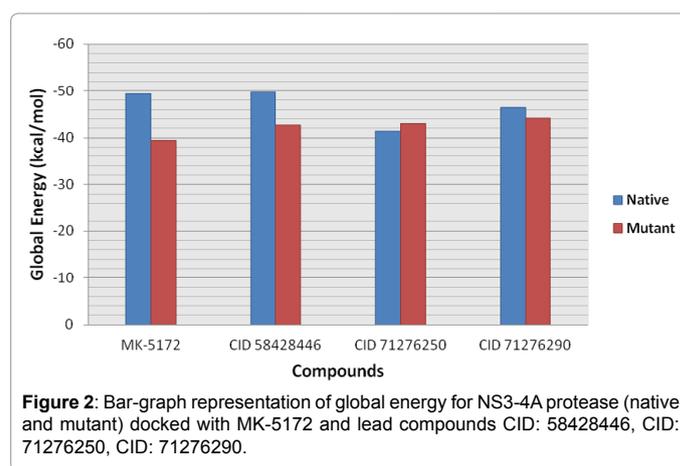


Figure 2: Bar-graph representation of global energy for NS3-4A protease (native and mutant) docked with MK-5172 and lead compounds CID: 58428446, CID: 71276250, CID: 71276290.

Molecular dynamics simulation

RMSD details were analyzed using GROMACS package 4.5.3. RMSD is the measure of the deviation of the mutant structure from their native structure. RMSD analysis can give an idea of how much the three-dimensional structure has deviated over time. Figure 7 shows the stable binding of CID 71276250 with respect to MK-5172. MK-5172 has deviated more than of CID 71276250 molecule. Native MK-5172 acquired ~0.20 nm at 200 ps during the simulations, while

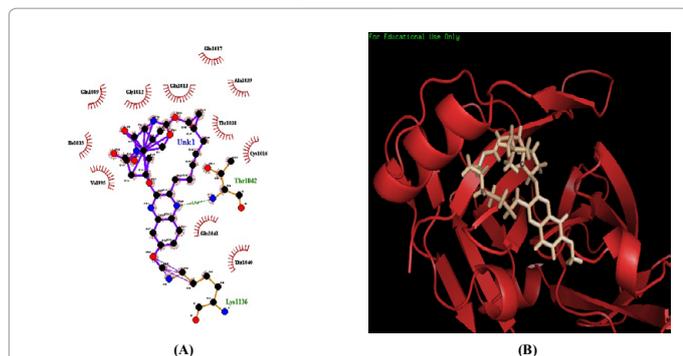


Figure 3: Complex structure of native NS3 protease with MK-5172 **A)** LigPlus view **B)** PyMol view.

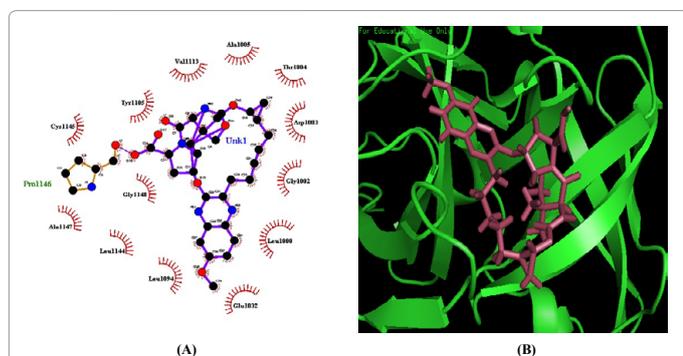


Figure 4: Complex structure of NS3 protease mutant (A156T) with MK-5172 **A)** LigPlus view **B)** PyMol view.

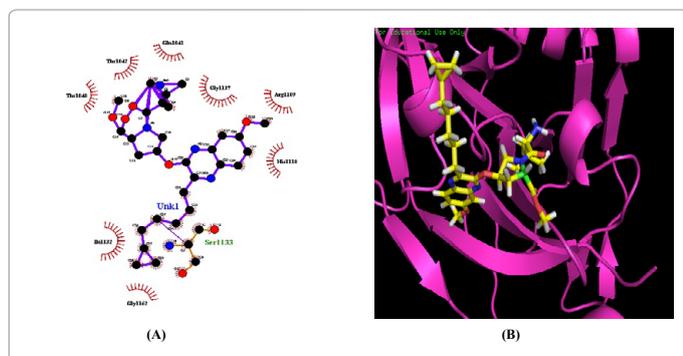


Figure 5: Complex structure of native NS3 protease with lead-compound CID: 71276250 **A)** LigPlus view **B)** PyMol view.

mutant MK-5172 structure acquired ~ 0.17 nm of backbone RMSD at 200 ps. Between a period of 400 and 800 ps, a mutant MK-5172 is able to maintain ~ 0.18 nm of backbone RMSD, while the native MK-5172 showed frequent changes (~ 0.23 to ~ 0.25 nm) in backbone RMSD. Native MK-5172 gained RMSD dominance over mutant and attained RMSD of ~ 0.24 nm at 1000 ps and the mutant MK-5172 achieved ~ 0.18 nm at 1000 ps. Native CID 71276250 acquired ~ 0.23 nm between a period of 400-800 ps while the mutant CID 71276250 acquired between ~ 0.17 to ~ 0.20 nm. At 1000 ps, native CID 71276250 attained RMSD of ~ 0.20 nm whereas mutant CID 71276250 attained RMSD of ~ 0.20 nm. This data suggested that MK-5172 had undergone more conformational changes throughout the simulation than the CID 71276250. Thus, indicating that MK-5172 has deviated more than the

CID 71276250. This shows that CID 71276250 is more stable than MK-5172. Further research still needs to be done to analyze their SAR. We envision that this preliminary study could lead to the development of potent, pan-genotypic and improved drugs against Hepatitis C virus.

Conclusion

Although there is no specific vaccine for HCV till now and the available standard therapy of Peg-IFN/Ribavirin is associated with many side effects (including depression, headache, fever, flu like symptoms, hemolytic anemia), there has been substantial progress in the research work and clinical improvement of novel antiviral drugs. NS3-4A protease is recognized as an important target because of its central role in the viral replication and survival inside patients and in confounding the innate immune response to viral infection. Developing inhibitors of the protease that can be orally available, have reduced toxicity and side effects, improved SVR, high mutant potency, pan-genotypic and available at low cost is the most challenging aspect. Study shows that currently, MK-5172 is the drug that satisfies the requisites to most extent and also has reduced cross-resistance evidences. Our research based on structure based drug design and similarity search, identified three lead compounds. Thus, we conclude that the inhibitors computationally studied here may be potent drug candidates and their potency may be increased against HCV NS3/4A protease with relatively simple structural changes. Further investigation of this molecule using experimental approaches would be an interesting future direction.

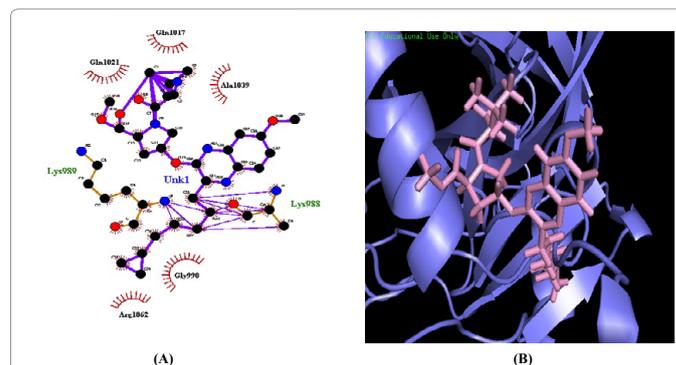


Figure 6: Complex structure of mutant NS3 protease (A156T) with lead-compound CID: 71276250 **A)** LigPlus view **B)** PyMol view.

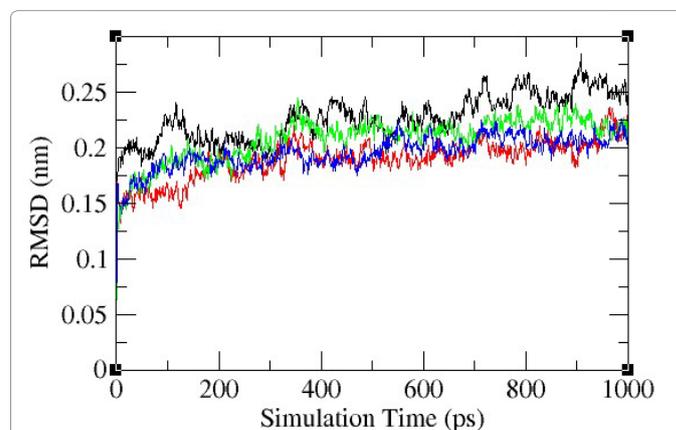


Figure 7: Root mean square deviations (RMSD) correspond to the native (black), mutant (red) type NS3-4A protease-MK5172 complex, native (green) and mutant type NS3-4A protease- CID 71276250 (blue) complex along the MD simulation at 300 K.

CID	Mutagenic	Tumorigenic	Irritant	Reproductive Effective	cLogP	Solubility	Drug Likeness	Drug Score
MK-5172 (46930991)	N	N	Y	N	3.91	-5.18	-29.04	0.13
59178018	N	N	Y	N	3.91	-5.18	-29.04	0.13
71278306	N	N	Y	N	3.91	-5.18	-29.04	0.13
71307589	N	N	Y	N	3.91	-5.18	-29.04	0.13
67304193	N	N	Y	N	4.18	-5.34	-29.09	0.12
67304412	N	N	Y	N	4.06	-5.47	-30.74	0.12
71276289	B	B	Y	B	4.41	-5.23	-10.36	0.06
68491802	N	N	Y	N	3.79	-5.31	-30.7	0.13
68491803	N	N	Y	N	3.79	-5.31	-30.7	0.13
59178019	N	N	Y	N	4.34	-5.31	-28.72	0.12
67304190	N	N	Y	N	4.34	-5.31	-28.72	0.12
71276290	N	N	Y	N	2.31	-4.19	-15.54	0.18
71278250	N	N	Y	N	4.83	-5.35	-23.24	0.1
72535811	N	N	Y	N	4.34	-5.31	-28.72	0.12
71278254	N	N	Y	Y	4.48	-5.8	-85.7	0.05
71278255	N	N	Y	N	5.39	-6.34	-88.03	0.04
71278248	N	N	Y	N	1.67	-5.32	-60.2	0.13
68236439	N	N	Y	N	4.16	-5.2	-33.67	0.12
59178016	N	N	Y	N	4.21	-5.44	-30.39	0.12
67304410	N	N	Y	N	4.21	-5.44	-30.39	0.12
71276250	N	N	N	N	4.37	-2.08	-20.19	0.27
72535809	N	N	Y	N	4.21	-5.44	-30.39	0.12
59178020	N	N	Y	N	5.07	-6	-32.16	0.09
67304070	N	N	Y	N	5.07	-6	-32.16	0.09
72535812	N	N	Y	N	5.07	-6	-32.16	0.09
71276269	B	B	Y	B	4.41	-5.23	-10.36	0.06
68236192	N	N	Y	N	4.59	-5.33	-33.35	0.11
71529660	N	N	Y	Y	5.39	-6.34	-88.03	0.04
71276247	N	N	Y	N	1.67	-5.32	-60.2	0.13
71532366	N	N	Y	N	1.67	-5.32	-60.2	0.13
68236191	N	N	Y	N	4.47	-5.47	-28.13	0.11
58428446	N	N	Y	N	4.09	-4.75	-9.5	0.14

Table 2: Toxicity analysis-OSIRIS data.

Acknowledgements

The authors express deep sense of gratitude to the management of Vellore Institute of Technology for all the support, assistance and constant encouragements to carry out this work.

References

- Avolio S, Robertson K, Hernando JI, DiMuzio J, Summa V (2009) Inhibitors of hepatitis C virus NS3/4A: α -Ketoamide based macrocyclic inhibitors. *Bioorganic & Medicinal Chemistry Letters* 19: 2295–2298.
- Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW, et al. (1989) Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* 244: 359–362.
- Houghton M, Weiner A, Han J, Kuo G, Choo QL (1991) Molecular biology of the hepatitis C viruses: implications for diagnosis, development and control of viral disease. *Hepatology* 14: 381–388.
- Chevaliez S, Pawlotsky JM (2007) Interferon-based therapy of hepatitis C. *Adv Drug Deliv Rev* 59: 1222–1241.
- Suzuki R, Sakamoto S, Tsutsumi T, Rikimaru A, Tanaka K, et al. (2005) Molecular determinants for subcellular localization of hepatitis C virus core protein. *J Virol* 79: 1271–1281.
- Bartenschlager R, Lohmann V, Wilkinson T, Koch JO (1995) Complex formation between the NS3 serine-type proteinase of the hepatitis C virus and NS4A and its importance for polyprotein maturation. *J Virol* 69: 7519–7528.
- Brass V, Berke JM, Montserret R, Blum HE, Penin F, et al. (2008) Structural determinants for membrane association and dynamic organization of the hepatitis C virus NS3-4A complex. *Proc Natl Acad Sci U S A* 105: 14545–14550.
- Schiering N, D'Arcy A, Villard F, Simic O, Kamke M, et al. (2011) A macrocyclic HCV NS3/4A protease inhibitor interacts with protease and helicase residues in the complex with its full-length target. *Proc Natl Acad Sci U S A* 108: 21052–21056.
- Love RA, Parge HE, Wickersham JA, Hostomsky Z, Habuka N, et al. (1996) The crystal structure of hepatitis C virus NS3 proteinase reveals a trypsin-like fold and a structural zinc binding site. *Cell* 87: 331–342.
- Melnikova I (2011) Hepatitis C--pipeline update. *Nat Rev Drug Discov* 10: 93–94.
- McInnes C (2007) Virtual screening strategies in drug discovery. *Curr Opin Chem Biol* 11: 494–502.
- Freceer V, Kabelác M, De Nardi P, Pricl S, Miertus S (2004) Structure-based design of inhibitors of NS3 serine protease of hepatitis C virus. *J Mol Graph Model* 22: 209–220.
- Ismail NS, El Dine RS, Hattori M, Takahashi K, Ihara M (2008) Computer based design, synthesis and biological evaluation of novel indole derivatives as HCV NS3-4A serine protease inhibitors. *Bioorg Med Chem* 16: 7877–7887.
- Duffy BC, Zhu L, Decornez H, Kitchen DB (2012) Early phase drug discovery: cheminformatics and computational techniques in identifying lead series. *Bioorg Med Chem* 20: 5324–5342.
- Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, et al. (2002) The Protein Data Bank. *Acta Crystallogr D Biol Crystallogr* 58: 899–907.
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, et al. (2000) The Protein Data Bank. *Nucleic Acids Res* 28: 235–242.
- Bolton E, Wang Y, Thiessen PA, Bryant SH (2008) PubChem: Integrated platform of small molecules and biological activities. *Annual Reports in Computational Chemistry* 4: 217–224.
- Gasteiger J, Rudolph C, Sadowski J (2009) Automatic generation of 3D-atomic

- coordinates for organic molecules. *Tetrahedron Computer Methodology* 3(6): 537-547.
19. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, et al. (2005) GROMACS: fast, flexible, and free. *J Comput Chem* 26: 1701-1718.
20. Shoichet BK (2004) Virtual screening of chemical libraries. *Nature* 432: 862-865.
21. Ewa B, Xavier L, Anna C, Joanna MK, Katarzyna HK, et al. (2011) Virtual screening strategies in drug design-methods and application. *Journal of Biotechnology, Computational Biology and Bionanotechnology* 92: 249-264.
22. Feldman HJ, Snyder KA, Ticol A, Pintilie G, Hogue CW (2006) A complete small molecule dataset from the protein data bank. *FEBS Lett* 580: 1649-1653.
23. Willet P, Barnard JM, Downs GM (1998) Chemical similarity searching. *J Chem Inf Comp Sci* 38: 983-996.
24. Ertl P, Rohde B, Selzer P (2000) Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *J Med Chem* 43: 3714-3717.
25. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* 46: 3-26.
26. Muegge I (2003) Selection criteria for drug-like compounds. *Med Res Rev* 23: 302-321.
27. Buntrock RE (2002) ChemOffice Ultra 7.0. *J Chem Inf Comput Sci* 42: 1505-1506.
28. Proudfoot JR (2002) Drugs, leads, and drug-likeness: an analysis of some recently launched drugs. *Bioorg Med Chem Lett* 12: 1647-1650.
29. Sander T. (2001) OSIRIS property explorer. Allschwil: Actelion Pharmaceuticals Ltd.
30. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ (2005) PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res* 33: W363-367.
31. Connolly ML (1983) Solvent-accessible surfaces of proteins and nucleic acids. *Science* 221: 709-713.
32. Zhang C, Vasmatazis G, Cornette JL, DeLisi C (1997) Determination of atomic desolvation energies from the structures of crystallized proteins. *J Mol Biol* 267: 707-726.
33. Andrusier N, Nussinov R, Wolfson HJ (2007) FireDock: fast interaction refinement in molecular docking. *Proteins* 69: 139-159.
34. Hess B, Kutzner C, Spoel D, Lindahl E (2008) GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput* 4: 435-447.
35. Meagher KL, Carlson HA (2005) Solvation influences flap collapse in HIV-1 protease. *Proteins* 58: 119-125.
36. Schüttelkopf AW, van Aalten DM (2004) PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr* 60: 1355-1363.
37. Darden T, Perera L, Li L, Pedersen L (1999) New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations. *Structure* 7: R55-60.
38. Van Gunsteren WF, Berendsen HJC (1977) Algorithms for macromolecular dynamics and constraint dynamics. *Mol Phys* 34: 1311-1327
39. Zhou S, Li Y, Hou T (2013) Feasibility of using molecular docking-based virtual screening for searching dual target kinase inhibitors. *J Chem Inf Model* 53: 982-996.
40. Sousa SF, Cerqueira NM, Fernandes PA, Ramos MJ (2010) Virtual screening in drug design and development. *Comb Chem High Throughput Screen* 13: 442-453.
41. Reddy AS, Pati SP, Kumar PP, Pradeep HN, Sastry GN (2007) Virtual screening in drug discovery—a computational perspective. *Curr Protein Pept Sci* 8: 329-351.
42. Karthick V, Ramanathan K, Shanthi V, Rajasekaran R (2013) Identification of potential inhibitors of H5N1 influenza A virus neuraminidase by ligand-based virtual screening approach. *Cell Biochem Biophys* 66: 657-669.
43. Milan Remko (2009) Theoretical study of molecular structure, pKa, lipophilicity, solubility, absorption, and polar surface area of some hypoglycemic agents. *Journal of Molecular Structure: THEOCHEM* 897: 73-82.
44. Renxiao Wang, Ying Fu, Luhua Lai (1997) A new atom-additive method for calculating partition coefficients. *Journal of Chemical Information and Computer Sciences* 37: 615-621.
45. Tetko IV (2005) Computing chemistry on the web. *Drug Discov Today* 10: 1497-1500.