Discovery of small molecule inhibitors against the NS3/4A serine protease of Hepatitis C virus genotype 3 via high-throughput virtual screening and *in vitro* evaluations

Sakhor, W.[1], Teoh, T.C.[2], Yusof, R.[1], Lim, S.K.[3] and Razif, M.F.M.[1*]

[1]Department of Molecular Medicine, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, Malaysia
[2]Institute of Biological Sciences, Faculty of Science, University of Malaya, 50603 Kuala Lumpur, Malaysia
[3]Department of Pharmaceutical Chemistry, Faculty of Pharmacy, University of Malaya, 50603 Kuala Lumpur, Malaysia
[*]Corresponding author e-mail: fazril.razif@um.edu.my

Abstract. The hepatitis C virus (HCV) consists of eight genotypes and 90 subtypes, with genotype (GT) 3 being the second most common globally and is linked to higher incidences of steatosis and rapid development of fibrosis and cirrhosis. The NS3/4A serine protease, a heterodimer complex of two HCV non-structural proteins, is an effective target for pharmaceutical intervention due to its essential roles in processing HCV polyproteins and inhibiting innate immunity. This study combines structure-based virtual screening (SBVS) of predefined compound libraries, pharmacokinetic prediction (ADME/T) and *in vitro* evaluation to identify potential low molecular weight (<500 Dalton) inhibitors of the NS3/4A serine protease (GT3). *In silico* screening of ZINC and PubChem libraries yielded five selected compounds as potential candidates. Dose-dependent inhibition of the NS3/4A serine protease and HCV replication in HuH-7.5 cells revealed that compound A (PubChem ID No. 16672637) exhibited inhibition towards HCV GT3 with an IC_{50} of 106.7µM and EC_{50} of 25.8µM, respectively. Thus, compound A may be developed as a potent, low molecular weight drug against the HCV NS3/4A serine protease of GT3.

INTRODUCTION

Hepatitis C is a global health concern that affects ~177.5 million people, representing nearly 2.5% of the world population (Petruziello *et al.*, 2016), and is one of the main causes of death and morbidity. The disease is caused by the hepatitis C virus (HCV), a blood borne virus that was first discovered in 1989 (Choo *et al.*, 1989). HCV is classified under the genus *Hepacivirus* and is a member of the *Flaviviridae* family (Simmonds *et al.*, 2017). Early stage diagnosis of the disease is difficult because acute infections are frequently asymptomatic. About ~75–85% of those with acute HCV infection will eventually develop a chronic infection, leading to fibrosis, cirrhosis and hepatocellular carcinoma (Pinzani *et al.*, 2005; Hajarizadeh *et al.*, 2013). HCV displays a high degree of genetic heterogeneity due to the error-prone nature of RNA-dependent RNA polymerase. At present, there are eight confirmed genotypes (GT) and 90 subtypes of HCV (Smith *et al.*, 2019). Amongst all genotypes, GT3 infection is relatively difficult to treat, especially in patients with prior HCV treatment failure or cirrhosis. GT3 infection is associated with higher incidences of steatosis (Hwang & Lee, 2011; Roingeard, 2013), rapid development of fibrosis and cirrhosis, and higher rates of hepatocellular carcinoma (Kanwal *et al.*, 2014; Chan *et al.*, 2017).

The HCV RNA genome encodes for a long polyprotein precursor which consists of four structural proteins (C, E1, E2 and p7)

and six non-structural proteins (NS2, NS3, NS4A, NS4B, NS5A and NS5B). This large polyprotein is cleaved into 10 individual proteins by the action of both cellular and viral proteases (Kim *et al.*, 1996). The NS3/4A serine protease cleaves at four junctions, between NS3/NS4A (self-cleavage), NS4A/NS4B, NS4B/NS5A and NS5A/NS5B (Baetenschlager & Lohmann, 2000) which is critical to the HCV replication system as viral replication can only begin after all individual proteins have been cleaved. The NS3/4A serine protease is also involved in the inhibition of innate immunity. Cleavage of the TIR-domain-containing adapter-inducing IFN-β and mitochondrial antiviral signalling protein prevents the activation of interferon type I genes and IFN-stimulated genes (Dustin & Rice, 2007) while cleavage of T-cell protein tyrosine phosphatase interferes with the growth signaling/Akt pathway Brenndörfer *et al.*, 2009). Its significant roles in viral replication and pathogenesis of HCV make the NS3/4A serine protease a fascinating target for therapeutic development.

Unlike both hepatitis A (HAV) and hepatitis B (HBV) virus, there are currently no vaccines against HCV. In May 2011, the United States Food and Drug Administration (U.S. FDA) approved the first generation of direct-acting antivirals (DAAs), Boceprevir (Chen & Njoroge, 2010) and Telaprevir (Kwong *et al.*, 2011), to treat HCV. Since then, numerous highly potent DAAs targeting the NS3/4A protease (Rosenquist *et al.*, 2014; Ng *et al.*, 2018), NS5A (Belema *et al.*, 2014; Yu *et al.*, 2018) and NS5B polymerase (Sofia, 2016; ElKassas *et al.*, 2017) have been developed. Yet, some are only effective against certain genotypes. At present, first-generation DAAs are no longer recommended due to their adverse side effects and reduced efficacy (World Health Organization, 2017). Earlier standard of care for GT3-infected patients was achieved via the combination therapy of Sofosbuvir-Ribavirin; however, the sustained virologic response is low compared to the other genotypes (Jacobson *et al.*, 2013; Lawitz *et al.*, 2013). Current treatment options recommended for GT3 infections, which consists of a combination of dual and triple DAAs (AASLD-IDSA, 2018; Pawlotsky *et al.*, 2018), have shown to provide greater efficacy, even in cirrhotic and treatment-experienced patients (Foster *et al.*, 2015; Wyles *et al.*, 2018). However, (i) the high costs associated with these newer therapies, (ii) the association of GT3 with greater liver complications, and (iii) the presence of resistance-associated substitutions that affect tolerances to DAAs (Vermehren & Sarrazin, 2012; Poveda *et al.*, 2014; Lontok *et al.*, 2015; Raj *et al.*, 2017) emphasizes the need for continuous DAA development.

Numerous computer-aided approaches have been employed to discover new potent HCV inhibitors (Li *et al.*, 2013; Kumar *et al.*, 2014; Basetto *et al.*, 2016; Ganesan & Barakat, 2017) by combing through large virtual compound databases, followed by *in vitro* and *in vivo* evaluations. Therefore, this study employs structure-based virtual screening (SBVS) of current predefined compound libraries to identify potential potent, low molecular weight inhibitors targeting the NS3/4A serine protease of wild type HCV GT3, followed by *in vitro* investigations into the potency of the selected compounds.

MATERIALS AND METHODS

Protein 3D structure modeling of GT3 NS3/4A serine protease
HCV GT3 NS3/4A serine protease protein 3D structure is needed for high-throughput virtual screening. Therefore, homology modeling by YASARA (www.yasara.org) was used to construct a model of the NS3/4A serine protease based on CASP-approved protocol. The HCV GT3 protein sequence was retrieved from UniProt (Id: I1Y879) and the templates for homology modeling were identified from the PDB database (PDB ID: 4A1X, 308B, 4A92, 3M5M and 4I33). The models generated underwent structure validation based on Z- scores calculated from molecular dynamic force field energies. The stereochemical qualities of selected model of NS3/4A serine protease (4A1X-B02) was further validated by generating Ramachandran plot using Procheck:

http://servicesn.mbi.ucla.edu/PROCHECK/ (Laskowski *et al.*, 1993). Compatibility of the 3D atomic models to the NS3/4A amino acid sequence was confirmed via Verify3D: http://servicesn.mbi.ucla.edu/Verify3D/ (Liithy *et al.*, 1992). The verified homology model was used as the template for virtual screening.

Virtual screening
Virtual screening was performed using Mobyle@RPBS and Idock webservers which provides a large database of compounds from PubChem and ZINC (Irwin *et al.*, 2012) respectively. Mobyle@RPBS provides virtual screening services under the MTiOpenScreen program that operates with Autodock Vina and Vina empirical scoring function (Alland *et al.*, 2005; Néron *et al.*, 2009). Idock is a tool for structure-based virtual screening for flexible ligand docking that is motivated from Autodock Vina (Li *et al.*, 2012). MTiOpenScreen (Labbé *et al.*, 2015) by Mobyle@RPBS provides predefined libraries curated from PubChem: i) 99 288 diverse chemical compound collections and ii) 51 232 focused chemical compound collections that target protein-protein interactions. In addition, there were a total of 23 129 083 compounds from ZINC available via Idock. All compounds were filtered according to conditions outlined by Lipinski's rule of 5 to obtain a lead-like compound.: no more than 5 hydrogen bond donors, no more than 10 hydrogen bond acceptors, molecular weight less than 500 Daltons and log P not greater than 5 (Lipinski *et al.*, 1997; 2001). Overall, there were ~90 000 compounds screened from the two webservers. Virtual screening with Mobyle@RPBS and Idock was performed by targeting the search space calculated from protein residues on the active site; particularly the catalytic triad (His57, Asp81, and Ser139) of the GT3 NS3/4A serine protease.

ADME/T prediction
Absorption, distribution, metabolism, excretion and toxicity (ADME/T) properties of the 300 selected compounds were predicted using Discovery Studio 4.5. A total of 217 compounds were removed from further analysis because they possessed no interactions with the wild type GT3 NS3/4A serine protease catalytic triad (His57, Asp81, and Ser139) and/or having unfavorable ADME/T properties. Evaluation of the toxicity potential for the remaining compounds were conducted based on the TOxicity Prediction by Komputer Assisted Technology (TOPKAT; Accelrys).

Redocking
83 shortlisted compounds were redocked locally using Autodock Vina (Trott & Olson, 2010) to observe its reproducibility. Simeprevir, a DAA used for the treatment of HCV GT1, was used as the standard compound in virtual screening and redocking. The crystal structure was obtained from RCSB Protein Data Bank (PDB) database: 3KEE and extracted out from protein-ligand complex with Discovery Studio 4.5. The NS3/4A protein and selected ligands structure were prepared and saved in pdbqt format using AutodockTools 1.5.6 software. Configuration file named "conf.txt" were prepared; the input includes the receptor and ligand file in pdbqt format, search space area to be docked by ligands in x, y, z dimensions; 22Å x 20Å x 20Å with x, y and z coordinates of the center 57.002, 74.518, -48.664 respectively. The exhaustiveness was set at 100 and the output file was named as "out.pdbqt". Docking was run following manual in http://vina.scripps.edu/tutorial.html

Redocking analysis
The outputs of docking by Autodock Vina were analysed by PyMOL (The PyMOL Molecular Graphics System, Version 1.3 Schrodinger, LLC). 2D interaction diagrams were examined via Discovery Studio 4.5 (Dassault Systèmes BIOVIA). Each compound was ranked based on its free binding energies as compared to Simeprevir, binding interactions with residue in the GT3 NS3/4A catalytic triad (His57, Asp81, Ser 139) and the Oxyanion hole (Gly137). Docking simulations had produced nine binding conformations by default for each compound; these binding conformations were scored based on their binding frequency to the

catalytic triad. The output lists 18 compounds that presented interactions with any of the three amino acids in the catalytic triad with binding conformations score ≥4 (Supplementary Material; Table S1). Five compounds named A, B, C, D and E (Table 2) were selected and purchased from Asinex, USA (A), ChemBridge, USA (B, E), Enamine, USA (C) and KeyOrganics, UK (D); with ≥90% purity to investigate their inhibitory potential against the NS3/4A serine protease of GT3 *in vitro*.

Recombinant GT3 NS3/4A protein expression and purification

A fusion gene containing a NS4A fragment linked to a NS3 serine protease domain by a GSGS-linker was synthesized in a pUC57 plasmid (Nanogene Solutions) with flanking *Bam*HI and *Hind*III restriction sites for cloning. The NS4A-GSGS-NS3 gene was subcloned into the pQE30 plasmid (Qiagen) and expressed as a fusion with a hexa-histidine tag in *Escherichia coli* Rosetta cells. Recombinant protein was expressed by the addition of 0.5mM IPTG followed by 5h incubation at room temperature. Cells were lysed by sonication in 50mM Hepes, 300mM NaCl, lysozyme, 10% glycerol, (pH 7.8). Lysates were centrifuged at 8 800 x g for 40 min at 4°C to obtain cell-free supernatant. Soluble protein fractions were incubated with pre-equilibrated Ni^{2+}-NTA (Gold Bio) resin for 2h at 4°C on a rocking platform. The Ni^{2+}-NTA resin was washed twice with 50mM Hepes, 300mM NaCl, 30mM imidazole (pH 7.8) and the NS3/4A protein was eluted with 50mM Hepes, 300mM NaCl, 150mM imidazole (pH 7.8). Eluted protein was determined by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and verified via Western blot analysis. The NS3/4A protein was stored in 50% glycerol at -80°C.

Protease activity assay

All five compounds were dissolved in 100% DMSO (10mM) and diluted with assay buffer (50mM Hepes, 300mM NaCl, 20% glycerol, 10mM DTT). Varying concentrations (triplicates) of each compound were incubated with 20nM of purified GT3 NS3/4A protein in a black 96-well plate for 10 min at 30°C with shaking. The reactions were initiated by addition of 5µM substrate (Ac-DE-D(EDANS)-EE-Abu-ψ-[COO]-AS-K(DABCYL)-NH2). Fluorescence signal was monitored for 30 min at 340/490 nm using the Infinite 200 microplate reader (Tecan). GraphPad Prism (GraphPad Software, Inc.) was used to determine the IC_{50} value of each compound.

Cell culture

HuH-7.5 cells were cultured in Dulbecco's Modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (Gibco). Transfected HuH-7.5 cells stably expressing S52/SG-Feo (SHI) were maintained in DMEM medium with 10% fetal bovine serum and 250µg/ml G418. Cells were incubated at 37°C in humidified atmosphere with 5% CO_2. Huh-7.5 cells were confirmed to be absent of mycoplasma contamination.

Cytotoxicity assay

Toxicity of each compound against Huh-7.5 cells was measured using Cell titer 96^R Aqueous One Solution Cell Proliferation Assay (Promega). 1 x 10^4 cells were seeded per well in a 96-well plate and incubated at 37°C in humidified atmosphere with 5% CO_2 overnight. Increasing concentrations of each compound (0µM, 6.25µM, 12.5µM, 25µM, 50µM, 100µM, 200µM) was added to triplicate wells and incubated for 72h. Cell viability was measured using Epoch Microplate Spectrophotometer (Biotek) at 490nm.

In vitro transcription and transfection

Neomycin-selectable subgenomic replicon of GT3 with firefly luciferase reporter gene in its construct, S52/SG-Feo (SHI) plasmid was transformed into *E. coli* DH5α strain. The plasmid DNA was isolated and used as template to synthesize RNA via *in vitro* transcription using RiboMAX™ Large Scale RNA Production System-T7 (Promega). Prior to transcription, the plasmid was linearized with FastDigest XbaI (Thermo Fisher) and purified with PCR Purification Mini kit (Favorgen) according to manufacturer's instructions. RNA concentration and purity

were determined using Thermo Scientific NanoDrop™ 2000/2000c Spectrophotometer. S52/SG-Feo (SHI) RNA was transfected into HuH-7.5 cells using Lipofectamine 3000 reagent (Thermo Fisher). Selection of colonies was achieved with 600 µg/ml G418 for two weeks. Subsequently, cells were maintained with 250 µg/ml G418.

Luciferase-based replicon assay

HuH-7.5 cells stably expressing S52/SG-Feo (SHI) were seeded at 1×10^4 cells per well in a white 96-well plate overnight. Compounds A and D were diluted in culture medium at varying concentrations and maintained in 1% DMSO. (Compound A: 5µM, 10µM, 15µM, 20µM, 30µM, 40µM; Compound D: 10µM, 50µM, 100µM, 150µM, 200µM). After 72h, viral RNA replication was evaluated using ONE-Glo™ EX Luciferase Assay System (Promega). Measurements were taken using the Infinite 200 microplate reader (Tecan).

Statistical analysis

Assays were conducted in triplicates in the presence of a vehicle control. GraphPad Prism software (GraphPad Software, Inc.) was used to determine the IC_{50} in the protease activity assay, CC_{50} values in the cytotoxicity assay and EC_{50} value in the luciferase assay as well as determining means, standard deviations and standard error. The statistical analysis of one-way ANOVA was conducted, results obtain has significant difference at $P < 0.05$.

RESULTS

The 3D structure of wild type GT3 NS3/4A serine protease was predicted via homology modeling; Figure 1a and Figure 1b. A position-specific scoring matrix (PSSM) was built by running PSI-BLAST of NS3/4A target sequences against UniProt. Potential templates were identified in the PDB database based on the PSSM profile. The templates were ranked based on alignment scores and structural quality. Out of the 46 hits found, five templates with the highest scores were selected namely PDB: 4A1X, 308B, 4A92, 3M5M and 4I33. These 5 templates are from HCV G1a and G1b. Nine protein models were generated from the five selected templates. Loops were modeled and the side chains built were optimized and fine-tuned. Combined steepest descent and simulated annealing minimization was applied to the modeled parts. All models were validated based on Z-scores of molecular dynamics force fields energies (Table 1). The model with the best overall quality Z-score, 4A1X-B02 was selected for further study.

The quality of the 4A1X-B02 model was further verified by generating a Ramachandran plot to assess the stereochemical qualities of the protein; Figure 1c. Based on the plot, this model was considered to possess acceptable stereochemical qualities as 92.5% of the residues were located in the most favored regions, with no residues detected in the disallowed regions. The structure was also confirmed via Verify3D to assess the compatibility of the 3D atomic model to its amino acid sequence. The results revealed that 90.31% of the residues had an average 3D-1D score of ≥ 0.2 in the 3D/1D profile.

A multi-step approach was employed in search for potential small molecule inhibitors against the NS3/4A serine protease of HCV GT3; Figure 1d. Two large databases of compounds currently accessible via ZINC and PubChem were explored. SBVS of compounds was accomplished using Mobyle@RPBS and Idock webservers, followed by prediction into their ADME/T properties and redocking via Autodock Vina. Following redocking, 18 compounds were shortlisted. Potential compounds were selected with conditions as follows: (i) molecular weight ≤ 500 Dalton, (ii) comparable binding affinity as the standard, Simeprevir (iii) interactions with at least one residue in the catalytic triad, (iv) binding conformations score ≥ 4 (v) favorable ADME/T properties. Five purchasable compounds were selected for further evaluation *in vitro*. The IUPAC name, identity number, binding affinities, conformation scores and molecular weight of these five compounds are summarized in Table 2. The 5 compounds are labeled as compound A, B, C, D and E.
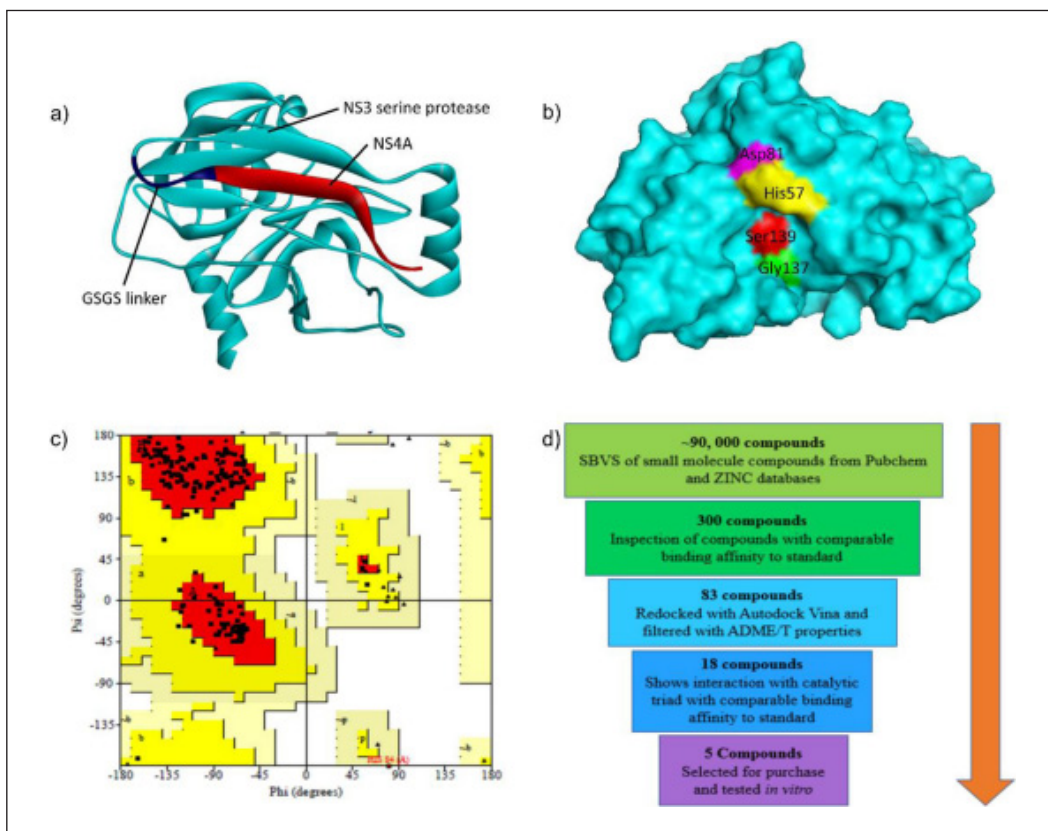
Figure 1. 3D structural modelling of the HCV (GT3) NS3/4A serine protease and inhibitor screening approach.
(a) Structural motif of the NS3/4A serine protease in solid ribbon (Cyan: NS3 serine protease; Blue: GSGS linker; Red: NS4A cofactor). (b) Surface display of the NS3/4A serine protease. The binding pocket of its active site is highlighted in yellow (His57), red (Ser139) and magenta (Asp81). The oxyanion hole is highlighted in green (Gly137; Cyan: NS3 serine protease). (c) Ramachandran plot of model 4A1X-B02 generated from PROCHECK. Red represents the most favoured regions; yellow represents additional allowed regions; cream represents generously allowed regions; white represents disallowed regions. A good quality model is expected to possess ≥90% amino acid residues in the most favoured regions. In this model, 92.5% of residues are located within the most favoured regions. (d) Multi-step approach employed *in silico* in search of potential small molecule inhibitors against the NS3/4A serine protease.

Table 1. Model ranking of the initial homology models generated from YASARA.
The models are ranked based on Z- scores of structural validations calculated from molecular dynamics force field energies. The model with the best overall quality Z-score; 4A1X-B02 was accepted as the final model.

Rank	Model ID	Dihedral Z-score	1D Z-score	3D Z-score	Overall quality Z-score	Comment
1	**4A1X-B02**	**2.137**	**-0.714**	**-1.281**	**-0.564**	**Good**
2	3O8B-A02	1.821	-0.745	-1.359	-0.658	Good
3	3O8B-A01	1.770	-0.830	-1.345	-0.692	Good
4	4A1X-B01	1.849	-0.972	-1.470	-0.795	Good
5	4A1X-B03	1.652	-1.019	-1.520	-0.865	Good
6	3M5M-B	1.997	-1.145	-1.610	-0.906	Good
7	4I33-A	1.465	-1.516	-1.230	-0.951	Good
8	4A92-~02	1.364	-1.060	-1.862	-1.081	Satisfactory
9	4A92-~01	1.073	-1.388	-2.124	-1.373	Satisfactory

From docking simulations on Autodock Vina, compound A possessed the same binding affinity (-7.8 Kcal/mol) as Simeprevir. The other four compounds exhibited comparable binding affinities (Table 2). These simulations also revealed that all five compounds successfully bound to the shallow groove within the active site of the GT3 NS3/4A serine protease. Binding conformations of Compound A and D scored 8/9 and 9/9, indicating that most of the conformations successfully docked inside the active site and interacted with key residues of the catalytic triad, namely Ser139 and His57 via hydrogen bonding and van der Waals interactions, respectively. Both compounds also possessed interactions with Gly137 of the oxyanion hole, and residues of the hydrophobic pockets. The structure and binding dynamics of each compound are shown in Figure 2.

All selected compounds obey Lipinski's rule of 5; however the selected standard, Simeprevir has two violations (i) molecular weight higher than 500 Dalton, and (ii) possess more than 10 hydrogen bond acceptors. ADME/T screening via Discovery Studio 4.5 predicted that only compound B has good solubility whereas the other compounds, including Simeprevir, have poor solubility. The oral bioavailability and intestinal absorption of all five compounds are predicted to be very high. Only compound D is predicted to inhibit the function of cytochrome P450 2D6. Evaluation of their toxicity potential using TOPKAT predicted that all compounds, including Simeprevir, are non-mutagenic and non-toxic.

Purified GT3 NS3/4A serine protease with a molecular mass of ~23 Kda (Du *et al.*, 2002) was obtained via overexpression in *E. coli* Rosetta and confirmed by SDS-PAGE and

Table 2. IUPAC name, binding affinity and ADME/T properties of five shortlisted compounds, with Simeprevir as a standard

Compound Identifier/IUPAC Name	Identity	Binding Affinity[a] (Kcal/mol)	Conformation Score[b]	Molecular Weight (Dalton)
A [(3S)-1-[5-(1,3-benzodioxol-5-yl)-1,2-oxazole-3-carbonyl] piperidin-3-yl] -(2-fluorophenyl) methanone	Pubchem 16672637	-7.8	8/9	422.41
B (3aR,4R)-3-cyclobutyl-4-(2-quinolyl)-3a,4,5,7-tetrahydropyrazolo [3,4-b] pyridin -6-one	ZINC 65463613	-7.6	6/9	318.37
C (4S)-2-oxo-N-[6-(trifluoromethyl)-2-pyridyl]-3,4-dihydro-1H-quinoline-4-carboxamide	ZINC 75126755	-7.6	7/9	335.28
D 2-[2-[2-(4-fluorophenyl)-1,3-thiazol-4-yl] ethyl] isoindole-1,3-dione	Pubchem 1473490	-7.4	9/9	352.38
E [(2R)-2-hydroxyspiro[indane-1,4'-piperidine]-1'-yl]-[3-(m-tolyl)-1H-pyrazol-4-yl] methanone	ZINC 72124233	-7.2	4/9	387.47
Simeprevir		-7.8	8/9	749.94

[a] Binding Affinity of best conformations predicted by Autodock Vina (Kcal/mol).
[b] Number of conformations that interact with amino acid residue of catalytic triad.
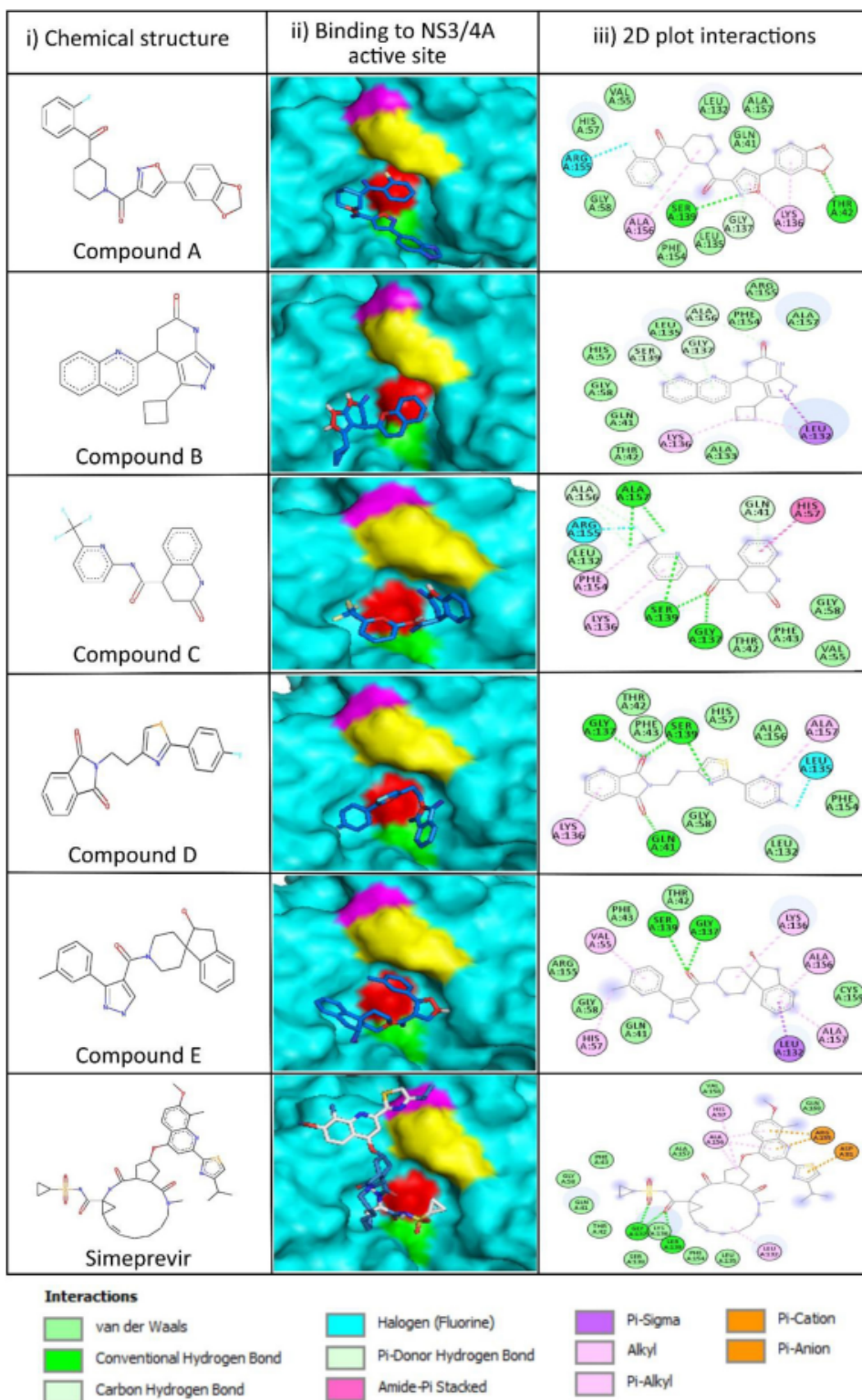
Figure 2. Chemical structures and binding dynamics of compound A, B, C, D, E and Simeprevir.
(i) Chemical structures of compound A, B, C, D, E and Simeprevir. (ii) Binding of each compound to the
NS3/4A active site; Red: Ser139; Green: Gly137; Yellow: His57; Magenta: Asp81; Cyan: NS3/4A. (iii) 2D
plot showing interactions of each compound with amino acids in the vicinity of the NS3/4A active site.

Western blot (Supplementary Material; Figure S1). A FRET-based assay was employed to evaluate the inhibitory activities of each compound against the NS3/4A serine protease. Initial screenings of all compounds at 100µM demonstrated that compound A (40.8%) and D (48.9%) exhibited the greatest inhibition against the NS3/4A serine protease (Supplementary Material; Figure. S2). On the contrary, compound B appears to enhance protease activity by 24.8%. Both compound A and D were subsequently selected for further analysis. Results confirmed that both compounds A and D exhibited dose-dependent inhibition of the GT3 NS3/4A serine protease, each possessing an IC_{50} of 106.7µM and 86µM, respectively; Figure 3a and 3b.

Next, cell-based assays were performed to determine the cytotoxicity and inhibitory potential of both compounds on HCV replication. The 50% cytotoxicity concentration (CC_{50}) of compounds A and D, as determined 72h after exposure in Huh-7.5 cells, is 139.5µM and >200µM, respectively; Figure 3c and 3d. Following this, Huh-7.5 cells stably transfected with S52/SG-Feo (SHI), a bicistronic subgenomic replicon for HCV GT3a with firefly luciferase reporter gene in its construct, was used to determine viral RNA levels in cells post-treatment (Saeed et al., 2012). Increasing concentrations of compounds A and D were added to the transfected cells and incubated for 72 h. Compound A successfully inhibited 81% of HCV replication at 40µM, with an 50% effective concentration (EC_{50}) of 25.8µM; Figure 3e. Compound D was only able to inhibit 45.6% of HCV replication at 200µM; Figure 3f. The EC_{50} of compound D could not be determined due to its limited solubility. Overall, compound A was able to inhibit HCV replication in Huh-7.5-S52/SG-Feo (SHI) with a selectivity index (CC_{50}/EC_{50}) of 5.40.

DISCUSSIONS

This study combines *in silico* and *in vitro* methods in search for potential small molecule inhibitor candidates for wild type HCV GT3 NS3/4A serine protease. *In silico* virtual screening of compounds require protein 3D structure of GT3 NS3/4A. There are currently 123 crystal structures of HCV NS3/4 serine proteases deposited in the Protein Data Bank (PDB). However, only one, a GT1a3a chimera created by substituting three active site polymorphisms (R123T, D168Q and I132L) into GT1a NS3/4A, has been determined for GT3. No other structures of non-GT1 NS3/4A have been determined. Among the reasons is difficulty in performing crystallography due to complications in large-scale expression and purification (Soumana et al., 2016). In our study, homology modeling was able to predict the 3D structure of HCV GT3 NS3/4A serine protease, based on the actual GT3 amino acid sequence, via a protocol that adopts critical assessment of structure prediction (CASP). Generating Ramachandran plots and Verify 3D further aided in assessing the quality of the 3D structures (Lüthy et al., 1992; Carugo et al., 2013). Based on the plot displaying dihedral angles psi, ψ against phi, φ of amino acid residues, the 3D protein structure of model 4A1X-B02 possessed adequate stereochemical qualities with > 90% residues located in the most favored regions and no residues detected in the disallowed regions. Residues that deviate far from the allowed regions imply there are some error with the structure (Carugo & Djinovic-Carugo, 2013). Nonetheless, there are studies that report estimation of 0.4% residues located in the disallowed regions (Pal & Chakrabarti, 2002). Our Verify 3D result showed high-profile scores of compatibilities between the GT3 NS3/4A serine protease 3D structure and its amino acid sequence. At least 80% of the amino acids scoring > 0.2 in the 3D/1D profile is required for an estimation of accurate protein structure matched to its amino acid sequence.

The catalytic site of the GT3 NS3/4A serine protease was chosen as a target in virtual screening to inhibit viral replication. The catalytic site comprises of three amino acid residues, His57, Asp81, and Ser139, also known as the catalytic triad; Figure 1(b) (Kim et al., 1996; Love et al., 1996; Lin, 2006). They function by performing catalytic acid-base reactions on its target peptides (Hedstrom,
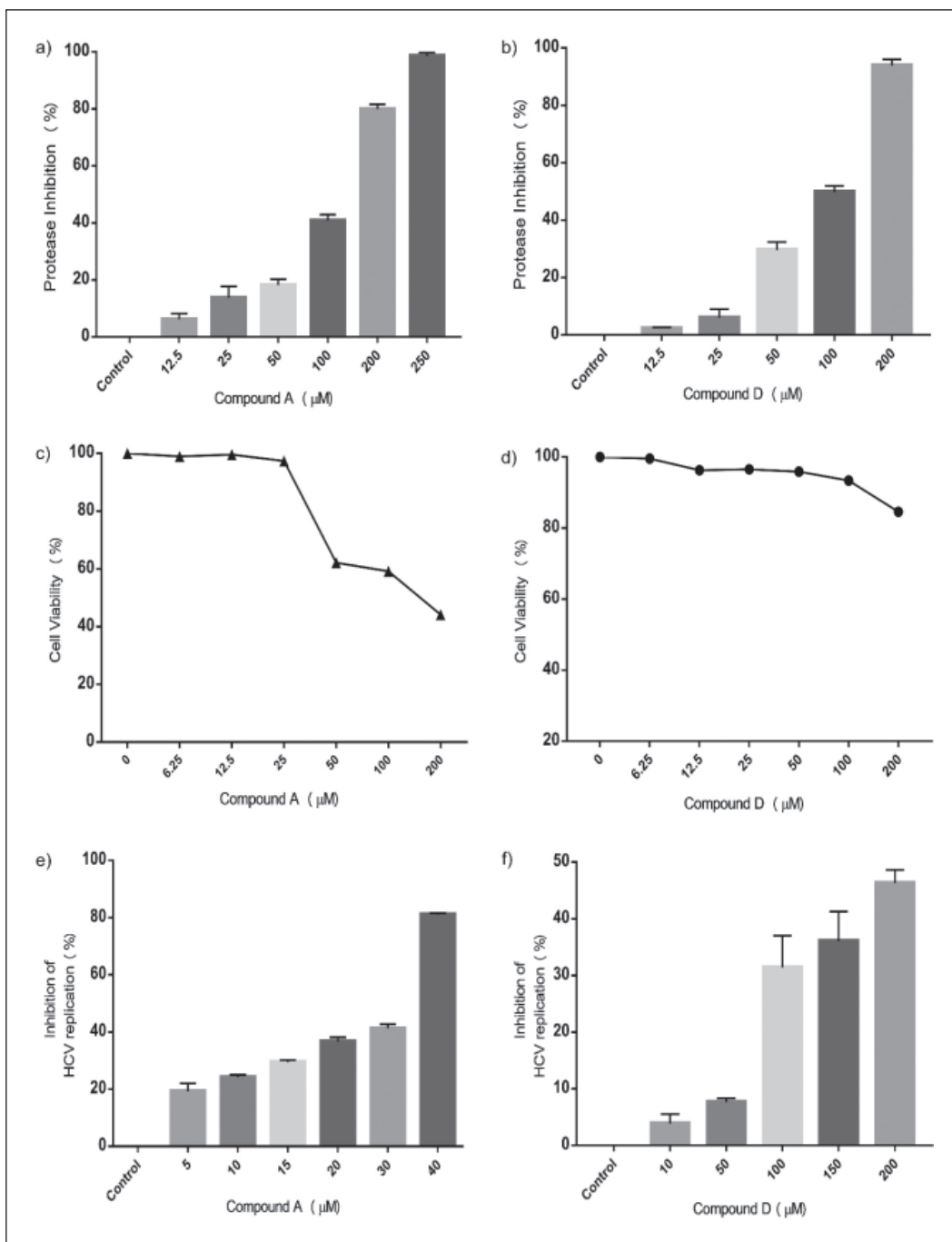
Figure 3. Inhibitory potential of compounds A and D via *in vitro* assays.
(a) Dose-dependent inhibition (%) of NS3/4A serine protease by compound A. (b) Dose-dependent inhibition (%) of NS3/4A serine protease by compound D. (c) Cell viability (%) of HuH-7.5 cells upon treatment with compound A. (d) Cell viability (%) of HuH-7.5 cells upon treatment with compound D. (e) Dose-dependent inhibition (%) of HCV replication by compound A in HuH-7.5 cells expressing S52/SG-Feo (SHI). (f) Dose-dependent inhibition (%) of HCV replication by compound D in HuH-7.5 cells expressing S52/SG-Feo (SHI). Bars represent the standard deviation of mean. Significant difference at P < 0.05.

2002). Studies have shown that site-directed mutagenesis of amino acids in the catalytic triad promotes the abolishment of proteolytic cleavage at the four junctions, which are NS3/NS4A, NS4A/NS4B, NS4B/NS5A and NS5A/NS5B (Bartenschlager *et al.*, 1993; Grakoui *et al.*, 1993). This highlights the importance of the catalytic triad in the viral replication process. X-ray crystallography has revealed that the catalytic site possesses a shallow substrate binding groove (Kim *et al.*, 1996), foreseeing difficulty in uncovering suitable small molecule inhibitors (Lin, 2006). Virtual screening and redocking reveal that the 5 selected compounds are able to dock to the catalytic site groove and interact with its residues with compound A possessing excellent binding affinity similar to Simeprevir (-7.8Kcal/mol).

Screening of compounds using Lipinski's rule of 5 and ADME/T properties was performed in order to obtain compounds with drug-like properties. There have been many prospective drug candidates that have proven ineffective in later stages of drug discovery and clinical studies due to unfavorable pharmacokinetic properties (Gombar *et al.*, 2003; Wang & Urban, 2004; Mandlik *et al.*, 2015).

In spite of the numerous benefits associated with structure-based virtual screening, there are a few drawbacks that may arise such as false-positive results. Kim *et al.* (2018) suggested that these may occur as a result of conformational changes of the receptor through ligand binding and also via docking algorithm preferences in selecting true-positives, which may require evaluation on a case-by-case basis. One way to prevent false positives is to conduct *in vitro* experiments, as both docking and ADME/T predictions do not warrant efficiency of the compounds.

In vitro evaluation via protease and luciferase cell-based assay results in IC_{50} of 106.7µM and EC_{50} of 25.8µM. Previous bioassay data on Simeprevir revealed that it is a very potent inhibitor of HCV, especially GT1. The median inhibition constant (IC_{50}) of Simeprevir against GT1a and GT1b is 0.5nM and 0.4nM, respectively. In comparison, this is markedly reduced against GT3, 37 nM (Lin *et al.*, 2009; Tanwar *et al.*, 2012). The EC_{50} of Simeprevir for GT1a and GT1b ranges between 8.1 nM to 28.4 nM (Lin *et al.*, 2009). However, the inhibition is reduced significantly by ~1014 folds in GT3 (Lenz *et al.*, 2013).

Combining *in silico* and *in vitro* methods narrowed down compound A as a potential compound to be developed as an HCV GT3 NS3/4A serine protease inhibitor. Compound A is a chiral compound, docked as a single enantiomer and tested *in vitro* as an enantiopure compound. Analysis of its binding interactions with GT3 NS3/4A serine protease revealed that the compound is theoretically able to dock inside the shallow groove of the active site (Figure 2) with a binding conformation score of 8/9. Compound A is predicted to inhibit catalysis by interacting with residues of the catalytic triad and those around the active site. The 2D diagram of the compound-protein interaction (Figure 2) illustrates that compound A forms hydrogen bonds with Ser139 through its phenyl group and oxazole and van der Waals interaction with His57. Ser139 is known to play key roles in NS3/4A serine protease catalysis, by attacking the carbonyl carbon of the substrate peptide's scissile bond assisted by His57 and Asp81 (Hedstrom, 2002; Raney *et al.*, 2010). Compound A also interacts with two amino acids in the oxyanion hole; forming hydrogen bond with Gly137 through the oxazole and van der Waals interaction with Ser139. The oxyanion hole; Gly137 and Ser139 assist in stabilizing negative charges formed on the peptide carbonyl oxygen during the catalytic process making it one of the important residues in serine protease mechanism (Raney *et al.*, 2010). Additionally, compound A interacts with four residues (Leu135, Phe154, Ala156, Ala157) located in the hydrophobic pockets of the active site. Similar patterns of interactions are formed from docking between Simeprevir to the NS3/4A serine protease of GT3; 4A1X-B02 and from crystal structure of Simeprevir to the NS3/4A serine protease of GT1 (Cummings *et al.*, 2010), through interactions with Ser139, His57 and Gly137, with more extensive interactions observed as it is a large macrocyclic compound (Supple-

mentary Material; Figure S3). The loss of potency of Simeprevir towards GT3 compared to GT1 can be attributed to the loss of hydrogen bond interactions in the active site and the presence of active site polymorphism D168Q (Soumana *et al.*, 2016). However, our analysis shows that binding of compound A appears to avoid any interactions with D168Q.

Thorough literature review of compound A has revealed only one previous bioassay data, as a modulator of the EP2 prostaglandin E2 receptor (NCBI, 2008). There are no reports on compound A against HCV, especially the NS3/4A serine protease for GT3. The oxazole structure in compound A is interesting as it is able to form hydrogen bonds with two important residues; Ser139 and Gly137. Oxazole-based derivative compounds, aryloxazole and benzoxazole, have been used to develop inhibitors against HCV replication by targeting viral entry and NS5B polymerase, respectively (Ismail *et al.*, 2013; He *et al.*, 2017). There has also been development of many medicinal drugs that use oxazole and its derivatives such as anti-fungal (Wani *et al.*, 2015), anti-bacterial (Prakash *et al.*, 2014), anti-cancer (Choi *et al.*, 2013), and anti-inflammatory (Pedada *et al.*, 2016).

We acknowledge the limitations of our study, which includes (i) our virtual screening did not include any known NS3/4A mutants of HCV GT3, (ii) no structural modifications were performed on the current set of compounds, and (iii) the viral specificity of Compound A was not determined as no further testing was conducted on other HCV genotypes and human serine proteases. However, given that the development of antivirals specifically targeting GT3 NS3/4A is limited, we believe that there is a potential for additional structural optimizations of compound A to further improve its inhibition of HCV GT3 replication. It is also of interest to determine if compound A possesses the ability to inhibit replication of other HCV genotypes given the conserved nature of the NS3/4A serine protease catalytic triad (Naeem & Waheed, 2017).

CONCLUSION

This study concludes compound A is a novel compound that has the potential to be developed as a small molecule inhibitor against the NS3/4A serine protease of HCV GT3 based on structure-based virtual screening methods in combination with protease and cell-based assays. Our data suggests that compound A inhibits viral replication via interactions with Ser139 and His57 in the catalytic site of the NS3/4A serine protease and Gly137 in the oxyanion hole. The low molecular weight and high bioavailability nature of this compounds makes it a favorable drug candidate. Modifications and structural changes may be able to increase its inhibitory potency. Overall, *in silico* screenings prior to biological investigations is a useful method for the rapid prediction, identification and selection of candidate compounds that can interact explicitly with the desired molecular target.

Supplementary Materials
All supplementary materials can be obtained from the corresponding author, upon request.

Conflict of interest
The authors report no conflicts of interest.

REFERENCES

AASLD-IDSA HCV Guidance Panel (AASLD-IDSA). (2018). Hepatitis C Guidance 2018 Update: AASLD-IDSA Recommendations for Testing, Managing, and Treating Hepatitis C Virus Infection. *Clinical Infectious Diseases* **67**: 1477-1492.

Alland, C., Moreews, F., Boens, D., Carpentier, M., Chiusa, S., Lonquety, M., Renault, N., Wong, Y., Cantalloube, H., Chomilier, J., Hochez, J., Pothier, J., Villoutreix, B.O., Zagury, J.F. & Tuffery, P. (2005). RPBS: a web resource for structural bio-informatics. *Nucleic Acids Research* **33**(suppl_2): W44-49.

Bartenschlager, R. & Lohmann, V. (2000). Replication of hepatitis C virus. *Journal of General Virology* **81**(7): 1631-1648.

Bartenschlager, R., Ahlborn-Laake, L., Mous, J. & Jacobsen H. (1993). Nonstructural protein 3 of the hepatitis C virus encodes a serine-type proteinase required for cleavage at the NS3/4 and NS4/5 junctions. *Journal of Virology* **67**(7): 3835-3844.

Bassetto, M., Leyssen, P., Neyts, J., Yerukhimovich, M.M., Frick, D.N. & Brancale, A. (2016). Computer-aided identification, synthesis and evaluation of substituted thienopyrimidines as novel inhibitors of HCV replication. *European Journal of Medicinal Chemistry* **123**: 31-47.

Belema, M., Lopez, O.D., Bender, J.A., Romine, J.L., St. Laurent, D.R., Langley, D.R., Lem, J.A., O'Boyle, D.R., Sun, J., Wang, C., Fridel, R.A. & Meanwell, N.A. (2014). Discovery and development of hepatitis C virus NS5A replication complex inhibitors. *Journal of Medicinal Chemistry* **57**(5): 1643-1672.

Brenndörfer, E.D., Karthe, J., Frelin, L., Cebula, P., Erhardt, A., Schulte am Esch, J., Hengel, H., Bartenschlager, R., Sallberg, M., Haussinger, D. & Bode, J.G. (2009). Nonstructural 3/4A protease of hepatitis C virus activates epithelial growth factor–induced signal transduction by cleavage of the Tcell protein tyrosine phosphatase. *Hepatology* **49**(6): 1810-1820.

Carugo, O. & Djinoviæ-Carugo, K. (2013). Half a century of Ramachandran plots. *Acta Crystallographica Section D: Biological Crystallography* **69**(8): 1333-1341.

Chan, A., Patel, K. & Naggie S. (2017). Genotype 3 Infection: The Last Stand of Hepatitis C Virus. *Drugs* **77**(2): 131-144.

Chen, K.X. & Njoroge, F.G. (2010). The journey to the discovery of boceprevir: an NS3–NS4 HCV protease inhibitor for the treatment of chronic hepatitis C. In: Progress in medicinal chemistry, Lawton, G. & Witty D.R. (editors) 1st edition. Amsterdam: Elsevier, pp. 1-36.

Choi, M.J., No, E.S., Thorat, D.A., Jang, J.W., Yang, H., Lee, J., Choo, H., Kim, S.J., Lee, C.S., Ko, S.Y., Lee, J., Nam, G. & Pae, A.N. (2013). Synthesis and biological evaluation of aryloxazole derivatives as antimitotic and vascular-disrupting agents for cancer therapy. *Journal of Medicinal Chemistry* **56**(22): 9008-9018.

Choo, Q.L., Kuo, G., Weiner, A.J., Overby, L.R., Bradley, D.W. & Houghton, M. (1989). Isolation of a cDNA Clone Derived from a Blood-Borne Non-A, Non-B Viral Hepatitis Genome. *Science* **244**(4902): 359-362.

Cummings, M.D., Lindberg, J., Lin, T.I., de Kock, H., Len, O., Lilja E., Fellander, S., Baraznenok, V., Nystrom, S., Nilsson, M., Vrang, L., Edlund, M., Rosenquist, A., Samuelsson, B., Raboisson, P. & Simmen, K. (2010). Inducedfit binding of the macrocyclic noncovalent inhibitor TMC435 to its HCV NS3/NS4A protease target. *Angewandte Chemie International Edition* **49**(9): 1652-1655.

Du, G.X., Hou, L.H., Guan, R.B., Tong, Y.G. & Wang, H.T. (2002). Establishment of a simple assay *in vitro* for hepatitis C virus NS3 serine protease based on recombinant substrate and single-chain protease. *World Journal of Gastroenterology* **8**(6): 1088.

Dustin, L.B. & Rice, C.M. (2007). Flying under the radar: the immunobiology of hepatitis C. *Annual Review of Immunology* **25**: 71-99.

El Kassas, M., Elbaz, T., Hafez, E., Wifi, M.N. & Esmat, G. (2017). Discovery and preclinical development of dasabuvir for the treatment of hepatitis C infection. *Expert Opinion on Drug Discovery* **12**(6): 635-642.

Foster, G.R., Afdhal, N., Roberts, S.K., Bräu, N., Gane, E.J., Pianko, S., Lawitz, E., Thompson, A., Shiffman, M.L., Cooper, C., Towner, W.J., Conway, B., Ruane, P., Bouliere, M., Asselah, T., Berg, T., Zeuzem, S., Rosenberg, W., Agarwal, K., Stedman, C.A.M., Mo, H., Dvory, Sobol, H., Han, L., Wang, J., McNaly, J., Osinusi, A., Brainard, D.M., McHutchinson, J.G., Mazzotta, F., Tran, T.T., Gordon, S.C., Patel, K., Reau, N., Mangia, A. & Sulkowski, M. (2015). Sofosbuvir and velpatasvir for HCV genotype 2 and 3 infection. *New England Journal of Medicine* **373**(27): 2608-2617.

Ganesan, A. & Barakat K. (2017). Applications of computer-aided approaches in the development of hepatitis C antiviral agents. *Expert Opinion on Drug Discovery* **12**(4): 407-425.

Gombar, V.K., Silver, I.S. & Zhao, Z. (2003), Role of ADME characteristics in drug discovery and their *in silico* evaluation: *in silico* screening of chemicals for their metabolic stability. *Current Topics in Medicinal Chemistry* **3**(11): 1205-1225.

Grakoui, A., McCourt, D.W., Wychowski, C., Feinstone, S.M. & Rice, C.M. (1993). Characterization of the hepatitis C virus-encoded serine proteinase: determination of proteinase-dependent polyprotein cleavage sites. *Journal of Virology* **67**(5): 2832-2843.

Hajarizadeh, B., Grebely, J. & Dore, G.J. (2013). Epidemiology and natural history of HCV infection. *Nature Reviews Gastroenterology & Hepatology* **10**(9): 553-562.

He, S., Li, K., Lin, B., Hu, Z., Xiao, J., Hu, X., Wang, A.Q., Xu, X., Ferrer, M., Southall, N., Zheng, W., Aube, J., Schoenen, F.J., Marugan, J.J., Liang, T.J. & Frankowski, K.J. (2017). Development of an Aryloxazole Class of Hepatitis C Virus Inhibitors Targeting the Entry Stage of the Viral Replication Cycle. *Journal of Medicinal Chemistry* **60**(14): 6364-6383.

Hedstrom, L. (2002). Serine protease mechanism and specificity. *Chemical Reviews* **102**(12): 4501-4524.

Hwang, S.J. & Lee, S.D. (2011). Hepatic steatosis and hepatitis C: still unhappy bedfellows? *Journal of Gastroenterology and Hepatology* **26**: 96-101.

Irwin, J.J., Sterling, T., Mysinger, M.M., Bolstad, E.S. & Coleman, R.G. (2012) ZINC: a free tool to discover chemistry for biology. *Journal of Chemical Information and Modeling* **52**(7): 1757-1768.

Ismail, M.A.H., Adel, M., Ismail, N.S.M. & Abouzid, K.A.M. (2013). Molecular Design, Synthesis and Cell Based HCV Replicon Assay of Novel Benzoxazole Derivatives. *Drug Research* **34**(03): 109-120.

Jacobson, I.M., Gordon, S.C., Kowdley, K.V., Yoshida, E.M., Rodriguez-Torres, M., Sulkowski, M.S., Shiffman, M.L., Lawitz, E., Everson, G., Bennet, M., Schiff, E., Al-Assi, M.T., Subramaniam, M., An, D., Lin, M., McNally, J., Brainard, D., Symonds, W.T., McHutchinson, J.G., Patel, K., Feld, J., Pianko, S. & Nelson, D.R. (2013). Sofosbuvir for hepatitis C genotype 2 or 3 in patients without treatment options. *New England Journal of Medicine* **368**(20): 1867-1877.

Kanwal, F., Kramer, J.R., Ilyas, J., Duan, Z. & ElSerag, H.B. (2014). HCV genotype 3 is associated with an increased risk of cirrhosis and hepatocellular cancer in a national sample of US Veterans with HCV. *Hepatology* **60**(1): 98-105.

Kim, H.H., Hyun, J.S., Choi, J., Choi, K.E., Jee, J.G. & Park S.J. (2018). Structural ensemble-based docking simulation and biophysical studies discovered new inhibitors of Hsp90 N-terminal domain. *Scientific Reports* **8**(1): 1-13.

Kim, J.L., Morgenstern, K.A., Lin, C., Fox, T., Dwyer, M.D., Landro, J.A., Chambers, S.P., Markland, W., Lepre, C.A., O'Malley, E.T., Harbeson, S.L., Rice, C.M., Murcko, M.A., Carong, P.R. & Thomson, J.A. (1996). Crystal structure of the hepatitis C virus NS3 protease domain complexed

with a synthetic NS4A cofactor peptide. *Cell* **87**(2): 343-355.

Kumar, A., Gupta, R., Verma, K., Iyer, K., Shanthi, V. & Ramanathan, K. (2014). Identification of novel hepatitis C virus NS3-4A protease inhibitors by virtual screening approach. *Journal of Microbial and Biochemical Technology* **6**(4): 1-7.

Kwong, A.D., Kauffman, R.S., Hurter, P. & Mueller, P. (2011). Discovery and development of telaprevir: an NS3-4A protease inhibitor for treating genotype 1 chronic hepatitis C virus. *Nature Biotechnology* **29**(11): 993.

Labbé, C.M., Rey, J., Lagorce, D., Vavruša, M., Becot, J., Sperandio, O. & Miteva, M.A. (2015). MTiOpenScreen: a web server for structure-based virtual screening. *Nucleic Acids Research* **43**(W1): W448-W454.

Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography* **26**(2): 283-291.

Lawitz, E., Mangia, A., Wyles, D., Rodriguez-Torres, M., Hassanein, T., Gordon, S.C., Schultz, M., Davis, M.N., Kayali, Z., Reddy, K.R., Jacobson, I.M., Kowdley, K.V., Nyberg, L., Subramanian, G.M., Hyland, G.H., Anterburn, S., Jiand, D., McNally, J., Brainard, D., Symonds, W.T., McHutchinson, J.G., Sheikh, A.M., Younossi, Z. & Gane, E.J. (2013). Sofosbuvir for previously untreated chronic hepatitis C infection. *New England Journal of Medicine* **368**(20): 1878-1887.

Lenz, O., Vijgen, L., Berke, J.M., Cummings, M.D., Fevery, B., Peeters, M., Smedt, G.D., Moreno, C. & Picchio, G. (2013). Virologic response and characterisation of HCV genotype 2–6 in patients receiving TMC435 monotherapy (study TMC435-C202). *Journal of Hepatology* **58**(3): 445-451.

Li, H., Leung, K.S. & Wong, M.H. (2012). idock: A multithreaded virtual screening tool for flexible ligand docking. In: Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), 2012 IEEE Symposium, pp. 77-84.

Li, J., Liu, X., Li, S., Wang, Y., Zhou, N., Luo, C., Luo, X., Zheng, M., Jiang, H. & Chen, K. (2013). Identification of novel small molecules as inhibitors of hepatitis C virus by structure-based virtual screening. *International Journal of Molecular Sciences* **14**(11): 22845-22856.

Lin, C. (2006). HCV NS3-4A serine protease. In: Hepatitis C viruses: Genomes and Molecular Biology, Tan S.L. (editor). Norfolk: Horizon Bioscience, pp. 163-206.

Lin, T.I., Lenz, O., Fanning, G., Verbinnen, T., Delouvroy, F., Scholliers, A., Vermeiren, K., Rosenquist, A., Edlund, M., Samuelsson, B., Vrang, L., de Kock, H., Wigerinck, P., Roboisson, P. & Simmen, K. (2009). *In vitro* activity and pre-clinical profile of TMC435350, a potent hepatitis C virus protease inhibitor. *Antimicrobial Agents and Chemotherapy* **53**(4): 1377-1385.

Lipinski, C.A., Lombardo, F., Dominy, B.W. & Feeney, P.J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advance Drug Delivery Reviews* **46**: 3-26.

Lipinski, C.A., Lombardo, F., Dominy, B.W. & Feeney, P.J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advance Drug Delivery Reviews* **23**(1-3): 3-25.

Lontok, E., Harrington, P., Howe, A., Kieffer, T., Lennerstrand, J., Lenz, O., McPhee, F., Mo, H., Parkin, N., Pilot-Matias, T. & Miller, V. (2015). Hepatitis C virus drug resistance–associated substitutions: State of the art summary. *Hepatology* **62**(5): 1623-1632.

Love, R.A., Parge, H.E., Wickersham, J.A., Hostomsky, Z., Habuka, N., Moomaw, E.W., Adachi, T. & Hostomska, Z. (1996). The crystal structure of hepatitis C virus NS3 proteinase reveals a trypsin-like fold and a structural zinc binding site. *Cell* **87**(2): 331-342.

Lüthy, R., Bowie, J.U. & Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. *Nature* **356**(6364): 83-85.

Mandlik, V., Bejugam, P.R. & Singh, S. (2015). Application of Artificial Neural Networks in Modern Drug Discovery. In: Artificial Neural Network for Drug Design, Delivery and Disposition, Puri, M., Pathak, Y., Sutariya, V.J., Tipparaju, S. & Moreno, W. (editors). London: Academic Press, pp. 123-139.

Naeem, A. & Waheed, Y. (2017). Sequence analysis of hepatitis C virus non-structural protein 34A serine protease and prediction of conserved B and T cell epitopes. *Biomedical Reports* **7**(6): 563-566.

National Center for Biotechnology Information (NCBI). PubChem BioAssay Database; AID=940. (2008) Retrieved from https://pubchem.ncbi.nlm.nih.gov/bioassay/940

Néron, B., Ménager, H., Maufrais, C., Joly, N., Maupetit, J., Letort, S., Carrere, S., Tuffery, P. & Letondal, C. (2009). Mobyle: a new full web bioinformatics framework. *Bioinformatics* **25**(22): 3005-3011.

Ng, T.I., Tripathi, R., Reisch, T., Lu, L., Middleton, T., Hopkins, T.A., Pithawalla, R., Irvin, M., Dekhtyar, T., Krishnan, P., Schnell, G., Beyer, J., McDaniel, K.F., Ma, J., Wang, G., Jiang, L., Or, Y.S., Kemp, D., Pilot-Matias, T. & Collins, C. (2018). *In vitro* antiviral activity and resistance profile of the next-generation hepatitis C virus NS3/4A protease inhibitor glecaprevir. *Antimicrobial Agents and Chemotherapy* **62**(1): e01620-17.

Pawlotsky, J.M., Negro, F., Aghemo, A., Berenguer, M., Dalgard, O., Dusheiko, G., Marra, F., Puoti, M. & Wedemeyer, H. (2018). EASL Recommendations on Treatment of Hepatitis C 2018. *Journal of Hepatology* **69**(2): 461-551.

Pal, D. & Chakrabarti, P. (2002). On residues in the disallowed region of the Ramachandran map. *Biopolymers: Original Research on Biomolecules* **63**(3): 195-206.

Pedada, S.R., Yarla, N.S., Tambade, P.J., Dhananjaya, B.L., Bishayee, A., Arunasree, K.M., Philip, G.H., Dharmapuri, G., Aliev, G., Putta, S. & Rangaiah, G. (2016). Synthesis of new secretory phospholipase A2-inhibitory indole containing isoxazole derivatives as anti-inflammatory and anticancer agents. *European Journal of Medicinal Chemistry* **112**: 289-297.

Petruzziello, A., Marigliano, S., Loquercio, G., Cozzolino, A. & Cacciapuoti, C. (2016). Global epidemiology of hepatitis C virus infection: An update of the distribution and circulation of hepatitis C virus genotypes. *World Journal of Gastro-enterology* **22**(34): 7824-7840.

Pinzani, M., Rombouts, K. & Colagrande, S. (2005). Fibrosis in chronic liver diseases: diagnosis and management. *Journal of Hepatology* **42**(1): S22-S36.

Poveda, E., Wyles, D.L., Mena, Á., Pedreira, J.D., Castro-Iglesias, Á. & Cachay, E. (2014). Update on hepatitis C virus resistance to direct-acting antiviral agents. *Antiviral Res*earch **108**: 181-191.

Prakash, T.B., Reddy, G.D., Padmaja, A. & Padmavathi, V. (2014). Synthesis and antimicrobial activity of amine linked bis-and tris-heterocycles. *European Journal of Medicinal Chemistry* **82**: 347-354.

Raj, V.S., Hundie, G.B., Schürch, A.C., Smits, S.L., Pas, S.D., Pogam, S., Janssen, H.L.A., de Knegt, R.J., Osterhaus, A.D.M.E., Najera, I., Boucher, C.A. & Haagmans, B.L. (2017). Identification of HCV resistant variants against direct acting antivirals in plasma and liver of treatment naïve patients. *Scientific Reports* **7**(1): 1-10.

Raney, K.D., Sharma, S.D., Moustafa, I.M. & Cameron, C.E. (2010). Hepatitis C virus non-structural protein 3 (HCV NS3): a multifunctional antiviral target. *Journal of Biological Chemistry* **285**(30): 22725-22731.

Roingeard, P. (2013). Hepatitis C virus diversity and hepatic steatosis. *Journal of Viral Hepatitis* **20**(2): 77-84.

Rosenquist, Å., Samuelsson, B., Johansson, P.O., Cummings, M.D., Lenz, O., Raboisson, P., Simmen, K., Vendeville, S., de Kock, H., Nilsson, M., Horvath, A., Kalmeijer, R., de la Rosa, G. & Beumont-Mauviel, M. (2014). Discovery and development of simeprevir (TMC435), a HCV NS3/4A protease inhibitor. *Journal of Medicinal Chemistry* **57**(5): 1673-1693.

Saeed, M., Scheel, T.K., Gottwein, J.M., Marukian, S., Dustin, L.B., Bukh, J. & Rice, C.M. (2012). Efficient replication of genotype 3a and 4a HCV replicons in human hepatoma cells. *Antimicrobial Agents and Chemotherapy* **56**(10): 5365-5373.

Simmonds, P., Becher, P., Bukh, J., Gould, E.A., Meyers, G., Monath, T., Muerhoff, S., Pletnev, A., Rico-Hesse, R., Smith, D.B. & Stapleton, J.T. (2017). ICTV Virus Taxonomy Profile: Flaviviridae. *Journal of General Virology* **98**(1): 2-3.

Smith, D.B., Bukh, J., Kuiken, C., Muerhof, A.S., Rice, C.M., Stapleton, J.T. & Simmonds, P. (2019). HCV classification: A web resource to manage the classification and genotype and subtype assignments of hepatitis C virus. In *Flaviviridae Study Group*. Retrieved from https://talk.ictvonline.org/ictv_wikis/flaviviridae/w/sg_flavi/56/hcv-classification

Sofia, M.J. (2016). Sofosbuvir: The Discovery of a Curative Therapy for the Treatment of Hepatitis C Virus. *Successful Drug Discovery* **2**: 163-188.

Soumana, D.I., Yilmaz, N.K., Ali, A., Prachanronarong, K.L. & Schiffer, C.A. (2016). Molecular and dynamic mechanism underlying drug resistance in genotype 3 hepatitis C NS3/4A protease. *Journal of the American Chemical Society* **138**(36): 11850-11859.

Tanwar, S., Trembling, P.M. & Dusheiko, G.M. (2012). TMC435 for the treatment of chronic hepatitis C. *Expert Opinion on Investigational Drugs* **21**(8): 1193-1209.

Trott, O. & Olson, A.J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multi-threading. *Journal of Computational Chemistry* **31**(2): 455-461.

Vermehren, J. & Sarrazin, C. (2012). The role of resistance in HCV treatment. *Best Practice & Research Clinical Gastroenterology* **26**(4): 487-503.

Wang, J. & Urban, L. (2004). The impact of early ADME profiling on drug discovery and development strategy. *Drug Discovery World* **5**(4): 73-86.

Wani, M.Y., Ahmad, A., Shiekh, R.A., Al-Ghamdi, K.J. & Sobral, A.J. (2015). Imidazole clubbed 1, 3, 4-oxadiazole derivatives as potential antifungal agents. *Bioorganic & Medicinal Chemistry* **23**(15): 4172-4180.

World Health Organization (WHO). Hepatitis C (2017). Retrieved from http://www.who.int/news-room/fact-sheets/detail/hepatitis-c

Wyles, D., Poordad, F., Wang, S., Alric, L., Felizarta, F., Kwo, P.Y., Maliakkal, B., Agarwal, K., Hassanein, T., Weilert, F., Lee, S.S., Kort, J., Lovell, S.S., Liu, R., Lin, C., Pilot-Matias, T., Krishnan, P. & Mensa, F.J. (2018). Glecaprevir/pibrentasvir for hepatitis C virus genotype 3 patients with cirrhosis and/or prior treatment experience: a partially randomized Phase 3 clinical trial. *Hepatology* **67**(2): 514-523.

Yu, W., Tong, L., Selyutin, O., Chen, L., Hu, B., Zhong, B., Hao, J., Ji, T., Zan, S., Yin, J., Ruck, R.T., Curry, S., McMonagle, P., Agrawal, S., Rokosz, L., Carr, D., Ingravallo, P., Bystol, K., Lahse, F., Liu, R., Chen, S., Feng, K., Cartwright, M., Asante-Appiah, E. & Kozlowski, J.A. (2018). Discovery of MK-6169, a Potent Pan-Genotype Hepatitis C Virus NS5A Inhibitor with Optimized Activity against Common Resistance-Associated Substitutions. *Journal of Medicinal Chemistry* **61**(9): 3984-4003.