Mol Divers (2015) 19:529–539
DOI 10.1007/s11030-015-9591-5

FULL-LENGTH PAPER

Combination of pharmacophore hypothesis and molecular docking to identify novel inhibitors of HCV NS5B polymerase

Amaravadhi Harikishore[1] · Enlin Li[1] · Jia Jun Lee[1] · Nam-Joon Cho[2] · Ho Sup Yoon[1,3]

Received: 19 October 2014 / Accepted: 25 March 2015 / Published online: 11 April 2015

Abstract Hepatitis C virus (HCV) infection or HCV-related liver diseases are now shown to cause more than 350,000 deaths every year. Adaptability of HCV genome to vary its composition and the existence of multiple strains makes it more difficult to combat the emergence of drug-resistant HCV infections. Among the HCV polyprotein which has both the structural and non-structural regions, the non-structural protein NS5B RNA-dependent RNA polymerase (RdRP) mainly mediates the catalytic role of RNA replication in conjunction with its viral protein machinery as well as host chaperone proteins. Lack of such RNA-dependent RNA polymerase enzyme in host had made it an attractive and hotly pursued target for drug discovery efforts. Recent drug discovery efforts targeting HCV RdRP have seen success with FDA approval for sofosbuvir as a direct-acting antiviral against HCV infection. However, variations in drug-binding sites induce drug resistance, and therefore targeting allosteric sites could delay the emergence of drug resistance. In this study, we focussed on allosteric thumb site II of the non-structural protein NS5B RNA-dependent RNA polymerase and developed a five-feature pharmacophore hypothesis/model which estimated the experimental activity with a strong correlation of 0.971 & 0.944 for training and test sets, respectively. Further, the Güner-Henry score of 0.6 suggests that the model was able to discern the active and inactive compounds and enrich the true positives during a database search. In this study, database search and molecular docking results supported by experimental HCV viral replication inhibition assays suggested ligands with best fitness to the pharmacophore model dock to the key residues involved in thumbs site II, which inhibited the HCV 1b viral replication in sub-micro-molar range.

Graphical Abstract HCV nonstructural protein NS5B RNA-dependent RNA polymerase (RdRP) mediates the catalytic role of viral RNA replication. Lack of host RNA-dependent RNA polymerase enzyme had made it an attractive and hotly pursued target for drug discovery efforts. In this study, we developed a five-feature pharmacophore (3D QSAR) model for thumb site inhibitors of HCV RdRP, which estimated the experimental activity with a strong correlation of 0.971 & 0.944 for training and test sets, respectively. Our database search and molecular docking results suggested that the compounds **1** and **2** with best fitness to the pharmacophore model were predicted to interact with key residues involved in thumbs site II and could inhibit the HCV RdRP activity. Further, the compounds **1** and **2** potently inhibited HCV 1b viral replication in sub-micro-molar range.

Electronic supplementary material The online version of this article (doi:10.1007/s11030-015-9591-5) contains supplementary material, which is available to authorized users.

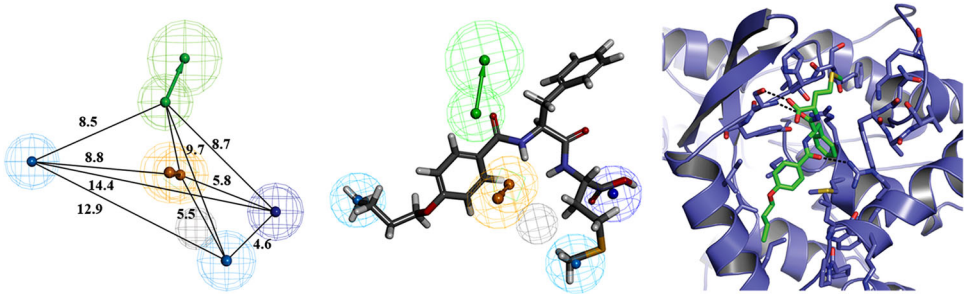
✉ Amaravadhi Harikishore
amaravadhi@ntu.edu.sg

✉ Ho Sup Yoon
hsyoon@ntu.edu.sg

[1] School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551, Singapore

[2] School of Materials Science & Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798, Singapore

[3] Department of Genetic Engineering, College of Life Sciences, Kyung Hee University, Yongin-si, Gyeonggi-do 446-701, Republic of Korea



Keywords HCV · NS5B polymerase · RDRP · Thumb site II inhibitor · Pharmacophore modeling

Introduction

Hepatitis C virus (HCV) belongs to *Flaviviridae* family, and its genome consists of 9.6-kb, single-stranded RNA with positive polarity and encodes a large molecular mass precursor polyprotein. The precursor molecule upon proteolytic processing yields three structural proteins (C, E1 and E2) and seven non-structural proteins (p7, NS2, NS3, NS4A, NS4B, NS5A and NS5B) [1,2]. Among these proteins, non-structural protein 5B (NS5B) mediates HCV RNA-dependent RNA polymerase (RdRp) and terminal transferase activities and regulates HCV viral replication. Recent trends have shown that HCV infections have afflicted more than 150 million people worldwide, which is approximately 3 % of the world population [3]. Every year, Hepatitis C-related liver diseases like liver cirrhosis and hepatocellular carcinoma kills more than 350,000 people [4]. The conventional HCV treatment during the past decade employed a combination of pegylated interferon and ribavirin which had imparted only a 40–50 % sustained viral response (SVR) in HCV-infected patients [5]. Though the introduction of direct-acting antiviral(s) (DAA) in combination with interferon and ribavirin had increased the SVR of the patients, it required a longer treatment regimen (48 weeks) and was often associated with concomitant adverse side effects on patients due to use of interferon [6]. Current DAA development was targeted on non-structural proteins such as NS3/4A serine protease, NS5A and NS5B RNA-RdRp which are shown to be essential for HCV replication [7]. More importantly, research efforts targeting NS5B polymerase led to the identification of sofosbuvir which has been recently approved by FDA for use in a HCV treatment that was free from interferon [8]. Adaptability of HCV virus to mutate key residues involved in drug binding and the ability to vary its composition to diverse HCV genetic strain populations as evidenced with the six strains such as 1a, 1b, 2, 4, 3 and 6 present the possibility of emergence of drug resistance. Therefore, continued efforts are needed to develop DAA that can overcome mutations at drug-binding sites and to obfuscate the emergence of drug resistance due viral mutations on drug-binding sites [9].

X-ray crystallographic studies have revealed that NS5B structure resembled a right hand and possessed a finger, thumb and palm domains. These studies have shown that there are at least four allosteric sites in NS5B polymerase for inhibitor binding, namely palm site I, palm site II, nucleoside or thumb site I and non-nucleoside or thumb site II besides the RNA binding site (Supplementary Figure 1) [10–12]. NS5B inhibitors are broadly classified as (i) nucleoside inhibitor that competes with nucleotides at the active site and (ii) non-nucleoside inhibitors that could bind either at the thumb site II or palm site (Chart 1). Non-nucleoside inhibitors are docked at the thumb site II that are involved in interaction with key residues such as S476, Y477, L419, L482, L497 and W528 (Supplementary Figure 2) and allosterically inhibit the viral RNA polymerase activity. NS5B has been shown to interact with cyclophilin and share its cyclophilin-binding region with that NS5A. It was indirectly inferred from NMR binding studies that NS5B thumb site region could be involved in its interaction with NS5A, and ligand binding at thumb site prevented the NS5B and NS5A interactions [13]. A recent study exemplified that ligand binding at the thumb site II induced the conformational changes, re-orient the C-terminal tail, β-loop structural elements into the palm site of NS5B polymerase. These conformational changes further stabilize the closed state and allosterically inhibit the RNA binding and NS5B polymerase activity [14]. In this study, using combined approaches of pharmacophore modelling and virtual screening methods, we present key features that are important for inhibition of RNA polymerase activity. Further, our database screening and evaluation of molecular interactions at the thumb site II led us to the identification of novel chemical entities that could bind to thumb site II of NS5B and inhibit NS5B polymerase activity.

Materials and methods

Data collection for pharmacophore generation

Three hundred and ten HCV NS5B thumb site II inhibitors were selected from the literature [5,15–44] with activity

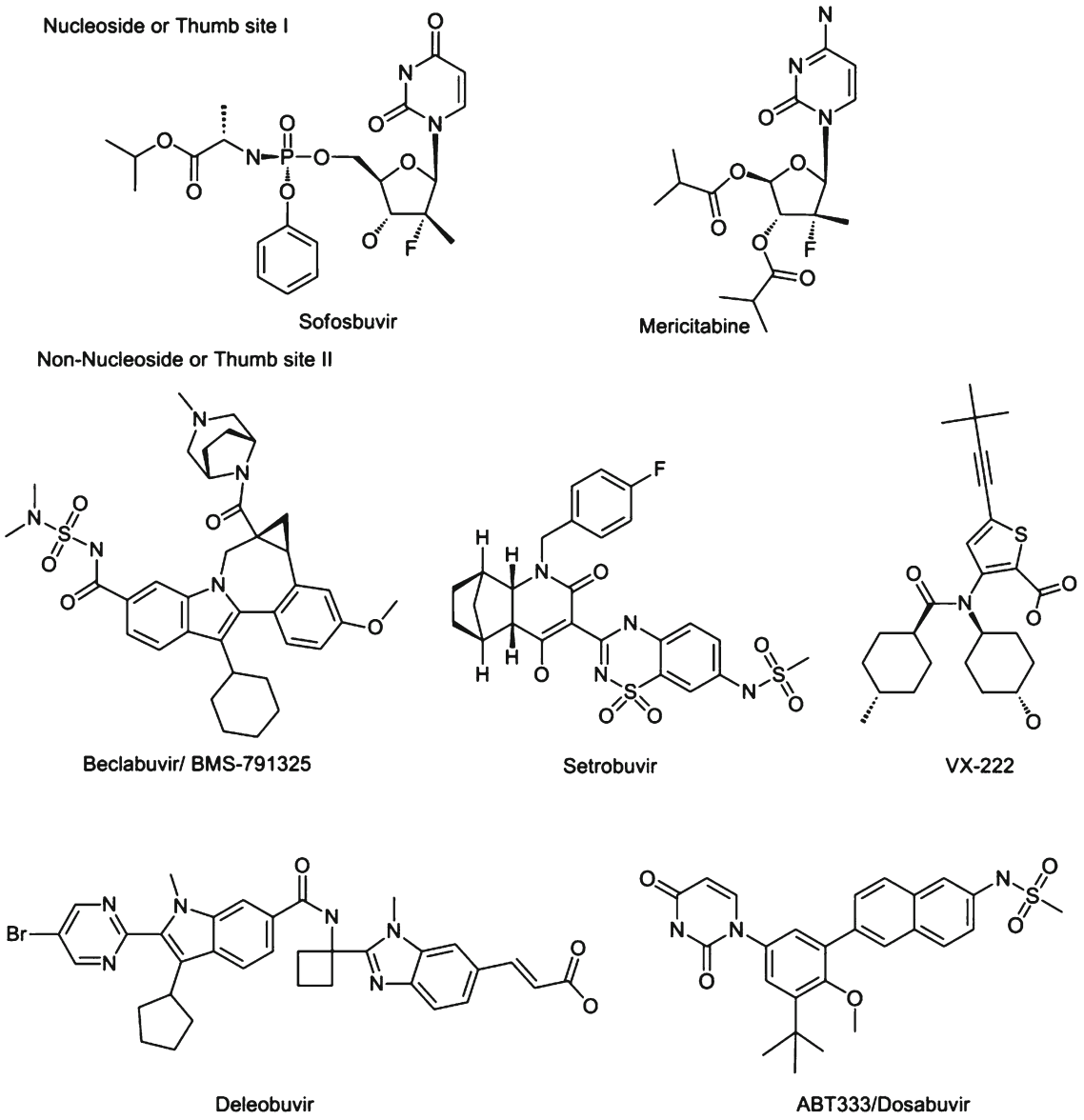


Chart 1 Nucleoside and non-nucleoside inhibitors that inhibit HCV RdRP

(IC_{50}, μM) determined from HCV genotype 1b NS5B RdRP inhibition assay [24]. Eighteen molecules [15,17,20,21,26,28–31,34,35,38,44] with diverse structures (Fig. 1) and a fourfold activity spread (0.002–92 μM) were used to develop a pharmacophore model. The remaining 292 compounds (Supplementary Data1) were used as test set for validation of the model. Most of the modelling simulation works were performed on linux workstation using Accelrys Discovery Studio 3.5 (DS3.5) modelling suite [45]. 2D structures for all the ligands were sketched in Accelrys Draw [46]. 3D coordinates were generated using 'Prepare Ligands' module, and the obtained 3D conformations were energy minimized using smart minimizer for 2000 steps with CHARMm forcefield using the 'Minimize Ligands' module. A maximum of 255 conformations were generated for each molecule using the BEST conformation method in the 'Generate Conformations' module. The BEST search method employed a poling algorithm [47] that samples diverse low-energy conformations which are far from local minimum.

Construction of pharmacophore model

The '3D QSAR Pharmacophore Generation' module (Hypogen program) [48] was used to generate a valid pharmacophore from 18 training set molecules with the following features hydrogen bond acceptor (0-3), hydrogen bond donor (0-3), hydrophobic aliphatic (0-3), negative ionizable (1-2)

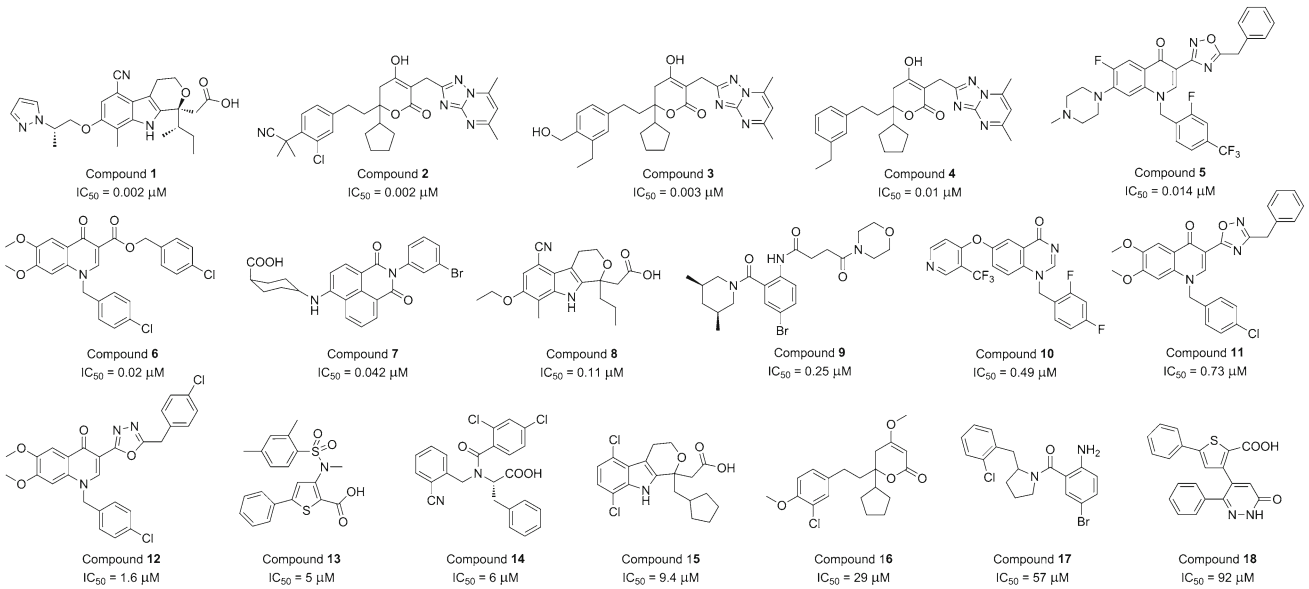


Fig 1 The 18 training set molecules that were used to develop NS5B thumb site II pharmacophore model

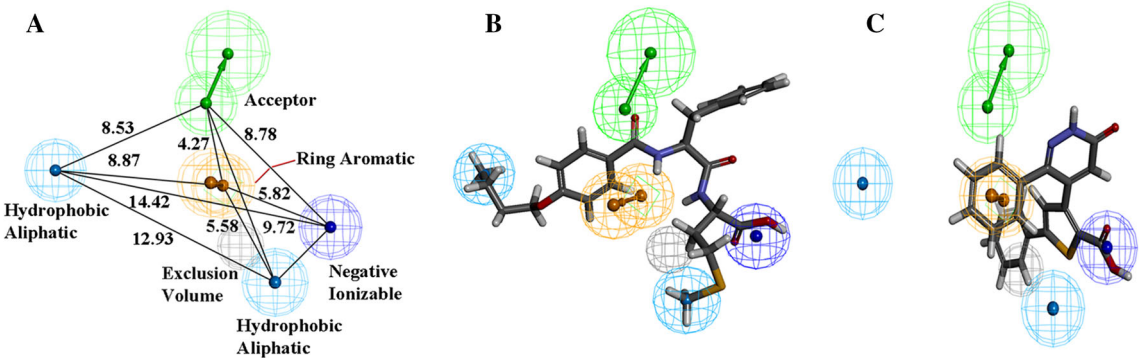


Fig 2 Best NS5B thumb site II pharmacophore model and its mapping to active and inactive. **a)** Pharmacophore model with 5 features, 1 negative ionizable feature (*blue*), 1 hydrogen bond acceptor (*green*), 1 ring aromatic feature (*orange*) and 2 hydrophobic aliphatic features (*cyan*). The *grey* sphere represents the exclusion volume. The distances (in Å) between the features are also shown. **b)** All features in the pharmacophore model are mapped to the most active molecule in the training set (**1**, 0.002 μM). **c)** Only 2 out of 5 features in the pharmacophore model are mapped to the most inactive molecule (**18**, 92 μM)

and ring aromatic (0-3) with interfeature distance of 3 Å as input parameters. Each run of the module generates ten best phamacophore models ranked based on the correlation between the experimental activity and the estimated activity predicted by each model. Hypogen algorithm used in '3D QSAR pharmacophore generation' uses the user input data i.e., training set, diverse conformational models, chemical features, activity and above-mentioned parameters to generate top ten pharmacophore models. Hypogen builds simplest model that can correlate the estimated activities with measured activities. The pharmacophore generation generally involves three phases: (1) a constructive phase—which identifies the features that are associated with most active compounds and builds hypotheses (models) that are common among the most active compounds, (2) subtractive phase—where the features that are common among the inactive compounds are removed from the initial hypotheses, and (3) optimization phase—where an attempt is made to refine the initial hypotheses model with slight refinement by either rotation or translation of feature type or add / remove certain feature and retains changes that lower the error involved during regression of activity estimate. The final model contains an ensemble or a set of generalized chemical features in three-dimensional space as well as regression information that helps predict the measured activity based on their fitness to these features [48]. In order for the model to be acceptable, the following control or cost metrics should be satisfied. Configuration cost should be always less than 17. The configuration cost of the model represents the entropy involved with the training set employed. Whenever the configuration cost

Table 1 Fitness scores and prediction of IC_{50} of training set molecules

Name	Fitness	Est IC_{50}	Exp IC_{50}	Error
Compound **1**	9.53	1.30e−03	2.00e−03	−1.6
Compound **2**	9.0	4.40e−03	2.00e−03	2.2
Compound **3**	9.13	3.20e−03	3.00e−03	1.1
Compound **4**	8.88	5.70e−03	0.01	−1.8
Compound **5**	8.12	0.033	0.014	2.4
Compound **6**	8.09	0.035	0.02	1.7
Compound **7**	8.08	0.036	0.042	−1.2
Compound **8**	7.08	0.36	0.11	3.3
Compound **9**	6.37	1.9	0.25	7.4
Compound **10**	6.74	0.78	0.49	1.6
Compound **11**	7.03	0.41	0.73	−1.8
Compound **12**	6.98	0.45	1.6	−3.5
Compound **13**	5.92	5.2	5	1
Compound **14**	6.29	2.2	6	−2.7
Compound **15**	6.15	3.1	9.4	−3.1
Compound **16**	5.28	22	29	−1.3
Compound **17**	5.14	32	57	−1.8
Compound **18**	4.67	92	92	1

was more than 17, one can only proceed further by changing the training set data. In addition to configuration cost, three other costs, namely the null cost, fixed cost and total cost, also determine the validity of the model [48]. The null cost represents the highest cost for the phamacophore run as the cost was computed on averaged value of activity of the training set molecules without the structural weight to activity. The fixed cost represents the cost from the simplest pharmacophore model that fits the training set molecules perfectly. Total cost refers to the cost computed by structural feature(s) contribution to activity for each of the compound in the training set. For a model to be significant, the null cost should be greater than the fixed cost by more than 40 bit units. A difference of 40–60 bit units corresponds to a 75–90 % probability for correlation between the experimental and estimated activity values. Further, the total cost (the cost of the pharmacophore model generated) should be close to the fixed cost in order for the model to be significant. Furthermore, to ensure that the best run is not random or a chance correlation, a randomization test using the Fisher validation (catscramble) test [48] with confidence level of 98 % was carried out under the '3D QSAR Pharmacophore Generation' module. The program randomizes the activity data of the training set molecules and generates 49 random pharmacophore models. None of these random models should have a higher correlation or lower total cost than the best model. Next, the predictive ability of the model to estimate the activity of 293 compounds (test set) that were not involved inthe training set was carried out using 'Ligand Pharmacophore Mapping' module.

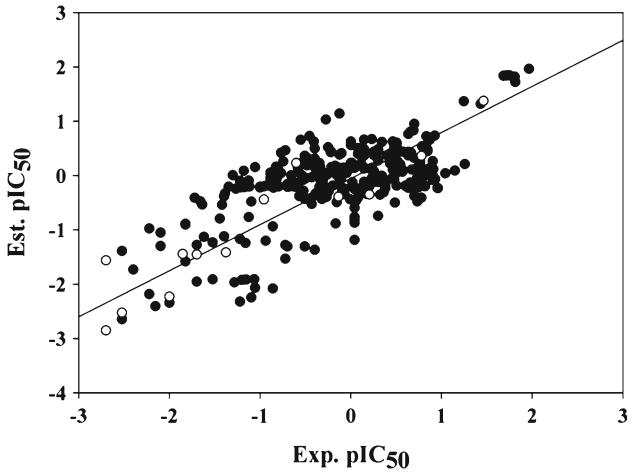


Fig 3 Estimated pIC_{50} against experimental pIC_{50} for compounds in the training and test set. The best model had a correlation (r) of 0.971 for training set and 0.944 for test set

Table 2 Güner Henry scoring

Sl. no	Parameter	No. of molecules
1	Total molecules in data set (D)	310
2	Total number of actives in data set (A)	55
3	Total hits (Ht)	11
4	Active hits (Ha)	9
5	Yield of actives in percentage [(Ha/Ht) × 100]	81.8
6	Coverage of actives in percentage [(Ha/A) × 100]	16.4
7	Enrichment factor (E) [(Ha × D)/(Ht × A)]	4.61
8	False negatives [A − Ht]	46
9	False positives [Ht − Ha]	2
10	Goodness of hit score [(Ha/4HtA)(3A + Ht) × (1 − ((Ht − Ha)/(D − A)))]	0.65

For the model to be robust for database search, there should be strong correlation between the estimated and the experimental activity of the test set compounds. The Güner-Henry (GH) score was also calculated for the best model to determine the quality of the database screen. It uses a scoring method that infers the quality of the yield and the coverage of active compounds during a database search. The score has a range of 0–1, with 0 being a null model and 1 being the ideal model.

Database and virtual screening

After the validation of the pharmacophore model, the model was used to screen the ChemDiv [49] and Interbioscreen

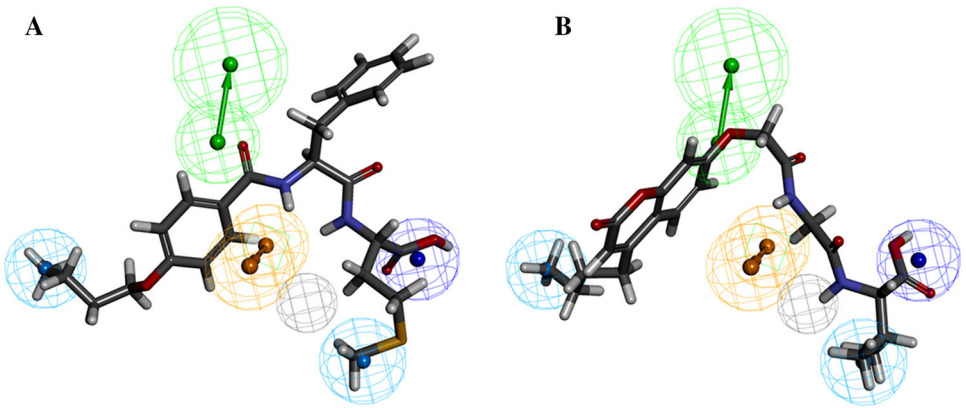


Fig 4 Pharmacophore mapping of best fitting novel compounds from interbioscreen (IBS) Library. As shown in A, B, most of the important features except ring aromatic feature (*orange*) in the best pharmacophore had fitting with corresponding atoms of the ligands

libraries [50] derived from natural products compounds. The database screening was done using 'Search 3D Database' module in Discovery Studio 3.5. The 'Ligand Pharmacophore Mapping'/'Screen Library' module was used to estimate the activity of the library. Thus obtained focused library or the best hits were further analysed with molecular docking studies to investigate the possible interaction with HCV NS5B at thumb site II.

Docking studies

The HCV NS5B crystal structure (PDB ID: 4EO6 [51]) that has resolution of 1.79 Å was used for molecular docking studies. The protein was prepared for simulation studies by deleting the heteroatoms, and charges and potentials were assigned using CHARMm force field in Discovery Studio 3.5. The missing atoms / residues were corrected using the 'Build and Edit Protein' module. The protein structure was energy minimized for 5000 steps (with the heavy atoms constrained) using the conjugate gradient algorithm with the 'Minimization' module in DS3.5. Genetic Optimization for Ligand Docking (GOLD) 5.0 program [52,53] was employed for evaluation of protein–ligand interaction energies. The active site of 5 Å from the co-crystallized ligand was defined as the binding site of thumb site II. The docking parameters were optimized by stimulating the binding mode of the co-crystallized ligand, and the same parameters were used to dock the focused library. The ChemPLP and GOLD scoring function(s) were used to evaluate the molecular interactions of the ligand–protein complexes.

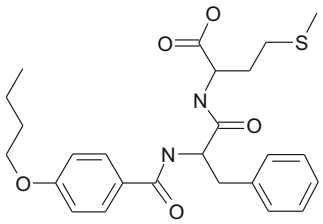
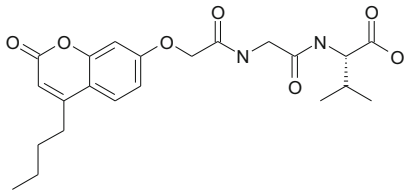
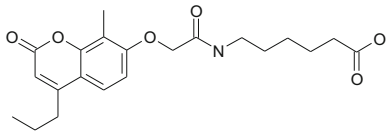
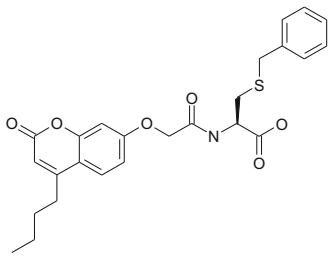
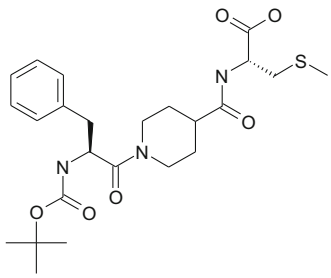
Cell culture and transfection

Huh.7 cells that contain the HCV genotype 1b sub-genomic replicon [54], designated as Huh.8 cells, were a generous gift from Prof. Charles M. Rice. Huh.8 cells were grown on T75 cell culture flasks in a DMEM medium supplemented with 10 % FBS. The cells were maintained in a humidified incubator with 5 % CO_2 at 37 °C. When the cells reached 80–90 % confluency, total cell count was determined using the haemocytometer and 200,000 cells per well were seeded into 6-well cell culture plates. The presence of HCV genotype 1b genome was verified by treating Huh.8 cells with Cyclosporine A (CsA) as a control to monitor the HCV 1b viral replication. Cells were incubated with inhibitors (CsA, compound 1-5(0–500 nM)) for 3 days and then harvested for downstream application. Experiments were performed in triplicate, and the IC_{50} were determined by non-linear regression with dose-response inhibition model using GraphPad Prism 5 software [55].

RNA extraction and RT-qPCR

Total RNA was extracted from treated cells using TRIZOL reagent (Ambion, Life Technologies) as previously described [56]. Briefly, cells were lysed using the appropriate amount of TRIZOL reagent. Chloroform was added to the TRIZOL solution and centrifuged to obtain three separate phases. RNA containing aqueous phase was isolated by precipitation with isopropanol. The RNA pellet was then washed with 75 % ethanol and re-suspended in DEPC-treated water. Synthesis of total cDNA from total RNA was performed using the Transcriptor First Strand cDNA Synthesis Kit (Roche) according to manufacturer's instructions. Random hexamer primers were used for the reaction. Real-Time quantitative PCR (RT-qPCR) was set up using SYBR® Select Master Mix (Applied Biosystems) according to the manufacturer's instructions on Applied Biosystems 7500 Real- Time PCR machine. A standardized amount of cDNA (400 ng), HCV genotype 1b primers, Forward primer (5′-CGGTAGCTCTGAATCGTCGGCTGT-3′) and Reverse primer (5′-CATCCTCACTAGCCTCTTCACTCACGG-3′) which detects the viral NS5A gene was employed to evaluate the amount of HCV 1b viral replication.

Table 3 Structures of 5 top ligands with best fitness to pharmacophore model and had good GOLD PLP scores

Sl. no	Structure	Estimate IC_{50} (μM)	Fitness	GOLD Score	PLP	Vendor ID	HCV 1b replication inhibition $IC_{50} \pm$ SEM (nM)
1		0.0053	8.90	76.88		STOCK1N-09906	272.8 ± 1.70
2		0.0096	8.65	62.04		STOCK1N-54222	36.00 ± 1.77
3		0.024	8.25	69.8		STOCK1N-40747	Not determined
4		0.031	8.14	69.72		STOCK1N-54416	Not determined
5		0.039	8.04	68.3		STOCK1N-47746	Not determined

Results and discussion

Our best pharmacophore model had five features—one negative ionizable (blue), one hydrogen bond acceptor (green), one ring aromatic (orange) and two hydrophobic aliphatic (cyan) features and are spaced at specific distances from each other (Fig. 2a). These features are necessary for binding to thumb site II of NS5B RdRP. The best pharmacophore model has a correlation (r) of 0.971 and a configuration cost of 10.4 bits. The configuration cost (10.4 bits) of the model is well within the specified 17 bits, indicating that the program was able to handle complexity involved with the training set. The null, fixed and total costs are 151, 72.1 and 78.2 bits, respectively. The total cost of the pharmacophore run was closer to the fixed cost, and the cost difference (78.9) between the null cost and fixed costs indicates that there is a probability of more than 90 % of true correlation. The Hypogen program (in 3D QSAR Pharmacophore Generation) computes a fitness

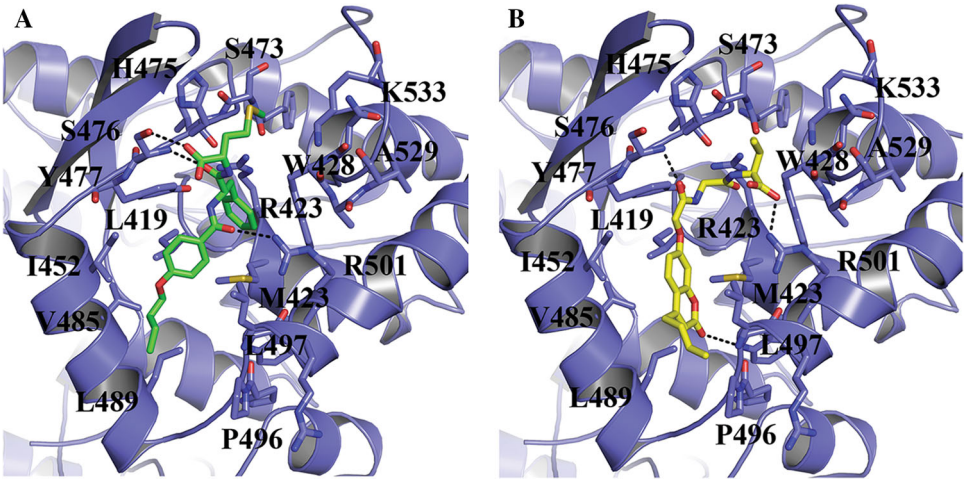


Fig 5 Binding mode of best ligands at thumb site II of NS5B RdRP. **a** The 4-butoxy-benzoyl amino fragment of the scaffold was engaged in hydrophobic interactions with groove residues L419, M423, I497, I482, V485, L489 and P496. Further, its carbonyl group was also involved in hydrogen bonding interaction with R501 residue. The carboxyl and carbonyl groups on '2-amino-acetyl amino' substituent were mainly involved in bonded interactions with side-chain oxygen and main-chain amide atoms of S472. Furthermore, the 4-methylsulfanyl moiety on butyric acid could project into crevices between His475 and K533 and stabilize the ligand binding. **b** Chromene nucleus in this ligand was docked at the hydrophobic groove lined by L419, M423, I497, I482, V485, L489 and P496 residues. Further, 2-oxo group was also engaged in hydrogen bonding to P496 residue. The carbonyl moieties on 2-[2-acetylamino]-acetylamino3-methyl butyric acid were mainly involved in bonding interactions with residues such as R501, W428 and S476, and stabilized the ligand binding at thumb site and inhibition of the HCV RdRP activity and viral replication

score for each ligand based on geometric fit of each ligand (in the training set) to the centre of the feature spheres. It estimated the activity of the ligands by simple regression of the fitness score to the activity. The higher the fitness score, the higher was the predicted activity (Table 1). Our best model was able to predict the experimental activity of the training set molecules with a strong correlation of 0.971. Likewise, the best model was also able to predict the activity of the test set molecules with a good correlation of 0.941 (Fig. 3). The most active compound **1** in the training set mapped to all the features in the best pharmacophore model (Fig. 2b). In contrast, the most inactive ligand, compound **18**, missed three features such as the hydrogen bond acceptor (green) and the two hydrophobic aliphatic features (cyan) (Fig. 2c), which explains its lack of activity. The Fisher validation or cat scramble test with 98 % confidence limits verified that the best pharmacophore model was a statistically significant and was not a chance correlation. None of the 13 random models that were generated had a correlation that was higher than the best pharmacophore model, and likewise none of the random models had a total cost value lower than the best model (Supplementary Figure 3a, 3b). Further, the predictive ability of the best model was also verified using 292 test set compounds. The best model was able to recognize the active and inactive compounds and estimate their experimental activities with a correlation of 0.944 (Fig. 3). The ability of the best model to discern actives and in-actives was further quantified using GH scoring, and our results (Table 2) indicate that the best model had a GH score of 0.65 and could differentiate between the actives (true positives) and the in-actives (false positives). 3D Database screening with an in-house virtual library prepared interbioscreen natural product-based collection gave a focused library of 321 molecules. Ligand pharmacophore mapping was done to evaluate the fitness of focused library to the best model and provided an estimate of their activity on HCV RdRP. The best fitting molecules S1N-09906 and S1N-54222 were predicted to have an IC_{50} of 0.0096 and 0.0053 μM, respectively. As shown in Fig. 4a and b, the best hits were shown to map most of the features except the ring aromatic feature and are estimated to possess a sub-micro-molar activity. The lack of fitness to ring aromatic feature might not be that critical to the activity prediction in comparison to lack of fitness to aliphatic features which hindered the activity to a greater extent (Fig. 3c, Supplementary Figure 4a, 4b).

Our molecular docking results with best fitting ligands had high interaction or fitness scores with thumb site residues of NS5B polymerase. The best fitting ligand (1, Stock1N-09906) bears a 4-butoxy-benzoylamino moiety linked to 2-amino-acetyl amino fragment via 4-methylsulfanyl butyric acid (Table 3). On the one hand, the 4-butoxy-benzoylamino scaffold was predicted to orient into the thumb site II pocket and was engaged in hydrophobic interactions with residues L419, M423, I497, I482, V485, L489 and P496. Further, its carbonyl group was also involved in hydrogen bonding interaction with R496 (Fig. 5A). On the other hand,

the carboxyl and carbonyl groups on the 2-amino-acetyl amino substitution were engaged in hydrogen bonded interactions with S472 side chain oxygen and main chain amide atoms.The 4-methylsulfanyl moiety on butyric acid could project into crevices between His475 and K533 and stabilize the ligand binding. Similarly, Stock1N-054222 (2, Table 3) bearing a '4-butyl-2-oxo-2H-chromen-7yloxy' nucleus was oriented into the hydrophobic groove (L419, M423, I497, I482, V485, L489, P496 residues) with strong hydrophobic interactions. Further, the 2-oxo group on chromene was also involved in hydrogen interaction with L497, while the 2-[2-acetylamino]-acetylamino-3-methyl butyric acid fragment on 7th position of chromene nucleus maintained the hydrogen bonding interactions at the thumb site. The carboxyl group on butyric acid was not only involved in hydrogen bonding interactions with R501 side chain atoms but also maintained close contacts with W528 main chain carbonyl atoms, respectively (Fig. 5B). Likewise, the carbonyl moieties on 2-[2-acetylamino]-acetylamino-3-methyl butyric acid were engaged in hydrogen bonding interactions with R423 and S476, respectively. Other best fitting ligands belonging to this scaffold (3, 4 and 6, Table 3) also showed similar interactions at thumb site II.

In order to assess the efficacy of our top ranking ligands, HCV 1b viral replication inhibition studies using HCV 1b replicon assay system were carried out and our results showed that the compounds **1** and **2**, which are within the top five compounds tested (Table 3), were able to potently inhibit HCV 1b viral replication at 1 μM concentration. Further, dose-response studies in comparison with cyclosporine have revealed that compounds **1** and **2** inhibited the HCV 1b viral replication in nanomolar range with an IC_{50} of 272.8 ± 1.70 (SEM) nM, 36.00 ± 1.77 (SEM) nM, respectively (Supplementary Figure 5). Recent studies have highlighted that the mutations (M423T or L419M) on thumb site of NS5B polymerase could lower the ligand binding affinity (1, 2) and are susceptible to drug resistance [56]. Our best compounds (**1**, **2**) bear no structural similarity to the ligands (VX-222 and ANA598) that lacked sensitivity to thumb site mutations [56]. Further, both the compounds **1** and **2** are not only involved in close van der Waals contacts with hydrophobic residues in the vicinity of M423 residue but also extend their hydrophobic interactions with other residues (His475 and K533) in the hydrophobic cavity. We speculate that these additional hydrophobic interactions could enable the compounds **1** and **2** to overcome the effect of mutations such at M423T or L419M as shown in previous studies [57]. Further, *in vitro* binding studies using compounds **1** and **2** on the wild-type and M423T HCV NS5B polymerases might provide further insights into drug resistance by M423T mutation at HCV thumb site. Taken together, both the hydrophobic as well as hydrogen bonding interactions with S476 and R501, which are involved at thumb site II, could firmly lock the ligands at thumb site II, stabilize the closed or inactive conformation and allosterically inhibit the RNA polymerase activity of NS5B.

Conclusions

Combined approaches of ligand- and structure-based drug development are often used to identify novel chemical entities in academia and industry. Our ligand-based 3D QSAR pharmacophore profiling on thumb site II inhibitors enabled us to identify key features that are required to bind thumb site II and allosterically inhibit the HCV RdRP-mediated viral replication. Our pharmacophore model not only predicted the experimental activity with good correlation but also enriched the probability of finding true positives. Together with virtual screening and viral replication inhibition studies, we were able to identify two scaffolds (Table 3) that are predicted to bind at thumb site II of NS5B polymerase and inhibit the viral replication in nanomolar range.

Acknowledgments We wish to thank the funding support from Ministry of Education (AcRF Tier 2 ARC25/12).

Conflict of interest The authors confirm that this article content has no conflict of interest.

References

1. Manns MP, von Hahn T (2013) Novel therapies for hepatitis C - one pill fits all? Nat Rev Drug Discov 12:595–610. doi:10.1038/nrd4050
2. Fusco DN, Chung RT (2012) Novel therapies for hepatitis C: insights from the structure of the virus. Annu Rev Med 63:373–387. doi:10.1146/annurev-med-042010-085715
3. Lavanchy D (2009) The global burden of hepatitis C. Liver Int 29(Suppl 1):74–81. doi:10.1111/j.1478-3231.2008.01934.x
4. Marinho RT, Barreira DP (2013) Hepatitis C, stigma and cure. World J Gastroenterol 19:6703–6709. doi:10.3748/wjg.v19.i40.6703
5. Alexopoulou A, Papatheodoridis GV (2012) Current progress in the treatment of chronic hepatitis C. World J Gastroenterol 18:6060–6069. doi:10.3748/wjg.v18.i42.6060
6. Tse MT (2013) All-oral HCV therapies near approval. Nat Rev Drug Discov 12:409–411. doi:10.1038/nrd4036
7. Kiser JJ, Flexner C (2013) Direct-acting antiviral agents for hepatitis C virus infection. Annu Rev Pharmacol Toxicol 53:427–449. doi:10.1146/annurev-pharmtox-011112-140254
8. Degasperi E, Aghemo A (2014) Sofosbuvir for the treatment of chronic hepatitis C: between current evidence and future perspectives. Hepatic Med 6:25–33. doi:10.2147/HMER.S44375
9. Plaza Z, Soriano V, Gonzalez Mdel M, Di Lello FA, Macias J, Labarga P, Pineda JA, Poveda E (2011) Impact of antiretroviral therapy on the variability of the HCV NS5B polymerase in HIV/HCV co-infected patients. J Antimicrob Chemother 66:2838–2842. doi:10.1093/jac/dkr385
10. Ago H, Adachi T, Yoshida A, Yamamoto M, Habuka N, Yatsunami K, Miyano M (1999) Crystal structure of the RNA-dependent RNA

polymerase of hepatitis C virus. Structure 7:1417–1426. doi:10.1016/S0969-2126(00)80031-3
11. Bressanelli S, Tomei L, Roussel A, Incitti I, Vitale RL, Mathieu M, De Francesco R, Rey FA (1999) Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. Proc Natl Acad Sci USA 96:13034–13039. doi:10.1073/pnas.96.23.13034
12. Lesburg CA, Cable MB, Ferrari E, Hong Z, Mannarino AF, Weber PC (1999) Crystal structure of the RNA-dependent RNA polymerase from hepatitis C virus reveals a fully encircled active site. Nat Struct Biol 6:937–943. doi:10.1038/13305
13. Rosnoblet C, Fritzinger B, Legrand D, Launay H, Wieruszeski JM, Lippens G, Hanoulle X (2012) Hepatitis C virus NS5B and host cyclophilin A share a common binding site on NS5A. J Biol Chem 287:44249–44260. doi:10.1074/jbc.M112.392209
14. Boyce SE, Tirunagari N, Niedziela-Majka A, Perry J, Wong M, Kan E, Lagpacan L, Barauskas O, Hung M, Fenaux M, Appleby T, Watkins WJ, Schmitz U, Sakowicz R (2014) Structural and regulatory elements of HCV NS5B polymerase-beta-loop and C-terminal tail-are required for activity of allosteric thumb site II inhibitors. PLoS One 9:e84808. doi:10.1371/journal.pone.0084808
15. Antonysamy SS, Aubol B, Blaney J, Browner MF, Giannetti AM, Harris SF, Hebert N, Hendle J, Hopkins S, Jefferson E, Kissinger C, Leveque V, Marciano D, McGee E, Najera I, Nolan B, Tomimoto M, Torres E, Wright T (2008) Fragment-based discovery of hepatitis C virus NS5b RNA polymerase inhibitors. Bioorg Med Chem Lett 18:2990–2995. doi:10.1016/j.bmcl.2008.03.056
16. Beaulieu PL, Bos M, Cordingley MG, Chabot C, Fazal G, Garneau M, Gillard JR, Jolicoeur E, LaPlante S, McKercher G, Poirier M, Poupart MA, Tsantrizos YS, Duan J, Kukolj G (2012) Discovery of the first thumb pocket 1 NS5B polymerase inhibitor (BILB 1941) with demonstrated antiviral activity in patients chronically infected with genotype 1 hepatitis C virus (HCV). J Med Chem 55:7650–7666. doi:10.1021/jm3006788
17. Beaulieu PL, Coulombe R, Duan J, Fazal G, Godbout C, Hucke O, Jakalian A, Joly MA, Lepage O, Llinas-Brunet M, Naud J, Poirier M, Rioux N, Thavonekham B, Kukolj G, Stammers TA (2013) Structure-based design of novel HCV NS5B thumb pocket 2 allosteric inhibitors with submicromolar gt1 replicon potency: discovery of a quinazolinone chemotype. Bioorg Med Chem Lett 23:4132–4140. doi:10.1016/j.bmcl.2013.05.037
18. Biswal BK, Wang M, Cherney MM, Chan L, Yannopoulos CG, Bilimoria D, Bedard J, James MN (2006) Non-nucleoside inhibitors binding to hepatitis C virus NS5B polymerase reveal a novel mechanism of inhibition. J Mol Biol 361:33–45. doi:10.1016/j.jmb.2006.05.074
19. Chan L, Das SK, Reddy TJ, Poisson C, Proulx M, Pereira O, Courchesne M, Roy C, Wang W, Siddiqui A, Yannopoulos CG, Nguyen-Ba N, Labrecque D, Bethell R, Hamel M, Courtemanche-Asselin P, L'Heureux L, David M, Nicolas O, Brunette S, Bilimoria D, Bedard J (2004) Discovery of thiophene-2-carboxylic acids as potent inhibitors of HCV NS5B polymerase and HCV subgenomic RNA replication. Part 1: Sulfonamides. Bioorg Med Chem Lett 14:793–796. doi:10.1016/j.bmcl.2003.10.067
20. Chan L, Pereira O, Reddy TJ, Das SK, Poisson C, Courchesne M, Proulx M, Siddiqui A, Yannopoulos CG, Nguyen-Ba N, Roy C, Nasturica D, Moinet C, Bethell R, Hamel M, L'Heureux L, David M, Nicolas O, Courtemanche-Asselin P, Brunette S, Bilimoria D, Bedard J (2004) Discovery of thiophene-2-carboxylic acids as potent inhibitors of HCV NS5B polymerase and HCV subgenomic RNA replication. Part 2: tertiary amides. Bioorg Med Chem Lett 14:797–800. doi:10.1016/j.bmcl.2003.10.068
21. Chan L, Reddy TJ, Proulx M, Das SK, Pereira O, Wang W, Siddiqui A, Yannopoulos CG, Poisson C, Turcotte N, Drouin A, Alaoui-Ismaili MH, Bethell R, Hamel M, L'Heureux L, Bilimoria D, Nguyen-Ba N (2003) Identification of N, N-disubstituted phenylalanines as a novel class of inhibitors of hepatitis C NS5B polymerase. J Med Chem 46:1283–1285. doi:10.1021/jm0340400
22. Ding Y, Smith KL, Varaprasad CV, Chang E, Alexander J, Yao N (2007) Synthesis of thiazolone-based sulfonamides as inhibitors of HCV NS5B polymerase. Bioorg Med Chem Lett 17:841–845. doi:10.1016/j.bmcl.2006.08.104
23. Gopalsamy A, Aplasca A, Ciszewski G, Park K, Ellingboe JW, Orlowski M, Feld B, Howe AY (2006) Design and synthesis of 3,4-dihydro-1H-[1]-benzothieno[2,3-c]pyran and 3,4-dihydro-1H-pyrano[3,4-b]benzofuran derivatives as non-nucleoside inhibitors of HCV NS5B RNA dependent RNA polymerase. Bioorg Med Chem Lett 16:457–460. doi:10.1016/j.bmcl.2005.08.114
24. Gopalsamy A, Lim K, Ciszewski G, Park K, Ellingboe JW, Bloom J, Insaf S, Upeslacis J, Mansour TS, Krishnamurthy G, Damarla M, Pyatski Y, Ho D, Howe AY, Orlowski M, Feld B, O'Connell J (2004) Discovery of pyrano[3,4-b]indoles as potent and selective HCV NS5B polymerase inhibitors. J Med Chem 47:6603–6608. doi:10.1021/jm0401255
25. Gopalsamy A, Shi M, Ciszewski G, Park K, Ellingboe JW, Orlowski M, Feld B, Howe AY (2006) Design and synthesis of 2,3,4,9-tetrahydro-1H-carbazole and 1,2,3,4-tetrahydrocyclopenta[b]indole derivatives as non-nucleoside inhibitors of hepatitis C virus NS5B RNA-dependent RNA polymerase. Bioorg Med Chem Lett 16:2532–2534. doi:10.1016/j.bmcl.2006.01.105
26. Jackson RW, LaPorte MG, Herbertz T, Draper TL, Gaboury JA, Rippin SR, Patel R, Chunduru SK, Benetatos CA, Young DC, Burns CJ, Condon SM (2011) The discovery and structure-activity relationships of pyrano[3,4-b]indole-based inhibitors of hepatitis C virus NS5B polymerase. Bioorg Med Chem Lett 21:3227–3231. doi:10.1016/j.bmcl.2011.04.052
27. Kucukguzel I, Satilmis G, Gurukumar KR, Basu A, Tatar E, Nichols DB, Talele TT, Kaushik-Basu N (2013) 2-Heteroarylimino-5-arylidene-4-thiazolidinones as a new class of non-nucleoside inhibitors of HCV NS5B polymerase. Eur J Med Chem 69:931–941. doi:10.1016/j.ejmech.2013.08.043
28. Kumar DV, Rai R, Brameld KA, Riggs J, Somoza JR, Rajagopalan R, Janc JW, Xia YM, Ton TL, Hu H, Lehoux I, Ho JD, Young WB, Hart B, Green MJ (2012) 3-heterocyclyl quinolone inhibitors of the HCV NS5B polymerase. Bioorg Med Chem Lett 22:300–304. doi:10.1016/j.bmcl.2011.11.013
29. Kumar DV, Rai R, Brameld KA, Somoza JR, Rajagopalan R, Janc JW, Xia YM, Ton TL, Shaghafi MB, Hu H, Lehoux I, To N, Young WB, Green MJ (2011) Quinolones as HCV NS5B polymerase inhibitors. Bioorg Med Chem Lett 21:82–87. doi:10.1016/j.bmcl.2010.11.068
30. LaPorte MG, Draper TL, Miller LE, Blackledge CW, Leister LK, Amparo E, Hussey AR, Young DC, Chunduru SK, Benetatos CA, Rhodes G, Gopalsamy A, Herbertz T, Burns CJ, Condon SM (2010) The discovery and structure-activity relationships of pyrano[3,4-b]indole based inhibitors of hepatitis C virus NS5B polymerase. Bioorg Med Chem Lett 20:2968–2973. doi:10.1016/j.bmcl.2010.03.002
31. Laporte MG, Jackson RW, Draper TL, Gaboury JA, Galie K, Herbertz T, Hussey AR, Rippin SR, Benetatos CA, Chunduru SK, Christensen JS, Coburn GA, Rizzo CJ, Rhodes G, O'Connell J, Howe AY, Mansour TS, Collett MS, Pevear DC, Young DC, Gao T, Tyrrell DL, Kneteman NM, Burns CJ, Condon SM (2008) The discovery of pyrano[3,4-b]indole-based allosteric inhibitors of HCV NS5B polymerase with in vivo activity. ChemMedChem 3:1508–1515. doi:10.1002/cmdc.200800168
32. Laporte MG, Lessen TA, Leister L, Cebzanov D, Amparo E, Faust C, Ortlip D, Bailey TR, Nitz TJ, Chunduru SK, Young DC, Burns CJ (2006) Tetrahydrobenzothiophene inhibitors of hepatitis C virus NS5B polymerase. Bioorg Med Chem Lett 16:100–103. doi:10.1016/j.bmcl.2005.09.047

33. Li H, Linton A, Tatlock J, Gonzalez J, Borchardt A, Abreo M, Jewell T, Patel L, Drowns M, Ludlum S, Goble M, Yang M, Blazel J, Rahavendran R, Skor H, Shi S, Lewis C, Fuhrman S (2007) Allosteric inhibitors of hepatitis C polymerase: discovery of potent and orally bioavailable carbon-linked dihydropyrones. J Med Chem 50:3969–3972. doi:10.1021/jm0704447
34. Li H, Tatlock J, Linton A, Gonzalez J, Borchardt A, Dragovich P, Jewell T, Prins T, Zhou R, Blazel J, Parge H, Love R, Hickey M, Doan C, Shi S, Duggal R, Lewis C, Fuhrman S (2006) Identification and structure-based optimization of novel dihydropyrones as potent HCV RNA polymerase inhibitors. Bioorg Med Chem Lett 16:4834–4838. doi:10.1016/j.bmcl.2006.06.065
35. Li H, Tatlock J, Linton A, Gonzalez J, Jewell T, Patel L, Ludlum S, Drowns M, Rahavendran SV, Skor H, Hunter R, Shi ST, Herlihy KJ, Parge H, Hickey M, Yu X, Chau F, Nonomiya J, Lewis C (2009) Discovery of (R)-6-cyclopentyl-6-(2-(2,6-diethylpyridin-4-yl)ethyl)-3-((5,7-dimethyl-[1,2,4]tr iazolo[1,5-a]pyrimidin-2-yl)methyl)-4-hydroxy-5,6-dihydropyran-2-one (PF-00868554) as a potent and orally available hepatitis C virus polymerase inhibitor. J Med Chem 52:1255–1258. doi:10.1021/jm8014537
36. Malancona S, Donghi M, Ferrara M, Martin Hernando JI, Pompei M, Pesci S, Ontoria JM, Koch U, Rowley M, Summa V (2010) Allosteric inhibitors of hepatitis C virus NS5B polymerase thumb domain site II: structure-based design and synthesis of new templates. Bioorg Med Chem 18:2836–2848. doi:10.1016/j.bmc.2010.03.024
37. Manfroni G, Cannalire R, Barreca ML, Kaushik-Basu N, Leyssen P, Winquist J, Iraci N, Manvar D, Paeshuyse J, Guhamazumder R, Basu A, Sabatini S, Tabarrini O, Danielson UH, Neyts J, Cecchetti V (2014) The versatile nature of the 6-aminoquinolone scaffold: identification of submicromolar hepatitis C virus NS5B inhibitors. J Med Chem 57:1952–1963. doi:10.1021/jm401362f
38. Ontoria JM, Rydberg EH, Di Marco S, Tomei L, Attenni B, Malancona S, Martin Hernando JI, Gennari N, Koch U, Narjes F, Rowley M, Summa V, Carroll SS, Olsen DB, De Francesco R, Altamura S, Migliaccio G, Carfi A (2009) Identification and biological evaluation of a series of 1H-benzo[de]isoquinoline-1,3(2H)-diones as hepatitis C virus NS5B polymerase inhibitors. J Med Chem 52:5217–5227. doi:10.1021/jm900517t
39. Reddy TJ, Chan L, Turcotte N, Proulx M, Pereira OZ, Das SK, Siddiqui A, Wang W, Poisson C, Yannopoulos CG, Bilimoria D, L'Heureux L, Alaoui HM, Nguyen-Ba N (2003) Further SAR studies on novel small molecule inhibitors of the hepatitis C (HCV) NS5B polymerase. Bioorg Med Chem Lett 13:3341–3344. doi:10.1016/S0960-894X(03)00670-X
40. Stammers TA, Coulombe R, Duplessis M, Fazal G, Gagnon A, Garneau M, Goulet S, Jakalian A, LaPlante S, Rancourt J, Thavonekham B, Wernic D, Kukolj G, Beaulieu PL (2013) Anthranilic acid-based thumb pocket 2 HCV NS5B polymerase inhibitors with sub-micromolar potency in the cell-based replicon assay. Bioorg Med Chem Lett 23:6879–6885. doi:10.1016/j.bmcl.2013.09.102
41. Stammers TA, Coulombe R, Rancourt J, Thavonekham B, Fazal G, Goulet S, Jakalian A, Wernic D, Tsantrizos Y, Poupart MA, Bos M, McKercher G, Thauvette L, Kukolj G, Beaulieu PL (2013) Discovery of a novel series of non-nucleoside thumb pocket 2 HCV NS5B polymerase inhibitors. Bioorg Med Chem Lett 23:2585–2589. doi:10.1016/j.bmcl.2013.02.110
42. Yan S, Appleby T, Larson G, Wu JZ, Hamatake RK, Hong Z, Yao N (2007) Thiazolone-acylsulfonamides as novel HCV NS5B polymerase allosteric inhibitors: convergence of structure-based drug design and X-ray crystallographic study. Bioorg Med Chem Lett 17:1991–1995. doi:10.1016/j.bmcl.2007.01.024
43. Yan S, Larson G, Wu JZ, Appleby T, Ding Y, Hamatake R, Hong Z, Yao N (2007) Novel thiazolones as HCV NS5B polymerase allosteric inhibitors: Further designs, SAR, and X-ray complex structure. Bioorg Med Chem Lett 17:63–67. doi:10.1016/j.bmcl.2006.09.095
44. Yang H, Hendricks RT, Arora N, Nitzan D, Yee C, Lucas MC, Yang Y, Fung A, Rajyaguru S, Harris SF, Leveque VJ, Hang JQ, Pogam SL, Reuter D, Tavares GA (2010) Cyclic amide bioisosterism: strategic application to the design and synthesis of HCV NS5B polymerase inhibitors. Bioorg Med Chem Lett 20:4614–4619. doi:10.1016/j.bmcl.2010.06.008
45. Accelrys (2007) Discovery Studio Modeling Environment, Release 3.5 Accelrys Software Inc, San Diego
46. Accelrys Draw. Accelrys Software Inc, San Diego
47. Smellie A, Teig SL, Towbin P (1995) Poling: promoting conformational variation. J Comput Chem 16:171–187. doi:10.1002/jcc.540160205
48. Guner OF (2000) Pharmacophore Perception, Development, and Use in Drug Design (Iul Biotechnology Series, 2). International University Line, La Jolla. doi:10.3390/50700987
49. ChemDiv, 6605 Nancy Ridge Dr, San Diego, CA 92121. www.chemdiv.com
50. InterBioScreen 119019 Moscow. www.ibscreen.com/indexhtm
51. Canales E, Carlson JS, Appleby T, Fenaux M, Lee J, Tian Y, Tirunagari N, Wong M, Watkins WJ (2012) Tri-substituted acylhydrazines as tertiary amide bioisosteres: HCV NS5B polymerase inhibitors. Bioorg Med Chem Lett 22:4288–4292. doi:10.1016/j.bmcl.2012.05.025
52. Jones G, Willett P, Glen RC (1995) Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. J Mol Biol 245:43–53. doi:10.1016/S0022-2836(95)80037-9
53. Jones G, Willett P, Glen RC, Leach AR, Taylor R (1997) Development and validation of a genetic algorithm for flexible docking. J Mol Biol 267:727–748. doi:10.1006/jmbi.1996.0897
54. Blight KJ, Kolykhalov AA, Rice CM (2000) Efficient initiation of HCV RNA replication in cell culture. Science 290:1972–1974. doi:10.1126/science.290.5498.1972
55. Nonlinear regression with Dose response inhibition model was performed using GraphPad Prism version 5.00 for Windows. GraphPad Software, 7825 Fay Ave #230, La Jolla, CA 92037. www.graphpad.com
56. Xue W, Jiao P, Liu H, Yao X (2014) Molecular modeling and residue interaction network studies on the mechanism of binding and resistance of the HCV NS5B polymerase mutants to VX-222 and ANA598. Antivir Res 104:40–51. doi:10.1016/j.antiviral.2014.01.006
57. Golub AG, Gurukumar KR, Basu A, Bdzhola VG, Bilokin Y, Yarmoluk SM, Lee JC, Talele TT, Nichols DB, Kaushik-Basu N (2012) Discovery of new scaffolds for rational design of HCV NS5B polymerase inhibitors. Eur J Med Chem 58:258–264. doi:10.1016/j.ejmech.2012.09.010