# Ligand and Structure Based Virtual Screening Strategies for Hit-Finding and Optimization of Hepatitis C Virus (HCV) Inhibitors

G. Melagraki* and A. Afantitis*

*Department of Cheminformatics, Novamechanics Ltd, Nicosia, Cyprus*

**Abstract:** Virtual Screening (VS) has experienced increased attention into the recent years due to the large datasets made available, the development of advanced VS techniques and the encouraging fact that VS has contributed to the discovery of several compounds that have either reached the market or entered clinical trials. Hepatitis C Virus (HCV) nonstructural protein 5B (NS5B) has become an attractive target for the development of antiviral drugs and many small molecules have been explored as possible HCV NS5B inhibitors. In parallel with experimental practices, VS can serve as a valuable tool in the identification of novel effective inhibitors. Different techniques and workflows have been reported in literature with the goal to prioritize possible potent hits. In this context, different virtual screening strategies have been deployed for the identification of novel Hepatitis C Virus (HCV) inhibitors. This work reviews recent applications of virtual screening in an effort to identify novel potent HCV inhibitors.

**Keywords:** Ligand-based and structure-based virtual screening, quantitative structure-activity relationships (QSAR), docking, Hepatitis C Virus (HCV) inhibitors, NS5B.

## INTRODUCTION

Hepatitis C Virus (HCV) is the etiological agent of non-A and non-B hepatitis. HCV is a serious public health threat as chronic HCV infections can cause liver cirrhosis and hepatocellular carcinoma [1, 2]. Although it is estimated that 200 million cases of HCV infections exist worldwide at present neither a vaccine nor an effective therapy with broad spectrum mode of action is so far available [3]. Currently the anti-HCV therapy is based primarily on the prescription of interferon-α (IFN) alone or in combination with the nucleoside inhibitor ribavirin (RBV) [4]. Due to low therapeutic effects and associated side effects of these strategies, such as depression and anemia, the development of more effective and safer anti-HCV agents has become an imperative necessity. At present around 50% of HCV genotype-1 infected patients fail to achieve a sustained virological response [5]. Intensive efforts have been devoted toward the discovery of innovative HCV antiviral agents that are less toxic and more potent.

Hepatitis C Virus (HCV) is a member of the *Flaviviridae* family and is a positive-sense single stranded RNA virus with a genome that encodes a single large polyprotein of approximately 3000 amino acids with four structural proteins located in the N terminus and six nonstructural proteins encoded in the remainder. The development of HCV antiviral agents has focused mainly on HCV non structural protein NS3/4A serine protease and NS5B RNA polymerase. NS5B is a RNA-dependent RNA-polymerase which has been extensively characterized at the biochemical and structural level [6, 7]. Multiple crystallographic structures of HCV NS5B have been solved to date and reveal a common polymerase 3D structure that resembles a right hand made up of fingers, thumb and palm domains. The non structural protein 5B (NS5B) is identified in many studies as an ideal drug target for combating HCV. A number of anti-NS5B compounds, categorized as nucleoside inhibitors (NIs) and non-nucleoside inhibitors (NNIs), have been identified by high throughput screening (HTS) assays [8]. NIs are modified nucleosides which can be used as substrates at the polymerase active site to compete with natural nucleosides acting as non-obligate chain terminators. The effectiveness of NIs has been reported to be compromised due to the generation of resistant mutants and the adverse side-effects. The NIs can be antagonistic when used with Ribavirin making them less viable candidates for such a combination therapy. On the other hand, NNIs bind to allosteric sites distinct from the active site and have begun to be identified through HTS and crystallographic analysis of the inhibitor-NS5B complex. NNIs are diverse small molecules. They are non-competitive inhibitors that target the alloenzyme free of substrate and uncomplexed with any other non-structural replicative proteins. Many different molecular classes have been reported to possess promising inhibition potency and include compounds such as benzimidazole, indole, thiophene, dihydropyranone, purrolidine, benzothiadiazine, thiazolone, coumestan, and benzylidene analogs [9, 10]. In the past few years a number of novel, specifically target antiviral therapies have entered clinical trials [11], including HCV protease inhibitors from Schering-Plough (boceprevir; SCH 503034) and Vertex (telaprevir; VX-950), and HCV polymerase inhibitors from Roche Laboratories (R1626).

Despite the fact that a great variety of molecular scaffolds has been explored the development of novel HCV inhibitors is still in ongoing progress with the aim to achieve the promising inhibition potency in both enzyme and replicon assays combined with favorable pharmacokinetics profiles and devoid of off-target activities.

For this purpose computational techniques can serve as a useful aid in minimizing the cost and time required to achieve this goal. In order to facilitate the drug discovery process, in combination with experimental practices, virtual screening is a very powerful method for the identification of lead compounds for target proteins [12-17]. *In silico* approaches have emerged as complementary approaches to costly and time-consuming experiments with the goal to either to prioritize active compounds from a library of druglike molecules or to design novel molecular scaffolds based on the explored structure – activity relationships [18-26].

## VIRTUAL SCREENING

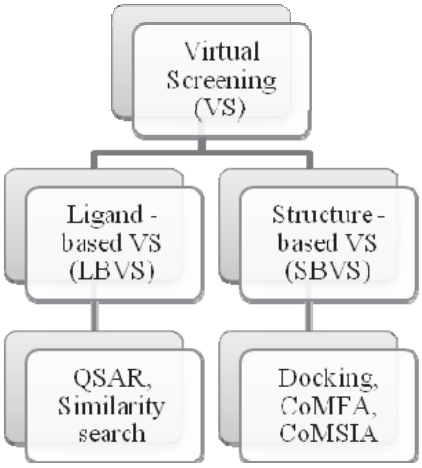
Virtual Screening is the computational analogue of High Throughput Screening and refers to the *in silico* evaluation of properties, such as activity, of different molecular scaffolds. Different applications of machine learning to virtual screening have been presented in the literature including both ligand-based similarity searching and structure-based docking. The main purpose of such applications is to prioritize databases of molecules as active against a particular protein target [27-31].

As presented by D.E. Clark [32] the earliest paper to use the phrase 'virtual screening' was published in 1997 and after that, an impressive growth of publications is demonstrated (data taken from 1997 until 2008) describing the development and application of virtual screening technologies. Although it is still too early for a conclusive analysis, some case studies presented in the same perspective suggest that VS has already played a significant role in

*Address correspondence to this author at the Department of Cheminformatics, Novamechanics Ltd, Nicosia, Cyprus; Tel: +357 99048039; Fax: +357 22347772; E-mails: melagraki@novamechanics.com, afantitis@novamechanics.com

0929-8673/11 $58.00+.00 

the discovery of some compounds that are now in the clinic or even on the market. This remark does not of course suggest that there are no challenges left to overcome by virtual screening practitioners. There is an increasing need of integrating several information resources and building a decision system to support the consensus outcome of several virtual screening strategies. The virtual screening continues to evolve and meanwhile it is wise to explore as many approaches as possible to maximize the likelihood of retrieving the greatest number of novel active chemotypes. Different machine learning algorithms are developed towards this goal with the aim to increase reliability and hit rates. When large databases with virtual structures or de novo design of compounds are involved in the virtual screening process special care should be given to the synthetic accessibility of proposed compounds. Some of the prioritized structures could be quite complicated and therefore difficult to synthesize if a demanding long-step synthesis with carefully controlled reactions is required. Synthetic chemists and chemoinformaticians working in close collaboration will ensure synthetic accessibility of proposed compounds and subsequent success of hit finding.




**Scheme 1.** Virtual screening techniques to develop new drug bullets.

Virtual Screening techniques used for the identification of HCV inhibitors can be subdivided in two main approaches: Ligand-Based Virtual Screening (LBVS) and Structure-Based Virtual Screening (SBVS) presented in Scheme **1** [33, 34]. For the first approach biological data are explored to identify known active or inactive compounds that will be used to retrieve other potentially active molecular scaffolds based on similarity measures, common pharmacophore or descriptor values. Machine Learning is quickly gaining popularity in LBVS as novel algorithms are proposed to build accurate and robust quantitative structure-activity relationships [35]. Different techniques are proposed and each method has its own advantages and disadvantages. Among these methods regression and classification methods such as Multiple Linear Regression, Nearest Neighbors, Naïve Bayesian Classification, Support Vector Machines, Neural Networks and Decision Trees have been successfully applied. These methods and their applications are reviewed in a recent publication by Melville *et al*. [36]. For the second approach the protein structure of interest is available and novel molecular patterns are explored by docking into the active site of the biochemical target using computer algorithms and scoring functions. Depending on the available information both methods can be used individually or in combination. None of these proposed approaches can be a priori considered superior to the other as it is suggested in the literature that depending on the case the one can outperform the other. Muegge and Oloff [37] highlighted some recent success stories of virtual screening [38] using both LBVS and SBVS by making the best use of several informational sources: i.e. pharmacophore models using known ligands are combined with homology models, QSAR models are combined with docking approaches. For both approaches there are active fields for improvements. For instance, in SBVS scoring functions, treatment of protein flexibility and treatment of water molecules are some of the current issues whereas in LBVS new descriptors that facilitate scaffold hopping and the definition of the applicability domain are active fields of research.

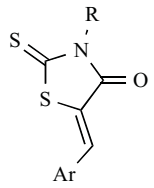
One of the basic preconditions of developing either of the aforementioned approaches is to ensure the availability of reliable bioassay datasets. Available libraries provide great opportunities and also challenges for computational studies to reveal relationships between chemical structures and biological activities in order to assist virtual screening and computer added drug design. Such a valuable library that was recently exploited by X-Q S Xie [39] for virtual screening applications is PubChem. PubChem is a valuable public molecular information resource repository that contains millions of compound records and bioassay data collections and has attracted many chemoinformaticians to carry out *in silico* studies for identifying novel hits with new chemical scaffolds. As reviewed in this work, PubChem has several advantages although it also exhibits some limitations. Novel data-mining cheminformatics tools and virtual screening algorithms are being developed and used to retrieve, annotate and analyze the large-scale and highly complex PubChem bioactivity data in order to facilitate computer-aided virtual screening.

## *IN SILICO* STRATEGIES – IDENTIFYING HCV INHIBITORS

Several studies have been reported in the literature with the aim of identifying novel potent HCV inhibitors using different *in silico* protocols and workflows. Some of the recent applications for the identification of HCV NS5B polymerase inhibitors are summarized below. The methods described follow either the Structure-based or the Ligand-based virtual screening approach using different libraries of compounds, computational tools and combinations which are analyzed below.

In their recent publication, Talele *et al*. [40] presented a virtual screening workflow to identify and further design novel HCV NS5B polymerase inhibitors. The proposed methodology employs structure-based virtual screening, synthesis and structure – activity relationship (SAR) optimization approach. A ChemBridge database including 260000 compounds was initially used as source to identify novel potent HCV inhibitors. As a first step this database was initially filtered according to several criteria such as Lipinski's Rule of Five, use of no tautomers, removal of metals from salts and one stereoistomer per ligand. These filtering criteria resulted in filtering out approximately 80% of the compounds included in the original dataset. The 3D structures of the remaining 52000 compounds were docked in allosteric pocket AP-1 of NS5B polymerase (PDB ID: 2DXS) employing the Glide's High Throughput Virtual Screening (HTVS) workflow. This docking screening filtered out several compounds and resulted in a set of 650 compounds. These compounds were then further analyzed in the context of different parameters that imply improbable docking orientations in AP-1 of NS5B, yielding a set of 23 compounds bearing structural diversity and good docking orientations for biological evaluation. An *in vitro* NS5B RNA-dependent RNA-polymerase (RdRp) was carried out and revealed that among these compounds, two were worth further SAR investigation. Among the two only compound **1**, presented in Table **1**, gave active analogs. In particular, commercially available rhodanine analogs were used for 3- and 5- position exploration leading to the identification of potent analog **2**. More non-commercially available rhodanine analogs were further designed, synthesized and tested and more potent compounds such as **3** and **4** were identified. Docking analysis of

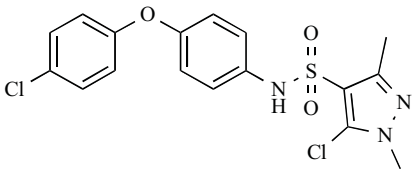
**Table 1. Potent HCV NS5B Inhibitors Proposed By Ref. [40]**



| Compound | R | Ar | $IC_{50}$(μM) |
|---|---|---|---|
| **1** | -$CH_2$COOH | 4-$CF_3$Ph | 55.2 ± 1.10 |
| **2** | -CH(COOH)$CH_2$Ph | 2,4-ClPh | 10.6 ± 1.50 |
| **3** | -$CH_2$COOH | 3,4-ClPh | 8.4 ± 1.50 |
| **4** | -$CH_2$COOH | 3-PhenoxyPh | 7.7 ± 1.40 |

compound **4** with the tetracyclic indole- and benzylidene- binding allosteric pockets (AP-1 and AP-3, respectively) of NS5B revealed topological similarities between these two pockets. Compound **4**, a novel rhodanine analog with NS5B inhibitory potency of the low micromolar level range is proposed as a promising lead for future development of more potent NS5B inhibitors.

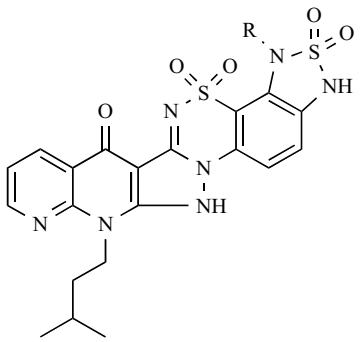
Kim *et al*. [41] performed a pharmacophore – guided virtual screening in order to identify novel anti-HCV diketoacid replacements through UNITY-based pharmacophore search of a database which was then used for docking-based virtual screening. Based on their previous work on the binding mode analysis, structure-based 3D-QSAR study and pharmacophore-guided docking, they contacted a rational strategy to find novel potent HCV inhibitors. LeadQuest compound library implemented in SYBYL package was used to initiate the proposed virtual screening procedure. Three pharmacophore models were used by UNITY search for filtering the compound library. UNITY module uses a conformationally flexible 3D-searching algorithm to result in rapid identification of molecules that match with the given pharmacophore. In this study using the LeadQuest database and the three generated pharmacophore models, 183, 32 and 12 molecules were selected. Among the 227, only 103 were successfully docked using the FlexX-Pharm module in SYBYL 7.2 package and a 3D crystal structure of HCV RdRp (PDB ID: 1GX6). These selected compounds with the reasonable poses were then scored using the CScore module of SYBYL which comprises five different scoring functions (Dock, Chem, FlexX, PMF and Gold) and the top 40 compounds with the highest total score were selected for testing. Among those, compound **5** shown in Scheme **2**, which was in perfect match with the pharmacophore model as well as the hydrophobic hole in the aryl-binding site, showed potent anti-HCV activity in the cell-based assay (1.2% relative cell viability in comparison with the control).



**Scheme 2.** Potent HCV NS5B inhibitor proposed by ref. [41].

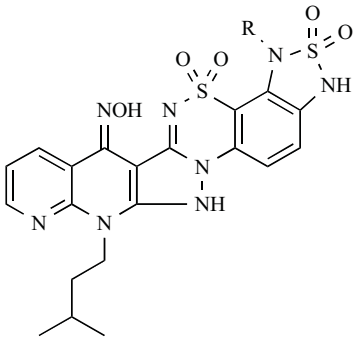
Melagraki *et al*. [42] presented the results of a ligand-based virtual screening optimized procedure on compounds evaluated as inhibitors of genotype 1 HCV polymerase (benzothiadiazines). A decision support system is presented for the design of synthetically accessible HCV inhibitors. As a first step, quantitative structure–activity patterns were investigated for the selected compounds. An accurate and reliable QSAR model involving five descriptors that is able to predict successfully the HCV inhibitory potency against genotype 1 HCV polymerase was presented. These five descriptors account for the electronic, topological and phsicochemical information of molecules and describe and model successfully the binding affinity of the small molecules contained in the initial dataset. The model was fully validated using different validation techniques. In a next step structural modifications were proposed to afford novel active patterns. The effects of various structural modifications on biological activity were investigated and biological activities of novel structures were estimated using the developed QSAR model. More specifically a search for optimized pharmacophore patterns by insertions, substitutions, and ring fusions of pharmacophoric substituents of the main building block scaffolds was described. The detection of the domain of applicability defined compounds whose estimations could be accepted with confidence. Some of the proposed active compounds (**6-15**) that fall within the domain of applicability are presented in Scheme **3**.

A new effort to identify novel molecular scaffolds previously untested as NS5B inhibitors is described in the recent work of Musmuca *et al*. [43] The proposed protocol includes 3D-QSAR, Ligand-based (LB) and Structure-based (SB) alignment methods, a LB-SB virtual screening (LB-SB-VS) and biological evaluation. First structure-based 3D-QSAR models were generated employing NS5B non-nucleoside inhibitors (NNIs). Among the 27 NS5B/NNIs complexes available in PDB, two groups were selected as training sets: 15 NNIs that bind to the enzyme thumb domain and 10 NNIs that bind to the enzyme palm domain. All the subsequent molecular docking simulations and ligand based alignments were carried with these training sets. First, the GRID/GOLPE method was used to define two final SB 3D-QSAR models that were internally and externally validated using several statistical criteria. Ligand based (LB) and structure based (SB) alignments methodologies were rigorously investigated with the aim to assess the reliability on the correct molecular alignment for unknown binding mode modeled compounds. For the LB approach the application of the concept of morphological similarity implemented in Surflex was used while for the SB approach Autodock program was selected. Both Surflex and Autodock programs were able to reproduce with minimal errors the experimental binding conformations of 24 experimental NS5B allosteric inhibitors. Eighty-one (thumb) and 223 (palm) modeled compounds taken from literature were LB and SB aligned and used as external validation sets for the development of 3-D QSAR models. Low error of prediction proved the 3-D QSARs to be useful scoring functions for the *in silico* screening procedure. Finally, the virtual screening of the NCI Diversity Set led to the selection for enzymatic assays of 20 top-scoring molecules for each final model. Among the 40 selected molecules, preliminary data yielded four derivatives exhibiting $IC_{50}$ values ranging between 45 and 75 μM.



R=Me, Et, n-Pr, i-Pr, n-Bu, s-Bu

**6-11**

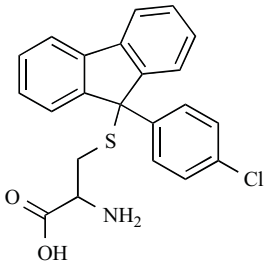
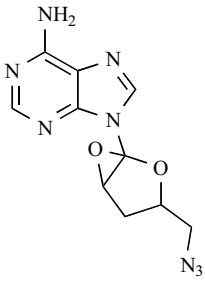
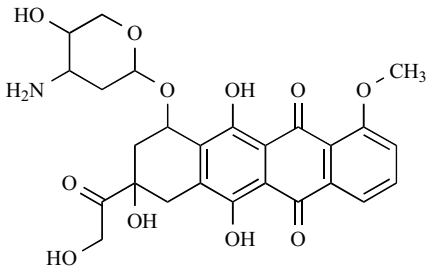
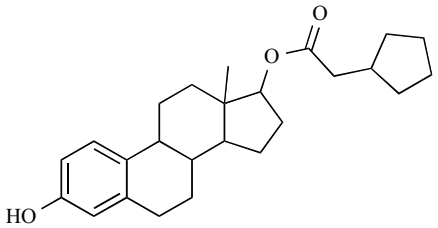


R=n-Pr, vinyl, n-Bu, s-Bu

**12-15**

**Scheme 3.** Potent HCV NS5B inhibitors proposed by ref. [42].

**Table 2. Potent HCV NS5B Inhibitors Proposed by Ref. [43]**

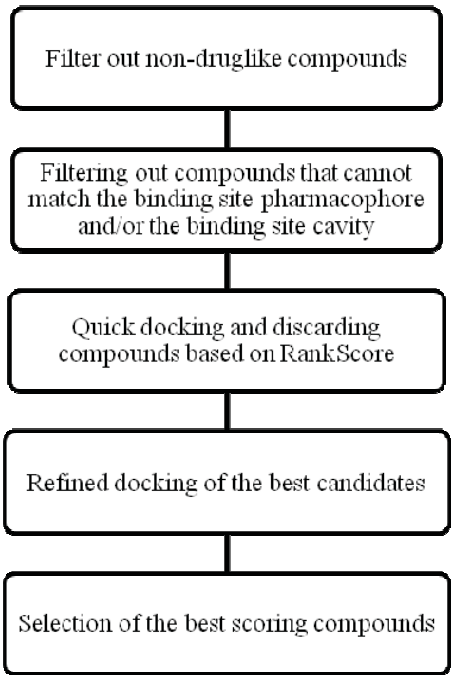
| **Thumb Domain** | |
|---|---|
|  |  |
| **16** (NSC 123526), $IC_{50}$ = 46μM | **17** (NSC 125626), $IC_{50}$ = 46μM |
| **Palm Domain** | |
|  |  |
| **18** (NSC 169534), $IC_{50}$ = 64.5μM | **19** (NSC 3354), $IC_{50}$ = 54.3μM |

The structures are shown in Table **2**. Binding mode analysis of hit compounds within the NS5B polymerase thumb domain showed that one of them, compound **16**, exhibited a docked conformation which was in good agreement with the thumb training set most active compound.

A recently reported docking program, FITTED v1.5, was used in the work of Corbeil *et al*. [44] in an effort to identify potential HCV polymerase inhibitors. The program was fully validated focusing on HCV inhibitors and the results revealed good accuracy mainly due to the consensus docking approach implemented which is expected to accurately filter out unreasonable poses. The validation also demonstrated the key role of protein flexibility accounted for by FITTED in the HCV polymerase context. To demonstrate the ability to identify active compounds within a large set, the Maybridge library of druglike molecules which was obtained from the ZINC Web site was seeded with 23 known actives ranging from nanomolar to micromolar activities and used based on the following protocol. The automated screening protocol included five steps is shown in Scheme **4**.

Based on this procedure, high computed enrichment factors indicated the importance of the method. Top scoring compounds from Maybridge Library were screened in HCV polymerase assays. This resulted in the identification of two active compounds which exhibited $IC_{50}$ values of 7 and 12 μM respectively.

In their recent work Ryu *et al*. [45] investigated potential inhibitor binding pockets of NS5B distinct from the nucleoside binding site and performed virtual screening of compounds that fit this binding pocket from an available database containing millions of compounds. As a result of their analysis they found that the thumb and palm domain interface is a promising inhibitor binding pocket and a pharmacophore map was generated for structure-based virtual screening. Feature-based pharmacophore models were constructed in which the pharmacophoric points are represented by chemical features such as hydrogen bond acceptors/donors or hydrophobic features. The pharmacophore model was used as a search query for identifying inhibitors from the 3-D small molecule database PharmoDB$^{TM}$ that contains 3.5million compounds. After filtering out the primary screened compounds the test compounds



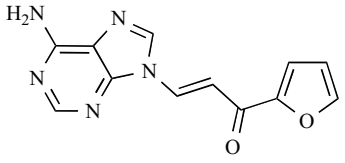

**Scheme 4.** Protocol adopted by Corbeil *et al.* to identify active HCV inhibitors.

were down sized from 3.5 million to 119. As a next step it was investigated whether the virtually screened compounds inhibit RdRp activity *in vitro*. From the pool of tested compounds, three promising compounds (**20-22**) shown in Scheme **5** exhibited $IC_{50}$ values of about 20μM and their structures are illustrated below. For a refined docking model of compound **20** into the HCV NS5B an automated docking simulation model was performed using AUTODOCK and X-ray crystal structure of NS5B of HCV genotype 1b from PDB (2AX0).

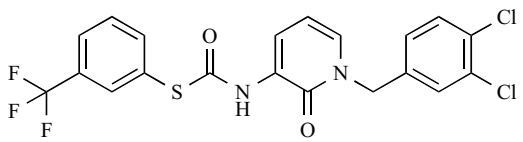
S. Louise-May *et al*. [46] identified novel potent diakyl substituted thiophene inhibitors of HCV *via* an *in silico* screening procedure. Available lead-like compounds were docked in the thumb domain of HCV NS5B in order to identify novel putative ligands. The 3D structure of HCV NS5B (1GX6) and Glide program were used for the docking and a combined database with lead-like compounds from the Specs & BioSpecs, Bionet, Microsource, Available Chemicals Directory (ACD), Nanoscale, ChemDiv, Orion, Asinex, Interbioscreen, Timtec and Cambridge databases was used for virtual screening. The database was refined according to several filters and finally a library of 90000 compounds was used. The top 1318 ligands with a gride score of -7.17 or better were visually inspected and finally a set of 50 structurally diverse ligands was purchased. The selected compounds were evaluated in an HCV replicase complex (RC) assay (cell free) and HCV replicon assay (cell-based). The most active compound (**23**, $IC_{50}$ 50-100μM) is shown in Scheme **6**.

This active compound **23** was followed-up by applying substructure and similarity searches and in-house synthesis of analogs. When the amide function was replaced by a cyano group (**24**) an improvement in potency was observed in the RC assay ($IC_{50}$ = 25μM). Ethyl ester and carboxylic acid at the 3-position were also tested but proven inactive. Loss of activity was also observed for simple substitutions of a methyl group for an ethyl at the 4 position. Activity of compound (**25**) was also found comparable with compound (**23**) in the range of 30-50μM. The compounds active in the RC assay were also evaluated in the HCV replicon assay. As suggested, lower than expected $EC_{50}$ values are partly derived from the cellular toxicity

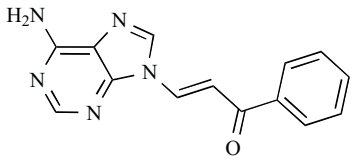
An interesting work was recently presented by Li *et al*. [47] Molecular dynamics simulations, free energy decomposition and docking were used in an effort to investigate and compare five different molecular scaffolds of HCV NS5B inhibitors. As many X-ray structures of HCV NS5B polymerase complexed with NNIs are publicly available, the authors selected five of them (1YVF, 2GC8, 2GIQ, 2JC0, 2QE5) based on the following criteria: the protein is HCV NS5B polymerase genotype 1b, the NNIs all bind to the same palm binding site and these five NNIs represent different scaffolds. In particular the NNIs included are: acrylic acid, proline sulfonamide, thiadiazine derivative, acryl pyrrolidine and athranilic acid derivative. The binding free energies were computed using the molecular mechanics generalized Brn surface area (MM/GBSA) method. A key finding of the research was that Tyr448 plays the most critical role in the binding of most inhibitors. Detailed analysis suggested that major contributions favorable to binding site are vdW and electrostatic energies, whereas polar solvation energies opposed the binding. Nonpolar salvation energies contribute slightly favorably. Moreover, the important residues for NNI binding were identified by the free energy decomposition. In a next step, authors explored the feasibility of docking-based drug design in the binding site. An optimized docking protocol using AUTODOCK program was presented based on cross-docking the five inhibitors in the palm binding site of the enzyme. The protein model of 2GC8 was selected as a representative protein for further screening research. Two diversity sets: NCI diversity set and HitFinder collection of Maybridge library were used for virtual screening based on the optimized docking protocol. After the analysis two criteria were proposed to select candidate hits. The first is that there are at least two stable hydrogen bonds formed



**20 (CVP09011)**

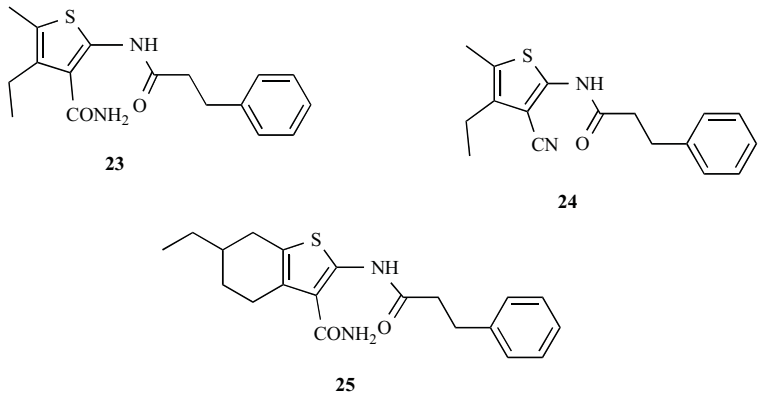


**21 (CVP09047)**



**22 (CVP09081)**

**Scheme 5.** Potent HCV NS5B inhibitors proposed by ref. [45].

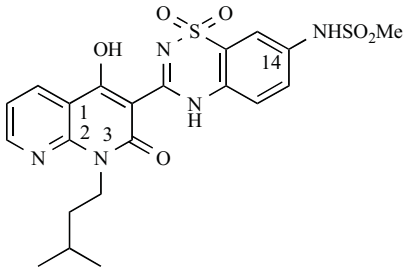


**23** **24** **25**

**Scheme 6.** Potent HCV NS5B inhibitors proposed by ref. [46].

between the compound and any two residues of Cys366, Ser367, Arg386, Tyr415, Gln446, Tyr448, Gln449 and Ser556. The second is that there should be tight binding in the deep hydrophobic pocket surrounded by Pro197, Arg200, Cys366, Met414 and Tyr448 such as strong vdW or electrostatic interaction.

In their recent work Wang *et al*. [48] presented a ligand- and receptor- based 3D-QSAR study of 239 benzothiadiazine inhibitors of HCV NS5B polymerase using CoMFA and CoMSIA tools. Several statistical criteria were used to prove the robustness and reliability of the produced models. The contour diagrams produced by CoMFA and CoMSIA approach gave a good insight about the intermolecular interactions of inhibitors with the surrounding environment. Molecular docking and Molecular dynamics simulations were also performed for the rational development of novel inhibitors using the X-ray structure of GT-1b NS5B (PDB 3H98). 3D-QSAR, docking and MD modeling led to the following key findings regarding the preferable substituents in different positions of compound **26** shown in Scheme **7**.



**Scheme 7.** Potent HCV NS5B inhibitor proposed by ref. [48].

For positions 1 and 2 substituents with the size and length of benzo group (i.e. isoster) are preferable whereas substituents with halogens or polar functionality are not. For position 3, substituents containing a linear or branched alkyl chain increase the activity whereas unsaturation and polar functionality limit the potency. Finally for position 14, a polar substituent is preferential.

All the aforementioned workflows proposed in literature combine various resources of known HCV inhibitors with different LBVS and SBVS strategies and available libraries. Although different paths were followed, these approaches were proven successful in prioritizing potent effective compounds that are commercially available or synthetically feasible. In many cases, *in vitro* evaluation confirmed the *in silico* screening results.

## CONCLUSIONS

Although some promising *in vitro* HCV NS5B inhibitors have been proposed in literature, more efforts are needed to rapidly achieve the goal of identifying more potent and less toxic HCV antiviral compounds. In this direction of identifying therapeutic weapons against HCV, structure-based drug design approaches, development of quantitative structure-based activity relationships, docking studies, library searching and *in silico* screening studies are ongoing with the hope to discover new active lead compounds. In this work we have reviewed some of these efforts by highlighting the materials and methods used for achieving this goal and the final promising structures proposed in each work.

As proven by many success stories, virtual screening methods for compound prioritization represent viable alternatives to traditional experimental screening methods in drug discovery. Although in the past several years there has been a tremendous increase in hit finding through virtual screening techniques, there is still an active field of research and several challenges that need to be considered. The availability of a pool of highly functional machine learning algorithms, statistical software and online tools together with abundant data of different origin and format highlight the necessity of integrating different resources. The preparation of compounds available in different libraries is demanding including preprocessing, clean up and standardization filters. Moreover, a great variety of available algorithms for modeling, similarity search, data mining, docking, validation, domain of applicability can be used in combination for a consensus approach depending on the problem addressed. Apart from advancing the algorithms used and proposing efficient computational protocols, automating the steps required in a virtual screening process is a very challenging task [49]. The automation is feasible through the use of data-pipelining technology such as KNIME (www.knime.org) which allows componentized manipulation of data. Compounds included in datasets pass different components which execute several functions that compose the overall workflow needed for compound prioritization. The ultimate goal is that the results are physically meaningful and the potent proposed structures are synthetically accessible [50].

## ACKNOWLEDGEMENT

This work was supported by funding under the Seven Research Framework Programme of the European Union, Project PATHOSYS (HEALTH-F5-2010-260429).

## REFERENCES

[1] Choo, Q. L.; Kuo, G.; Weiner, A. J.; Overby, L. R.; Bradley, D. W.; Houghton, M. Isolation of a CDNA clone derived from a blood-borne non-A, non-B viral-hepatitis genome *Science,* **1989**, *244*, 359-362.

[2] Chander, G.; Sulkowski, M. S.; Jenckes, M. W.; Torbenson, M. S.; Herlong, H. F.; Bass, E. B.; Gebo, K. A. Treatment of chronic hepatitis C: a systematic review *Hepatology,* **2002**, *36*, S135-S144.

[3] Jenny-Avital, E. R. Hepatitis C. *Curr. Opin. Infect. Dis.,* **1998**, *11*, 293-299.

[4] Wang, Q.; Heinz, B.A. Recent advances in prevention and treatment of hepatitis C virus infections. *Prog. Drug. Res.,* **2000**, *55*, 1-32.

[5] Manns, MP.; McHutchison, JG.; Gordon SC.; Rustgi, VK.; Shiffman, M.; Reindollar, R.; Goodman, ZD.; Koury, K.; Ling, M-.; Albrecht, JK. Peginterferon alfa-2b plus ribavirin compared with interferonalfa-2b plus ribavirin for initial treatment of chronic hepatitis C: A randomised trial. *Lancet*, **2001**, 358, 958-965.

[6] Legrand-Abravanel, F.; Nicot, F.; Izopet, J.New NS5B polymerase inhibitors for hepatitis C *Expert. Opin. Invest. Drugs.,* **2010**, 19, 963-975.

[7] Beaulieu, P. L.Non-nucleoside inhibitors of the HCV NS5B polymerase: progress in the discovery and development of novel agents for the treatment of HCV infection *Curr. Opin. Invest. Drugs.*, **2007**, 8, 614-634.

[8] Betzi, S.; Eydoux, C.; Bussetta, C.; Blemont, M.; Leyssen ,P.; Debarnot, C.; Ben-Rahou, M.; Haiech, J.; Hibert, M.; Gueritte, F.; Grierson, DS.; Romette, J.; Guillemot, J.; Neyts, J.; Alvarez, K.; Morelli, X.; Dutartre, H.; Canard, B. Identification of allosteric inhibitors blocking the hepatitis C virus polymerase NS5B in the RNA synthesis initiation step. *Antivir. Res.*, **2009**, 84, 48-59.

[9] Wu, J.Z.; Yao, N.; Walker, M.; Hong, Z. Recent Advances in Discovery and Development of Promising Therapeutics against Hepatitis C Virus NS5B RNA-Dependent RNA Polymerase. *Mini-Rev. Med. Chem.,* **2005,** *5(12)*, 1103-1112.

[10] Koch, U.; Narjes, F. Recent progress in the development of inhibitors of the hepatitis C virus RNA-dependent RNA polymerase. *Curr. Top. Med. Chem.*, **2007**, *7(13)*, 1302-29.

[11] Sulkowski, M.S. Specific targeted antiviral therapy for hepatitis C. *Curr. Gastroenterol. Rep.,* **2007**, 9, 5-13.

[12] Afantitis, A.; Melagraki, G.; Koutentis, PA.; Sarimveis, H.; Kollias, G. Ligand - based virtual screening procedure for the prediction and the identification of novel β-amyloid aggregation inhibitors using kohonen maps and counterpropagation artificial neural networks. *Eur. J. Med. Chem.*, **2011**, *46*, 497-508.

[13] Melagraki, G.; Afantitis, A.; Sarimveis, H.; Koutentis, PA.; Kollias, G.; Igglessi-Markopoulou, O. Predictive QSAR workflow for the *in silico* identification and screening of novel HDAC inhibitors. *Mol. Divers.*, **2009**, *13*, 301-311.

[14] Melagraki, G.; Afantitis, A.; Sarimveis, H.; Koutentis, PA.; Markopoulos, J.; Igglessi-Markopoulou, O. Optimization of biaryl piperidine and 4-amino-2-biarylurea MCH1 receptor antagonists using QSAR modeling.; classification techniques and virtual screening. *J. Comput. Aided. Mol. Des.*, **2007**, *21*, 251-267.

[15] Chen, C.; Chiou, C.; Chen, G.S.; Chen, S.; Hu, C.; Chi, W.; Chu, Y.; Hwang, L.; Chen, P.; Chen, D.; Liaw, S.; Chern, J.; Structure-based discovery of triphenylmethane derivatives as inhibitors of hepatitis C virus helicase. *J. Med. Chem.* **2009**, *52*, 2712-2723.

[16] Walker, T.; Grulke, CM.; Pozefsky, D.; Tropsha, A. Chembench: A cheminformatics workbench. *Bioinformatics*, **2010**, 26, 3000-3001.

[17] Peterson, YK.; Wang, XS.; Casey, PJ.; Tropsha, A. Discovery of geranylgeranyltransferase-I inhibitors with novel scaffolds by the means of quantitative structure-activity relationship modeling.; virtual screening.; and experimental validation. *J. Med. Chem*., **2009**, 52, 4210-4220.

[18] Fernandes, MB.; Rossini, M.; Stefani, R.; Tavares, L.; De Paulo Emerenciano, V.; Scotti, L.; Da Silva, MS.; Scotti, MT. Volsurf descriptors to analyse anti-HCV and cytotoxic activities of sesquiterpene lactones from asteraceae family. *Lett. Drug. Des. Discov.*, **2011**, *8*, 44-58.

[19] Fatemi, MH.; Dorostkar, F. QSAR prediction of D2 receptor antagonistic activity of 6-methoxy benzamides. *Eur. J. Med. Chem.*, **2010**, *45*, 4856-4862.

[20] Toropov, AA.; Toropova, AP.; Benfenati, E.; Manganaro, A. QSAR modelling of the toxicity to tetrahymena pyriformis by balance of correlations. *Mol. Divers.,* **2010**, *14*, 821-827.

[21] Afantitis, A.; Melagraki, G.; Sarimveis, H.; Koutentis, PA.; Igglessi-Markopoulou, O.; Kollias, G. A combined LS-SVM & MLR QSAR workflow for predicting the inhibition of CXCR3 receptor by quinazolinone analogs. *Mol. Divers.*, **2010**, *14*, 225-235.

[22] De Melo, EB.; Martins, JPA.; Jorge, TCM.; Friozi, MC.; Ferreira, MMC. Multivariate QSAR study on the antimutagenic activity of flavonoids against 3-NFA on salmonella typhimurium TA98. *Eur. J. Med. Chem.,* **2010**, *45*, 4562-4569.

[23] Shahlaei, M.; Sabet, R.; Ziari, MB.; Moeinifard, B.; Fassihi, A.; Karbakhsh, R. QSAR study of anthranilic acid sulfonamides as inhibitors of methionine aminopeptidase-2 using LS-SVM and GRNN based on principal components. *Eur. J. Med. Chem.,* **2010**, *45*, 4499-508.

[24] Melagraki, G.; Afantitis, A.; Sarimveis, H.; Igglessi-Markopoulou, O.; Alexandridis, A. A novel RBF neural network training methodology to predict toxicity to vibrio fischeri. *Mol. Divers.*, **2006**, 10 213-221.

[25] Toropova, AP.; Toropov, AA.; Lombardo, A.; Roncaglioni, A.; Benfenati, E.; Gini, G. A new bioconcentration factor model based on SMILES and indices of presence of atoms. *Eur. J. Med. Chem.,* **2010**, *45*, 4399-4402.

[26] Melagraki, G.; Afantitis, A.; Sarimveis, H.; Igglessi-Markopoulou, O.; Koutentis, PA.; Kollias, G. *In silico* exploration for identifying structure-activity relationship of MEK inhibition and oral bioavailability for isothiazole derivatives. *Chem. Biol. Drug. Des.*, **2010**, *76*, 397-406.

[27] Melagraki, G.; Afantitis, A.; Sarimveis, H.; Koutentis, PA.; Markopoulos, J.; Igglessi-Markopoulou, O. A novel QSPR model for predicting θ (lower critical solution temperature) in polymer solutions using molecular descriptors. *J. Mol. Model.*, **2007**, 13, 55-64.

[28] Gadhe C.G.; Lee S.H.; Madhavan T.; Kothandan G.; Choi D.; Cho SJ. Ligand based CoMFA.; CoMSIA and HQSAR analysis of CCR5 antagonists. *Bull. Kor. Chem. Soc.,* **2010**, *31*, 2761-2770.

[29] Goodarzi, M.; Freitas, MP.; Ghasemi, N. QSAR studies of bioactivities of 1-(azacyclyl)-3-arylsulfonyl-1H-pyrrolo[2.; 3-b]pyridines as 5-HT6 receptor ligands using physicochemical descriptors and MLR and ANN-modeling. *Eur. J. Med. Chem,.* **2010**, *45*, 3911-3915**.**

[30] Mouchlis, VD.; Mavromoustakos, TM.; Kokotos, G. Molecular docking and 3D-QSAR CoMFA studies on indole inhibitors of GIIA secreted phospholipase A2. *J. Chem. Inf. Model.,* **2010**, 50, 1589-1601.

[31] Mouchlis, VD.; Mavromoustakos, TM.; Kokotos, G. Design of new secreted phospholipase A2 inhibitors based on docking calculations by modifying the pharmacophore segments of the FPL67047XX inhibitor. *J. Comput. Aided. Mol. Des.*, **2010** 24 107-115.

[32] Clark, D. E. What has virtual screening ever done for drug discovery, *Expert Opin. Drug Discov. Perspect,*. **2008**, *3*, 841-851.

[33] Kim, K.; Kim, ND.; Seong, B. Pharmacophore-based virtual screening: A review of recent applications. *Expert Opin. Drug Discov.,* **2010**, 5, 205-222.

[34] Përez-Nueno, VI.; Ritchie, DW.; Rabal, O.; Pascual, R.; Borrella, JI.; Teixidó, J. Comparison of ligand-based and receptor-based virtual screening of HIV entry inhibitors for the CXCR4 and CCR5 receptors using 3D ligand shape matching and ligand-receptor docking. *J. Chem. Inf. Model.,* **2008**, 48, 509-533.

[35] Singh, J.; Shaik, B.; Singh, S.; Agrawal, VK.; Khadikar, PV.; Deeb, O.; Supuran, CT. Comparative QSAR study on para-substituted aromatic sulphonamides as CAII inhibitors: Information versus topological (distance-based and connectivity) indices. *Chem. Biol. Drug. Des.,* **2008**, 71, 244-259.

[36] Melville, J. L.; Burke, E. K.; Hirst, J. D. Machine Learning in Virtual Screening. Comb. *Chem. High Through. Screen.,* **2009**, *12*, 332-343.

[37] Muegge, I.; Oloff, S. Advances in virtual screening. *In silico* techniques, *Drug Discov. Today.: Technologies*. **2006**, *3*, 405-411.

[38] Afantitis, A.; Melagraki, G.; Sarimveis, H.; Koutentis, PA.; Markopoulos, J.; Igglessi-Markopoulou, O. Investigation of substituent effect of 1-(3.;3-diphenylpropyl)-piperidinyl phenylacetamides on CCR5 binding affinity using QSAR and virtual screening techniques. *J. Comput. Aided. Mol. Des*., **2006**, 20, 83-95.

[39] Xie, X-S. Exploiting PubChem for virtual screening. *Expert Opin. Drug Discov.,* **2010,** *5*, 1205-1220.

[40] Talele, T. T.; Arora, P.; Kulkarni, S. S.; Patel, M. R.; Singh, S.; Chudayeu, M.; Kaushik-Basu, N. Structure-based virtual screening, synthesis and SAR of novel inhibitors of hepatitis C virus NS5B polymerase. *Bioorg. Med. Chem.*, **2010**, *18*, 4630-4638.

[41] Kim, J.; Kim, K.-S.; Kim, D.-E.; Chong, Y. Identification of Novel HCV RNA-dependent RNA polymerase Inhibitors Using Pharmacophore- Guided Virtual Screening. *Chem. Biol. Drug Des.*, **2008**, *72*, 585-591.

[42] Melagraki, G.; Afantitis, A.; Sarimveis, H.; Koutentis, P.A.; Markopoulos, J.; Igglessi -Markopoulou, O. Identification of a series of novel derivatives as potent HCV inhibitors by a ligand -based virtual screening optimized procedure. *Bioorg. Med. Chem.*, **2007**, *15*, 7237-7247.

[43] Musmuca, I.; Caroli, A.; Mai, A.; Kaushik-Basu, N.; Arora, P.; Ragno, R. Combining 3-D Quantitative Structure-Activity Relationship with Ligand Based and Structure Based Alignment Procedures for *in Silico* Screening of New Hepatitis C Virus NS5B Polymerase Inhibitors. *J. Chem. Inf. Model.,* **2010**, *50*, 662-676.

[44] Corbeil, C.R.; Englebienne, P.; Yannopoulos, C. G.; Chan, L.; Das, S.K.; Bilimoria, D.; L'heureux, L.; Moitessier, N. Docking ligands into flexible and solvated macromolecules. 2. Development and application of fitted 1.5 to the virtual screening of potential HCV polymerase inhibitors. *J. Chem. Inf. Model.,* **2008**, *48*, 902-909.

[45] Ryu, K.; Kim, N. D.; Choi, S. I.; Han, C.K.; Yoon, J. H.; No, K. T.; Kim, K.-H.; Seong, B. L. Identification of novel inhibitors of HCV RNA-dependent RNA polymerase by pharmacophore-based virtual screening and *in vitro* evaluation. *Bioorg. Med. Chem.*, **2009**, *17*, 2975-82.

[46] Louise-May, S.; Yang, W.; Nie, X.; Liu, D.; Deshpande, M. S.; Phadke, A. S.; Huang, M.; Agarwal, A. Discovery of novel dialkyl substituted thiophene inhibitors of HCV by *in silico* screening of the NS5B RdRp. *Bioorg. Med. Chem. Lett.*, **2007**, *17*, 3905-3909.

[47] Li, T.; Froeyen, M.; Herdewijn, P. Insight into ligand selectivity in HCV NS5B polymerase: molecular dynamics simulations, free energy decomposition and docking. *J. Mol. Model.* **2010**, *16*, 49-59.

[48] Wang, X.; Yang, W.; Xu, X.; Zang, H.; Li, Y.; Wang, Y. Studies of Benzothiadiazine Derivatives as Hepatitis C Virus NS5B Polymerase

Inhibitors Using 3D-QSAR, Molecular Docking and Molecular Dynamics. *Curr. Med. Chem.*, **2010**, *17*, 2788-2803.

[49] McInnes, C. Virtual screening strategies in drug discovery. *Curr. Opin. Chem. Biol.* **2007**, *11*, 494-502.

[50] Gasteiger, J. De novo design and synthetic accessibility. *J. Comput. Aided. Mol. Des.*, **2007**, *21*, 307-309.

**Received: February 14, 2011** **Revised: May 14, 2011** **Accepted: May 15, 2011**