Università degli Studi di Cagliari

DOTTORATO DI RICERCA

Sviluppo e sperimentazione dei farmaci antivirali

Ciclo XXIII

TITOLO TESI

Identification and Development of Novel Inhibitors of Viral Polymerases: An *In-silico* Approach

BIO/19

Presentata da:	Saumya Shukla
Coordinatore Dottorato:	Prof.ssa. Alessandra Pani
Tutor:	Prof. Paolo La Colla
	Dr. Attilio V. Vargiu
	Prof. Paolo Ruggerone

Esame finale anno accademico 2009-2010

University of Cagliari
Department of Biomedical Science and Technology
Section of General Microbiology and Virology
& Microbial Biotechnologies

Research Doctorate in
Development and Evaluation of Antiviral Drugs
Coordinator of the Doctorate
Prof. Alessandra Pani

Identification and Development of Novel Inhibitors of Viral Polymerases: An *In-silico* Approach

Supervisor:
Prof. Paolo La Colla

Co-Supervisor:
Dr. Attilio V. Vargiu
Prof. Paolo Ruggerone

Candidate:
Saumya Shukla(Pandey)

XXIII Cycle 2008-2010

Acknowledgments

I have been inspired and encouraged by numerous people who have helped me in immeasurable ways to complete this thesis. To acknowledge all those people who have contributed to the work described in this thesis is an impossible task, given the many people that have helped to design, implement, apply, criticize, sponsor and evangelize the work. I am going to try anyway, and if your name is not listed, rest assured that my gratitude is not less than for those listed below.

First and foremost, I would like to thank my thesis advisor, Prof. Paolo Lacolla, who has provided generous support throughout my stay at Cagliari. I am very thankful to him for providing me the intellectual freedom to work on whatever research I found interesting. I consider myself very fortunate to have been a part of his research group.

I take this opportunity to extend my gratitude towards Prof. Paolo Ruggerone and Dr Atillio Vargiu who have been my joint supervisors. I owe them lots of gratitude for having shown me the right way of research. Their overly enthusiasm and integral view on research and mission for providing 'only high-quality work and not less' have made a deep impression on me.

I would like to thank my friend and colleague of five years, Dr Shailendra Asthana. We shared a great coordination and have together faced and accepted failures and success alike. It's because of his motivation, help and constant support that I have been able to come this far.

Successful completion of a thesis requires loads of work from many people, some contribute to it professionally and others by creating an atmosphere suitable for research and high efficiency. This nest was decorated by many of my friends. Amit Khairnar, Amit Kumar, Shaji Varghese, Paola Nigreddu, Massimiliano Ventroni, Ajita Ratani, Brajesh Barse, Amita Barse and Pradip Hilwale beautifully decorated this nest. After four years, Cagliari has become a home away from home, all because of my wonderful friends.

I am blessed to have two wonderful sisters. They have always tolerated me and boosted my enthusiasm. I can never thank you girls enough for your phenomenal support during the touch patches I went through. Here, I would also like to convey my gratitude towards my brothers in law. They have always been around when I am in need.

I would like to thank my parents for all their efforts and support all through my life. I feel a deep sense of gratitude for my father and mother who formed part of my vision and taught me the good things that really matter in life. They molded me to be what I am now. They are my source of strength. They has always supported me, inspired me, motivated me, and tolerated me. They have set my standard very high and trained me to achieve higher. I owe all my successes and smiles to them. I would also like to thank my parents-in law for their understanding and support specially in last one year.

Lastly but most importantly, the credit for whatever I have achieved goes to my husband. If not for him, it is just not practically possible to have completed my Ph.D. I want to thank my husband for all the support that he gave me.

Summary

Our ability to answer some of the most fundamental questions of drug development process, such as how proteins undergo conformational changes to bind a ligand, what are key protein residues which play pivotal role for the binding, why a compound is active or inactive, or how its activity might be improved, how protein dynamics is changed on introduction of resistant mutations, how resistant mutations render protein resistant against inhibitors, what the mechanism of resistance changes in the binding free energy of a particular drug candidate or the mechanisms and energetic consequences of conformational changes in a protein etc, is limited often by our inability to appropriately catch dynamic features, including interactions with the solvent, in a crystallographic structure which is just an average structure and thus lacks dynamics.

Here is where various biophysics based computational techniques come to our rescue, which provide huge wealth of information related to protein dynamics and protein-ligand recognition. These methods have grown in their effectiveness because of advances in algorithms, representations, and mathematical procedures for studying such processes and are in the position of providing improved understanding of the basic science, the biological events and molecular interactions that define a target for therapeutic intervention. However, the approximations introduced in the computational approaches at different levels make a single technique in many cases inappropriate to tackle the dauntingly complex problems related to the drug design. Thus, a continuous feedback with experiments is required but also the appropriate combination of different computational strategies might be helpful.

This work represents the application of computational techniques, such as docking, molecular dynamics, algorithms to calculate free energy of binding of ligands into the binding pocket (MM-PBSA) and algorithms to study rare events (like unbinding of ligand from the binding site, Metadynamics, etc.) to explore, at microscopic level, the key pattern of interaction between protein and ligand, to understand the effect of mutations, to get an insight of the full and undocking path and to calculate binding energetics.

In this thesis, polymerases of three positive strand RNA viruses, viz, Hepatitis C Virus (HCV), Bovine Viral Diarrhea Virus (BVDV) and Human Immunodeficiency Virus (HIV) have been targeted with an aim to come out with more potent antivirals effective against wide range of resistant mutations. These viruses have created havoc by effecting Human race directly or indirectly and thereby creating an urgent need for further refinement and new development of antivirals drugs. Herein, plethora of computational techniques has been applied on the three protein systems to answer fundamental questions of drug resistance, drug inhibition and providing clues for improvement of antivirals.

Contents:

I)- Background and Computational Approaches

- **Introduction**
  - Antivirals and theoretical aspects
  - Motivation of the work
  - Outline of the thesis
  - Thesis in Flow chart
- **The Actors: Polymerases and Antivirals**
  - **Polymerases**
    - Genome organization(HCV, BVDV and HIV)
    - Domain Architecture (HCV, BVDV and HIV)
  - **Antivirals, a small journey**
    - Nucleoside Analogs (NIs)
    - Non-Nucleoside Analogs (NNIs)
- **Methods**
  - **Molecular Docking**
  - **Molecular Dynamics**
    - Integration of Newton equations of Motion
    - Multiple Time Step Integrator
    - The Interaction Potential
    - Constraints for Hydrogen
    - Boundary Conditions
    - Statistical Ensembles
  - **Long Time Scale Simulations**
    - Metadynamics
    - The Algorithm
    - How to Choose Cvs
    - Estimation of Error
  - **Binding Affinity: MM-PBSA**

III)- Different binding sites of the benzimidazole compound 227G on HCV and BVDV RdRps revealed by MD simulations

- Abstract
- Introduction
- Material and Methods
  - Biological
    - Cell and viruses
    - HCV replicon assay.
    - Expression and purification of BVDV and HCV1b RdRps (NS5B).

- RNA-Dependent RNA Polymerase Assays.

■ Computational
- Structure of Receptors (x-ray)
- Structure of the complex
  - Docking
- MD Simulation protocol
  - Parametrization
  - Dynamics
  - Analysis of structures and dynamics.
- Structure of Complexes
  - Docking
- Analysis of structures and dynamics
- Area Analysis
- Results
- Experimental
  - Antiviral activity of 227G compound.
  - In vitro RdRp assays.
- Computational
- Docking
- MD Simulations
  - BVDV Complex
  - HV Complex
- Comparison With APO-Protein
  - RMSD-Time wise
  - RMSD-Residue-Wise
- Area with respect to APO-protein
- Discussion
  - BVDV
  - HCV
- Conclusion

IV)- Quinoline Tricyclic Derivatives. Design, Synthesis and Preliminary In vitro and In silico Antiviral Activity Against Flaviviridae Family of Three New Classes of Virus-Encoded RNA-Dependent RNA Polymerase (RdRp) Inhibitors

- Abstract
- Introduction
- Material and Methods
  - MM-PBSA Calculation
  - Metadynamics
  - Putative Binding cavity
- Results
  - Docking

- MD Simulation
- Metadynamics
  - PS999
    - Mini-1
    - Mini-2
    - Undocking Mechanism
    - Escape
  - PS1036
    - Mini-1
    - Mini-2
    - Undocking Mechanism
    - Escape
- Binding Free Energy Calculation
- Drug Resistance.

(V)Identification of binding cavity for novel Non-Nucleodside inhibitors of HCV RdRp: A molecular docking study

- Abstract
- Introduction
- Methods
  - Receptor
  - Complexes
    - Docking
    - Assessment of Docking
  - MM-PBSA
- Results
  - Docking Analysis
  - MD Analysis
    - PS1126
      - Arg200 as a trigger to the switch for opening of the otherwise buried cavity
    - PS1101
    - PS1097
  - Mechanisms of Inhibition
    - Mechanism of Inhibition of PS1101
    - Mechanism of Inhibition of PS1097
    - Mechanism of Inhibition of PS1126
  - Comparison of Calculated and Experimental Binding Affinities.
  - Conclusion

VI)-Conclusion

List of Figures :

1)Introduction:

Figure1: Global variations in age adjusted incidence rates of liver cancer, prevalence of chronic HCV infection and chronic HBV infection. Adapted from: Yang, J. D. & Roberts, L. R. (2010) Hepatocellular carcinoma: a global view Nat, Rev. Gastroenterol. Hepatol.

Figure2: A global View of HIV infection, 33.3 million people [31.4-35.3] living with HIV, 2009. Adapted From: Joint United Nations Programme on HIV/AIDS(UNAIDS) and World Health Organization (WHO) 2009 AIDS epidemic update

Figure3:HCV genome organization (top) and polyprotein processing (bottom). Structural biology of hepatitis C virus.Penin F, Dubuisson J, Rey FA, Moradpour D, Pawlotsky JM.

Figure4: Cartoon representation of HCV RdRp. Different domains of RdRp viz thumb in blue, finger in green, and palm in magenta.

Figure5: Cartoon representation of BVDV RdRp. Different domains of RdRp viz thumb in blue, finger in green, palm in magenta and unique N-terminal is yellow in color.

Figure6: Cartoon representation of HIV-RT. Two subunits P66 (colored according to the subdomains) and P51(white) are depicted here. Finger is represented in blue, Palm in red, Thumb in green, the connection in magenta and RNaseH in yellow. The template is also represented here with the two strands being colored in orange and cyan.

Figure7: The structures of NM 283 (left) and NM 107 (right)

Figure8: Ribbon representation of the overall structure of HCV NS5B polymerase representing four NNIs binding site in different domains

Chapter 3)

Figure 1: The structure of HCV (B) and BVDV (C) RdRps are represented with their different subunits and domains: the thumb is colored in dark green, fingers are colored in ice blue and the palm domain along with unique N-Terminal domain is colored in purple. The surface view areas highlight the binding site of 227G in HCV (thumb) and BVDV(finger) respectively. (A) Ball-and-stick representation of 227G with all atoms labeled according to the notation of the force field. 227G is well characterized as ring I, II , arms, anchor and its tail. All images are generated by VMD [Humphrey 1996].

Figure2: Two different orientation of 227G in BVDV and HCV RdRp. Orientations are taken on the basis of the criteria mention in material and methods. During MD simulation we observed, one orientation of 227G is stable enough during whole length of trajectory while other one resulted into protein-227G dissociation in each case. (A) stable state of 227G in case of BVDV while B)- is unstable state. C)- stable state of 227G in HCV and D)- unstable state respectively. . The instability of 227G is because its failure to acquire its arm and tails in to their respective pockets. In both cases the key residues are same and their mode of interaction is almost same, but stability of only one pose clearly explained the specificity of 227G orientation.

Figure3 : HCV RdRp, surface view representing the different pockets of cavity. Pocket1 and pocket2 are colored yellow and blue respectively while the extended hydrophobic cavity is represented in orange. A)depicts the active state of the protein wherein the Lys30 and Lys31 residues of the lambda1 loop sub optimally occupy the sub pocket1 and pocket2 of the cavity. B)227G accommodates itself in the HCV RdRp cavity. Arms of 227G occupy the sub pockets which in the apo protein is occupied by residues of lambda1 loop. C)- Binding cavity of HCV RdRp, indicating three different sub-pockets, I, II and extended pocket. Different arms of inhibitor

occupying different pockets of the cavity. The CMF (white), 2DSX (pink) and 227G are green in color. The CMF and 2DSX correspond to their respective PDB ids.

Figure 4: Interaction map of BVDV RdRp and HCV RdRp in complex with 227G. (A). The spiral rings indicate hydrogen bonds (HB) involving residues Asp126, Ser533 and Ala221, as well as a salt-bridge between Asp126 and Asn264. Four loops L_1, L_2, L_3 and L_4 are represented in pink while Motif-I is in blue. Residues contributing in HpH interactions are in represented by VdW surface and are coloured white. (C) the helices (H_1, H_2 and H_3) outlining the cavity are coloured pink. Residues outlining the cavity are shown in VdW surface view as well as indicated by ball and stick representation, (B & D)Statistics of HpH contacts formed between 227G-BVDV 227G-HCV complexes respectively. Labels on the X-axis refer to the atoms of 227G as in Fig. 1c while on the Y-axis the interaction lifetime is reported as fraction of the entire simulation time.

Figure5: Hydrogen Bond Interaction network: Inter (227G-Asp126), Intra (Glu128-Asp126) & (Asp126-Asn264), water mediated interaction (227G:N1-WAT) & (WAT:Glu128), as mentioned in Table3. This network is strong enough having occupancy more than 80 %.

Figure6: RMSD of whole system as a function of time relative to the starting structures in both HCV and BVDV, APO and COMPLEX

Figure7: Average RMSD during the simulation for each residue with respect to a reference snapshot. We observed huge fluctuations in the wild type as compared to the complex. Maximum peaks were obtained in the N-terminl region, residue surrounding motifI and motifII and loop L3 in Apo which got stabilised on 227G binding .

Figure8: Top view of the template entrance channel of BVDV RdRp is indicated by blue and the rest protein is in ribbon and green color. The putative binding cavity encircled by Loops L_1, L_2, L_3 and L_4 lye in the template entrance channel. (A)- The cavity view in crystal structure, indicates enough space for the entrance of template, while in (B)- 227G, indicated in yellow color, totally blocks the entrance to the channel. (C) Depicts the superposition of the averaged structures of the loops in complex (yellow) and Apo (violet) while the crystal is presented in red color. For the calculation of area, we considered L1, L2 and L4 which showed maximum variation between complex and Apo protein. The distances between residues R127(L_4-white), F224 (L_2-yellow) and A392 (L_1-orange) sampled during the simulations were used to calculate the area of the cavity in the complex and in apo-protein.(D) Distribution of the area during the simulations for the Apo-protein is shown by the green curve while the red indicates the complex system. The black line indicates the value extracted from the X-ray structure.

Figure 9: Superposition of HCV and BVDV RdRp complexes, taking their average pdb from their respective simulation in each cases. HCV RdRp is in green in color while BVDV RdRp is red in color. 227G in each case is rendered in VDW and lies in thumb and finger domain of HCV and BVDV system respectively.

Chapter4)

Figure1: Complete overview of the location of binding site of PS999 and PS1036 complex systems. The snapshot taken were the average structures for both the systems. The compounds are render in VdW, binding site residues are in ball and stick model, while binding cavity is in surface view. Whole protein rendered in ribbon, coloring in secondary structure form.

Figure2: Expanded view of interaction between key residues of binding cavity of PS999 (A)and PS1036(B). Snapshot taken from the deepest minima in both PS999 and PS1036. Binding site residues are shown here using ball-sticks representation(basic:blue, acidic:red, hydrophobic:white). In panel (A) Interaction map of PS999, residues involved in making HB with PS999 are Pro262 of motif-I and Arg127 of loop L1, while in panel (B) Interaction map of PS1036, Pro262 and Asn264, both from motif-I.

Figure3: Average ligand-residue interaction energies in kcal/mol over compound PS999 and PS1036 for the residues that contribute most to the ligand -surrounding. In the upper panel the interaction energy of WT and Mutant with PS999. In lower panel interaction energy of PS1036. The residues highlighted in bold are resistant mutant which reduces considerable in mutant (Red bar) with comparison to WT (black bar).

Figure4: A)-The two dimensional Free energy surface(FES) of PS999'S undocking process, as a function of *gdist* and *HpH* contacts. Each line correspond to 1kcal/mol. where each color and each line correspond to 1kcal/mol A)-this panel represent the FES of PS999, representation two deep minima while panel **B)**-Evolution of the center of mass of PS999 during the metadynamics simulations of the undocking process. The residues of the putative binding site are highlighted in licorice. C)- Pictorial representation of closing of binding site in mini2 (green in color), as compared to the mini1(purple and transparent).

Figure5: A)-The two dimensional Free energy surface(FES) of PS1036'S undocking process, as a function of *gdist* and *HpH* contacts. Each line correspond to 1kcal/mol. where each color and each line correspond to 1kcal/mol A)-this panel represent the FES of PS1036, representation two deep minima while panel 2B)-represent the FES of PS1036 with only one minima. **B)**- Evolution of the center of mass of A)- PS1036 during the metadynamics simulations of the undocking process. The residues of the putative binding site are highlighted in licorice.

Figure6: Electrostatic potential map of the PS999-binding pocket of BVDV RdRp. A)- PS999 in WT B) PS999 in A392E mutant (positive in blue; negative in red and neutral in white in color). The major notable difference on electrostatic map of PS999 is pointed with arrow in both cases, shown the effect of mutation A392E

Chapter5)

Figure1. Ribbon representation of the overall structure of HCV NS5B polymerase in blue along with transparent surface view in white. Solid surface representations of four allosteric binding sites are also shown. The dark red colored surface corresponds to the thumb allosteric binding pocket of indole group "siteA"; the dark green colored surface corresponds to the 2nd thumb allosteric binding site "site B", the dark yellow colored surface corresponds to the allosteric binding site situated in the palm domain "site C" and the dark skyblue colored surface correspond to 2nd palm domain binding site "site D".

Figure 2: Depicts the final docking conformations taken for each compound in each of the four binding site. Residues defining the cavity, as obatined by our analysis of X ray structures are presented as licorice. It should be noted here that while in case of PS1126 we were able to get good clusters in all the four site, on the contrary , there docking was not succefull for site D in case of PS1101 and PS1097.

Figure3: Expanded view of interaction between key residues of binding cavity of PS1126. Snapshot taken from the average structure from the equilibrium of 20ns MD simulation. The Binding site residues are shown here using sticks representation(basic:blue, acidic:red, hydrophobic:white) while compounds are in ball and stick in cyan color. In panel (A) and (B)Interaction map and interaction energy of PS1126 Average ligand-residue interaction energies in kcal/mol over compound PS1126 for the residues that contribute most to the ligand -surrounding.

Figure4: In the panel one; the three key residues Arg200, Asn316 and Cys366 shown in stick. The arrangement of these residues, " *hinge*", with the spiral ring represent the HB between Arg200 and Asn316, which opens in presence of PS1126. In the panel two, the possible population states for torsions of residue Arg200. The atoms forming ramachandran torsions (Φ, Ψ) are : Φ torsion are:CB-CG-CD-NE; and Ψ torsion are : CG-CD-NE-CZ. The Ramachandran dihedral illustrating the change in the side chain dihedral angles upon ligand binding. In panel (C)-for Φ torsion(Apo) [-80:-25], [30:120] and [150:225] and for complex [-100:-40], [130:225]. In panel (D)- Ψ torsion are [-180:-50] for Apo and for complex [150:250]. The panel C is the population states of Cys366 in presence of ligand. The distribution for Apo are [-100: -40], [50:110] and [150:215], while for complex [-100:-40] and [160:230].

Figure5: Expanded view of interaction between key residues of binding cavity of PS1097, PS1101 and PS1126. Snapshot taken from the average structure from the equilibrium of 20ns MD simulation. The Binding site residues are shown here using sticks representation(basic:blue, acidic:red, hydrophobic:white) while compounds are in ball and stick in cyan color. In panel (A) and (B)Interaction map and interaction energy of PS1097; panel (C) and (D) Interaction map and interaction energy of PS1101 and panel (E) and (F) are Interaction map and interaction energy of PS1126. Average ligand-residue interaction energies in kcal/mol over compound PS1097, PS1101 and PS1126 for the residues that contribute most to the ligand -surrounding.

Figure6: Expanded view of interaction between key residues of binding cavity of PS1097, PS1101 and PS1126. Snapshot taken from the average structure from the equilibrium of 20ns MD simulation. The Binding site residues are shown here using sticks representation(basic:blue, acidic:red, hydrophobic:white) while compounds are in ball and stick in cyan color. In panel (A) and (B)Interaction map and interaction energy of PS1097; panel (C) and (D) Interaction map and interaction energy of PS1101 and panel (E) and (F) are Interaction map and interaction energy of PS1126. Average ligand-residue interaction energies in kcal/mol over compound PS1097, PS1101 and PS1126 for the residues that contribute most to the ligand -surrounding.

Figure7: Key structure playing crucial role in polymerization. These five structurally and functionally mportant region shown as β hairpin in brown, C-terminal in blue, inner core of thumb domain in green , primer-grip in yellow and catalytic triad, red in color. The compounds PS1097(sky blue), PS1101 (mauve), and PS1126(yellow)in wireframe into their specific binding mode in their binding site. The resistance mutation of that corresponding residues are shown in stick (res-type).

List of TABLES:

Introduction:

Table1:Approved antiviral drugs against HCV. Adapted from: HCV Advocate; July 24, 2010;Hepatitis C treatment in current clinical development; Alan Franciscus.

Table2:Approved antiretroviral drugs. Adapted from: Drugs Used in the Treatment of HIV Infection, U.S. FDA, http://www.fda.gov/oashi/aids/virals.html. Drugs are listed in order of FDA approval within each class.

Chapter3

Table1: System Details

Table 2: Comparative evaluation of: i) antiviral activity of 227G against BVDV (cell-based assay) and against HCV-1b (Huh7 cell-based HCV replicon system assay); ii) inhibitory activity of 227G in enzymatic assays on recombinant purified RdRps of BVDV and HCV-1b.

Table 3: Hydrogen Bond Interaction Network

Table: 4 Residues wise interaction energy of 227G A)- BVDV RdRp-227G B)- HCV RdRp-227G.

Chapter4

Table1: Interaction Energy of key residues in PS999-RdRp Complex (Kcal/mol).

Table2: Interaction Energy of key residues in PS1036-RdRp complex (Kcal/mol)

Table3: MMPBSA

Chapter 5

Table:1 Interaction Energy of key residues in PS1126-RdRp complex (Kcal/mol)

Table:2 Interaction Energy of key residues in PS1101-RdRp complex (Kcal/mol)

Table:3 Interaction Energy of key residues in PS1097-RdRp complex (Kcal/mol)

Table:4 Comparison of Calculated and Experimental Binding Affinities.

1)
Background and Computational Approaches

Introduction

The journey for Development of antiviral has always been a roller coaster ride. For nearly two decades,(around 1947-1950) futile attempts on the use of bacterial antibiotics on a number of viral infection led to the dogma that viruses are not susceptible to antibiotics and that selective toxicity for these obligate intracellular parasites was unattainable. Clearly, 1950 experienced something of a FALSE DAWN for the concept of specific antiviral therapy. In 1959 a major breakthrough occurred in the field of antivirals, with the discovery of 5-iodo-2'-deoxyuridine (idoxuridine,IDU) by Dr Bill Prusoff and the identification of its antiviral properties[1,2,3]. In 1972, another landmark, ribavarin,[4,5] was reported as the first broad-spectrum antiviral acting against many different virus families, notably the negative RNA strand Virus, RNA Syncytial virus. Then came the advent of acyclovir[6,7], which is acknowledged as the major milestone in antiviral drug discovery. Noteworthy, its particular mechanism of HSV inhibition and its pharmacology made it one of the safest drug of all times with almost no side effect described during three decades of use.

Many decades after their birth, antiviral have, at last definitely come of age, 37 antivirals have been approved for the treatment of viral diseases with their application primarily aimed at therapy of herpesvirus as well as HIV, HBV and HCV. Much has been done, yet a lot have to be done.

For example, even after 21 years from the discovery of hepatitis C virus (HCV), 170 to 200 million people worldwide are believed to suffer from chronic HCV infection, a blood-born disease that targets the liver and progresses to organ cirrhosis and primary cancer in a significant proportion of patients. To give the numbers, hepatitis C (officially recognized as the “silent epidemic”) infected approximately 180 million people worldwide, and it is the leading reason for liver transplantation in the United States [Fig .1] [8,9]. No vaccine or effective therapy broadly targeting all genotypes of HCV[10] is disposable at present; treatment options against HCV include pegylated intereferon α alone or in combination with ribavirin, a broad spectrum antiviral agent. Antiviral treatment is unlikely to induce significant population-wide reductions, and mortality is expected to continue to increase, at least in the next 10-20 years.

Similar story exist for human immunodeficiency virus (HIV) infection, of which the appearance was first reported in 1983. Now, 25 years after the inhibitory effects of the first anti-HIV drug, azidothymidine (AZT)1 on the replication of HIV were described(then called HTLV-III (human T-cell lymphotropic virus type III) or LAV (lympha- denopathy-associated virus) more than 20 anti-HIV drugs, belonging to 7 classes, have been added to the armamentarium of HIV anitiviral database. These efforts have transformed what used to be a rapid and lethal infection into a chronic condition that can be controlled for many years through combination therapies with different classes of antiviral drugs — known as highly active antiretroviral therapy (HAART) but complete eradication is not yet possible. According to the Joint United Nations Programme on HIV/AIDS (UNAIDS) and World Health Organization (WHO) 2009 AIDS epidemic update, approximately 33.3 million people globally are living with HIV infection, 2.3 million of whom are children (Figure 2). One in 155 adults and adolescents are infected with HIV globally, with an astonishing rate of one in 21 in Sub-Saharan Africa.[11] In the USA alone, more than half a million people were living with HIV in 2007, and 44,000 new cases were reported in the same year.[12] We should therefore guard against triumphalism, for much remains to be done: the pandemic’s toll still rises, and we cannot be certain when it will reach the high water mark.

Another viral pathogen, Bovine viral diarrhoea virus (BVDV), has created havoc by affecting cattles and resulting into wide assortment of disease manifestations which include resorption, mummification, or abortion of the dead fetus while those fetuses who survive early infection may be malformed or blind, may have skeletal defects, respiratory problems, underdeveloped brain or weak immune system.[13] This not only cause high mortality in cattle but also create huge economic burden.

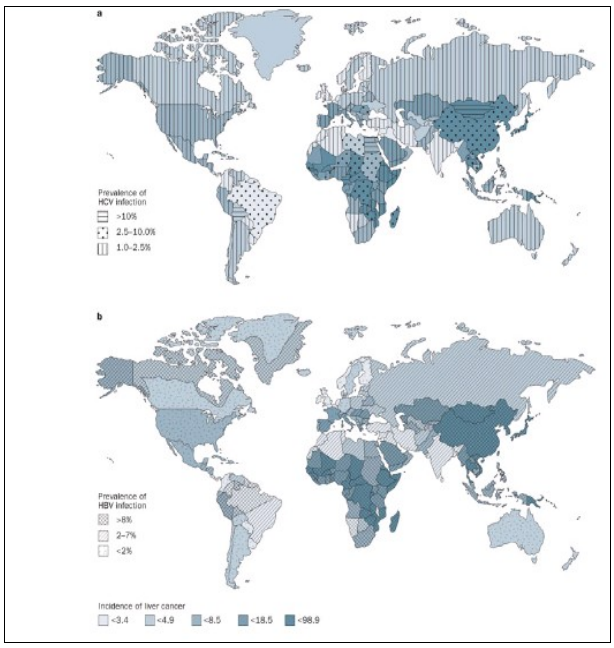


Figure1: Global variations in age adjusted incidence rates of liver cancer, prevalence of chronic HCV infection and chronic HBV infection. Adapted from: Yang, J. D. & Roberts, L. R. (2010) Hepatocellular carcinoma: a global view Nat, Rev. Gastroenterol. Hepatol. doi:10.1038/nrgastro.2010.100

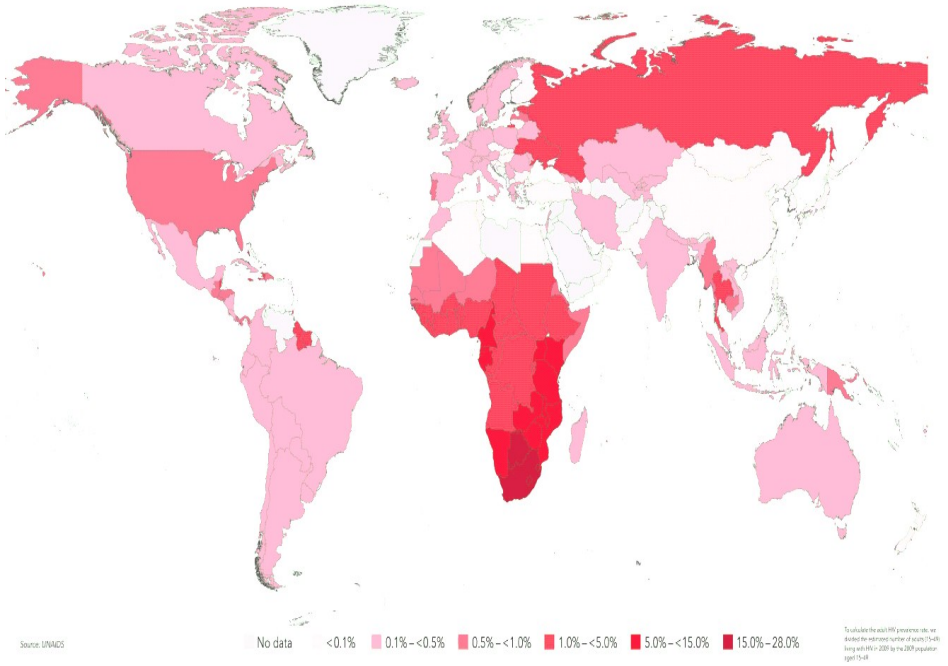


Figure2: A global View of HIV infection, 33.3 million people [31.4-35.3] living with HIV, 2009. Adapted From: Joint United Nations Programme on HIV/AIDS(UNAIDS) and World Health Organization (WHO) 2009 AIDS epidemic update.

All this above mentioned statistic put forth the gravity of antiviral infection in human and animal alike thereby creating an urgent need to develop effective antivirals. A combined effort coming from chemistry, pharmacology, microbiology, biochemistry and biophysics is needed for the development of broad spectrum or specific antivirals.

One of the major obstacles to antiviral therapy is the development of resistant mutation and, as a matter of fact, all active antiviral agents have shown to select for resistance mutations. Virus replication is an error prone process resulting in a large number of variants (quasispecies) in patients. XX Resistance is the cause and/or the consequence of treatment failure. Viral infection is characterized by a very high replication rate, with the production of 1 to 10 billion new virus particles per day in an untreated infected individual . Moreover, viruses lacks proof-reading functionality, and this results in an average error rate per detectable nucleotide incorporated of 1/1700 . Hence focus is needed on the development of antivirals effective against wide range of resistant mutation as well as coming up with combination therapy as the 'High active anti retroviral Therapy' (HAART) in case of HIV infection.

Life cycle of viral life cycle include various steps, all of which can be targeted for the development of antiviral therapy. These steps broadly encompass viral attachment, entry, and fusion; viral RNA translation; post translational processing; HCV replication; and viral assembly and release. Among the other targets, Viral polymerase, which is strictly required for viral replication, has been the focus of intense drug discovery activity as they are considered as significant targets for therapeutic intervention. Several reasons that makes viral polymerase an excellent target for antiviral investigation include: 1) Successes stories in targeting the polymerases from other viral systems such as cytomegalovirus (CMV), herpes simplex virus (HSV), hepatitis B virus (HBV); 2) The most promising antivirals target viral proteins or processes that are not endogenous to host cells. Close structural homologs of the polymerase do not exist within the uninfected host cell and hence it makes it a interesting target for fighting viral infection. Compared to the host RNA and DNA polymerases, polymerase has distinct subcellular localization at the interface of the endoplasmic reticulum (ER) membrane and cytoplasm, a novel catalytic mechanism and many unique structural features, all of which make it an attractive target for developing effective and selective anti-Viral therapeutics.

From a biochemical point of view, the polymerase works in three stages: 1) initiation, 2) transition, and 3) elongation phase. In theory, polymerase inhibitors could potentially target all these steps either by acting at one or several of them.

Polymerase inhibitors falls into two categories: Nucleoside and non-Nucleoside inhibitors. Nucleoside analogue inhibitors (NI), discovered through the rational search for substrate analogues, target the active site of polymerase, either competing with natural NTP substrates and/or acting as “chain terminators”, or causing a mutational “error catastrophe” by being incorporated into the elongating nascent RNA molecule. NIs are often endowed with a broad-spectrum antiviral activity, due to the fact that, during evolution, RNA viruses have remained unchanged certain sequences of their RdRp, by virtue of the key role in the process of transcription and replication of viral genomes. However, the effectiveness of nucleoside analogues has been reported to be compromised due to the generation of resistant mutants and the adverse side-effect, necessitating the development of new inhibition targets and inhibitors of NS5B. Accordingly, the non-nucleoside inhibitors binding to allosteric sites distinct from the active site have begun to be identified through high throughput screening (HTS) and crystallographic analysis of the inhibitor–NS5B complex. NNI target the alloenzyme free of substrate and uncomplexed with any other non-structural replicative proteins. They are inactive when the enzyme has entered the processive RNA elongation phase , which suggests that they target preferentially the initiation phase, by blocking the enzyme, thereby preventing conformational change needed for the polymerization.In fact, with the allosteric polymerase inhibitors, the non nucleoside inhibitors, development of antiviral polymerase inhibitors has reached a stage where one could truly talk about rational design of polymerase inhibitors.

Considering the gravity of viral infections, with special reference to BVDV, HCV and HIV viral pathogens, the main objective of my thesis is to address the challenges encountered in the development and improvement of effective antivirals against the former mentioned pathogens. Recent success obtained in combating these viruses by targeting NNI against the viral polymerases have prompted me to study and explore in depth the

molecular recognition of the NNIs against the viral polymerases. This work represents the application of several computational techniques, such as docking, molecular dynamics, algorithms to calculate free energy of binding of ligands into the binding pocket (MM-PBSA) and algorithms to study rare events, such as unbinding of ligand from the binding site (Metadynamics) to explore, at microscopic level, the key pattern of interaction between protein and ligand, to understand the effect of mutations, to gain an insight of the full docking and undocking path and to calculate binding energetics, all with an aim to develop and improve effective NNI antivirals against the former mention three viral polymerase systems.

Due to the recent developments in the field of chemistry, pharmacology, microbiology, biochemistry and biophysics, development of new antivirals is no longer confined to the imagination of chemist but results from a direct dialogue between biologist, chemist and computational biologist. Another ground breaking development was provided by the human genome project[14,15] as well as further projects to resolve the genome of organisms which led to an abundant amount of information about the amino acid sequences of proteins in these organisms. The wealth of knowledge about genetic sequences and sequences of amino acid, is invaluable for the modern taxonomic identification.[16,17] Thus, such DNA bar coding complements the classical way of determining the taxonomy of organisms by their phenotype with a close inspection of the genotype as well.

However, the amino acid sequence of proteins yields only limited information about the structure and function of biological macromolecules. Processes such as 1) molecular recognition with respect to protein-ligand interaction, 2) catalytic function of enzymes, gating mechanisms in ion channels, 3) conformational change required by the protein to accommodate the ligands into binding site, 4) mechanisms of inhibition and drug resistance, 5) why a compound is active or inactive, or how its activity might be improved, 6) change in the protein dynamics upon introduction of resistance mutations, 7) understanding the dynamics of proteins at different timescales, from fast internal motions to slow conformational changes, or and 8) signal transduction or immune response pathways cannot be explained by these sequences.

To gain insight into such processes, the three-dimensional structure of a folded amino acid sequence, a functional protein, is necessary. The structure determination of proteins or nucleic acids is based on experimental techniques such as X-ray crystallography[18,19] or nuclear magnetic resonance (NMR),[20,21] which have made remarkable progress in solving high-resolution structures over the past years. Such structures are archived and accessible via the RSCB Protein Data Bank.[22] Furthermore, theoretical approaches for structure prediction from amino acid sequences are under constant development and have been monitored by CASP (Critical Assessment of Structure Prediction) over the last 14 years.

A particular snapshot of a protein structure is a point in configuration space. Accordingly, a fluctuating structure of a protein is described by a region on a high-dimensional complex free energy landscape, which is dynamically explored.[23] Such protein structure dynamics is the key to obtain insight into the function of bio-molecules. A structure derived by classical X-ray crystallography is just an average and thus lacks dynamics. Moreover, the free energy minimum of the structure derived with the help of such a protein crystal is governed by the crystallization conditions and possibly changed by the typically low temperature usage. However, experimental methods to probe the structural dynamics with atomic detail are available, even though with intrinsic limitations. Time-resolved X-ray crystallography[24, 25] reveals conformational protein motions on the picosecond time scale, but wide-spread use is impeded due to the massive experimental effort involved. NMR relaxation measurements[26, 27, 28, 29] have been used to probe protein dynamics on pico- to nanosecond and micro- to millisecond time-scales, and thus fail to reveal dynamics on the nano- to microsecond time scale.

Such time-resolved dynamics is needed to understand molecular recognition mechanisms, which enable specific binding of drugs to target proteins, or binding of polymerases to antivirals leading to inhibition of the polymerase process. Moreover, as has been shown in a recent study of ubiquitin with a combination of residual dipolar coupling (RDC) NMR and Molecular Dynamics (MD) simulations,[30] conformational changes in proteins, and, in this context, the determination between induced fit and conformational selection mechanisms can be observed on the nano to microsecond time scale.

Experimental techniques like replicon, cell-based and enzymatic assays can be used to show that a compound effectively inhibits protein target by providing CC_{50}, EC_{50} and IC_{50} values. An approximate estimation of binding affinity can be obtained from the IC_{50} values. Also, experimental techniques, such as fluorescence spectroscopy[31] or surface plasmon resonance spectroscopy[32, 33] are used to obtain binding affinities of receptor/ligand complexes via the equilibrium dissociation constant (KD), which is directly related to the absolute free energy difference of the binding process.. However, measuring these binding affinities can be limited by experimental conditions, especially if large ligand concentrations are needed .
Computational approaches to assess the binding affinity of receptor/ligand complexes are thus very helpful when experimental measurements are brought to their limits, and therefore aim to have a predictive function to assess such binding affinities. Methods such as ligand-docking have nowadays become almost routine techniques in medicinal chemistry, and are widely used in both the lead-discovery and the lead optimization phases in rational drug design.

The general docking problem consists in the prediction of the best matching between two or more molecular entities that form an intermolecular complex.[34] In this context, ligand-docking methods attempt to find the most favorable configuration, which maximizes the interaction and minimizes the binding energy, of a small molecule inside a binding site in a, typically, rigid protein. The ligand is fitted into this binding site by flexible rotation of functional groups within the ligand, and rotation of the ligand itself. Electrostatic and van der Waals interactions are calculated for the different conformations and a scoring function evaluates the ligand conformations, which energetically fit best. Such a challenging issue might in principle be addressed at different levels of complexity for the physical description of the system under investigation. Nevertheless, molecular docking is always concerned with two closely related sub-problems, [35,34] namely, the sampling of the configurational space (pose generation), and the evaluation of the stability of the intermolecular complex generated at the previous step (pose scoring). The most popular docking programs employed by the pharmaceutical community usually trade accuracy for speed.

Several applications with different scoring functions, such as Autodock,[36] FlexX,[37], Gold,[38] or GLIDE,[39] have been developed over the past years. It is common to these applications that an implicit solvent environment is used and the flexibility of the protein is often neglected. Despite the advantage to screen large libraries of ligands in a short amount of time, the different scoring functions often lead to different and inconsistent results [40]. This inconsistency could be attributed to several limitations that are summarized as follows:

(i) limited coverage of the conformational degrees of freedom,
(ii) only partial accounting for the role of the solvent,
(iii) approximate estimation of the binding free energy,
(iv) lack of insight on the docking path and the transition state,
(v) approximate treatment of target and ligand flexibility

Nevertheless, despite of the several limitations of the procedure, molecular docking is a powerful tool to provide the first guess of the binding site (when no information available), the initial starting orientation for further refinement (e.g by molecular dynamics simulations), and a clue about the binding affinities of the compound into the binding site.

Molecular Dynamics (MD) is a powerful theoretical method routinely used to simulate the dynamics of complex physical, chemical and biochemical systems. Its success stems from the improvement of the force field describing the interactions among the many constituents of the systems, the development of a battery of algorithms and the availability of powerful computers. The dynamic properties and processes of molecules can thus be investigated by researchers in a high number of fields such as structural biochemistry, biophysics, molecular biology and pharmaceutical issues . Using MD simulations the thermodynamic properties and time-dependent phenomena (i.e. kinetic) can be studied and this allows an understanding of various dynamic aspects of biomolecular structure, protein-ligand recognition and function. The major strengths of MD simulations are the possibility of its combination with statistical mechanics, which connects microscopic simulation with macroscopic observables. Statistical mechanics can provide a rigorous framework of mathematical expressions that can relate the distribution and motion of atoms with macroscopic observables such as temperature, pressure, heat capacity and free energies. In this way we are able to predict, for instance, changes in the binding free energy of a particular drug candidate or the mechanisms and energetic

consequences of conformational changes in a protein along with study of macromolecular stability [41], the role of dynamics in enzyme activity [43,42], molecular recognition and the properties of complexes [44] and small molecule transport [45], etc.

Despite being powerful technique MD has its own set of limitations. MD simulations allow investigating processes occurring on timescales of ~100ns. However, many interesting and relevant biological process happen on time scales that are orders of magnitude larger, and are therefore termed as rare events. For example, the association and dissociation of ligand from its binding cavity (~μs), protein folding (μs-few seconds), protein-protein interactions, transport of molecules across membrane channels (order ~μs) and many others. Over the years, we have observed an astounding increase in computer power (Blue gene, DESRES), which promise to increase utility of MD simulations to investigate more and more complex systems on μs timescale. However, these supercomputing machines require and rely they power also on the development of new algorithms.: Examples of such effort are the new algorithms that allow overcoming the time problem keeping the atomistic description intact by accelerating rare events to timescales reachable in MD simulations.

One of such approaches able to predict rare events within the framework of an all-atom description is metadynamics.[45,46] Metadynamics [45, 47] has been shown to be a useful method in docking simulations, [48,49] allowing different binding poses to be efficiently explored, and providing the free energy profile of the docking and undocking mechanism. In particular, metadynamics was successfully applied in this thesis to our protein-compound complexes in order 1) identify alternative binding poses other than the ones predicted by docking, 2) to compute the relative free energy of the poses obtained as well as to study the unbinding process of compounds with respect to the target proteins.

However, metadynamics have some limitations, which make it a not very practical choice for the docking. In particular, it requires a careful choice of collective variables (CVs). In order to reconstruct the FES, the Cvs must describe all the slow events that are relevant to the ligand binding/unbinding. In absence of a relevent CV, the reconstruction of the FES fails. This often implies the need to run multiple metadynamics simulations to understand which CVs are needed for each protein–ligand complex. A second limitation of metadynamics concerns the determination of the binding free energy that will requires the simulation via metadynamics of the binding process besides the unbinding. However, frequently the CVS that are appropriate for the study of the unbinding are useless for the association, making the use of metadynamics cumbersome and not well suited for the specific issue. Herein, to address the evaluation of the binding free energy we utilized an other methodology namely, molecular mechanics Poisson–Boltzmann surface area (MM/PBSA), to dissect the free energy at a quantitative level into its enthalpic and entropic contributions of the poses obtained from metadynamics. MM/PBSA is held to be one of the more computationally tractable means of obtaining reasonable estimates of the binding free energy of a complex system. In comparison with other methodologies reported in the literature, like free energy perturbation [49] and linear interaction energy [50], MM/PBSA is more validated in details. The MM/PBSA approach is attractive as 1) it does not contain any parameters that vary for different ligand receptor systems and it involves a set of physically well-defined terms and 2) the solvent is treated implicitly and the electrostatic components are obtained from a dielectric continuum model with a dielectric constant for the solute and the solvent.

Aim of the work:

The aim of this thesis is to carry out a comprehensive usage of the methodology mentioned above in order to explore, at microscopic level, the ligand-target recognition process, binding energetics, the effect of mutation on the dynamics of the protein as well as to understand the mechanism of drug inhibition and drug resistance, all this, with a view to provide useful hints in guiding the rational design for development of more selective and potent anti-viral compounds.

The plethora of computational methods was used to study action of antiviral compounds against polymerases of BVDV, HCV and HIV viruses.

Organization of Thesis

This thesis is organized as follows. Chapter 2 includes details on the main ACTORS of our study which include Polymerases of BVDV, HCV and HIV viruses and the promising Antiviral compounds. Thereafter, in Chapter 3 , extensive description of the methods used to solve the biological problem is provided. The rest of the thesis has the following chapters:

Chapter3: Different binding sites of the benzimidazole compound 227G on HCV and BVDV RdRps revealed by MD simulations.

Key-abstract
The virally encoded RdRp has emerged as a prime target in search for specific HCV and BVDV antiviral. Benzimidazole class of compound has been extensively pursued as potent Non-Nucleoside Inhibitors (NNI's) against both classes of viruses. For the first time, Screening efforts in our lab have come out with potent NNI belonging to benzimidazole class of compound which is active against both HCV and BVDV RdRp. Our tested compound was confirmed to strongly inhibit BVDV RdRp activity in a dose-dependent manner and they also showed inhibition on HCV1b-NS5B.

While, the HCV RdRp structure in complex with the related analogs showed that these inhibitors bind at the surface of the thumb, the resistant mutation and newly identified binding site for the same analog were found to lye in the finger domain of BVDV RdRp. Thus, the molecular mechanism of inhibition of the inhibitor found to be different in the otherwise closely related polymerases.
Molecular dynamics simulations helped us to gain a deeper insight into the interaction pattern, mechanism of inhibition of inhibitor onto the two RdRps. The multidisciplinary approach allowed us to (i) critically assess the dogma of antiviral therapy; (ii) to screen efficiently the possible information valuable for designing effective and selective inhibitors by identifying at a molecular level the interaction pattern of the compound with its viral targets; (iii) to assess the use of BVDV as surrogate for the development of antivirals against HCV

Results:
Experimental

1. The benzimidazole compound 227G is active against HCV and BVDV RdRps with low CC_{50}, EC_{50} and IC_{50} values. First time a benzimidazole is active against both RdRps.
2. Crystal structures of HCV in complex with indole compounds have been reported in the literature (other's works), while here, we have performed mutagenesis studies on BVDV showing RdRp to be the target of 227G.

Computational

1. Site identification: 227G, in HCV bind on thumb domain while in BVDV, a **new** binding site was identified in the finger domain.
2. In HCV the binding mode of 227G turns out to be very similar to those of indole compounds. The binding of 227G disrupts the connection between the fingertip and the thumb domains, leading the enzyme to assume an open (and inactive) conformation. This is likely to be the mechanism of action by 227G.
3. In BVDV the binding site is different, as well is the mechanism of action. Indeed the drug binds to the finger domain and by interacting with the four loops from finger and thumb region completely closes the entrance of the channel to the RNA template. Performing different analyses has validated this finding.

Conclusion:

1. The benzimidazole derivative 227G is reported to be active against both HCV and BVDV at low micro-molar range.
2. Different Binding site :Despite, 227G, being active against both RdRp, its binding site on both the HCV and BVDV RdRp is different. In HCV RdRp it binds in a thumb domain while in BVDV RdRp it was found to bind in the finger domain of RdRp.
3. Mechanism of inhibition: We hypothesized the mechanism of inhibition of 227G in BVDV and HCV RdRp. In HCV RdRp, 227G binds in a thumb domain, disrupting the connection between the thumb and finger domain and thereby stop polymerization while in BVDV RdRp it blocks the entrance to the template channel and ultimately the inhibits te polymerisation.
4. The inventory of key interaction pattern between 227G and HCV and BVDV RdRp were obatined with an aim of providing possible hints at improving potency of 227G.

Chapter 4: Quinoline Tricyclic Derivatives Design, Synthesis and Preliminary *In vitro* and *In silico* Antiviral Activity Against Flaviviridae Family of Three New Classes of Virus-Encoded RNA-Dependent RNA Polymerase (RdRp) Inhibitors.

Key-abstract:

In this preliminary study three new classes of RNA-dependent RNA polymerase (RdRp) of Flaviviridae inhibitors, the linear N-tricyclic systems derived by condensation of the quinoline nucleus with 1,2,3-triazole, imidazole or pyrazine (obtaining triazolo[4,5-g]quinolines, imidazo[4,5-g]quinolines and pyrido[2,3-g]quinoxalines respectively), has been discovered. In particular the activity of the title compounds were evaluated in cell culture systems against YFV (as viruses representative of Flaviviruses) and BVDV (Pestiviruses). Furthermore the cytotoxicity was evaluated in parallel cell-based assays. 1:1 Mixture of bis-triazoloquinolines (1m), imidazoquinolines (2e,h) and pyridoquinoxalines (4h,j and 5n) showed anti-BVDV activity in the range 1-5 μM in cell-based assays. Mutation experiments versus both NS5b and NS3 enzymes of resistant strains of BVDV of two selected compounds (2h and 5m) confirmed the inhibition of the enzyme target. Overall, by these experiments the imidazo[4,5-g]quinoline (2h) emerged as a potent BVDV polymerase inhibitor endowed with IC50 = of 0.06 μM. Herein, we have utilized several biophysical computational approaches, ranging from Docking, standard molecular dynamics and metadynamics , in order to predict the binding site, to identify key determinants of ligand binding, the energetics of unbinding as well as the escape mechanism of the lead compounds PS999 and PS1036 in BVDV RdRp. MM/PBSA techniques were used to calculate the binding free energies of 227G in both HCV and BVDV RdRp. Calculated binding free energies were then converted to computational IC_{50} which were then compared with the experimentally available IC_{50} values.

Biological Results:

1. Compounds belonging to triazolo[4,5-*g*]quinolines, imidazo[4,5-*g*]quinolines and pyrido[2,3-*g*]quinoxalines were discovered and their activity were evaluated against YFV and BVDV viruses.
2. Mutation experiments confirmed BVDV RdRp as the target for the compounds

Computational results

1. Identified Binding cavity for leads belonging to pyridoxoquinoxaline(PS1036) as well as imidazoquinoline(PS999) derivative in BVDV RdRp. The choice for the study of the leads was based on the knowledge of the resistant mutations against these compound.
2. MD simulations on the complexes obtained from docking were performed. Key interaction responsible for the binding of compounds in the putative binding cavity were determined
3. Metadynamics runs were performed to identify alternate binding poses other than that obtained from docking. Energetics of unbinding was estimated and pathway for the dissociation of the compounds was computed.
4. Binding energy of the stable poses obtained from metadynamics were calculated. The calculated IC_{50} value coincided well with the experimental IC_{50} values.

Conclusion

1. Conglomeration of resistant mutation of different classes of compounds(in this work, pyridoxoquinoxalines and imidazoquinoline) in the finger domain of BVDV RdRp indicated that the binding site of NNI may lye in the finger domain. We performed extensive computational studies and identified stable poses of the compounds in the finger domain of BVDV RdRp. Interstingly, 227G, a benzimidazole compound,(chapetr 1) was found to bind in the same cavity. Thus we proposed that finger domain can be a HOTSPOT for NNI binding.
2. We proposed a similar hypothesis for mechanism of inhibition as in case of 227G. Since the compound target same binding site hence they can stop polymerisation by blocking the entrance to the template channel.
3. Extracted key binding features of compounds in binding pocket can provide basis for the development of more efficient antivirals.

Chapter-5: Identification of binding cavity for novel Non-Nucleodside inhibitors of HCV RdRp: A molecular docking study

Abstract:
The virus encoded RdRp has emerged as a prime target in the search for specific HCV antivirals. Identification and successes of NNI, which are non-competitive diverse small molecules, against other polymerase, have encouraged them to be used as effective antivirals against HCV RdRp. Screening studies have came out with compounds belonging to Pyridoquinoxaline(PS1097) Phenanthroline(PS1101) and Imidazoquinolines (PS1126), effective against HCV RdRp at a very low micromolar range. Information regarding the binding cavity of these compounds are still lacking. Herein, we have utilized docking procedure to investigate binding sites, binding modes as well as binding affinity of these compounds in HCV RdRp.To rule out any bias, each of the three reported HCV NS5B NNI binding site represented by (A) Benzofurans (B) N,N-disubstituted phenylalanine and (C) benzothiadiazine inhibitors was examined for the binding of the compounds. We then used all atom standard molecular dynamics (MD) simulations to investigate the most probable binding site for the compounds as in both the complexes, there was a dissociation in two out of the three cavities under investigation. Therefore, Identifying the binding cavities and orientation of our compounds can help us to build a microscopically well-funded picture to elucidate the mode of action of compounds and thereby providing clues for rational drug design.

Results:

1. A blind docking procedure was applied in all four well known binding sites reported for different classes of NNIs in HCV RdRp, to obtained the most prominent binding pocket for compound PS1097 (Pyridoxiquinoline) and compound PS1126 (Imidazoquinoline) and PS1101(Phenanthrolines).
2. We selected the most stable docking pose for each binding site to do MD simulation for better characterization of their binding affinity as well as to refine the docking poses .
3. We calculated the binding free energy of these compounds. From the binding free energy we obtained a IC_{50} value in very good agreement with the experimental one.

Conclusion

1. We identified probable binding cavity for NNIs belonging to three different classes Pyridoxoquinoxaline, Imidazoquinoxaline and phenonthrolines. This was an encouraging performance, given that all molecular modeling studies were performed in the absence of any available crystal structure of the protein in complex with these classes of inhibitors.
2. Key binding pattern of the compounds in the cavity were identified with an aim of further optimisation of the lead compounds in order to increase their potency.

3. Calculated binding free energies were found to be in agreement with the experimental data.
4. Furthermore, we were successful in throwing light on the probable mechanism of inhibition of the three compounds reported in this paper.

Work not presented here

Chapter 1: Point mutation I261M affects the dynamics of BVDV and its interaction with benzimidazole antiviral 227G.

abstract

Bovine viral diarrhea virus (BVDV) is a *Pestivirus* of the *Flaviviridae* family and represents a major viral pathogen in cattle and other ruminants. Infection with BVDV can result in a wide assortment of disease manifestations including resorption, mummification, or abortion of the dead fetus. Extensive study of BVDV is required not only because it causes heavy agronomic losses but also it is considered to be a valuable surrogate model for the study of the hepatitis C virus (HCV). Recently the point mutation I261M on the thumb domain was shown to confer resistance to BDVD against 227G and other benzimidazole compounds.

Here we investigated the role of this mutation by using a non conventional approach, which is not based on binding free energy calculations on structures of the mutated complex which are taken a priori similar to those of the wild one. Namely, we firstly performed MD simulations on the wild and mutated BVDV RdRp proteins in aqueous solution. Then, we selected representative equilibrium conformations by performing a cluster analysis, and ran docking calculations of 277G on representative of the 5 most populated clusters of each protein. Finally, high-score poses were subjected to MD simulations to assess structural and dynamical differences between wild and mutated 227G-protein adducts.

Interestingly, the mutation affects the structure and the dynamics of the protein, particularly in the region of binding of the ligand, and this results in different binding sites of 227G onto the two proteins. Moreover, while 227G closes the entrance for the RNA primer in the case of the wild protein, in presence of the mutation a gate and a channel leading to the catalytic site are still present. These results could furnish a possible molecular explanation of the resistance mechanism by mutation I261M.

In a related paper of us, we reported mutagenesis experiments aimed at confirming that BVDV RdRp is the target of our lead benzimidazole compound 227G. Here, we performed a multidisciplinary computational study to furnish a molecular-level explanation of effect of the resistant mutation. Our results are fully compatible with the mechanism of action furnished in the previous paper. Namely, we see that the mutation I261M alters the binding mode of the ligand, leaving the channels for the entrance of the template RNA, of NTP and the exit of double strand RNA all open. The dynamics of the protein is also altered to a less extent as compared to the wild type.

Results:

Experimental:
1. Site directed mutagenesis was done to identify the resistant mutation of 227G.
2. The mutated protein retains activity.

Computational:
1. Extensive docking studies were performed to identify the best orientation of 227G in COMPLEXm (227G in mutated BVDV). The effect of mutations was not calculated as in previous published works (i.e. in a static manner), but from a dynamical point of view.
2. Long MD simulations (total time ~120 ns) were performed to identify and to compare the effect of mutation on the structure of the APO and HOLO proteins with respect to the wild type one. The

simulated systems are (30ns in each case): APO (apo-protein WT), APOm (apo-protein mutated), COMPLEX (WT with 227G) and COMPLEXm (mutated protein with 227G).

3. Concerning the APO proteins, our simulations reveal:
    1. Significant structural changes in the region of binding, namely the linker region turns away from the binding pocket. This will have important consequences for the binding and the stability of the compound.
    2. The overall dynamics as well as the "functional channels" of the protein are similar, which is consistent with experiments showing that the mutant retains enzymatic activity.
4. Concerning the complexes:
    1. At opposite with COMPLEX, due to the movement of the linker (probably consequence of the presence of the long side chain of M261) in COMPLEXm, 227G is not stable in the initial docking pose, but moves and reorients itself many times until finds a stable position.
    2. In this position the drug is more solvent exposed, and accordingly to this its binding free energy is less. Moreover, in this new conformation the interaction between the drug and the linker and another loop (L3), which leaves the channel for the RNA template entrance opened (as well as the other channels). Thus, in principle the enzyme is still able to work, which still is consistent with experiments and can furnish a molecular-level rationale thereon.
    3. To further strengthen this hypothesis, we verified that also the dynamics of the APO and HOLO proteins are very similar in the mutant, while this is not true in the wild type, where the binding of the drug has a drastic effect.
    4. Concluding, only in the wild type 227G is able to alter structure and dynamics, so altering the functionality.

Chapter2 : Inhibition of viral RNA polymerase investigated by computer simulations

Abstract:
The drug resistant of various RdRp inhibitors has been studied using a new computational protocol, that is, extensive molecular docking on cluster basis, leading to valuable insight into the resistance mechanisms and structure –resistance corrections of the RdRp inhibitors associated with mutation lying in the strategic location inside the motif-I. By using the cluster basis molecular docking method, the calculated mutation caused shifts of the binding free energies linearly correlate very well with those derived from the corresponding experimental data, suggesting that the newly proposed protocol may be used as generalized approach to predict drug resistance associated with mutation I261M. Because it is essentially important for understanding the structure-resistance correlation and for structure based drug design to develop an effective computational protocol for drug resistance prediction, the reasonable and computationally efficient protocol for drug resistance prediction should be valuable for future structure-based design and discovery of anti-resistance drugs in various therapeutic areas.

Result :

1. We develop a new protocol to identify and to validate a proper binding mode using molecular docking, molecular dynamics, clustering, MM-PBSA and Metadynamics techniques.
2. We tried to explore the rare events, as to pull out the drug from their binding site "escape mechanism" through a new algorithm "metaD". During escape or dissociation of drug from the cavity, we observed some local transient binding site, which was barely noticed by x-ray and other techniques. The trapping of the drug into these local transient binding pockets could be a reason of increase of residence time of drug into binding cavity.
3. We compare the escape mechanism of 227G from wild-227G complex and Mutated-227G complex, and we observed there is a very low barrier ~5kcal/mol was required for the 227G to come out from the binding site, which is ~12kcal/mol in case of wild-227G complex, could be a loss of entropy in wild-227G complex.

Chapter3 :Combining Docking, Molecular Dynamics to Predict Binding Modes and Affinities for Non-

nucleoside Inhibitors to HIV-1 Reverse Transcriptase

Abstract:

The reverse transcriptase of human immunodeficiency virus (HIV) catalyzes a series of reactions to convert the single-stranded RNA genome of HIV into double-stranded DNA for host-cell integration, This task requires the reverse transcriptase from the viral genome. Docking, scoring, molecular dynamics (MD), methods are used here to predict binding modes and affinities for a set of 4 non-nucleoside inhibitors to HIV-1 reverse transcriptase. The location of each drug outside the binding pocket was determined by an automated docking program, and steering into the binding pocket followed a route that is likely to represent the actual entrance pathway. The comparison of the dynamics of DABO with crystal structure of structurally similar compounds has provide information on possible molecular mechanisms of ligand binding, specificity and regulation of RT. The present calculations provide a validation of the combination of docking, MD as a powerful tool in structure-based drug design, and the methodology is easily scalable for attaining a higher throughput of compounds. This coherent picture strongly suggests that attempts to understand through the structure-based drug design may be considerably more successful if dynamic structural aspects of the type studied here are considered,particularly in those region which is more flexible but functionally important.

Results:

1)-To validate, how the binding of MC1220 at an allosteric site affects the conformational dynamics of RT such that it inhibits DNA polymerization.
2)-A comparative analysis have been performed to understand the triggering role of “X” and “Y” position in DABO series.
3)- The probable mechanism of inhibition was gel well with the hypothesis of inhibition of DABOs.
4)- MD simulation has been done in order to under stand the key determinants and binding pattern of DABO and its analogs.

Thesis in Flow-chart

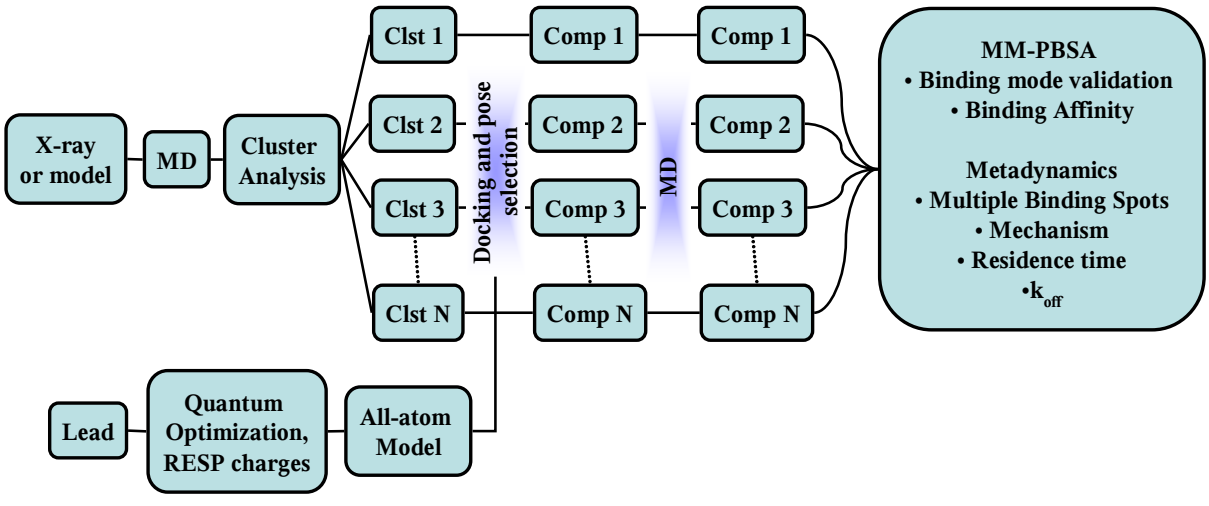
Antiviral Drug Development

Biological System (Polymerases)

BVDV(1S48) **HCV**(1NB4, 2BRK) **HIV** (1RT1,1DLO)

Class	Inhibitors	Class	Inhibitors	Class	Inhibitors
Benzi	$227G^{0.002}$	**Benzi**	$227G^{0.4}$	**DABOs**	$MC1220^{0.004}$
	$ASCC170^{3.0}$				$MC1332^{0.008}$
	$ASCC169^{12}$	**Imido**	$PS1126^{2.0}$		$MC1346^{0.005}$
Imido	$PS999^{0.6}$	**Pyrido**	$PS1097^{0.6}$		$MC1195^{0.05}$
Pyrido	$PS1036^{1.0}$	**Phenan**	$PS1101^{1.0}$		

Receptors and Antivirals; IC_{50} values in case if BVDV and HCV, while EC_{50} in Case of HIV-RT are in superscript form. The X-ray structures mainly used are written as superscript on receptors.



Computational methodology used to understand protein-ligand recognition.

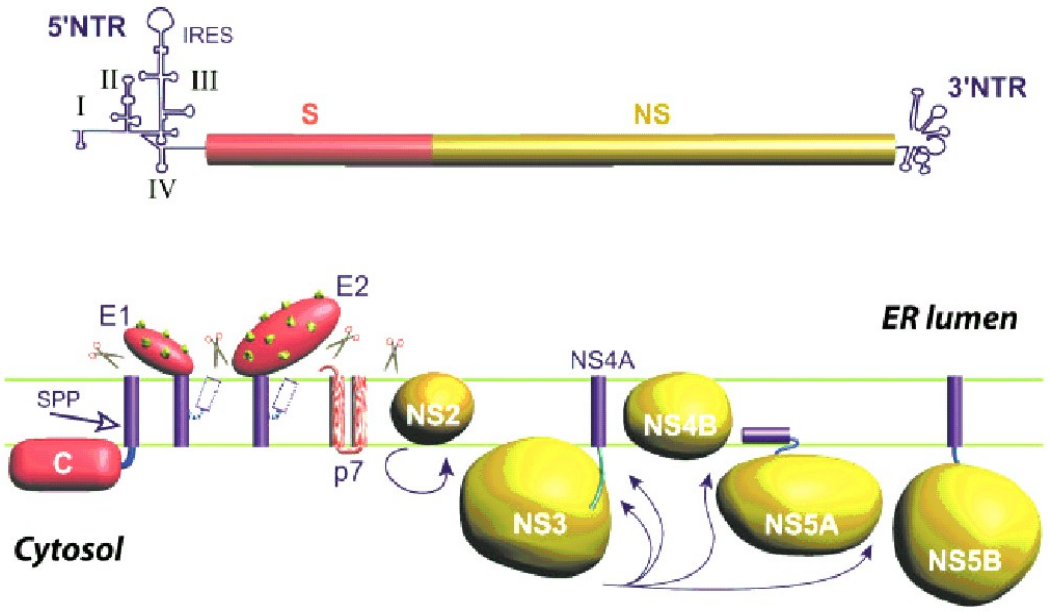
The Actors: Polymerases and Antivirals

Polymerases
(HCV, BVDV and HIV)

"*Infections by RNA viruses continue to exist as significant public health problems worldwide. So there is an urgent need for safer and more efficacious treatment option against infections caused by RNA viruses, researchers have devoted significant efforts over the last two decades to discovering and developing new antiviral agents*" *by Todd Apple*

Genomic organization:

Hepatitis C virus is a positive-sense single-stranded RNA virus with a genome of initiation of viral RNA translation. The HCV genomic RNA encodes a polyprotein of 3010–3011 amino acids which undergoes co-translational and post-translational proteolytic processing in the cytoplasm or in the endoplasmic reticulum of the infected cell to give rise to four structural and six non-structural (NS) proteins [51]. Fig 3 summarizes the information concerning the identification and the function of individual gene products. The structural proteins consist of the capsid or core protein C, two envelope glycoproteins E1 and E2, and small hydrophobic polypeptide p7. The non-structural proteins are NS2, NS3, NS4A and NS4B, NS5A and NS5B. NS2 has a zinc-stimulated protease activity that cleaves the NS2/NS3 junction [52]. NS3, a 70 kDa protein, has been subject of intensive study because it is a multifunctional molecule with a trypsin-like serine protease catalytic domain at the N-terminal 181 residues [53], and a nucleoside triphosphatase (NTPase)/RNA about 9.6 kb [54]. It consists of a single open-reading frame (ORF) flanked by 5' and 3' non-translated regions (NTRs) [55]. The 5' NTR contains the internal ribosome-entry site (IRES). NS4A is a small protein (8 kDa) that is an NS3 protease cofactor [56]. NS4B is a hydrophobic, 27 kDa protein, and it has been suggested that it may function as an anchor to secure part of the HCV replication apparatus to the endoplasmic reticulum (ER) membrane [57]. Two cytoplasmic phosphoproteins, p56 and p58 (56 kDa and 58 kDa, respectively) are produced from the NS5A region of the HCV genome [58,59]. Both proteins are phosphorylated at serine residues in the region between amino acids 2200 and 2250 and in the C-terminal region of NS5A. Although the function of NS5A in viral replication is unknown, it is possible that NS5A might be involved in the resistance of HCV to a-interferon therapy [60]. The purified enzyme NS5B exhibits a highly processive primer-independent RNA-dependent RNA polymerase (RdRp) activity capable of copying in vitro transcribed full-length genomic HCV RNA [61].



Amino acids	Protein	Functional
1-191	C	nucleocapsid
192-383	E1	Envelope glycoprotein
384-746	E2	Envelope glycoprotein
747-809	P7	virioporin
810-1026	NS2	Zn-activated NS2/3 auto-protease
1027-1657	NS3	Ser protease (aa 1-180)
		RNA helicase (aa 181-631)
1658-1711	NS4A	NS3 Ser protease co-factor (aa 21-34)
1712-1972	NS4B	Induces the formation of intracellular membrane ves.
1973-2420	NS5A	alpha-interferon resistance
2420-3010	NS5B	RNA-dependent-RNA-polymerase

Figure3: HCV genome organization (top) and polyprotein processing (bottom). The 5 untranslated region (UTR) consists of four highly structured domains and contains the internal ribosome entry site (IRES). The 3 UTR consists of stable stem-loop structures and an internal poly(U)/polypyrimidine tract. The central 9.6-kb ORF codes for a polyprotein of slightly more than 3000 aa depending on the HCV genotype. S and NS correspond to regions coding for structural and nonstructural proteins, respectively. The polyprotein processing and the location of the 10 HCV proteins relative to the ER membrane are schematically represented. Scissors indicate ER signal peptidase cleavage sites; cyclic arrow, autocatalytic cleavage of the NS2-NS3 junction; black arrows, NS3/NS4A proteinase complex cleavage sites; intramembrane arrow, cleavage by the signal peptide peptidase. The transmembrane domains of E1 and E2 are shown after signal-peptidase cleavage and reorientation of the respective C-terminus hydrophobic stretches (dotted rectangles). Green spots denote glycosylation sites of the E1 and E2 envelope proteins. Adapted from : Hepatology. 2004 Jan;39(1):5-19.Structural biology of hepatitis C virus.Penin F, Dubuisson J, Rey FA, Moradpour D, Pawlotsky JM.

BVDV Genome

Pestiviruses have a positive sense single stranded RNA genome [ss (+) RNA]. The BVDV genome has a size of approximately 12'500 nucleotides, the sequence of which is known since 1988. Some cp strains integrate small variable segments of viral nucleic acid or the host cell genome in certain places of their genome (in NS2 or between NS2 and NS3). Others show duplicates of specific protein coding regions (Npro, NS3), consequently their genome size rises to about 16.5 kb. The genomic RNA has one open reading frame of about 4000 codons, which encompasses most of the viral genome. Translation of the BVDV genome yields one precursor poly protein, which is cleaved co- and postranslationally by viral and host cell encoded proteases (”processing”). Most of the virally encoded cleaving is catalysed by a serine protease domain within NS3 and generates the non-structural proteins NS3 to NS5B, whereas the structural proteins are believed to be cleaved by cellular proteases.

HIV Genome

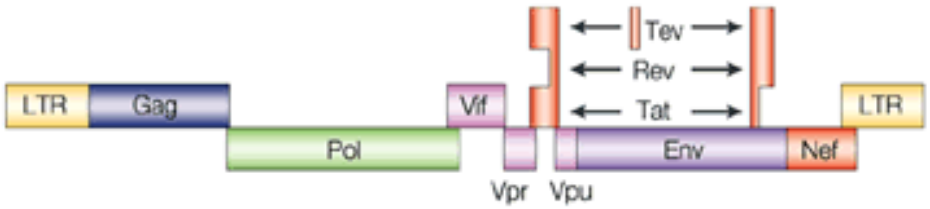


Figure4: Figure depicts HIV genome organization.

Human Immunodeficiency Virus type 1 (HIV-1) is a single stranded +RNA lentivirus from the Retroviridae family. Genomic organization is listed as below (Fig. 4):

GAG The genomic region encoding the capsid proteins (group specific antigens). The precursor is the p55 myristoylated protein, which is processed to p17 (MAtrix), p24 (CApsid), p7 (NucleoCapsid), and p6 proteins, by the viral protease. Gag associates with the plasma membrane, where virus assembly takes place. The 55-kDa Gag precursor is called assemblin to indicate its role in viral assembly.

POL The genomic region encoding the viral enzymes protease, reverse transcriptase, and integrase. These enzymes are produced as a Gag-Pol precursor polyprotein, which is processed by the viral protease; the Gag-Pol precursor is produced by ribosome frameshifting near the 3' end of *gag*.

ENV Viral glycoproteins produced as a precursor (gp160), which is processed to give a noncovalent complex of the external glycoprotein gp120 and the transmembrane glycoprotein gp41. The *env* gene does not actually code for gp120 and gp41, but for a precursor to both, gp160. During HIV reproduction, the host cell's own enzymes cleave gp160 into gp120 and gp41. The host cell protease that cleaves Env into gp120 and gp41 is Furin.

TAT Transactivator of HIV gene expression. Tat vastly increases the level of transcription of the HIV dsRNA. Before Tat is present, a small number of RNA transcripts will be made, which allow the Tat protein to be produced. Tat then binds to cellular factors and mediates their phosphorylation, resulting in increased transcription of all HIV genes, providing a positive feedback cycle.

REV The second necessary regulatory factor for HIV expression. A 19-kD phosphoprotein, localized primarily in the nucleolus/nucleus, Rev acts by binding to RRE and promoting the nuclear export, stabilization, and utilization of the viral mRNAs containing RRE. Rev is considered the most functionally conserved regulatory protein of lentiviruses. Rev cycles rapidly between the nucleus and the cytoplasm.

VIF Viral infectivity factor, a basic protein typically 23 kD. Promotes the infectivity but not the production of viral particles. In the absence of Vif, the produced viral particles are defective, while the cell-to-cell transmission of virus is not affected significantly. Found in almost all lentiviruses, Vif is a cytoplasmic protein, existing in both a soluble cytosolic form and a membrane-associated form. The latter form of Vif is a peripheral membrane protein that is tightly associated with the cytoplasmic side of cellular membranes.

VPR Vpr (viral protein R) is a 96-amino acid (14-kD) protein, which is incorporated into the virion. It interacts with the p6 Gag part of the Pr55 Gag precursor. Vpr detected in the cell is localized to the nucleus. Proposed functions for Vpr include the targeting the nuclear import of preintegration complexes, cell growth arrest, transactivation of cellular genes, and induction of cellular differentiation. In HIV-2, SIV-SMM, SIV-RCM, SIV-MND-2, and SIV-DRL the Vpx gene is apparently the result of a Vpr gene duplication event, possibly by recombination.

VPU Vpu (viral protein U) is unique to HIV-1, SIVcpz (the closest SIV relative of HIV-1), SIV-GSN, SIV-MUS, SIV-MON and SIV-DEN. There is no similar gene in HIV-2, SIV-SMM, or other SIVs. Vpu is a 16-kd (81-amino acid) type I integral membrane protein with at least two different biological functions: (a) degradation of CD4 in the endoplasmic reticulum, and (b) enhancement of virion release from the plasma membrane of HIV-1-infected cells. Env and Vpu are expressed from a bicistronic mRNA. Vpu probably possesses an N-terminal hydrophobic membrane anchor and a hydrophilic moiety.

NEF A multifunctional 27-kd myristoylated protein produced by an ORF located at the 3' end of the primate lentiviruses. Other forms of Nef are known, including nonmyristoylated variants. Nef is predominantly cytoplasmic and associated with the plasma membrane via the myristoyl residue linked to the conserved second amino acid (Gly).

VPX A virion protein of 12 kD found in HIV-2, SIV-SMM, SIV-RCM, SIV-MND-2, and SIV-DRL and not in HIV-1 or other SIVs. This accessory gene is a homolog of HIV-1 vpr, and viruses with vpx carry both vpr and vpx. Vpx function in relation to Vpr is not fully elucidated; both are incorporated into virions at levels comparable to Gag proteins through interactions with Gag p6. Vpx is necessary for efficient replication of SIV-SMM in PBMCs. Figure 4. depicts HIV genome organization.

Domain Architecture (HCV, BVDV and HIV)

Viral polymerase

It has been already mentioned in the introduction that Polymerase (RdRp in BVDV and HCV, RT in HIV) are considered as one of the most attractive target for the development of effective antivirals, and is a focus of this thesis. Herein this section, we are providing details on the structural organization of the protein with special reference to the different domains. Information regarding the catalytic site and special structural features are also provided here. An in depth understanding of the structural organization of protein is fundamental for a rational approach to designing of new antivirals.

Structure of polymerase subdomains: **HCV**

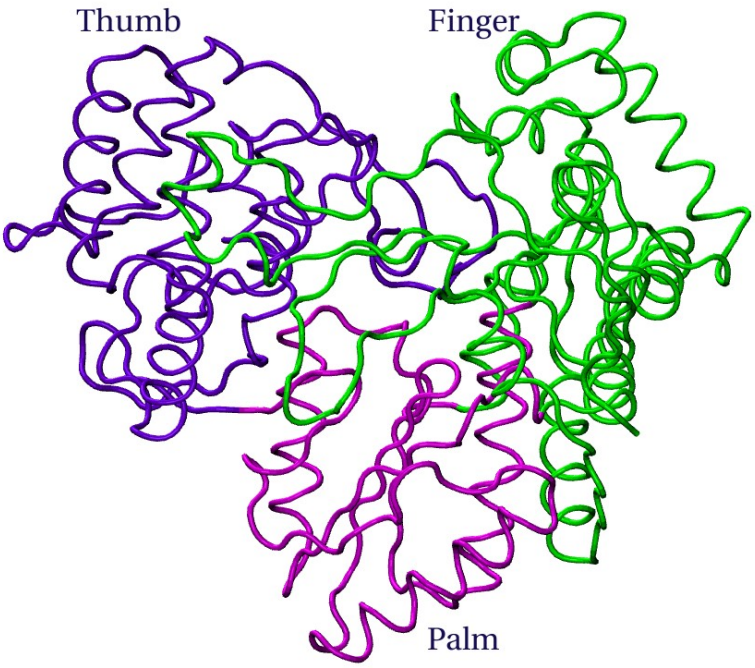


Figure5: Cartoon representation of HCV RdRp. Different domains of RdRp viz thumb in blue, finger in green, and palm in magenta.

A major advance in the understanding of the NS5B polymerase was provided by the resolution of the three-dimensional structures of several truncated forms of the apoenzyme and of complexes with nucleotides or RNA templates [62-67]. The NS5B has the canonical 'right hand' shape, with the characteristic fingers, palm and thumb sub-domains (Figure 5). Similarly to other RdRps, the HCV polymerase has a compact shape due to the presence of two extended loops, the fingertips, that connect the fingers and thumb domains and completely encircle the active site cavity, to which the RNA template and the nucleoside triphosphate (NTP) substrates have access through two positively charged tunnels [64].

The **palm** sub-domain (residue 188 to 227, 287 to 370) constitutes the catalytic site of the molecule and contains the residues responsible for the nucleotidyl transfer reaction within the D-(X)4-D and the GDD motif. The two-metal-ion **catalytic site** is conserved essentially in all known viral polymerases.[68-70] The magnesium ions assist phosphodiester bond formation by polarizing the hydroxyl group at the 3′ terminus of the growing RNA chain, facilitating nucleophilic attack upon the dNTP substrate in addition to stabilizing the transition state in which the phosphorus of the dNTP is linked to five oxygen atoms. The residues essential for metal binding are Asp220 and the carbonyl of the Thr221 peptide backbone, as well as Asp318 and Asp319.

Residues 1 to 187 and 228 to 286 constitute the **finger** sub-domain. The unique shape of the HCV NS5B RNA-dependent RNA polymerase (RdRp) arises from extended packing interactions between the fingers and thumb sub-domains. This inter domain linkage is provided by insertions within the fingers sub-domain. These insertions form two loops that pack against the thumb sub-domain (Figure5) The longer loop is located near the N-terminus and consists of residues 11–45; the shorter includes residues 139–160. These loops regulate the coordinated movements of the fingers and thumb during the polymerase reaction cycle.[62]

The **thumb** sub-domain of NS5B includes residues (371–528). Two structural elements peculiar to the NS5B structure are a **β-hairpin** (β-loop), protruding from the thumb into the active site, and a **C- terminal** region, located immediately before the transmembrane domain, that folds from the surface of the thumb towards the active site and establishes a series of hydrophobic interactions with a shallow pocket comprised between the palm and thumb sub-domains [66]. These elements (β-hairpin and C-terminus) are probably involved in positioning the 3′ terminus of template RNA and are considered crucial for template selection.[62]

BVDV

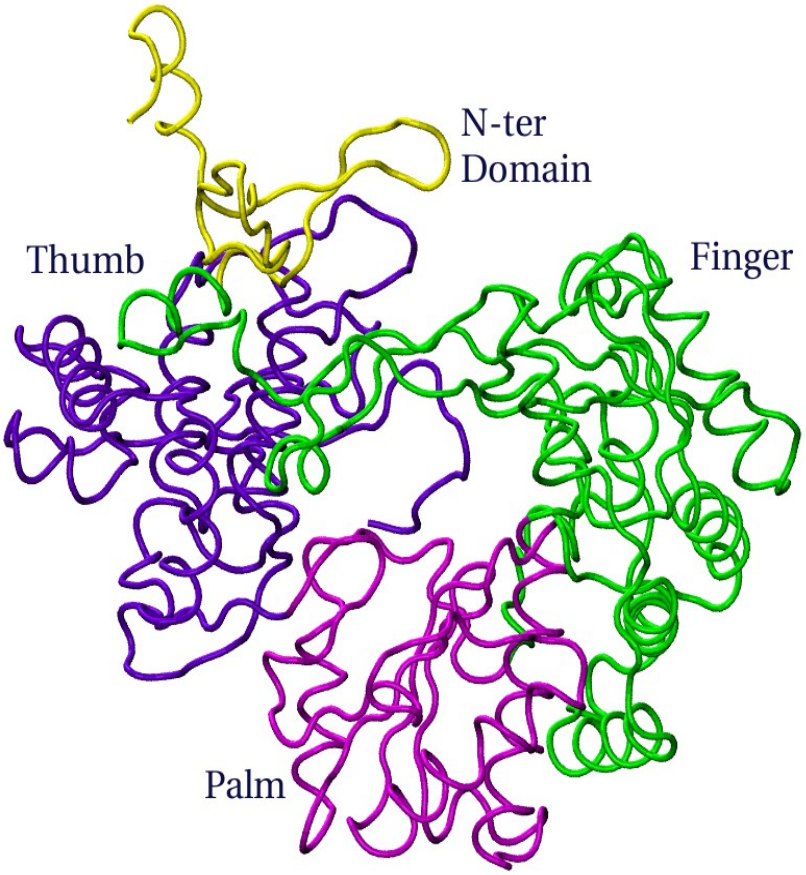


Figure6: Cartoon representation of BVDV RdRp. Different domains of RdRp viz thumb in blue, finger in green, palm in magenta and unique N-terminal is yellow in color.

The BVDV polymerase core (residues 139–679) has a roughly spherical shape. Similar to the HCV RdRp, BVDV RdRp comprises of three sub-domain, that is, finger, palm and thumb (Fig. 6). In addition, there is an N-terminal domain (residues 71–138), which is considered to be unique in BVDV RdRp.

The **finger** domain (residues 139–313 and 351–410) consists of 12 α-helices and 11 β-strands (ref1). In BVDV RdRp, as in other viral RdRps, the N terminus of the finger domain, together with a long insert in the fingers domain (residues 260–288), forms a fingertip region that associates with the thumb domain [71]. Among RdRps the conformation of the polypeptide chain in the fingertip region differs. For example, whereas BVDV polymerase has a three-stranded fingertip region, HCV [72-76] and calicivirus polymerases [77] have a four-stranded, and ϕ6 polymerase a six-stranded fingertip region [78]. The finger and thumb domains are linked through the fingertip region, thus the conformational change induced by template binding into the central channel may be limited [73,74]. The remainder of the finger domain is comprised of a β-strand-rich region (β-fingers) and an α-helix-rich region (α-fingers) close to the palm domain. Although there is little sequence identity, the topology of the fingers domain, excluding the finger-tips, is similar to that of other viral polymerases, making possible the structural alignment of the BVDV, HCV, calicivirus, and ϕ6 polymerase fingers domains. As expected, BVDV polymerase has greater structural similarity to HCV polymerase, indicating a closer evolutionary link between theses two viruses than with the other viruses.

The **palm** domain is the catalytic domain and shows the greatest structural conservation among all known polymerases. The palm domain of BVDV polymerase (residues 314–350, 411–500) consists of a four-stranded β-sheet flanked by two α-helices on one side and an additional α-helix on the other side of the β-sheet, nested between the palm and thumb domains. Compared to HCV polymerase, BVDV polymerase has a small α-helix (residues 433–439) inserted immediately before the first β-strand [77] of the β-sheet. A comparative analysis of RNA polymerase sequences in positive-strand RNA viruses has identified eight conserved sequence motifs, I-VIII [79,80]. Five of these motifs are in the palm domain, including the Gly-Asp-Asp (GDD) motif VI, which is essential for catalytic activity [81].

The **thumb** domain, consisting of the C-terminal region of the polypeptide chain, is the most diverse feature among the known polymerase structures. The thumb domain of BVDV polymerase (residues 501–679) contains eight α-helices and five β-strands and is larger than the thumb domain of other viral RdRps Although there is some structural similarity between the HCV and BVDV thumb domains, the overall topology is rather different. The BVDV polymerase thumb domain possesses a structural element ("the β-thumb region") in which two β-strands and their connecting loops form a layer that occludes the active site cavity. The β-thumb region interacts with the fingers and palm domains through a long C-terminal loop (residues 670–679), tucked between the palm and thumb domains. Together with a loop in the thumb domain, the β-thumb reduces the volume of the template channel. A similar protrusion into the active site, formed by a β-hairpin, is present in the HCV polymerase structure, although the protrusion originates from a different part of the thumb domain .

N-Terminal Domain.
At its N terminus the BVDV RdRp is 130 residues longer than HCV RdRp. The function of this N-terminal domain is not known, although up to 90 residues can be truncated from BVDV polymerase without loss of polymerase activity in vitro [81]. It has been suggested that the N-terminal region in other polymerases is required for binding proteins in the replication complex [83]. The ordered part of the N-terminal region (residues 92–138) folds into a separate domain. The N-terminal domain is situated over the thumb domain, interacting with the fingertip region and thumb domain through a β-hairpin motif. The β-hairpin motif points toward the template channel and partially occludes the channel entrance, possibly explaining why some N-terminal truncated proteins have greater polymerase activity than wild type [81]. The entrance to the template-binding channel created by the hairpin motif of the N-terminal domain is highly positively charged. An equivalent positively charged surface at the entrance to the channel was found in ϕ6 polymerase, where it was proposed to act as a "plough" that unwinds the double-stranded RNA, thus aiding the function of a helicase [78]. Similarly in BVDV polymerase, this positively charged surface might be used to open up RNA secondary structural hairpins before the ss-RNA template entering the active center.

HIV-RT

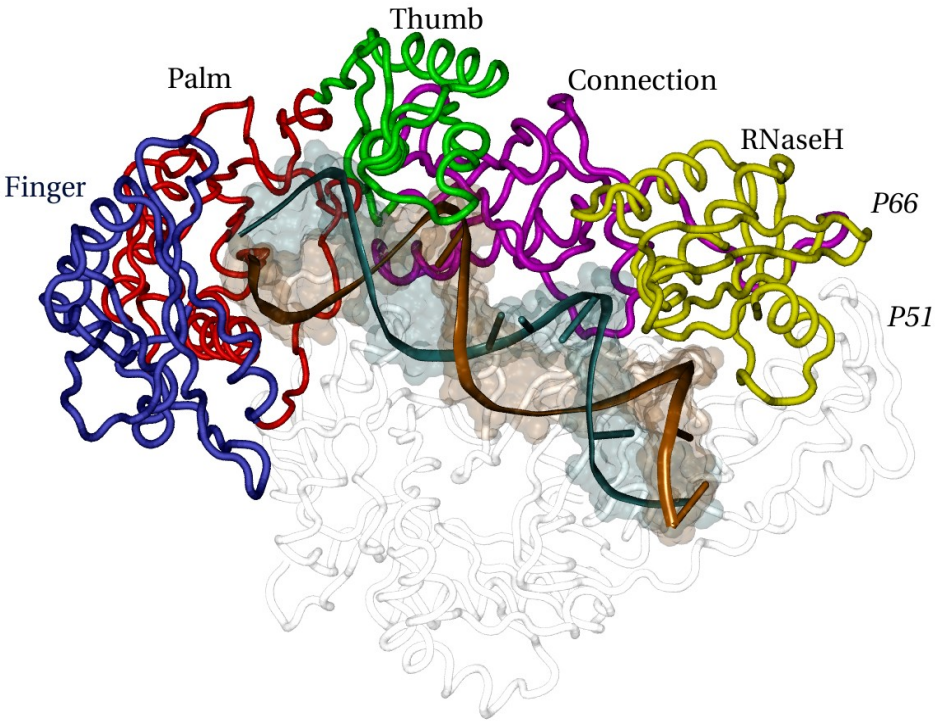


Figure7: Cartoon representation of HIV-RT. Two subunits P66 (colored according to the subdomains) and P51(white) are depicted here. Finger is represented in blue, Palm in red, Thumb in green, the connection in magenta and RNaseH in yellow. The template is also represented here with the two strands being colored in orange and cyan.

HIV-1 RT catalyzes the transcription of the single-stranded RNA viral genome into a double-stranded DNA form, which can be integrated into the human genome as the provirus. A wealth of crystallographic studies has been performed on RT, yielding ~ 60 high-resolution structures of the protein in a variety of states (as reviewed in Ref. [84] and documented in Ref. 85]). These include the “apo” form (no substrate and no inhibitor, e.g., Protein Data Bank (PDB) code 1DLO [86] and 1HMV[87]), substrate-bound forms (binary complexes of protein with nucleic acid substrate, and the ternary complex of protein with nucleic acid and nucleoside triphosphate, e.g., PDB codes 2HMI[88] and 1RTD[89]), and NNRTI-bound forms (co-crystallized NNRTIs, but no substrate, e.g., PDB codes 1BQM[90] and 1EP4[91]). RT is a 1000-amino-acid heterodimer of p66 and p51 subunits, each of them composed of “fingers,” “thumb,” “palm,” and “connection” subdomains (see Figure 7). The spatial arrangement of these subdomains is very different between the two subunits and it is thought that the p51 subunit mainly plays a structural role, with polymerization occurring at the p66 subunit.[84] Together, the fingers, palm, and thumb subdomains of p66 resemble a right hand and form a “clamp” that holds the double-stranded template–primer in position. Notably, the palm subdomain contains the “catalytic triad” of three aspartates (Asp110, Asp185, and Asp186), which are essential for the addition of nucleotide to the growing primer strand, and the so-called “primer grip,” which is though to be required for correct positioning of the 3′ end of the primer.[92] The p66 subunit contains an additional subdomain, known as the RNase H (RNH) domain, which is responsible for the other enzymatic activity of the protein degradation of the RNA strand during polymerization.

Antivirals, a small journey

HCV/BVDV

To date, the majority of studied antiviral drugs targets viral polymerases as the primary mechanism of action. Both biochemical and cell-based replicon assays have been employed in the identification and optimization of novel HCV and BVDV RdRp inhibitors. RdRp inhibitors reported can be divided into two groups: (i) nucleoside analogues, and (ii) non-nucleoside inhibitors (NNIs). Nucleoside analogues prevent replication of HCV RNA, blocking the elongation of new viral RNA strands, they generally target the polymerase active site in a competitive manner and tipically exhibit broader spectrum activity. Non-nucleoside counterparts inhibit the HCV polymerase enzyme itself and they have much greater specificity, acting by either interfering directly with the active site or binding to the allosteric site and preventing the initiation process.

Nucleoside inhibitors

Nucleoside inhibitors, whether chain terminators or non-chain terminators, can be effective in inhibiting the virus replication. Upon entry into the cells, nucleoside analogues are first converted to nucleotide triphosphates (NTP). The unnatural nucleoside inhibitors can serve as competitive substrates for the polymerase and can be incorporated into the nascent chain by the viral polymerase. This incorporation can lead to premature termination of the elongation process. The incorporated nucleotides may cause base mismatch in subsequent rounds of replication, resulting in accumulated mutations in the viral genome and then the so-called error catastrophe. Ribavirin-is a nucleoside analogue and it is the only one currently in use in HCV therapy, though its precise mechanism of action remains unclear and ribavirin used as monotherapy has no durable antiviral efficacy. Several sugar-modified and base-modified nucleoside analogues have been reported to inhibit the RdRp enzymatic activity and block HCV replication in the replicon systems. The nucleoside analogues, inhibiting NS5B polymerase activity, have mostly 2′-methylribose structures and act as chain terminator [93,94]. Chain terminators have proven to be effective against viral DNA polymerases, and a number of these compounds have been approved for treatment of HIV and hepatitis B. Nucleoside inhibitors are less prone to drug resistance induced by mutations than non-nucleoside inhibitors because they bind to the enzyme's active site, which is more conserved and less genetically variable than the allosteric (non-catalytic) sites bound by non-nucleosides. Anyway mutation may not be the only resistance mechanism. Researcher, led by Matthias Gotte at McGill University in Montreal, just published evidence for a new mechanism of viral resistance to nucleoside inhibitors: pyrophosphorolytic excision.The RNA polymerase, in the presence of cellular pyrophosphatase, can literary cut the nucleoside analogue drug off the chain, allowing the chain to resume growth. How much this happens in HCV is not yet known.

Several compounds, based on dioxolane triphosphate and nucleoside phosphate, have been reported [95]. β-D-2′-methyribofuranosyl guanosine is the most prominent, since it can be phosphorylated *in vivo* and orally administered. The only nucleoside analogue shown to have efficacy in a clinical study is NM 283 (Fig. 8, left panel) (Valopicitabine; University of Cagliari/Idenix/Novartis) [97]. A component of NM 283 is the orally administrable 3'-O-valine ester of 2′-C-methyl-cytidine, NM 107 (Fig 8 right panel). NM 107, once phosphorylated intracellularly to its 5' triphosphate metabolite, is initially known to be effective against BVDV. Later, a replicon cell lines-based assay revealed it to be effective against HCV. NM 283 is effective against replicons originating from different HCV strains . Replicons resistant to 2'-C-methylcytidine contain the S282T mutation in the viral polymerase and show a reduced fitness. NM 283, combined with pegylated IFN, exhibited improved efficacy and adequate tolerability compared to current therapies for chronic hepatitis C patients.

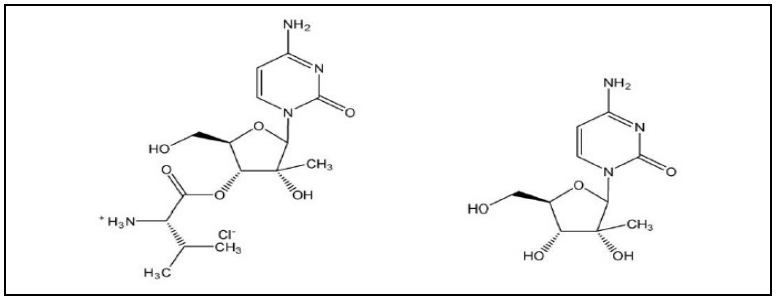


Figure8: The structures of NM 283 (left) and NM 107 (right)

Apart from NM283 other 2′-methyl nucleosides that selectively inhibit of HCV replication are 2′-O-methyl-cytidine, 2′-C-methyl-adeno-sine, 2′-C-methyl-guanosine, and 7-deaza-2′-C-methyl-adeno-sine. These are clinically less advanced, but some of them have been extensively characterized in vitro as well as in preclinical models of HCV infection [98-100]. In particular, the high inhibitory potency and reduced cellular toxicity of 7-deaza-2′-C-methyl-adenosine, combined with a very promising pharmacokinetic profile in preclinical animal species, makes this compound an attractive candidate for clinical development [101].

Lastly, RG7128 (Roche/Pharmasset), a 2′-modified nucleo-side analogue, has recently entered early clinical trials. RG7128 is an oral prodrug of 2′-deoxy-2′-fluoro-2′-C-methyl-cytidine, a pyrimidine nucleoside analogue that was shown to act as a non-obligate chain terminator of the NS5B polymerase activity [102] and to efficiently inhibit the replication of HCV RNA in hepatoma cell lines[Table 1] [103]
A 4′-modified nucleoside, 4′-azido-cytidine, was recently reported as a potent competitive inhibitor of NS5B-dependent RNA synthesis and hepatitis C virus replication in cell culture [104]. This compound was shown to inhibit RNA synthesis in a competitive manner and to be a moderately potent inhibitor of the replicon system. R1626 will enter a phase II trial at the two lower doses in combination with PEG-IFN with or without ribavirin.

Non-nucleoside inhibitors

Non-nucleoside inhibitors are the most diverse of the known HCV NS5B inhibitors. They act mostly as allosteric inhibitors and interfere with the conformational transition during the initiation of RNA synthesis [105]. In fact, besides the active site, the X-ray derived co-crystal structures of compounds bound to NS5B revealed distinct allosteric regulatory sites that are located distant to the active site and are targets for developing antiviral agents. Accordingly, compounds that interact either with the RdRp active site or the allosteric site could potentially interfere with substrate binding and/or conformational changes, thus effectively inhibiting initiation.

Table1:

Drug Name / Category	**Drug Name / Category**	**Pharmaceutical Company**	**Clinical Phase**
RG7128 (Polymerase Inhibitor)	RG7227 (ITMN-191) (Danoprevir) Protease Inhibitor	Genentech in collaboration with Pharmasset & InterMune	Phase I
Telaprevir Protease Inhibitor	VX-222 Polymerase Inhibitor	Vertex	Phase II
IDX375	Polymerase Inhibitor	Idenix	Phase I
ABT-072	Polymerase Inhibitor	Abbott	Phase I
MK-3281	Polymerase Inhibitor	Merck	Phase I
PSI-7851	Polymerase Inhibitor	Pharmasset	Phase I
ABT-333	Polymerase Inhibitor	Abbott	Phase I
VX-916	HCV Polymerase	Vertex	Phase I
Filibuvir (PF-00868554)	HCV Polymerase Inhibitor	Pfizer	Phase I
RG7128	Polymerase Inhibitor	Pharmasset/Genentech	Phase I
VX-222	Polymerase Inhibitor	Vertex	Phase II
IDX184	Polymerase Inhibitor	Idenix	Phase II
ANA598	Polymerase Inhibitor	Anadys Pharmaceuticals	Phase II
GS 9190	Polymerase Inhibitor	Gilead	Phase II
VX-759	Polymerase Inhibitor	Vertex	Phase II
PSI-7977	Polymerase Inhibitor	Pharmasset	Phase IIa

Table.1: Approved antiviral drugs against HCV. Adapted from: HCV Advocate; July 24, 2010;Hepatitis C treatment in current clinical development; Alan Franciscus.

Several classes of structurally distinct NNIs of the HCV RdRp have been identified and disclosed; these included but are not limited to benzothiadiazines, benzimidazoles/diamides, disubstituted phenylalanine/thiophene amides and substituted pyranones. Benzimidazole derivatives were the first non-nucleoside inhibitors that demonstrated to be active against the HCV polymerase by Japan Tobacco[106]. One representative compound showed an EC_{50} value in the submicromolar range. Investigation of the mechanism of action indicated that the compound did not compete with incorporation of NTP. Moreover, mutations, conferring resistance to these compounds, were mapped to proline residue 495, which is located on the surface of the polymerase thumb domain and away from the active site, suggesting that the compounds acted as allosteric inhibitors, by blocking the activity of the polymerase prior to the elongation step. The first NNIs of HCV, that entered clinical trials, were JTK-109 and JTK-003 (Japan Tobacco). These benzimidazole derivatives act as allosteric inhibitors and block the polymerase before elongation. Resistance of HCV replicons to this class of compounds maps to P495. Co-crystallization studies of the HCV polymerase, with a related analogue, suggests that these compounds bind on the surface of the thumb domain at a position that is in the 'closed' enzyme occupied by one of the fingertips. The binding of the compound to this cavity forces the enzyme in the 'open' configuration, annihilating its functionality. Boeringer Ingelheim reported a series of benzimidazole-containing heterocycles, by extending the original derivatives to topologically related scaffolds and incorporating an amide moiety inside the molecule. Interestingly, subsequent substitution of the benzimidazole with a pyrazolopyrimidine moiety led to active compounds with an EC_{50} less than 1 μM.

Thiophene derivatives are reversible allosteric inhibitors of the enzyme. Replicons resistant to the compounds contain mutation at position 419 and 423 at the base of the thumb, which is different from the position where the JT-compounds bind. Binding of the thiophene analogues to this allosteric site also results in a conformational change of the enzyme that may explain the antiviral activity.

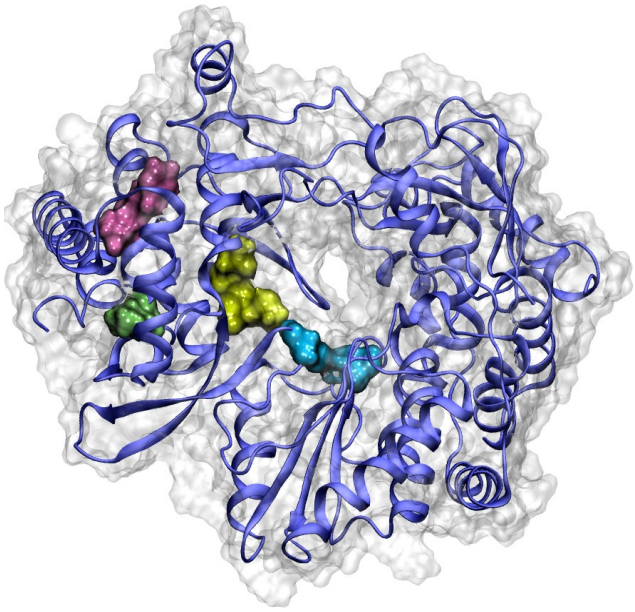


Figure9. Ribbon representation of the overall structure of HCV NS5B polymerase in blue along with transparent surface view in white. Solid surface representations of four allosteric binding sites are also shown. The dark red colored surface corresponds to the thumb allosteric binding pocket of indole group "siteA"; the dark green colored surface(thiophene) corresponds to the 2nd thumb allosteric binding site "site B", the dark yellow colored surface corresponds to the allosteric binding site situated in the palm domain "site C" and the dark skyblue colored surface(benzofuran) correspond to 2nd palm domain binding site "site D".

In the literature several Benzothiadiazines were reported that target the viral polymerase. Although this class of compounds induce a variety of mutation in the enzyme, only methionine 414 is believed to be part of the inhibitor binding site. This residue is located in the inner surface of the thumb domain, close to the catalytic site of the enzyme. The allosteric binding site of the Benzothiadiazine is thus again different from that of Thiophene and Benzimidazole derivatives. Cross-resistance studies and synergistic inhibition of HCV polymerase, by combination of a Benzimidazole and a Benzothiadiazine, further confirmed that these two structurally distinct classes of inhibitors had non-overlapping binding sites and thus acted with different mode of action. Heterocyclic Benzo-1,2,4-thiadiazine derivatives are specific inhibitors of RNA synthesis in HCV replicon systems [107]. GlaxoSmithKline (formerly SmithKline Beecham) disclosed a novel class of Benzothiadiazine derivatives, from which a representative compound displayed good potency both in biochemical assay, with an EC_{50} value of 80 nM and HCV replicon assay with an EC_{50} value of 500 nM. Moreover, the Benzothiadiazine derivatives were shown to be highly selective for the HCV RdRp, failing to inhibit other viral and mammalian polymerases. Treatment with the compound and IFN-α resulted in a highly synergistic effect in the replicon system. Further study demonstrated that the compounds are non-competitive for NTP incorporation and act to arrest *de novo* initiation of RNA synthesis prior to the elongation phase, possibly through interacting with the functionally critical NS5B polymerase active site. Substitution of the quinolinone moiety with pyrrolone group led to compounds with greater potency.

Researcher at Shire Biochem disclosed two classes of HCV polymerase inhibitors, including phenylalanine and thiophene carboxylate derivatives. Further studies demonstrated that the inhibitors bound to a allosteric site in the thumb domain, disticnt from the binding site of benzimidazole. In addition, other promising non-nucleoside inhibitors are in different phases of clinical trials. R803, a small molecule HCV RdRp inhibitor, developed by Rigel Pharmaceuticals, was found to be active in the replicon system, with EC_{50} below to 10 μM. R803 entered into a multi-dose phase I/II clinical study for chronic HCV infection, but recently the study has been cancelled. HCV-086, another orally available small molecule inhibitor of NS5B polymerase, was co-developed by Viropharma and Wyeth. Results from a phase 1b study demonstrated that HCV-086 possessed favourable pharmacokinetics and was generally safe and well tolerated. However, the overall antiviral activity of HCV-086 did not warrant further development. The follow-on compound HCV-796, a novel non-nucleoside HCV polymerase inhibitor, is being evaluated in ongoing clinical trials in combination with PEG-IFN. HCV-796 is well tolerated with no dose limiting toxicities and it displayed dose-dependent antiviral activity across multiple genotypes. The most significant and sustained reduction in viral load from baseline was observed on patient infected with genotype 1b HCV. Anyway, in August Viro Pharma and its development partner Wyeth announced discontinuation of phase 2 dosing of the polymerase inhibitor HCV-796, because of elevated liver enzymes in some patients. During 2007, companies, including seven biotechs, abandoned or suspended at least eight antiviral drugs in clinical development. The disappointments are not new: drugs specifically targeting the hepatitis C virus have been in the works since the early 1990s, but no compound of this kind has yet advanced to phase 3 trials. Four of the failed drugs were inhibitors of NS5B RNA polymerase. GlaxoSmithKline, in London, and XTL Biopharmaceuticals, of Valley Cottage, New York, discontinued their polymerase inhibitors in phase 1. These failures are a serious setback for the field, but they do not indicate a general drawback of the entire class of compounds. In fact, each drug has its own problem and has its own characteristics. At least three polymerase inhibitors remain in clinical development. A part from the molecules that are known to have entered clinical trials, several other NNIs of HCV are in advanced preclinical development and early results for the newer polymerase inhibitors look promising. A series of α,γ-diketoacid compounds, as inhibitor of HCV polymerase, has been identified. Further optimization led to identification of a potent HCV NS5B polymerase inhibitor, with an EC_{50} of 45 nM. Compounds of dihydroxypyrimidine carboxylic acid class are believed to chelate the two catalytic Mg^{2+} ions in the active site as diketoacid compounds. However, no data of *in vivo* activity for the pyrophosphate mimics are available and the high ionic nature of these compounds may raise concerns, such as low bioavailability and toxicity. The binding sites for (2s)-2-[(2,4-dichloro-benzoyl)-(3-trifluoromethylbenzyl)-amino]-3-phenyl-propionic acid and (2s)-2-[(5-Benzofuran-2-ylthiophen-2-ylmethyl)-(2,4-dichlorobenzoyl)-amino]-3-phenylpropionic acid are located in the thumb subdomain 35 Å far from the active site. A benzimidazole inhibitor and N,N disubstituted phenylalanine inhibitor also induced an allosteric interaction[108]. There are 3 classes of inhibitors that bind to a pocket in the thumb domain. Phene 2-carboxylic acids, phenylalanine derivatives and cyclopentyl dihydropyran-2-ones all bind to methionine-423[108]. Other non-nucleoside inhibitors, known to bind to methionine 423 in the palm and thumb domain of the active site cleft, include benzothiadiazine[107] and benzimidazole 5-carboxamide. Benzimidazole 5-carboxamide inhibits the initiation phase and acts in a non-

competitive manner with respect to NPT incorporation [108].

Natural and synthetic coumarins have multiple biological activities. For instance, they have been claimed to be useful as anticoagulant, antibacterial, anti-inflammatory, anticancer and anti-HIV agents. Very few are reported to be able to counteract HCV. Only some coumarin derivatives like Osthole (7-methoxy-8-prenylcoumarin) is capable of inhibiting HCV replication and/or proliferation. In particular, Osthole is able to counteract the progression of hepatitis C into hepatocarcinoma. Since the great variability of pharmacological response of coumarins is connected to large modification of their substituents, Mazzei and co-workers (Department of Pharmaceutical Sciences, Genova, Italy) found interesting to synthesize some coumarin Mannich bases in order to examine their ability to inhibit HCV replication. This idea proved to be a winning strategy: in fact, a number of synthesized compounds showed moderate, but significant activity against Flaviviridae family, paving the way for new coumarins to be used in the hepatitis C infection .

Non-nucleoside active against BVDV

VP32947 [109], mizoribine [110], BPIP[111], Acridones [112], AG110 [113], SC-560 [114], iminosugar derivatives [115], LZ37 [116], BIT225 [117]. γ-carboline[118,119,120]

HIV-RT

In a journey of 25 years, HIV-1, the retrovirus responsible for the acquired immunodeficiency syndrome (AIDS), has gone from being an "inherently untreatable" infectious agent to one eminently susceptible to a range of approved therapies. HIV epidemic fueled development of new antiviral drug classes, based on advances in the understanding of the viral life cycle, have transformed what used to be a rapid and lethal infection into a chronic condition that can be controlled for many years through combination of therapies with different classes of antiviral drugs — known as highly active antiretroviral therapy (HAART).

It all started in 1985, two years after the identification of human immunodeficiency virus (HIV) [121] and one year after the initial evidence about its etiological link to AIDS was reported [122]. Samuel Broder's group at the National Cancer Institute together with collaborators from Burroughs-Welcome company, identified 3-azidothymidine (AZT,zidovudine) as the first nucleoside inhibitor with in vitro anti-HIV activity.

The discovery of the anti-HIV activity of AZT was a landmark in the development of antivirals, providing the first proof of concept that the replication of HIV could be controlled by chemotherapy and thereby establishing the foundation of antiretroviral drug discovery research [123]. Thus, AZT became the first nucleoside HIV reverse transcriptase inhibitor (NRTI). Over the course of 25 years that followed after this seminal discovery, seven nucleosides have been approved by the United States Food and Drug Administration(FDA) for the treatment of HIV infection starting with the approval of AZT in 1987 and followed by didanosine (ddI), zalcitabine (ddC), stavudine (d4T), lamivudine (3TC), abacavir (ABC), tenofovir disoproxil fumarate (Table 2) Despite the approval of NRTIs as effective therapies for HIV-1/AIDS, side-effects of these antiretrovirals were found to be real and certainly not to be discounted.

Members of the of NRTIs were eventually joined by nonnucleoside RT inhibitors (NNRTIs), which were discovered in 1990, and interact with an allosteric binding site on HIV-1 reverse transcriptase that becomes exposed upon ligand binding[124]. NNRTIs are a key part of typical HAART regimes for treatment-naive patients (two NRTIs and one NNRTI), owing to their potency, favorable safety profile and ease of dosing. However, the relatively rapid emergence of resistance, resulting from mutations of residues that surround the NNRTI binding site (in particular K103N and Y181C), is a serious limitation. In addition to the three NNRTIs — nevirapine, delavirdine and efavirenz — that have been approved for the treatment of HIV, a few more are in clinical development[126], including rilpivirine, etravirine and dapivirine. Their mechanism of action is similar to that of the approved NNRTIs in that they interact with a specific binding site of the reverse transcriptase, thereby blocking the enzyme's activity. It has recently been demonstrated that NNRTI binding to the polymerase domain of the reverse transcriptase interferes with RNase H activity, and that mutations in the

NNRTI binding site (K103N, Y181C, Y188L and K103N/Y181C) reduce the potency of RNase H inhibition[125].

Despite progress in the treatment of HIV, there is still considerable room for improvement and expansion of antiviral drugs. Antiretroviral therapy has brought about a substantial decrease in the death rate due to HIV-1 infection, changing it from a rapidly lethal disease into a chronic manageable condition, compatible with very long survival. This has special implications within the classic boundaries of public health around the world, but at the same time in certain regions may also affect a cycle of economic and civil instability in which HIV-1/AIDS is both cause and consequence. Many challenges remain, including 1) the life-long duration of therapy; 2) the ultimate role of pre-exposure prophylaxis (PrEP); 3) the cardiometabolic side-effects or other toxicities of long-term therapy; 4) the emergence of drug-resistance and viral genetic diversity (non-B subtypes); 5) the specter of new cross-species transmissions from established retroviral reservoirs in apes and Old World monkeys; and 6) the continued pace of new HIV-1 infections in many parts of the world. All of these factors make refining current therapies and developing new therapeutic paradigms essential priorities. Fortunately, there are exciting new insights into the biology of HIV-1, its interaction with cellular resistance factors, and novel points of attack for future therapies. The current science will lead to new therapeutic strategies with far-reaching implications in the HIV-1/AIDS pandemic.

Brand name	Generic name(s)	Manufacturer name	Approval date	Time to approval
Nucleoside reverse transcriptase inhibitors (NRTIs)[a,b]				
Retrovir	Zidovudine, azidothymidine, AZT, ZDV	GlaxoSmithKline (original sponsor Burroughs-Wellcome)	19 March 1987	3.5 months
Videx	Didanosine, dideoxyinosine, ddI	Bristol Myers-Squibb	9 October 1991	6 months
Hivid	Zalcitabine, dideoxycytidine, ddC (no longer marked as of December 31, 2006)	Hoffmann-La Roche	19 June 1992	7.6 months
Zerit	Stavudine, d4T	Bristol Myers-Squibb	24 June 1994	5.9 months
Epivir	Lamivudine, 3TC	GlaxoSmithKline	17 November 1995	4.4 months
Combivir	Lamivudine and zidovudine	GlaxoSmithKline	27 September 1997	3.9 months
Ziagen	Abacavir sulfate, ABC	GlaxoSmithKline	17 December 1998	5.8 months
Videx EC	Enteric coated didanosine, ddI EC	Bristol Myers-Squibb	31 October 2000	9 months
Trizivir	Abacavir, zidovudine, and lamivudine	GlaxoSmithKline	14 November 2000	10.9 months
Viread	Tenofovir disoproxil fumarate, TDF	Gilead Sciences	26 October 2001	5.9 months
Emtriva	Emtricitabine, FTC	Gilead Sciences	02 July 2003	10 months
Epzicom	Abacavir and lamivudine	GlaxoSmithKline	02 August 2004	10 months
Truvada	Tenofovir disoproxil fumarate and emtricitabine	Gilead Sciences	02 August 2004	5 months
Nonnucleoside reverse transcriptase inhibitors (NNRTIs)[c]				
Viramune	Nevirapine, NVP	Boehringer Ingelheim	21 June 1996	3.9 months
Rescriptor	Delavirdine, DLV	Pfizer	4 April 1997	8.7 months
Sustiva	Efavirenz, EFV	Bristol Myers-Squibb	17 September 1998	3.2 months
Intelence	Etravirine	Tibotec Therapeutics	18 Jane 2008	6 months

Table. 2: Approved antiretroviral drugs. Adapted from: Drugs Used in the Treatment of HIV Infection, U.S. FDA, http://www.fda.gov/oashi/aids/virals.html. Drugs are listed in order of FDA approval within each class.

Reference :

1. Seth A, Misra A, Umrigar D (2004). "Topical liposomal gel of idoxuridine for the treatment of herpes simplex: pharmaceutical and clinical implications.". Pharm Dev Technol 9 (3): 277–89.
2. Otto S (1998). "Radiopharmaceuticals (Strontium 89) and radiosensitizers (idoxuridine).". J Intraven Nurs 21 (6): 335–7.
3. Fauth E, Zankl H (1999). "Comparison of spontaneous and idoxuridine-induced micronuclei by chromosome painting.". Mutat Res 440 (2): 147–56.
4. Torriani FJ, Rodriguez-Torres M, Rockstroh JK, et al. (July 2004). "Peginterferon Alfa-2a plus ribavirin for chronic hepatitis C virus infection in HIV-infected patients". N. Engl. J. Med. 351 (5): 438–50
5. Glanville AR, Scott AI, Morton JM, et al. (December 2005). "Intravenous ribavirin is a safe and cost-effective treatment for respiratory syncytial virus infection after lung transplantation". J. Heart Lung Transplant. 24 (12): 2114–9.
6. de Clercq, Erik; Field, Hugh J (5 October 2005). "Antiviral prodrugs — the development of successful prodrug strategies for antiviral chemotherapy". British Journal of Pharmacology (Wiley-Blackwell) 147 (1): pp. 1–11. January 2006
7. O'Brien JJ, Campoli-Richards DM. (1989). "Aciclovir. An updated review of its antiviral activity, pharmacokinetic properties and therapeutic efficacy". Drugs 37 (3): 233–309.
8. Shi, S. T., K. J. Herlihy, et al. (2008). "In Vitro Resistance Study of AG-021541, a Novel Nonnucleoside Inhibitor of the Hepatitis C Virus RNA-Dependent RNA Polymerase." Antimicrob. Agents Chemother. 52(2): 675-683
9. Ghany MG, Strader DB, Thomas DL, Seeff LB; American Association for the Study of Liver Diseases. Diagnosis, management, and treatment of hepatitis C: an update. Hepatology 49, 1335–1374 (2009).
10. Magdalena Sarasin-Filipowicz, Markus H Heim.Interferon-induced gene expression in chronic hepatitis C. Future Virology 2010 5:1, 25-31
11. World Health Organization. World Health Statistics 2009: Cause of spe cific mortality and morbidity. Available at: http://www.who.int/whosis/ whostat/EN_WHS09_Table2.pdf. Accessed on Dec 6, 2009.
12. Wheeler W, Mahle K, Bodnar U, et al. Antiretroviral drug resistance mutations and subtypes in drug-naïve persons newly diagnosed with HIV-1 infection, United States, Mar 2003–Oct 2006. 14th Conference on Retroviruses and Opportunistic Infections (CROI 2007). 2007 Feb 25–28; Los Angeles, CA.
13. E.J Richey.(2009).http://www.cattlenetwork.com/Cattle-Health--Treatment--Prevention---Control-Of-BVD/2009-06-01/Article.aspx?oid=496856&fid=VN-ANIMAL_HEALTH-BVD-ARTICLES
14. Venter, J. et al. The Sequence of the Human Genome. Science 291, 1304–1351 (2001).
15. Collins, F. et al. Finishing the euchromatic sequence of the human genome. Nature 431, 931–945 (2004).
16. Hebert, P., Cywinska, A., Ball, S. & R deWaard, J. Biological identifications through DNA barcodes. Proceedings- Royal Society of London. Biological sciences 270, 313–321 (2003).
17. Stoeckle, M. Taxonomy, DNA, and the Bar Code of Life. BioScience 53, 796–797(2003).
18. Hoppe, W., Lohmann, W., Markl, H. & Ziegler, H. Biophysik (1982).
19. Jung, J. & Lee, W. Structure-based Functional Discovery of Proteins: Structural Proteomics. Journal of Biochemistry and Molecular Biology 37, 28–34 (2004).
20. Bringer, A. & Nilges, M. Computational challenges for macromolecular structure determination by X-ray crystallography and solution NMR-spectroscopy. Quarterly Reviews of Biophysics 26, 49–125 (1993).
21. Nilges, M. Structure calculation from NMR data. Current Opinion in Structural Biology 6, 617–623 (1996).
22. Berman, H. et al. The Protein Data Bank. Nucleic Acids Research 28, 235–242 (2000).
23. Frauenfelder, H., Sligar, S. & Wolynes, P. The energy landscapes and motions of proteins. Science 254, 1598–1603 (1991).
24. Moffat, K. The frontiers of time-resolved macromolecular crystallography: movies and chirped X-ray pulses. Faraday Discussions 122, 65–77 (2003).
25. Schotte, F. et al. Watching a Protein as it Functions with 150-ps Time-Resolved X-ray Crystallography. Science 300, 1944–1947 (2003).
26. Eisenmesser, E., Bosco, D., Akke, M. & Kern, D. Enzyme Dynamics During Catalysis. Science 295, 1520–

1523 (2002).
27. Kempf, J. & Loria, J. Protein dynamics from solution NMR: theory and applications. Cell Biochemistry and Biophysics 37, 187–211 (2003).
28. Brischweiler, R. New approaches to the dynamic interpretation and prediction of NMR relaxation data from proteins. Current Opinion in Structural Biology 13, 175–183 (2003).
29. Mittermaier, A. & Kay, L. New Tools Provide New Insights in NMR Studies of Protein Dynamics. Science 312, 224–228 (2006).
30. Lange, O. et al. Recognition Dynamics Up to Microseconds Revealed from an RDC Derived Ubiquitin Ensemble in Solution. Science 320, 1471 (2008).
31. Seifert, U. Entropy Production along a Stochastic Trajectory and an Integral Fluctuation Theorem. Physical Review Letters 95, 40602 (2005).
32. Malmqvist, M. BIACORE: an affinity biosensor system for characterization of biomolecular interactions. Biochem Soc Trans 27, 335–40 (1999).
33. Englebienne, P., Van Hoonacker, A. & Verhas, M. Surface plasmon resonance: principles, methods and applications in biomedical sciences. Spectroscopy 17, 255–273 (2003).
34. Halperin, I.; Ma, B.; Nussinov, R. Proteins 2002, 47, 409.
35. Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. Proteins 2006, 65,15.
36. Morris, G. et al. Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. Journal of Computational Chemistry 19, 1639–1662 (1998).
37. Rarey, M., Kramer, B., Lengauer, T. & Klebe, G. A Fast Flexible Docking Method using an Incremental Construction Algorithm. Journal of Molecular Biology 261, 470–489 (1996).
38. Jones, G., Willett, P., Glen, R., Leach, A. & Taylor, R. Development and validation of a genetic algorithm for flexible docking. Journal of Molecular Biology 267, 727–748 (1997).
39. Friesner,R.A., Banks,J.L., Murphy,R.B., Halgren,T.A., Klicic,J.J., Mainz,D.T., Repasky,M.P., Knoll,E.H., Shelley,M., Perry,J.K. et al. (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J. Med. Chem., 47, 1739–1749.
40. de Graaf, C. et al. Catalytic site prediction and virtual screening of cytochrome P450 2D6 substrates by consideration of water and rescoring in automated docking. Journal of Medicinal Chemistry 49, 2417–2430 (2006).
41. Tiana, G.; Simona, F.; De Mori, G.M.S.; Broglia, R.A.; Colombo, G. Understanding the determinants of stability and folding of small globular proteins from their energetic. Protein Sci., 2004, 13, 113-123.
42. Wang, A.; Donini, O.; Reyes, C.M.; Kollman, P.A. Biomolecular simulations: recent developments in force fields, simulations of enzyme catalysis, protein-ligand, protein-protein, and protein-nucleic acid noncovalent interactions. Ann. Rev. Biophys. Biomol. Struct., 2001, 30, 211-220.
43. Warshel A. Computer simulations of enzyme catalysis: methods, progress, and insights. Ann. Rev. Biophys. Biomol. Struct. 2003, 32, 425-443.
44. Brooijmans, N.; Kuntz, I.D. Molecular recognition and docking algorithms. Ann. Rev. Biophys. Biomol. Struct., 2003, 32, 355-373.
45. Roux, B. Computational studies of the gramicidin channel. Acc. Chem. Res., 2002, 35, 366-375. b) Bond, P.J.; Sansom, M.S.P. The simulation approach to bacterial outer membrane proteins. Mol. Membr. Biol., 2004, 21, 151-161.
46. Laio, A.; Parrinello, M. Proc. Natl. Acad. Sci. U.S.A. 2002, 99, 12562.
47. Iannuzzi, M.; Laio, A.; Parrinello, M. Phys. ReV. Lett. 2003, 90,23802.
48. Gervasio, F.; Laio, A.; Parrinello, M. J. Am. Chem. Soc. 2005, 127, 2600.
49. Branduardi, D.; Gervasio, F.; Cavalli, A.; Recanatini, M.; Parrinello,M. J. Am. Chem. Soc. 2005, 127, 9147.
50. Hansson T, Marelius J, Aqvist J (1998) Ligand binding affinity prediction by linear interaction energy methods. J Comput Aided Mol Des 12:27–35. doi:10.1023/A:1007930623000
51. Hijikata M, Kato N, Ootsuyama Y, Nakagawa M, Shimotohno K. (1991) Gene mapping of the putative structural region of the hepatitis C virus genome by in vitro processing analysis. Proc Natl Acad Sci USA.; 88: 5547–51.
52. Love RA, Parge HE, Wickersham JA, Hostomsky Z, Habuka N, Moomaw EW, Adachi T, Hostomska Z. (1996) The crystal structure of hepatitis C virus NS3 proteinase reveals a trypsin-like fold and a structural zinc binding site. Cell.; 87: 331–42.

53. Tomei L, Failla C, Santolini E, De Francesco R, La Monica N. (1993) NS3 is a serine protease required for processing of hepatitis C virus polyprotein. J Virol.; 67: 4017–26.
54. Takamizawa A, Mori C, Fuke I, Manabe S, Murakami S, Fujita J, Onishi E, Andoh T, Yoshida I, Okayama H. (1991) Structure and organisation of the hepatitis C virus genome isolated from human carriers. J Virol.; 65: 1105–113.
55. Tanaka T, Kato N, Cho M-J, Shimotohno K. (1995) A novel sequence found at the 3'terminus of hepatitis C virus genome. Biochem Biophys Res Commun.; 215: 744–9.
56. Failla C, Tomei L, De Francesco R. (1994) Both NS3 and NS4A are required for proteolytic processing of hepatitis C virus nonstructural proteins. J Virol.; 68: 3753–60.
57. Lundin M, Monne M, Widell A, Von Heijne G, Persson MA. (2003) Topology of the membrane-associated hepatitis C virus protein NS4B. J Virol.; 77: 5428–38.
58. Kaneko T, Tanji Y, Satoh S, Hijikata M, Asabe S, Kimura K, Shimotohno K. (1994) Production of two phosphoproteins from the NS5A region of the hepatitis C viral genome. Biochem Biophys Res Commun.; 205: 320–6.
59. Neddermann P, Quintavalle M, Di Pietro Ch, Clementi A, Cerretani M, Altamura S, Bartholomew L, De Francesco R. (2004) Reduction of hepatitis C virus NS5A hyperphosphorylation by selective inhibition of cellular kinases activates viral RNA replication in cell culture. J Virol.; 78: 13306–14.
60. Tan S-L, Katze MG. (2001) How hepatitis C virus counteracts the interferon response: The jury is still out on NS5A.Virology.; 284: 1–12. dhanak
61. Oh J-W, Ito T, Lai MC. (1999) A recombinant hepatitis C virus RNA-dependent RNA polymerase capable of copying the full-length viral RNA. J Virol.; 73: 7694–702.
62. Adachi T, Ago H, Habuka N, Okuda K, Komatsu M, Ikeda S & Yatsunami K (2002) The essential role of C-terminal residues in regulating the activity of hepatitis C virus RNA-dependent RNA polymerase. Biochemica et Biophysica Acta 1601:38–48.
63. 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.
64. Bressanelli S, Tomei L, Rey FA & De Francesco R (2002) Structural analysis of the hepatitis C virus RNA polymerase in complex with ribonucleotides. Journal of Virology 76:3482–3492.
65. 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. Proceedings of the National Academy of Sciences, USA 96:13034–13039.
66. 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. Nature Structural Biology 6:937–943.
67. O'Farrell D, Trowbridge R, Rowlands D & Jager J (2003) Substrate complexes of hepatitis C virus RNA polymerase (HC-J4): structural evidence for nucleotide import and de novo initiation. Journal of Molecular Biology 326:1025–1035.
68. Hansen, J. L.; Long, A. M.; Schultz, S. C. Structure of the RNA-dependent RNA polymerase of poliovirus. Structure 1997, 5, 1109-1122.
69. Doublie, S.; Sawaya, M. R.; Ellenberger, T. An open and closed case for all polymerases. Struct. Fold. Des. 1999, 7, R31-R35.
70. O'Farrell, D.; Trowbridge, R.; Rowlands, D.; Jager, J. Substrate complexes of hepatitis C virus RNA polymerase (HC-J4): structural evidence for nucleotide import and de-novo initiation. J. Mol. Biol. 2003, 326, 1025-1035.
71. Ago, H., Adachi, T., Yoshida, A., Yamamoto, M., Habuka, N., Yatsunami, K. & Miyano, M. (1999) Structure (Cambridge, U.K.) 7, 1417–1426.
72. Lesburg, C. A., Cable, M. B., Ferrari, E., Hong, Z., Mannarino, A. & Weber, P. C. (1999) Nat. Struct. Biol. 6, 937–942.
73. Bressanelli, S., Tomei, L., Roussel, A., Incitti, I., Vitale, R. L., Mathieu, M., De Francesco, R. & Rey, F. A. (1999) Proc. Natl. Acad. Sci. USA 96, 13034–13039.
74. Bressanelli, S., Tomei, L., Rey, F. A. & De Francesco, R. (2002) J. Virol. 76, 3482–3492.
75. O'Farrell, D., Trowbridge, R., Rowlands, D. & Jager, J. (2003) J. Mol. Biol. 326, 1025–1035.
76. Ng, K. K. S., Cherney, M. M., Vazquez, A. L., Machı A., Alonso, J. M. M., Parra, F. & James, M. N. G. (2002) J. Biol. Chem. 277, 1381–1387.
77. Butcher, S. J., Grimes, J. M., Makeyev, E. V., Bamford, D. H. & Stuart, D. I. (2001) Nature 410, 235–240.
78. Koonin, E. V. (1991) J. Gen. Virol. 72, 2197–2206.

79. Poch, O., Sauvaget, I., Delarue, M. & Tordo, N. (1989) EMBO J. 8, 3867–3874.
80. Lai, V. C. H., Kao, C. C., Ferrari, E., Park, J., Uss, A. S., Wright-Minogue, J., Hong, Z. & Lau, J. Y. N. (1999) J. Virol. 73, 10129–10136.
81. Hong, Z., Cameron, C. E., Walker, M. P., Castro, C., Yao, N., Lau, J. Y. N. & Zhong, W. (2001) Virology 285, 6–11.
82. Shirako, Y., Strauss, E. G. & Strauss, J. H. (2000) Virology 276, 148–160.
83. Carroll S.S., Tomassini J.E., Bosserman M., Getty K., Stahlhut M.W., Eldrup A.B., Bhat B., Hall D., Simcoe A.L, LaFemina R., Rutkowski C.A., Wolanski B., Yang Z., Migliaccio G., De Francesco R., Kuo L.C., MacCoss M., Olsen D.B. Inhibition of hepatitis C virus RNA replication by 2'-modified nucleoside analogs. J Biol Chem. Apr 4; 278 (14), 11979-84 (2003).
84. Das K, Sarafianos SG, Arnold E, Hughes SH, William JL, Lane MD. Encyclopedia of Biological Chemistry. Elsevier; New York: 2004. HIV-1 Reverse Transcriptase Structure; pp. 388–392.
85. Lawtrakul L, Beyer A, Hannongbua S, Wolschann P. Quantitative structural rearrangement of HIV-1 reverse transcriptase on binding to non-nucleoside inhibitors. Monatsh Chem. 2004;135:1033–1046.
86. Arnold E, Das K, Ding J, Yadav PN, Hsiou Y, Boyer PL, Hughes SH. Targeting HIV reverse transcriptase for anti-AIDS drug design: structural and biological considerations for chemotherapeutic strategies. Drug Des Discov. 1996;13:29–47. [PubMed]
87. Rodgers DW, Gamblin SJ, Harris BA, Ray S, Culp JS, Hellmig B, Woolf DJ, Debouck C, Harrison SC. The structure of unliganded reverse transcriptase from the human immunodeficiency virus type 1. Proc Natl Acad Sci USA. 1995;92:1222–6. [PMC free article] [PubMed]
88. Ding J, Das K, Hsiou Y, Sarafianos SG, Clark AD, Jr, Jacobo-Molina A, Tantillo C, Hughes SH, Arnold E. Structure and functional implications of the polymerase active site region in a complex of HIV-1 RT with a double-stranded DNA template-primer and an antibody Fab fragment at 2.8 A resolution. J Mol Biol. 1998;284:1095–111.
89. Huang H, Chopra R, Verdine GL, Harrison SC. Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance. Science. 1998;282:1669–75.
90. Hsiou Y, Das K, Ding J, Clark AD, Jr, Kleim JP, Rosner M, Winkler I, Riess G, Hughes SH, Arnold E. Structures of Tyr188Leu mutant and wild-type HIV-1 reverse transcriptase complexed with the non-nucleoside inhibitor HBY 097: inhibitor flexibility is a useful design feature for reducing drug resistance. J Mol Biol. 1998;284:313–23.
91. Ren J, Nichols C, Bird LE, Fujiwara T, Sugimoto H, Stuart DI, Stammers DK. Binding of the second generation non-nucleoside inhibitor S-1153 to HIV-1 reverse transcriptase involves extensive main chain hydrogen bonding. J Biol Chem. 2000;275:14316–20.\
92. Das K, Ding J, Hsiou Y, Clark AD, Jr, Moereels H, Koymans L, Andries K, Pauwels R, Janssen PA, Boyer PL, Clark P, Smith RH, Jr, Kroeger Smith MB, Michejda CJ, Hughes SH, Arnold E. Crystal structures of 8-Cl and 9-Cl TIBO complexed with wild-type HIV-1 RT and 8-Cl TIBO complexed with the Tyr181Cys HIV-1 RT drug-resistant mutant. J Mol Biol. 1996;264:1085–100.
93. Carroll S.S., Tomassini J.E., Bosserman M., Getty K., Stahlhut M.W., Eldrup A.B., Bhat B., Hall D., Simcoe A.L, LaFemina R., Rutkowski C.A., Wolanski B., Yang Z., Migliaccio G., De Francesco R., Kuo L.C., MacCoss M., Olsen D.B. Inhibition of hepatitis C virus RNA replication by 2'-modified nucleoside analogs. J Biol Chem. Apr 4; 278 (14), 11979-84 (2003).
94. Witkowski J.T., Robins R.K., Sidwell R.W., Simon L.N. Design, synthesis, and broad spectrum antiviral activity of 1- -D-ribofuranosyl-1,2,4-triazole-3-carboxamide and related nucleosides. J Med Chem. Nov;15(11),1150-4 (1972).
95. Deval J.,. Powdrill M.H., D'Abramo C.M., Cellai L., Gottel M. Pyrophosphorolytic Excision of Nonobligate Chain Terminators by Hepatitis C Virus NS5B Polymerase ANTIMICROBIAL AGENTS AND CHEMOTHERAPY, Aug p. 2920–2928 (2007)
96. Benzaria S., Bardiot D., Bouisset T., Counor C., Rabeson C., Pierra C., Storer R., Loi A.G., Cadeddu A., Mura M., Musiu C., Liuzzi M., Loddo R., Bergelson S., Bichko V., Bridges E., Cretton-Scott E., Mao J., Sommadossi J.P., Seifer M., Standring D., Tausek M., Gosselin G., La Colla P. 2'-C-methyl branched pyrimidine ribonucleoside analogues: potent inhibitors of RNA virus replication. Antiviral Chem & Chemother., 18, 225-242 (2007).
97. Storer B.E. An evaluation of phase I clinical trial designs in the continuous dose-response setting. Stat Med. Aug 30; 20 (16), 2399-408 (2001).
98. A.B. Eldrup, C.R. Allerson, C.F. Bennett, S. Bera, B. Bhat, N. Bhat, M.R Bosserman, J. Brooks, C. Burlein,

S.S. Carroll, P.D. Cook, K.L. Getty, M. MacCoss, D.R. McMasters, D.B. Olsen, T.P. Prakash, M. Prhavc, Q. Song, J.E. Tomassini, J. Xia, Structure-activity relationship of purine ribonucleosides for inhibition of hepatitis C virus RNA-dependent RNA polymerase, J. Med. Chem. 47 (9) (2004) 2283–2295.

99. A.B. Eldrup, M. Prhavc, J. Brooks, B. Bhat, T.P. Prakash, Q. Song, S. Bera, N. Bhat, P. Dande, P.D. Cook, C.F. Bennett, S.S. Carroll, R.G. Ball, M.Bosserman, C. Burlein, L.F. Colwell, J.F. Fay, O.A. Flores, K. Getty, R.L LaFemina, J. Leone, M. MacCoss, D.R. McMasters, J.E. Tomassini, D. VonLangen, B. Wolanski, D.B. Olsen, Structure-activity relationship of heterobase-modified 2′-C-methyl ribonucleosides as inhibitors of hepatitis C virus RNA replication, J. Med. Chem. 47 (21) (2004) 5284–5297.
100. S.S. Carroll, J.E. Tomassini, M. Bosserman, K. Getty, M.W. Stahlhut, A.B. Eldrup, B. Bhat, D. Hall, A.L. Simcoe, R. LaFemina, C. Rutkowski, B. Wolanski, Z. Yang, G. Migliaccio, R. De Francesco, L.C. Kuo, M. MacCoss, D.B. Olsen, Inhibition of hepatitis C virus RNA replication by 2¢-modified nucleoside analogs, J. Biol. Chem. 278 (14) (2003)11979–11984.
101. D.B. Olsen, A.B. Eldrup, L. Bartholomew, B. Bhat, M.R. Bosserman, A.Ceccacci, L.F. Colwell, J.F. Fay, O.A. Flores, K.L. Getty, J.A. Grobler, R.L.LaFemina, E.J. Markel, G. Migliaccio, M. Prhavc, M.W. Stahlhut, J.E.Tomassini, M. MacCoss, D.J. Hazuda, S.S. Carroll, A 7-deaza-adenosine analog is a potent and selective inhibitor of hepatitis C virus replication with excellent pharmacokinetic properties, Antimicrob. Agents Chemother.48 (10) (2004) 3944–3953.
102. S.S. Carroll, M.-E. Davies, L. Handt, K. Koeplinger, R. Zhang, S.W. Ludmerer, M. MacCoss, D.J. Hazuda, D.B. Olsen, Robust suppression of viral replication by a nucleoside polymerase inhibitor in chimpanzees infected with hepatitis C virus, Hepatology 44 (Suppl 1) (2006) 535A.
103. P.A. Furman, E. Murakami, H. Bao, J. Symons, M.J. Otto, Inhibition of HCV replication by PSI-6130:mechanism of biochemical activation and inhibition, J. Hepatol. 46 (Suppl 1) (2007) S224.
104. W.R. Jiang, S. Ali, S. LePogam, C. Daniel, S. Chiu, T. Kretz, I. Najera, P.A. Furman, N. Cammack, J. Symons, Inhibition of HCV replication by PSI-6130: characterization of activity in the HCV replicon system, J. Hepatol.46 (Suppl 1) (2007) S228
105. Dhanak D., Duffy K.J., Johnston V.K., Lin-Goerke J., Darcy M., Shaw A.N., Gu B., Silverman C., Gates A.T., Nonnemacher M.R., Earnshaw D.L., Casper D.J, Kaura A., Baker A., Greenwood C., Gutshall L.L., Maley D., Del Vecchio A., Macarron R., Hofmann G.A., Alnoah Z., Cheng H.Y., Chan G., Khandekar S., Keenan R.M., Sarisky R.T. Identification and biological characterization of heterocyclic inhibitors of the hepatitis C virus RNA-dependent RNA polymerase. J Biol Chem. Oct 11; 277 (41), 38322-7 (2002).
106. Nguyen T.T., Gates A.T., Gutshall L.L., Johnston V.K., Gu B., Duffy K.J., Sarisky R.T. Resistance profile of a hepatitis C virus RNA-dependent RNA polymerase benzothiadiazine inhibitor. Antimicrob Agents Chemother. Nov; 47 (11), 3525-30 (2003).
107. Tomei L., Altamura S., Bartholomew L., Bisbocci M., Bailey C., Bosserman M., Cellucci A., Forte E., Incitti I., Orsetti L., Koch U., De Francesco R., Olsen D.B., Carroll S.S., Migliaccio G. Characterization of the inhibition of hepatitis C virus RNA replication by nonnucleosides. J Virol. Jan; 78 (2), 938-46 (2004).
108. Tomei L., Altamura S., Bartholomew L., Biroccio A., Ceccacci A., Pacini L., Narjes F., Gennari N., Bisbocci M., Incitti I., Orsetti L., Harper S., Stansfield I., Rowley M., De Francesco R., Migliaccio G. Mechanism of action and antiviral activity of benzimidazole-based allosteric inhibitors of the hepatitis C virus RNA-dependent RNA polymerase. J Virol. Dec; 77 (24),13225-31 (2003).
109. Baginski, S.G., Pevear, D.C., Seipel, M., Sun, S.C., Benetatos, C.A., Chunduru, S.K., Rice,C.M., Collett, M.S., 2000. Mechanism of action of a pestivirus antiviral compound. Proc. Natl. Acad. Sci. U.S.A. 97, 7981–7986.
110. Yanagida, K., Baba, C., Baba, M., 2004. Inhibition of bovine viral diarrhea virus (BVDV) by mizoribine: synergistic effect of combination with interferon- . Antiviral Res. 64, 195–201.
111. Paeshuyse, J., Leyssen, P., Mabery, E., Boddeker, N., Vrancken, R., Froeyen, M., Ansari,I.H., Dutartre, H., Rozenski, J., Gil, L.H., Letellier, C., Lanford, R., Canard, B.,Koenen, F., Kerkhofs, P., Donis, R.O., Herdewijn, P., Watson, J., De Clercq, E.,Puerstinger, G., Neyts, J., 2006. A novel, highly selective inhibitor of pestivirus replication that targets the viral RNA-dependent RNA polymerase. J. Virol. 80,149–160.
112. Tabarrini, O., Manfroni, G., Fravolini, A., Cecchetti, V., Sabatini, S., De Clercq, E., Rozenski, J., Canard, B., Dutartre, H., Paeshuyse, J., Neyts, J., 2006. Synthesis and anti-BVDV activity of acridones as new potential antiviral agents. J. Med. Chem. 49, 2621–2627.
113. Paeshuyse, J., Chezal, J.M., Froeyen, M., Leyssen, P., Dutartre, H., Vrancken, R., Canard, B., Letellier, C., Li, T., Mittendorfer, H., Koenen, F., Kerkhofs, P., De Clercq, E., Herdewijn, P., Puerstinger, G., Gueiffier, A., Chavignon, O., Teulade, J.C., Neyts, J., 2007. The imidazopyrrolopyridine analogue AG110 is a novel, highly selective inhibitor of pestiviruses that targets the viral RNA-dependent RNA polymerase at a hot

spot for inhibition of viral replication. J. Virol. 81, 11046–11053.
114. Okamoto, M., Sakai, M., Goto, Y., Salim, M.T.A., Baba, C., Goto, K., Watashi, K., Shimotohno, K., Baba, M., 2009. Anti-bovine viral diarrhea virus and hepatitis C virus activity of the cyclooxygenase inhibitor SC-560. Antiviral Chem. Chemother. 20, 47–54.
115. Chang, J., Wang, L., Ma, D., Qu, X., Guo, H., Xu, X., Mason, P.M., Bourne, N., Moriarty, R., Gu, B., Guo, J.T., Block, T.M., 2009. Novel imino sugar derivatives demonstrate potent antiviral activity against flaviviruses. Antimicrob. Agents Chemother. 53,1501–1508.
116. Paeshuyse, J., Letellier, C., Froeyen, M., Dutartre, H., Vrancken, R., Canard, B., De Clercq, E., Gueiffier, A., Teulade, J.C., Herdewijn, P., Puerstinger, G., Koenen, F.,Kerkhofs, P., Baraldi, P.G., Neyts, J., 2009. A pyrazolotriazolopyrimidinamine inhibitor of bovine viral diarrhea virus replication that targets the viral RNA dependent RNA polymerase. Antiviral Res. 82, 141–147.
117. Luscombe, C.A., Huang, Z., Murray, M.G., Miller, M., Wilkinson, J., Ewart, G.D., 2010. A novel Hepatitis C virus p7 ion channel inhibitor BIT225, inhibits bovine viral diarrhea virus in vitro and shows synergism with recombinant interferon- -2b and nucleoside analogues. Antiviral Res. 86, 144–153.
118. Sako, K., Aoyama, H., Sato, S., Hashimoto, Y., Baba, M., 2008. -Carboline derivatives with anti-bovine viral diarrhea virus (BVDV) activity. Bioorg. Med. Chem. 16, 3780–3790.
119. Aoyama, H., Sako, K., Sato, S., Nakamura, M., Miyachi, H., Goto, Y., Olamoto, M., Baba, M., Hashimoto, Y., 2009. Polymethylated -carbolines with potent anti-bovine viral diarrhea virus (BVDV) activity. Heterocycles 77, 779–785.
120. Mohammed TA Salim, Mika Okamoto, Shinnosuke Hosoda, Hiroshi Aoyama, Yuichi Hashimoto, Masanori Baba; Anti-bovine viral diarrhoea virus activity of novel diphenylmethane derivatives;Antiviral Chemistry & Chemotherapy, 2010;20:193-200
121. Barre-Sinoussi, F., Chermann, J.C., Rey, F., Nugeyre, M.T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C., Rozenbaum, W., Montagnier, L., 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science 220, 868–871.
122. Gallo, R.C., Salahuddin, S.Z., Popovic, M., Shearer, G.M., Kaplan, M., Haynes, B.F., Palker, T.J., Redfield, R., Oleske, J., Safai, B., et al., 1984. Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. Science 224, 500–503.
123. Mitsuya, H., Weinhold, K.J., Furman, P.A., St. Clair, M.H., Lehrman, S.N., Gallo, R.C., Bolognesi, D., Barry, D.W., Broder, S., 1985. 3 -Azido-3 -deoxythymidine (BW A509U): an antiviral agent that inhibits the infectivity and cytopathic effect of human T-lymphotropic virus type III/lymphadenopathy-associated virus in vitro. Proc. Natl. Acad. Sci. U.S.A. 82, 7096–7100.
124. Hang, J. Q. et al. Substrate-dependent inhibition or stimulation of HIV RNase H activity by non-nucleoside reverse transcriptase inhibitors (NNRTIs). Biochem. Biophys. Res. Commun. 352, 341–350 (2007).
125. Martin, J., Hitchcock, M.J., De Clercq, E., Prusoff, W., 2010. A brief history of the first generation nucleoside HIV reverse transcriptase inhibitors. Antiviral Res. 85, 34–38.

II-Methods

1 METHODS

1.1 DOCKING

Molecular docking is a multidimensional optimization problem, requiring efficient samplaing across the entire space of positional, orientational, and confermational possibilities available to both a receptor and a ligand it recognizes. Recent flexible docking studies and the results from the docking section of the critical assessment of protein.Structure (CASP2) prediction contest [14] suggest that a suitable combination of homology modeling, docking strategies and experimental studies can be used to reliably predict protein-ligand structures.

Protein-ligand interactions, described by the induced-fit model of molecular recognition, involve structural alterations in the receptor binding site and the ligand. Thus, it is critical that the model allows for flexibility in both the ligand and the receptor. The compouter programs AUTODOCK, GOLD, GLIDE, FLEX etc, which we have now integrated into our studies of protein-ligand interactions, allow partial flexibility of the receptor and full flexibility of the ligand . flexibility of the ligand and receptor is computationally cumbersome to include in the docking simulations.If the ligand and the receptor are completly flexible during the simulation, even if we exclude everything else but the receptor binding site, the computational cost of the docking problem becomes so great. Proper ways to accomodate flexibility in the docking simulations without increasing the computational effort too much is one of the most important issues in the development of new metodologies at the moment.Computer-aided ligand-receptor docking has become an important research tool in structural biology.

1.1.1 *Autodock*

AUTODOCK (current version 3.0 and 4.0) docks flexible small moleculae ligand to rigid macromolecular recepor and with some residue flexibility of receptor. In the docking simulations the number of rorsional degrees of freedom in the ligand can be varied, but some angles and bond distances are maintained rigid. AUTODOCK uses Monte-Carlo simulated annealing and/or Lamarckin genetic algorithm minimization scheme combined with a rapid, atomic resolution, grid based method of energy evaluation utilizing the Amber force field. In order to find low

energy confermations of ligands in the receptor binding site.The overall interaction energy between chemical species (i.e between atoms of the ligand and the receptor) is estimated by considering both Lennard-Jones atom-atom potential and electrostatic effects, summed for the individual interactions between atoms. A distance-dependent dielectric constant or special solvation energy grid can be used to account for the solvent effects.

Stoddard and Koshland were pioneering users of AUTODOCK in a clever prediction of a structure of a peotein-protein complex by docking seprate domains. apartm from many uses of AUTODOCK in various ways of protein ligand interaction, it also helped in invention of a drug which is in clinical trials for HIV. McComman et al in 2007 used AUTODOCK in conjunction with the Relaxed Complex Method. Merck Pharmaceutical company used McCommon's et al group protocol of AUTODOCK to design new drugs that target integrase, which lead in October 2007 to the first clinically approved HIV Integrase inhibitor:Isentress.

The docking of the ligands was explored with AUTODOCK and the docking energy produced with the AUTODOCK force field was corrected with the energy needed to solvate the ligands

1.1.2 *Protocol*

Autodock 4 was used for all docking calculations(Morris GM et al journal of computational chemistry; 2009). The AutoDockTools(ADT) package was employed to generate the docking files and analyze the docking results. In the abscence of literature data for the binding of different class of compounds, we performed molecular docking taking carben alpha atom. of the respective resistant mutant of the corresponding site as in case of HCV and the resistant mutation isolated from our lab in case of BVDV, as a grid center, for each compounds. The interaction pattern and key residues of crystallized ligand (table1), which correspond to the binding site A, B, C, and D were used for the identification of proper orientation of our compounds into the different binding sites. Two different grid boxes, one for each binding site, were centered on the average mass center of the ligands. Thus, a grid box of 91 X 91 X 86 points and a grid spacing of 0.375Å was set in order to accommodate the NNIs that bind on to the different sites of NS5B polymerase.

The second grid box (spacing,0.375Å) of 62 X 75 X 75 points was implemented in such a way to accommodate the NNIs into the more buried area of NS5B. The grid maps were generated for each atom probe, describing its interactions with the compounds. Autogrid 4, as implemented in the Autodock software package, was used to generate grid maps. The Lamarckian genetic algorithm (LGA)62 was employed to generate orientations

or conformations of the ligands within the binding site. The global optimization started with a population of 150 randomly positioned individuals, a maximum of 2.5 X 108 energy evaluations, and a maximum of 27,000000 generations, as mentioned the protocol for blind docking. A total of 200 runs was performed, while all the remaining run parameters were maintained at their default settings. A cluster analysis was carried out using 2Å as the root-mean-square deviation tolerance. Docking experiments were also tried using a single grid box comprising the two binding sites. Although Autodock performed quite well in selectively positioning the ligands in the right pockets, the results in term of root-mean-square-deviation (rmsd) values were not fully satisfying. This is likely due to the fact that with only 100 runs such a large region is not sufficiently explored. Nevertheless, a higher number of conformation sampling, may not assured better results in the absence of required no. of runs. So we went for 200 runs although it was computationally demanding.

1.2 ASSESSMENT OF DOCKING

1.2.1 *Redocking*

. Further refinement of docking was performed, aiming to check out the more realistically binding mode conformations of compounds with no experimental data. In order to check for the reliability of the docking protocol, we performed docking studies on NS5B polymerase complexes. The docking results were evaluated through a comparison of the predicted docked positions of the ligand and the experimental ones. As a measure of docking reliability, the rmsd between the positions of heavy atoms of the ligand in the calculated and experimental structures was considered.

The choice of the best conformation was based on the assumption that, although for high-throughput screening protocols, only the first ranked conformations should be considered (that is, the conformer characterized by the lowest estimated free energy of binding)63. In other cases, the lowest energy conformer of the most populated cluster should also be taken into account64. We applied the second protocol for investigation of binding mode of newly identified potent lead of different class of compounds. While the best docked is the lowest energy docked conformations of the first autodock generated cluster.the best cluster is the lowest energy docked conformation of the most populated cluster and best fitted cluster is the lowest energy docked confermation of the cluster showing the lowest rmsd value.In the ideal case, i.e., the error free docking program, the three conformations would have coincided.

The fact that Autodock reproduces the correct conformations (best fitted cluster) by 100 percent reflects the intrinsic lack of accuracy (at least for NS5B) of the implemented scoring function in selecting the right pose, and this is at the basis for the further analysis using molecular dynamics.

Flexibility of the ligand and the receptor is

1.3 MOLECULAR DYNAMICS

Molecular Dynamics (MD) simulation is a technique founded upon the basic principles of classical mechanics that provide a dynamical picture of the individual particles of the system at a microscopic level. Using this technique successive configuration of the molecular system (in the phase space of coordinates and momenta) is generated by integrating Newton's law of motion. The result is a trajectory, which contains the microscopic time evolution of the system in the phase space [**? ?**]. From the trajectory generated, one can compute the dynamical properties such as absorption spectra, rate constants and transport properties. Further, on combining MD with statistical mechanics as a mean of sampling, one can compute equilibrium properties such as average thermodynamics quantities, structure, and free energies along the reaction path seen as a union of all possible states of the system [**? ?**]. For instance, the statistical ensemble average of an observable **A** can be obtained as:

$$\langle A \rangle = \sum_{t=1}^{\tau \to \infty} A(t) \tag{1.1}$$

The assumption made here is called the ergodic hypothesis, i.e given an infinite amount of time, ensemble average of observable **A**, is equivalent to its time average. The main aspect in atomistic MD simulations are:

- An algorithm that samples the phase space
- The choice of the interaction potential, $V(\mathbf{r})$, between the atoms of the system.

Several simulations approaches were developed in the last decades that differs in the method to sample the phase space. The most fundamental form used to describe equation of motion is the Lagrangian form:

$$\frac{\mathbf{d}}{\mathbf{dt}}\left(\frac{\partial \mathbf{L}}{\partial \dot{\mathbf{q}}_\mathbf{j}}\right) - \left(\frac{\partial L}{\partial q_j}\right) = 0 \tag{1.2}$$

where L(q,$\dot{q}$) is Lagrangian defined as the difference between the kinetic and potential energies L=K-V, where q_j are general-

ized coordinates and $\dot{q}_j$ are the associated time derivative. The momentum conjugate to coordinate q_j is given as:

$$\mathbf{p}_j = \frac{\partial L}{\partial \dot{q}_j} \tag{1.3}$$

On substitution with the usual definition of kinetic and potential terms with cartesian coordinates r_i, equation 1.2 becomes:

$$\mathbf{F}_i = m_i \ddot{\mathbf{r}}_i \text{ with } \mathbf{F}_i = -\frac{\partial V(\mathbf{r})}{\partial \mathbf{r}_i} \tag{1.4}$$

where $V(\mathbf{r})$, the potential, is a function of the atoms positions and F_i represents the total force on atom i. In this equation one assumes that the nuclear motion of constituent particles obeys the laws of classical mechanics. This is an excellent approximation if the distance in the energetic (translational, rotational and vibrational) levels of the involved degrees of freedom is $<< kT$, where k is the Boltzmann constant and T the temperature.

In the Hamiltonian form the equation of motion for the cartesian coordinates is given by:

$$\dot{\mathbf{r}}_i = \frac{p_i}{m_i} \text{ and } \qquad \dot{\mathbf{p}}_i = \frac{\partial V(r)}{\partial r_i} \tag{1.5}$$

1.3.1 *Integration of Newton equations of Motion*

Under the influence of a potential, the motions of atoms are strongly coupled to each other giving rise to many-body problems that cannot be solved analytically [**?**]. Therefore, in MD calculation an iterative numerical procedure is employed to obtain an approximate solution for the equations of motion.

The two important properties of the equations of motion to be noted are:

- They must be time reversibe (t=-t).
- Conservation of total Energy (Hamiltonian) of the system.

For the first point, as the Newton equations are time reversible also the algorithm used is supposed to satisfy the same time reversal symmetry. The algorithms that are not time reversible do not normally preserve the phase space volume, i.e. they do not satisfy the Liouville theorem. For the second point, conservation of Hamiltonian is equivalent to conservation of total energy of the system and provides an important link between MD and statistical mechanics. The energy conservation condition H(p,r) = E, defines a hypersurface in the phase space called the constant energy, imposing a restriction on system to remain on this surface

[**?**]. A good way to check the accuracy of the algorithm is to follow the temporal evolution of an observable A that should be conserved (e.g. the total energy). In general a good algorithm must be such that:

$$\frac{|\, A(t_n) - A(t_0)\,|}{\langle A(t)\rangle} \ll 1, \qquad \text{for } (t_n - t_0) \gg \Delta t \tag{1.6}$$

there is no drift in the total energy.
The MD integration of the Newton's equation which have a continuous form, are based on assumption that position, velocities and other dynamical properties can be discretized using the Taylor series expansion:

$$r(t+\delta t) = r(t) + \Delta t v(t) + \frac{1}{2}\Delta t^2 a(t) + \frac{1}{6}\Delta t^3 b(t) + ... \tag{1.7}$$

$$v(t+\delta t) = v(t) + \Delta t a(t) + \frac{1}{2}\Delta t^2 b(t) + \frac{1}{6}\Delta t^3 c(t) + ... \tag{1.8}$$

The choice of the integration method depends on the degree of accuracy of problem at hand. One of the most useful form used is the velocity verlet algorithm [**?**], a variant of verlet algorithm [**?**]. The advantage is using velocity verlet method is that positions, velocities and acceleration are well synchronized that allow to calculate the kinetic energy contribution to the total energy at same time, from which potential energy is determined. The equations are:

$$\mathbf{r}_i(t+\Delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(\Delta t)\Delta t + \mathbf{a}_i\frac{1}{2}\Delta t^2 + \mathbf{O}(\Delta t^3) \tag{1.9}$$

$$\mathbf{v}_i(t+\Delta t) = \mathbf{v}_i(t) + [a_i(t) + a_i(t+\Delta t)]\frac{1}{2}\Delta t + \mathbf{O}(\Delta t^3) \tag{1.10}$$

where $\mathbf{a}_i$,$\mathbf{r}_i$,$\mathbf{v}_i$ are respectively the acceleration on the atom i, the atom position and the atom velocity. The algorithm has an accuracy of O (Δt^3) for the variables and it is reversible in time.

Together with conservation of energy and time-reversibility another important feature of an integrating algorithm is to permit long time steps Δt. It is expected that the numerical Newtonian trajectory will diverge from the "true" Newtonian trajectory. However, it is important that the integrating algorithm maintains a well defined energy tolerance ΔE throughout the simulation time. The error (ΔE) is known to decrease on decreasing the time step Δt. The aim here is to find a balance between using the largest possible time step and maintaining an acceptable ΔE all along the simulation. A large time step would lead to faster exploration, but energy would fluctuate widely with the possibility of the

simulation being catastrophically unstable, on other hand too short time step would lead to computation being needlessly slow. The choice of an integration step is determined by the nature of forces acting on the system. The golden rule is to choose time step ($\Delta t \sim 10^{-15}$s) such that the fastest motion of the system can be integrated accurately. This requirement is a severe restriction, particularly as high frequency motions are relatively of less interest and have minimal effect on the overall behavior of the system. One suggested approach is to freeze out such vibrations by constraining the appropriate bonds to their equilibrium values. Details of this approach is discussed in subsection 1.3.4, or to use multiple time step approach which is discussed in subsection 1.3.2.

1.3.2 *Multiple Time Step Integrator*

One of the approaches to accelerate the integration of equations of motion, is to use "multi-time" step algorithm such as reverse reference system propagation algorithm (r-RESPA)[**?**]. In the algorithm (r-RESPA), the molecular system is classified into number of groups according to how rapidly the forces varies over time. The starting point is the Liouville operator formulation, which can cast the equations for the Hamiltonian system (see equation1.5) in a general form:

$$\dot{\mathbf{x}} = iL\mathbf{x} \tag{1.11}$$

where **x** is the phase vector and iL is the Liouville operator. Consider a molecular system containing N atoms (or 3N degrees of freedom) with $\mathbf{x} = \{r_i, p_i\}$ representing a point in the phase space. The Liouville operator in cartesian coordinates is defined as:

$$i\mathbf{L} = \{..., H\} \equiv \sum_{i=1}^{3N} \left[\frac{\partial H}{\partial p_i} \cdot \frac{\partial}{\partial r_i} - \frac{\partial H}{\partial r_i} \cdot \frac{\partial}{\partial p_i} \right] \tag{1.12}$$

On subsituting equation 1.5 into equation 1.12, we get:

$$i\mathbf{L} = \{..., H\} \equiv \sum_{i=1}^{3N} \left[\frac{p_i}{m_i} \cdot \frac{\partial}{\partial r_i} + \mathbf{F}_i \cdot \frac{\partial}{\partial p_i} \right] \tag{1.13}$$

where F_i is the force on i^{th} degree of freedom, and {...,...} is the poisson bracket. The classical time propagator U(t) is unitary and defined as e^{iLT}, and the evolution of system Eq. 1.11 is expressed as:

$$\mathbf{x}(t) = e^{iLt}\mathbf{x}(0) \tag{1.14}$$

The action of operator U(t) on x(o) cannot be determined analytically, however the operator can be decomposed using Trotter theorem, such that the action of U(t) on x(o)for each part can be evaluated analytically. Applying the Trotter theorem we get:

$$\begin{aligned} \mathbf{e}^{i(L_1+L_2)t} &= \left[e^{i(L_1+L_2)t/P}\right]^P = \left[e^{i(L_1+L_2)\Delta t}\right]^P \\ &= \left[e^{iL_1(\frac{\Delta t}{2})}e^{iL_2\Delta t}e^{iL_1(\frac{\Delta t}{2})}\right]^P + O(t^3/P^2) \quad (1.15) \end{aligned}$$

where $\Delta t = t/P$. For finite P, the numerical iteration procedure is accurate to the second order in the time step at long times. From equation 1.15, for the three exponential terms, we define the discrete time propagator (U_1, U_2) as:

$$\begin{aligned} \mathbf{G}(\Delta t) &= U_1(\frac{\Delta t}{2}) + U_2(\Delta t) + U_1(\frac{\Delta t}{2}) + O(t^3/P^2) \\ &= e^{iL_1(\frac{\Delta t}{2})}e^{iL_2\Delta t}e^{iL_1(\frac{\Delta t}{2})} + O(t\Delta t^2) \quad (1.16) \end{aligned}$$

Since the three exponential terms in G Δt are separately unitary, G (Δt) is also unitary i.e $G^{-1}(t) = G^{\dagger}(t) = G(-t)$. Lets us now consider the propagator generated by subdivison as:

$$\begin{aligned} iL_1 &= \sum_{i=1}^{N} \frac{p_i}{m_i} \cdot \frac{\partial}{\partial r_i} \\ iL_2 &= \sum_{i=1}^{N} F_i \cdot \frac{\partial}{\partial p_i} \quad (1.17) \end{aligned}$$

The operator $U_1(\frac{\Delta t}{2})$ becomes a translation operator on the positions: $r_i \rightarrow r_i + \Delta t(\frac{p_i}{m_i})$, and operator $U_2(\Delta t)$ becomes a translational operator of momenta: $p_i \rightarrow p_i + (\frac{\Delta t}{2})F_i(r)$. On combining these two facts to action of operators in equation 1.16 on complete set of positions and momenta, yields the approximate evolution:

$$\begin{aligned} \mathbf{r}_i(\Delta t) &= r_i(0) + \Delta t v_i(0) + \frac{\Delta t^2}{2m_i}F_i(0) \\ \mathbf{v}_i(\Delta t) &= v_i(0) + \frac{\Delta t}{2m_i}[F_i(0) + F_i(\Delta t)] \quad (1.18) \end{aligned}$$

which is the famous velocity verlet **[?]** integrator derived using the operator formulation. The power of the operator based approach is its symplectic property which ensures no drift in the total energy, resistance to increase in time steps and allows generating stable long trajectories.

r-RESPA algorithms have been successfully employed to incorporate motions on more than two time scales. Let us consider a system with three characteristics time scales, a reference force F_i^{ref}, and two corrections F_i^{del} and F_i^{Del}, such that F_i=F_i^{ref}+F_i^{del}+F_i^{Del}. We define their Liouville operators as iL^{ref}, $iL^{(del)}$ and $iL^{(Del)}$ and the corresponding timescales δt, Δt and $\Delta\mathcal{T}$ respectively. The three time step propogator can then be written as:

$$\exp(iL\Delta\mathcal{T}) = \exp\left(iL^{(Del)}\frac{\Delta\mathcal{T}}{2}\right)\left\{\exp\left(iL^{(del)}\frac{\Delta t}{2}\right)\left[\exp(iL_2^{(ref)}) \times \exp(iL_1^{(ref)}\delta t)\exp\left(iL_2^{(ref)}\frac{\delta t}{2}\right)\right]^n \exp\left(iL^{(del)}\frac{\Delta t}{2}\right)\right\}^m \times \exp\left(iL^{Del}\frac{\Delta\mathcal{T}}{2}\right) \quad (1.19)$$

Thus, the correction due to slowest time scale is applied every m×n timesteps, and the intermediate time scale correction is applied every n steps. Such numerical procedure lead to considerable saving in the CPU time to perform a MD simulation.

1.3.3 *The interaction potential*

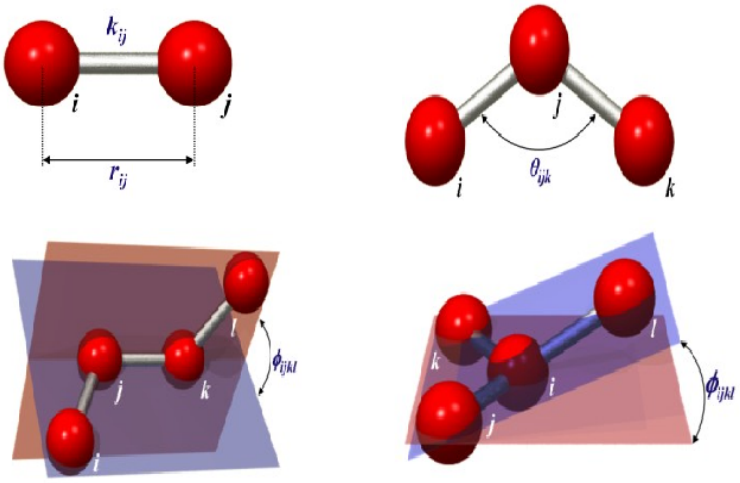


Figure 1: force field

The potential function $V(\mathbf{r})$ from which the forces used in MD are derived depends on the atomic coordinates $\mathbf{r}_i$.

$V(\mathbf{r})$ used in this thesis has the following expression:

$$
\begin{aligned}
V(r_1, r_2, \ldots, r_N) &= \sum_{\text{bonds}} \frac{1}{2} K_d (d - d_0)^2 \\
&+ \sum_{\text{angles}} \frac{1}{2} K_\theta (\theta - \theta_0)^2 \\
&+ \sum_{\text{improper dihedrals}} \frac{1}{2} K_\xi (\xi - \xi_0)^2 \\
&+ \sum_{\text{dihedrals}} K_\phi \left[1 + \cos(n\phi - \delta)\right] \\
&+ \sum_{\text{ij LJ}} \left[\left(\frac{C_{ij}^{12}}{r_{ij}^{12}} - \frac{C_{ij}^{6}}{r_{ij}^{6}} \right) \right] \\
&+ \sum_{\text{ij coulomb}} \frac{q_i q_j}{4\pi\varepsilon_0 \varepsilon_r r_{ij}}
\end{aligned}
\tag{1.20}
$$

The first two terms (two and three body interactions respectively) represent the bonds and angles potentials, that are approximated by harmonic functions. The third and fourth term describe four body interactions. Improper dihedral terms are typically described by an harmonic function. Instead proper dihedrals are described by periodic functions (i.e. cosine functions) of a given periodicity n. The last two terms are a Lennard-Jones (LJ) potential and a coulomb potential between pair (ij) of atoms. The LJ terms reproduce the Van der Walls interactions, while the coulomb potential terms reproduce the electrostatic properties of a protein. These interaction are modelled using the two-body approximation which doesnot explicitly account for the polarization effects, but on a average. The parameters used in this kind of potentials are typically obtained from quantum chemical calculations and experimental data (e.g. crystallographic data, spectroscopic data, etc). Among the popular sets of parameters (force fields) for MD simulations of proteins we can cite for example AMBER, GROMOS, CHARMM and OPLS. They all use the potential function expression given above for all the atoms of the simulated system except for the GROMOS(and CHARMM19 force field) force field in which a united atom description is used for non-polar hydrogens.

In MD simulations the description of the solvent (water for most of the biologically interesting systems) can be explicit or implicit. In the first case solvent molecules with a full atomistic force field description are added in the simulation box at the experimental density. In the implicit solvent description the solvent is treated as a dielectric medium in which the system is embedded. This is clearly a more approximated description but it is also computationally much more efficient since in many practical cases the solvent constitutes the majority of the atoms. In this

thesis we used for all the MD simulations the AMBER-95 and AMBER03 force field with an explicit description of the solvent given by the TIP3P[**?**] water model.

1.3.4 *Constraints for Hydrogen*

Constraints are used in MD to fix bonds to their equilibrium value. This allows increasing the simulation time step Δt. Constraining the bond lenght does not alter significantly the statistics as these are quantum degrees of freedom being mostly in their ground state at the normal simulation temperature. Using the bonds constraints it is possible to use $\Delta t \sim 2fs$ [**?**] (2-4 times larger than the one that can be used without constraints). A common method to introduce constraints is the algorithm SHAKE [**?**], in which after each time step the atoms positions iteratively are modified in order to satisfy the constraint.

SHAKE may have convergence problems when applied to large planar groups and its implementation could hinder the efficiency of computing. To improve these aspects the LINCS algorithm was recently introduced. For water molecules it is also possible to use an analytic solution of SHAKE called SETTLE.

1.3.5 *Boundary conditions*

To simulate a finite size system, boundary conditions are needed to avoid artifacts near the border of the simulation box. Typically periodic boundary conditions (PBC) are used. In this scheme short range non bonded interactions are calculated using the minimal image convention (only the nearest replica is considered). Typically a cut-off radius (R_c) is used for LJ interactions of the order of 10 Å. To avoid interactions between a particle and its periodic image each box side must be larger than $2R_c$.

The coulomb energy is instead treated considering the full periodicity of the system. For a periodic lattice made by N particles it is given by:

$$E = \frac{1}{8\pi\epsilon_0} \sum_{|n|=0}^{\infty}{}^{\star} \left[\sum_{i=0}^{N} \sum_{j=0}^{N} \frac{q_i q_j}{|r_{ij} + n|} \right] \tag{1.21}$$

where n indicates the periodic images, i,j the particles and the symbol $\star$ indicates that the summation does not contain the term with $i = j$ if $n = 0$.

The periodicity of the system speeds up the evaluation of the coulombic term. Although convenient, effective, and simple to apply, certain subtle problems arise for long range forces (electrostatics), whose spatial range may extend beyond the boundaries

of the container into surrounding images, present a challenge. Long range forces can only be correctly calculated by summing over all the periodic replicas of the original system. However, the associated computational effort is considerable. Fortunately, methods have been developed to treat this problem. Specifically, the Ewald summation technique, developed originally to treat Coulomb interactions and later extended to treat general interactions of the form $\frac{1}{r_n}$ for $n \leqslant 3$ has proved enormously successful. The basic idea behind the technique is to divide the relevant part of the potential into a short range and a long range contribution. For the Coulomb potential, $1/r$, for example, this can be achieved via the identity

$$\frac{1}{r} = \frac{erf(\alpha r)}{r} + \frac{erfc(\alpha r)}{r} \tag{1.22}$$

where erf(x) and erfc(x) are the error function and complementary error function, respectively (erf(x) + erfc(x)= 1). The variable, R, is a convergence parameter, which can be optimized for each system studied. The short range term, $erfc(\alpha r)/r$, is treated as an ordinary short range interaction, i.e., using a spherical cutoff to truncate the interaction at large spatial distances where the potential is small. The long range term, $erf(\alpha r)/r$, is Fourier transformed into reciprocal space, where it takes the short-ranged form, $\exp(-g^2/4\alpha^2)$, and can be evaluated accurately by summing over only a small number of reciprocal space vectors of the simulation cell. Such reciprocal space sums can be evaluated with high a degree of efficiency (N log N) using particle-mesh methods(PME)[**?**]. An extension of PME is the smooth PME. With respect to PME, this method uses a fixed cuttoff in the direct sum and uses the B-spline interpolation of the reciprocal space structures onto a rectangular grid, permitting the use of fast Fourier transforms to efficiently calculate the reciprocal sum. In this thesis we use SPME method to evaluate the electrostatic energies[**?**].

1.3.6 *Statistical Ensembles*

Molecular dynamics can be performed in different statistical ensembles. The traditionally used ensemble to perform MD is the micro-canonical ensemble (NVE), where the number of particles (N), the volume (V), and the total-energy (E) of the system are fixed to a constant value.
The simple extension of NVE ensemble is the canonical one (NVT), where the number of particles, the volume and the temperature are fixed to a constant value. The temperature T, in contrast to the number of particles N and volume V, is an inten-

sive parameter. The temperature T is related to the time average of the kinetic energy given as:

$$\mathbf{T} = \frac{2}{3}\frac{E_{kin}}{Nk_B} = \frac{1}{3Nk_B}\sum_{i=1}^{N}\frac{{p_i}^2}{m} \tag{1.23}$$

where, E_{kin} is the kinetic energy, k_B is the Boltzmann constant. The simplest way to control the temperature, is to rescale the velocities at each step by the factor $\lambda = \sqrt{\frac{T_{req}}{T_{curr}}}$, where T_{curr} is the current temperature calculated from the kinetic energy and T_{req} is the desired temperature (for instance 300 K). However, an alternative way to maintain is to couple the system to an external heat bath that is fixed at the desired temperature. The bath acts as a source of thermal energy, supplying or removing heat from the system as appropriate. This thermostat is named as the "Berendsen" thermostat. It is extremely efficient for relaxing a system to the target temperature, but once the system has reached equilibrium, it might be more important to probe a correct canonical ensemble.

Extended system methods, was originally introduced for performing constant MD simulation by Nosè in 1984, and subsequently developed by Hoover in 1985. The idea of the method was to reduce the effect of an external system, acting as a heat reservoir, to an additional degree of freedom s. This reservoir has a potential energy $(f+1)k_B T \ln s$, where f is the number of degrees of freedom in the physical system and T is the desired temperature. The kinetic energy of the reservoir is given as $(\frac{Q}{T})(\frac{ds}{dt}^2)$, where Q is considered as the fictitious mass of the extra degree of freedom. The magnitude of Q determines the coupling between the reservoir and the real system and so influences the temperature fluctuation. If Q is large then the energy flow is slow; in the limit of infinite Q, conventional molecular dynamics is regained. However, if Q is small then the energy oscillates, resulting in equilibrium problems. It has been suggested that Q should be proportional to $fk_B T$.

Another ensemble we discuss here it the NPT ensemble, an extension of NVT ensemble, where together with temperature the pressure of the system is maintained to a constant vaue. As most experimental measurements are usually made under conditions, which include a fixed pressure P, temperature T, and number of atoms N (constant-NPT ensemble), and so simulations in the isothermal-isobaric ensemble are the most directly relevant to experimental data. A simulation in NPT ensemble maintains the constant pressure by changing the volume of the simulation cell.

The amount of volume fluctuation is related to the isothermal compressibility, κ

$$\kappa = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T \tag{1.24}$$

An alternative to maintain constant pressure is to couple the system to a "pressure" bath, analogous to the temperature bath. The rate of change of pressure is given by:

$$\frac{dP(t)}{dt} = \frac{1}{\tau_p}(P_{bath} - P(t)) \tag{1.25}$$

where τ_p is the coupling constant, P_{bath} is the pressure of the 'bath', and P(t) is the actual pressure at time t. The volume of the simulation box is scaled by a factor λ, which is equivalent to scaling the atomic coordinates by a factor $\lambda^{1/3}$. Thus:

$$\lambda = 1 - \kappa\frac{\delta t}{\tau_p}(P - P_{bath}) \tag{1.26}$$

and the new position are given by:

$$r_i{}^{new} = \lambda^{1/3} r_i \tag{1.27}$$

In the extended pressure-coupling systems, an extra degree of freedom, corresponding to the volume of the box, is added to the system. The kinetic energy associated with this degree of freedom (which can be considered to be equivalent to piston acting on the system), is $(\frac{1}{2}Q)(\frac{dV}{dt}^2)$, where Q is the 'mass' of the piston. The piston also has a potential energy PV, where P is the desired pressure and V is the volume of the system. The volume varies in the simulation with the average volume being determined by the balance between the internal pressure of the system and the desired external pressure. In this thesis, we have performed MD simulation in both NVT and NPT ensembles.

1.3.7 *Solvation effect*

In most cases, we are interested in the properties of molecules in solution, usually in aqueous solution. therefore, it is essential to calculate not only the interaction between the atoms of biomolecules, but also the solvent effect on the interaction. there are two different approches to include the solvent effects in molecular dynamics. One is explicit solvent model, the other one is the continuum solvent model, also known as the implicit solvent model.

In explicit solvent models we need to calculate the interaction between the particles and every solvent molecule individully, therefore it is a very "computationally expensive " job, and the

cost increase with the number of interaction sites in the water model. Algorithms such as Particle Mesh Ewald (PME) method, particle-particle/particle-mesh(P3M) method are developed to accelrate the simulations.

Even with advanced algorithms, one obvious drawback of the explicit solvent model is the large system size due to numerous solvent molecules. An alternative approch is using the continuum solvent models, such as the Generalized Born(GB) model. This model represents the solvent implicity as a continuum with dielectric properties of water, and also includes the charge screening effects of salt. this lowers the calculation expense in two aspects: first, it removes the calculation of the interactions and motions involving solvent molecules;second, the absence of solvent friction accelerates the dynamics of the solute. However, the computing cost of GB model scales up with the size of the system. for large systems, the computing cost may be greater than using explicit solvent model.

Both explicit solvent and implicit solvent models have certain strengths and weakness. Calculations using explicit solvent generally yield more accurate results. In some cases, such as simulations involving water bridges, explicit water molecules are essential for the calculation. However, systems using explicit solvent have many more atoms (in most case mostly water molecules), demanding additional computing resources. Due to the fraction force from water molecules, the dynamics of the solute is also slower in the explicit solvent. this may be useful for understanding the real time scale of a biological event. However, for some studies where rate is not a factor, but the results of the motion are important, the implicit solvent model will be more efficient. Another important advantage of implict solvent model is that it calculates solvent free energy, instead of energy. In the explicit solvent model, water atoms are considered to be part of the system, and solvent-solute interactions are included in the non-bonded calculations. There is no independent solvation term. In implicit solvent the simulated system only includes solute. the molecular mechanic interactions are calculated first in vacuum. then an additional term, the solvation free energy is calculated and included in the energy of the system. the implicit solvent mimics the average effects of the water molecules, therefore it genrates solvation free energies. it is essentail for methods such as Molecular Mechanics-Poisson Boltzmann (Generalized Born) Surface Aea (MM-PB(GB)SA) approch, which needs an explicit solvation free energy term in calculation. just a note, MM-PB(GB)SA is a post-processing analysis, which means it is still recommended to genrate the structures using explicit solvent first, then using implicit solvent model to analyze the solvation free energy on existing structures.

Molecular Dynamics (MD) simulations allow investigating processes occurring on timescales of ~100ns. However, most interesting and relevant biological process happen on time scales that are orders of magnitude larger, and are therefore termed as rare events. For example, protein folding (μs-few seconds), protein-protein interactions, transport of molecules across membrane channels (order ~ μs) and many others. Over the years, we have observed an astounding increase in computer power (Blue gene, DESRES), which promise to increase utility of MD simulations to investigate more and more complex systems on μs timescale. However, these supercomputing machines are not available to all the research groups. Therefore another approach to overcome the timescale problem is to renounce the all atom approach and to use coarse grained models. This would retain the essential characteristics, however you require a detailed knowledge of system, that is often not available.
For systems, where its important to maintain the atomistic description, one can exploit methodology aimed at accelerating rare events to timescales reachable in MD simulations. Notable success has been achieved is using the accelerating methodology in diverse fields of interest. From their scope and range of applicability, they are classified in four categories [**?**]:

1. Methods aimed at improving sampling, in a subspace of few predefined collective variables (CVs), that allow reconstructing the probability distributions as a function of chosen CVs. Examples of these methods include thermodynamic integration [**? ?**], free energy perturbation [**?**], umbrella sampling [**?**], conformational flooding [**?**], weighted histogram [**? ? ?**], steered MD [**?**], Jarzynski's identity based methods [**?**] and adaptive force bias [**?**]. The power of these methods in highly dependent on judicious choice of CVs, and computational performance degrades as a function of the number of variables.

2. Methods aimed at exploring the transition mechanism. Examples in these catogeries are transition path sampling[**? ?**], finite temperature string method [**? ?**], transition interface sampling [**?**] and forward flux methods [**?**]. These methods do not require in most cases, an explicit definition of a reaction co-ordinate, but require a priori knowledge of initial and final states of process under investigation.

3. Methods for exploring the potential energy surfaces and localizing the saddle points that corresponds to a transition state. Examples in these catogeries are dimer method [**?**], hyperdynamics [**?**], multiple time scale accelerated MD [**?**]

and event based relaxation [**?**]. The power of these methods is limited to low dimensionality, and reliability degrades with the complexity of system.

4. Methods in which the phase space is explored simultaneously at different values of temperatures, are parallel tempering [**?**] and replica exchange [**?**], or as a function of the potential energy, such as multicanonical MD [**?**] and Wang-Landau [**?**].

1.4.1 *Metadynamics*

Metadynamics method encompasses several features of techniques mentioned earlier and provides a unified frame work for computing free energies and accelerating rare events. It is a powerful algorithm, based on dimensional reduction that is used for accelerating rare events in system described by complex Hamiltonians, at a classical or quantum level [**?**]. Before to use the algorithm, the requirements are to identify a set of CVs, which are assumed to describe well the process of our interest. The power of the algothim lies in treating the CVs simultaneously and in its flexibility: the method can be proficientely used both for reconstructing free energy and for accelerating rare events.

1.4.2 *The algorithm*

Consider a system described by a set of co-ordinates x and a potential V(x) evolving under the action of molecular dynamics, whose equilibrium distribution is canonical at a temperature $T(\frac{1}{\beta})$. Since our interest is to describe the system using a set of CVs, $S_\alpha(x)$, α=1,d where d is a small number assuming they provide a good coarse-grained description [**?**]. The equilibrium behavior of these CVs in defined by probability distribution:

$$P(s) = \frac{\exp\left(-(\frac{1}{T})F(s)\right)}{\int ds \exp\left(-(\frac{1}{T})F(s)\right)} \tag{1.28}$$

where s denotes the d dimensional vector (s_1,....,s_d), with the free energy given by:

$$F(s) = -T \ln\left(\int dx \exp\left(-\frac{1}{T}V(x)\right)\delta(s - S(x))\right) \tag{1.29}$$

in equation 1.29, capital S is used for representing the function of the coordinates S(x), while lower case s is used for denoting the value of the CVs.
Consider now a trajectory x(t) of the system at temperature T.

On computing the trajectory for a very long time, probability P(s) can be obtained by taking the histogram of CV s along this trajectory. At time t, P(s) ~ $\frac{1}{t}\int_0^t dt'\delta$ (S(x(t') - s). If the system displays metastability, the motion of s will be bound in some local minimum of F(s) (i.e., in local maximum of P(s)), and will escape from this minimum with a very low probability on the timescale determined by the potential V(x) alone. In Metadynamics, the metastability is eliminated by modifying the underlying potential V(x), by adding a history dependent term consisting of Gaussians centered along the trajectory in s space during the evolution of the system [**? ?**]. Each time a new gaussian is added at time τ_G, the biasing potential at time t is given as:

$$V_G(S(x),t) = \omega \sum_{\substack{t'=\tau_G,\tau_{2G}... \\ t'<t}} \exp(-\frac{(S(x)-s(t'))^2}{2\delta s^2}) \qquad (1.30)$$

In equation 1.30 capital S is used for denoting the function of the co-ordinates S(x), while lower case s is used for denoting the value of the CVs, s(t)=S(x(t)) is the value taken by the CV at time t, ω is gaussian width, δs is the gaussian width and frequency τ_G at which gaussians are added. The basic assumption of metadynamics is that V_G(s,t) defined in equation 1.30 after a sufficiently long time provides an estimate of the underlying free energy:

$$\lim_{t\to\infty} V_G(s,t) \sim -F(s). \qquad (1.31)$$

Here, in equation 1.31, a equilibrium quantity free energy (F(s)) is estimated by a non-equilibrium dynamics. Equation(1.31) can be qualitatively understood as slow 'deposition' (i.e. $\omega \to 0$), and in this limit V_G(s,t) varies very slowly, with the probability to observe s is the approximately proportional to exp[$-\frac{1}{T}F(s) + V_G$ (s,t))]. If the function F(s)+V_G(s,t) has some local minimum, s will be preferentially localized in the of this minimum and Gaussians will be deposited, till the minimum is filled. If we consider the situation where F(s) ~ $-V_G$(s,t) in a region Ω(s). In this case the probability distribution would be flat and the locations of the gaussians will not be affected by the bias. Hence, if $\omega \to 0$, only corrugations in the free energy that are not flattened by the dynamics will be the order of the size of the newly added gaussians.

In most cases, it is not sufficient to describe the process of interest using a single CV, therefore in cases where we use more than one CV at the same time, the metadynamics potential is given by

$$V_G(S(x),t) = \omega \sum_{\substack{t'=\tau_G,\tau_{2G},... \\ t'<t}} \exp\left(-\sum_{\alpha=1}^{d} \frac{(S_\alpha(x)-s_\alpha(t'))^2}{2\delta s_\alpha^2}\right)$$

(1.32)

and it is necessary to choose metavariable δs_α for each CV. The time required to escape from a local minimum is determined by number of Gaussians needed to fill the basin. This number is proportional to $(\delta s_\alpha)^d$ where d is the number of CV chosen to describe the system. The efficiency of the method scales exponentially with number of CV. The added gaussians can reproduce features of the FES on a scale larger than δs. The quality and accuracy of the free energy construction is highly influenced by the parameters of metadynamics. The parameters need to be chosen wisely, striking a balance between accuracy and efficiency.

1.4.3 *How to Choose CVs*

Reliability of Metadynamics is highly dependent on the choice of CVs, and this in turn depends on the process of investigation. Therefore if not essential it is important to have good knowledge of the process under investigation, which aids in chosing the right CVs. Important features that CVs must satisfy are:

1. They should be slow variables able to describe well the process of interest.
2. They should be able to distinguish between the initial, intermediate and final states.

Practical Example of choosing CVs

For example, lets consider the diffusion of antibiotics through OmpF channel. The OmpF channel has an axis of symmetry along Z-axis, with a hourglass shape. Here we want to follow the exit of antibiotic from the channel. In this case, one appropriate CV would be distance: defined as difference between the center of mass (com) of the antibiotic (a_1) and the center of mass of OmpF (p_1) along z-axis. The CV distance, s_{dist} is then defined as:

$$s_{dist} = \frac{\sum_{i \in a_1} m_i z_i}{\sum_{i \in a_1} m_i} - \frac{\sum_{i \in p_1} m_i z_i}{\sum_{i \in p_1} m_i} \tag{1.33}$$

We make a assumption in equation 1.33, com of OmpF (p_1) is constant respect to com of the antibiotic(a_1). Therefore using this CV s_{dist} we caputure the location of antibiotic in the OmpF channel.

1.4.4 *Estimation of Error*

Error estimation in Free energy calculations is very much dependent on the meta variables ($\omega, \delta s, \tau_G$). In [**?**] it has been shown

that the error on reconstruction profile is determined by the ratio of (ω/τ_G) and not by ω and τ_G alone. For instance, adding Gaussian of height ω=1.0 Kcal mol^{-1} every τ_G = 4ps, is equivalent to adding Gaussian of height ω=0.5 Kcal mol^{-1} every τ_G = 2 ps, with a requirement that τ_G remains much shorter than the time required to fill the free energy basin. The error ϵ in metadynamics is a measure of expected deviation of V_G(s,t) from -F(s):

$$\varepsilon^2(s) = \left\langle (V_G(s) + F(s))^2 \right\rangle_M = \tag{1.34}$$

$$= \left\langle (V_G(s) - \langle V_G(s) \rangle_M)^2 \right\rangle_M \tag{1.35}$$

Remarkably, the error does not depend on $F(s)$. An alternative expression for for the error, dependent on parameters: w/τ_G, T, D and S, is given as:

$$\bar{\varepsilon}^2_{\text{approx}} = C_d \frac{S^2 w T}{D\tau_G} \frac{\delta s}{S} \tag{1.36}$$

where C_d is a constant that depends only on the dimensionality. The two expressions for the error share the same functional dependence on w/τ_G, T, D and S. The ratio between the two expressions is approximately a constant as a function of δs only for $d = 1$ and $d = 2$, while significant deviations are observed in higher dimensions.

The dependence of the error on the simulation parameters becomes more transparent if $\bar{\varepsilon}$ is expressed as an explicit function of the total simulation time. Consider in fact a free energy profile $F(s)$ that has to be filled with Gaussians up to a given level F_{max} , for example the free energy of the lowest saddle point in $F(s)$. The total computational time needed to fill this profile can be estimated as the ratio between the volume that has to be filled and the volume of one Gaussian times τ_G :

$$t_{sim} \approx \tau_G \frac{F_{max}}{w} \left(\frac{S}{\delta s}\right)^d \tag{1.37}$$

Substituting in Eq. 3.36 yields

$$\bar{\varepsilon}^2 \approx \frac{\tau_S}{t_{sim}} F_{max} T f_d(\frac{\delta s}{S}) \tag{1.38}$$

where $\tau_S \doteq \frac{S^2}{D}$ is the average time required for the CVs to diffuse on a distance S and

$$f_d(\frac{\delta s}{S}) = (2\pi)^{d/2} \sum_k \frac{1}{k^2\pi^2} \exp\left(-\frac{k^2\pi^2}{2}\left(\frac{\delta s}{S}\right)^2\right)$$

is a function of $\frac{\delta s}{S}$ and of the dimensionality alone. Eq. 1.38 states that the error of a metadynamics reconstruction is inversely proportional to the square root of the total simulation time, measured in units of the diffusion time. The error will be large for slowly diffusing systems, in which the walker takes a long time to explore the CVs space.

1.5.1 *MM-PBSA*

one of the major goals of the computational chemistry is to develop methods to accurately predict the binding eergy of a ligand to a protein. This is of central interest in medicinal chemistry, because the action of most drugs (inhibition, activation etc.) is caused by the binding of the drug to its target receptor. However, many biochemical problems can be treated in a similar way For example, the reactivity of an enzyme can be estimated by comparing the free energy of the reactant and tranisition states of the active site in protein. Therefore, many methods have been developed with this aim [**?** **?**]. MM-PB(GB)SA method is another free energy calculation method(107). Unlike other free energy calculation methods like free energy perturbation (FEP) or Thermodynamic Integration(TI) method, this is the post-processing method, which means it calculates the free energy based on existing structure snapshots. MM-PBSA restrict the molecular simulations to the states before and after the binding process.It is an attractive approch because it does not contain any parameters that var for different ligands-receptor systems and it involves a set of physically well defined term: The binding affinity is estimated from the free energies of the receptor, the ligand and the complex. In MM-PB(GB)SA calculation, the free energy of the system is divided into three parts: molecular mechanics energy(MM), and non polar solvation free energy (SA). The MM energy usually consists of electrostatic energy and vander walls energy of the system in vacuum. In some applications the bond energies, angle energies, and dihedral angle energies are also included in MM energy. the entropy of the system can be included by using normaol mode analysis or other methods.

In this method to estimate the free eergy of the complex system, one carries out a molecular dynamics simulation, typically in a periodic box with water and counter ions, and correct representation of long-range electrostatic effects, saving a set of representative structures. then one post process these structures, remove any solvent and counter ions molecules, and calculates the free energy, G, accordign to the following equation:

$$\triangle G_{bind} = \triangle E_{MM} + \triangle G_{PBSA} - TS_{MM} \tag{1.39}$$

or

$$\triangle G_{bind} = G_{complex} - (G_{receptor} + G_{ligand}) \tag{1.40}$$

where $\triangle G_{bind}$ is the calculated average free energy, and $\triangle E_{MM}$ is the average molecular mechanical energy, $G_{complex}$, $G_{receptor}$,

and G_{ligand} are the free energies of the complex, the protein, and the ligand, respectively.

For each system, the free energy can be estimated in terms of molecular mechanic potential energy E_{MM}, the solvation free energy G_{solv}, and the entropic contribution (TS) as fallows:

$$G_{bind} = E_{MM} - TS + G_{solv} \tag{1.41}$$

$$E_{MM} = E_{internal} + E_{elec} + E_{vdW} \tag{1.42}$$

$$E_{internal} = \triangle E_{bond} + \triangle E_{angle} + \triangle E_{tors} \tag{1.43}$$

$$G_{solv} = G_{elec,solv} + G_{nonpolar,solv} \tag{1.44}$$

where these correspond to the bond, angle, torsion, van der waals, and electrostatic terms in molecular mechanical force field, evaluated with no non-bonded cutoff. All the terms in Eq. 1.42 are averages of energies obtained fron a number of snapshots taken from the MD simulations [**?**]. In order to reduce the the time consumption and to obtain the stable energies, the same geometry is normally used for all three reactants (complex, receptor,ligand), i.e only the complex is explictily simulated by MD.There by, $E_{internal}$ cancels in the clculation of $\triangle G_{bind}$. The MM-PBSA methos has been successfully applied to many systems [**? ?**], [**? ? ?**]. G_{pbsa} is the solvation free energy calculated with the numerical solution of the Poisson-Boltzmann equation and an estimated of the non-polar free energy with the simple surface area term. The nonpolar solvation contribution is estimated as

$$G^{S}_{nonpolar,solv} = \gamma A + b \tag{1.45}$$

Where A is solvent-accessible surface area and the solvent parameters, γ and b, are 0.00542 kcal/mol $Å^2$ and 0.92 kcal/mol, respectively.the probe radius of the solvent was set to 1.4 . the atomic radii of of the solute were taken from PARSE parameter set [**? ?**]. $-TS_{MM}$ is the solue entropy, which can be estimated by quasi harmonic analysis of the trajectory or by using normal-mode analysis (108). the solvant entropy is included tin the solvation free energy term. The free energy due to ionic strength effects can be added with a continuum approch.

The ability to accuratly calculate $\triangle G$, the average free energy, we can calculate $\triangle G$ for a given macromolecular system in various different confermations or structures, adds a very important methodology to our computational arsenal. This has been possible before with free energy perturbation but only for small

systems and very limited conformational or topological changes. By using continuum model, it is implicity integrating out all the solvent coordinates and simplifying the problem. Also, by calculating the absolute free energy directly with eq 1.2 between the two "end points" instead of calculating the relative free energy along a mapping coordinate, we are avoiding computations of eq 1.1 will have intrinsically much larger errors than free energy perturbation/thermodynamic integration calculations. What is surprising is that, despite these larger uncertainties, we can often calculate $\triangle G$ in respectable agreement define with experiment.(107)

Chapter3

Different binding sites of the benzimidazole compound 227G on HCV and BVDV RdRps, revealed by MD simulations

Abstract

The virally encoded RdRp has emerged as a prime target in search for specific HCV and BVDV antiviral. Benzimidazole class of compound has been extensively pursued as potent Non-Nucleoside Inhibitors (NNI's) against both classes of viruses. For the first time, Screening efforts in our lab have come out with potent NNI belonging to benzimidazole class of compound which is active against both HCV and BVDV RdRp. Our tested compound was confirmed to strongly inhibit BVDV RdRp activity in a dose-dependent manner and they also showed inhibition on HCV1b-NS5B.

While, the HCV RdRp structure in complex with the related analogs showed that these inhibitors bind at the surface of the thumb, the resistant mutation and newly identified binding site for the same analog were found to lye in the finger domain of BVDV RdRp. Thus, the molecular mechanism of inhibition of the inhibitor found to be different in the otherwise closely related polymerases.

Molecular dynamics simulations helped us to gain a deeper insight into the interaction pattern, mechanism of inhibition of inhibitor onto the two RdRps. The multidisciplinary approach allowed us to (i) critically assess the dogma of antiviral therapy; (ii) to screen efficiently the possible information valuable for designing effective and selective inhibitors by identifying at a molecular level the interaction pattern of the compound with its viral targets; (iii) to assess the use of BVDV as surrogate for the development of antivirals against HCV

Introduction

Viruses belonging to the Flaviviridae family, which encloses the three genera Hepacivirus, Flavivirus and Pestivirus, are responsible for severe disease in humans and animals[1-4] . Among the Hepaciviruses, Hepatits C Virus (HCV) is the frequent cause of acute and chronic hepatitis, liver cirrhosis and hepatocellular carcinoma worldwide, and creates a significant burden to healthcare systems due to mortality, morbidity and treatment costs [5-8]. To give the numbers, hepatitis C (officially recognized as the "silent epidemic") infected approximately 180 million people worldwide, and it is the leading reason for liver transplantation in the United States [9-11]. Unfortunately, more than 20 years after discovery of HCV [5] therapeutic options remain limited, and there are no vaccine or effective therapy broadly targeting all genotypes of HCV [12]

The current standard therapy consists in administration of pegylated alpha interferon along with nucleoside analogue inhibitor (NI) ribavirin[12-14], which is effective in only about half of patients who suffer from chronic HCV infection, and often produces toxicity and significant side effects .[15,16] Given the high prevalence of HCV, there is an urgent need to develop more effective and well tolerated therapies. [17,18]. In this respect, new therapeutic lines are following the concept of specifically targeted antiviral therapy for HCV (STAT-C), aiming at directly targeting HCV RNA and viral enzymes or influence host-virus interactions (at opposite with standard anti-HCV protocol which enhances natural immune responses to the virus) at influencing host-virus interactions [19-21] Namely, novel treatment options now in development are focused on inhibitors of HCV-specific enzymes NS3 protease [22-23] and NS5B RNA-dependent RNA polymerase (RdRp)[24]. However, the emergence of resistant mutations that limit the use of these compounds in monotherapy complicates the regimens. Thus, HCV infections are better treated by combination of drugs with distinct mechanisms of action and non-overlapping resistance profiles. This has shown to increase the effectiveness of antiviral responses across all genotypes, with shorter treatment duration and better tolerability .[25-29]

According to the expertises acquired in our group polymerases seem to be good target for inhibitors.[30-33] Among them two major classes can be identified on the basis of chemical structure and mode of action.

The first is constituted by Nucleoside analogue Inhibitors (NIs) , which act as chain terminators targeting the catalytic active site of polymerase. They are potentially effective against different genotypes, and polymerases feature a relatively high genetic barrier in developing resistance to them.[34-35] The second category is that of Non-Nucleoside analogs Inhibitors (NNIs), which achieve polymerase inhibition during the initiation phase and do not require activation like in case of NIs .[36-38]

NNIs bind to one of the at least four allosteric sites of the enzyme free of (its natural) substrate and uncomplexed with any other non-structural protein; they cause various conformational changes that inactivate the enzyme, or trap it in a functional conformation impeding an essential structural transition between initiation and elongation. [39-40]

The benzimidazole series of NNIs represents a promising group of compounds showing activity against HCV and Bovine Viral Diarrhoea Virus (BVDV). [41-43]These compounds were originally screened as specific inhibitors of HCV NS5B activity in vitro, but they exhibited marginal potency in cell-based assays, probably due to poor cellular permeability resulting from the presence of a 5-COOH group that is ionisable at physiological pH . A 1,2-disubstituted benzimidazole-5-carboxylic acid scaffold was identified as the minimum core for biological activity, thus paving the way for optimization and subsequent generation of analogues capable of inhibiting HCV RdRp with an increased potency.[44]

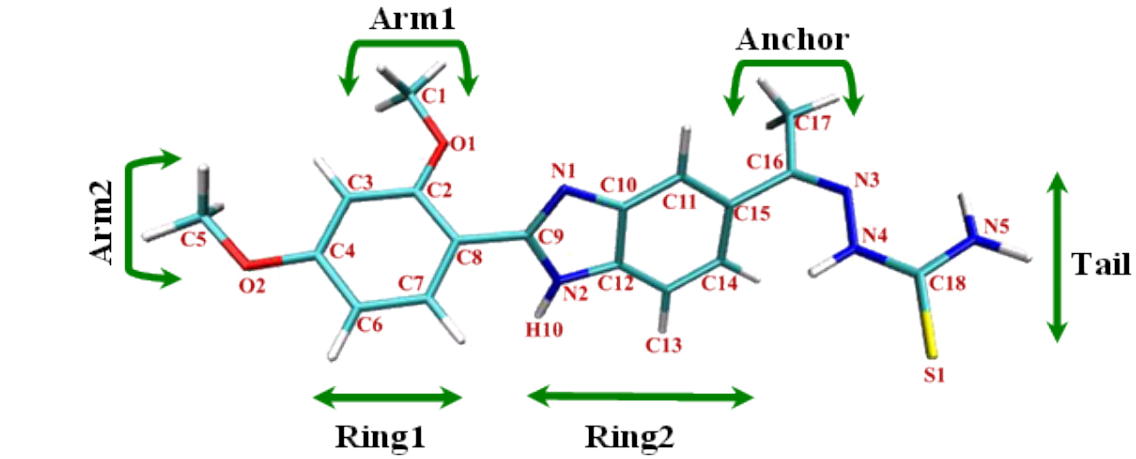
In a previous work we showed that our lead benzimidazole compound 227G (Fig.1C) is a highly potent and selective inhibitor of the replication of BVDV . BVDV is considered since long time an attractive and predictive surrogate model for identifying HCV inhibitors, for several practical reasons and due to the biological similarities between these two viruses [45-46]. However, although BVDV remains to date a popular surrogate model, the concordance between drugs capable of inhibiting its replication and those inhibiting HCV is still largely unknown. In other words, a complete validation of BVDV as surrogate virus model has not yet been demonstrated.[47-49]

In this work we used replicon, cell-based and enzymatic assays to show that 227G effectively inhibits HCV RdRp, although with slightly larger values of IC_{50}, EC_{50} and CC_{50} than in BVDV. These results contrast with the classical dogma characterizing the field of antiviral agents: they have high potency and selectivity, but a reduced spectrum of activity, which is often limited to a single viral genus. Nonetheless, exceptions to this dogma have already been reported in the literature. Concerning NIs, the 2'-methyl-derivatives represent the very first class of broad-spectrum compounds able of inhibiting the multiplication of RNA viruses belonging to different families .[50] With regards to NNIs, a similar exception is represented by imidazopyridine and acridone derivatives, both active against both HCV and BVDV.[51]

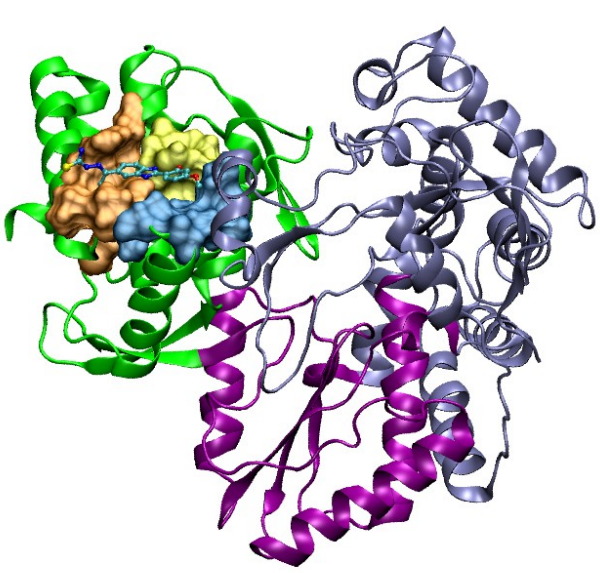
To our knowledge, 227G is thus the first benzimidazole compound showing antiviral activity against HCV and BVDV. Furthermore, it is the first among this family of compounds showing sub-micromolar values of IC_{50} for both viruses.

Despite these intriguing findings, no thorough and microscopically detailed studies have been performed aimed at comparing the molecular mechanism of action of NNIs in the two RdRps. In an effort at shedding light in this direction, we have used a multidisciplinary approach that combines biological assays and computational results to assess the molecular basis of the molecular interaction of 227G with the two viral RdRps.

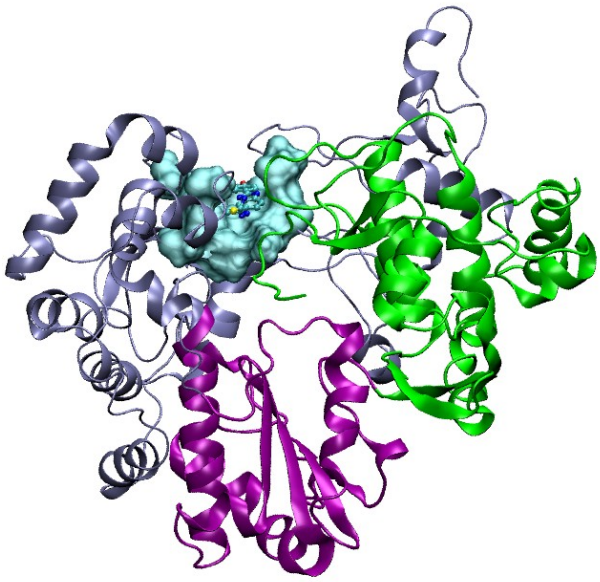
According to our results, the binding sites of 227G on the two polymerases are different, being on the thumb domain in HCV and on the finger domain in BVDV; in this latter case, a new binding site has been found, lying very close to motif-I of the finger-tip region. On the basis of these results we propose a different molecular mechanism of inhibition of BVDV and HCV by 227G: while in HCV the compound opens up the polymerase to an inactive conformation [52-54] in BVDV it could bind at the entrance of the template channel, blocking the insertion of the polynucleotide chain. The identification of different molecular interaction patterns in 227G-HCV vs 227G-BVDV reported for the first time here points to the fact that, even in the case where a compound is active against both RdRps, the equivalence between the mechanisms of action is not guaranteed. The microscopically based comparison constitutes a valuable starting point for the design of more effective and selective benzimidazole inhibitors. Such an ambitious but unavoidable goal can be reached only via a investigation strategy that combines biological relevant data with biophysical and biochemical pictures of the interaction mechanisms.



A)



B)



C)

Figure 1: The structure of HCV (B) and BVDV (C) RdRps are represented with their different subunits and domains: the thumb is colored in dark green, fingers are colored in ice blue and the palm domain along with unique N-Terminal domain is colored in purple. The surface view areas highlight the binding site of 227G in HCV (thumb) and BVDV(finger) respectively. (A) Ball-and-stick representation of 227G with all atoms labeled according to the notation of the force field. 227G is well characterized as ring I, II , arms, anchor and its tail. All images are generated by VMD [Humphrey 1996].

Systems and Methods

1)- Experimental:

Cells and viruses.
Madin-Darby Bovine Kidney (MDBK) cells were used for growth and selection of mutants resistant of Bovine Viral Diarrhea Virus (BVDV). Either cells (CCL-34) and viruses (VR-534 strain NADL) were purchased from American Type Colture Collection (ATCC). The absence of mycoplasma contamination was checked periodically by the Hoechst staining method.

HCV replicon assay.
A human hepatoma cell line bearing the HCV genotype 1b replicon (GS4.1 cells), kindly provided by C. Seeger (Fox Chase University, Philadelphia, PA, USA), **were** grown in DMEM supplemented with 10% foetal bovine serum (FBS), 2 mM L-glutamine, 110 mg/L sodium pyruvate, 0.1 mM non-essential amino acids, 100 U/mL penicillin–streptomycin and 0.5 mg/mL G418 (Invitrogen). For dose-response testing, the cells will be seeded in 96-well plates at $7.5x10^3$ cells/well in a volume of 50 µL of 10 two-fold serial dilutions of compounds (highest concentration, 75 µM) will be added and cell cultures will be incubated at 37 °C/5% CO_2 in the presence of 0.5% DMSO. Alternatively, compounds were tested at a single concentration of 15 µM. In all cases, Huh-7 cells will serve as a negative control. The cells were incubated in the presence of compounds. The inhibition of HCV replication is measured by quantification of viral NS4A protein using an enzyme-linked immunosorben assay (ELISA). After three days, the plates will be then fixed for 1 min with 1:1 acetone–methanol, washed twice with PBS containing 0.1% Tween 20, blocked for 1 h at room temperature with TNE buffer containing 10% FBS and then incubated for 2 h at 37 °C with the anti-NS4A mouse monoclonal antibody A-236 (ViroGen, Watertown, MA, USA) diluted in the same buffer. After washing three times with PBS containing 0.1% Tween 20, the cells were incubated for 1 h at 37 °C with anti-mouse immunoglobulin G–peroxidase conjugate in TNE buffer with 10% FBS. After washing as described above, the reaction were developed with O-phenylenediamine (Zymed, San Francisco, CA, USA). The reaction were stopped after 30 min with 2 NH_2SO_4 and the absorbance will be read at 492 nm. EC_{50} values were determined from the inhibition versus concentration data using a sigmoidal non-linear regression analysis based on four parameters with Tecan Magellan software. For cytotoxicity evaluation, GS4.1 were treated with compounds as described above and cellular viability will be monitored using the Cell Titer 96 Aqueous one solution cell proliferation assay (Promega). CC_{50} values will be determined from the % cytotoxicity versus concentration data with Tecan Magellan software as described above.

Expression and purification of BVDV and HCV1b RdRps (NS5B).

By inserting the truncated genes between the BamHI and XhoI cloning sites of the pET21 expression plasmid (Novagen) we generated plasmids for expression of a truncated form of HCV NS5B protein (C-terminal 21-amino-acid-deleted) and of BVDV NS5B (C-terminal 21-amino-acid-deleted). Recombinant RdRps were fused in frame with a C-terminal peptide containing six tandem histidine residues and expression of the His-tagged proteins was driven by a T7 RNA polymerase promoter. The resulting plasmids were transformed into the *Escherichia coli* strain BL21 (DE3) Rosetta pLysS (Novagene) and the transformants were then cultured in 5 ml of LB medium with 100 mg/ml ampicillin at 30 °C overnight. Cultures were diluted into 0.25 l of LB medium with 25 mg/ml kanamycin and 30 mg/ml chloramphenicol and incubated at 30 °C until the A_{600} reached 0.6-0.7. These cultures were then induced overnight with 2mM isopropyl- b-D-thiogalactopyranoside. The cells (from 0.25 l) were harvested by centrifugation, and cell lysis was achieved by the addition of 10 ml of CelLytic B (Sigma). Any insoluble material was removed by centrifugation at 11000 rpm (4 °C) for 60 min.

The soluble extracts were applied to a 5-ml column of nickel-nitrilotriacetic acid-agarose (Qiagen) that had been equilibrated with buffer A (50mM NaH2PO4, 300mM NaCl 10mM imidazole pH 8.0). The columns were washed with the same buffer and then eluted stepwise with buffer A containing, 50, 70, 130, 170, and 250mM imidazole. The polypeptide composition of the column fractions was monitored by SDSPAGE. The recombinant RdRps were retained on the column and recovered in the 130-250mM imidazole eluate. These fractions were dialyzed against buffer B (25mM Tris–HCl, pH 7.5, 2.5mM MgCl2, 1mM dithiothreitol, 50% glycerol). The protein concentrations were finally determined by the Bio-Rad dye binding method with bovine serum albumin as the standard.

RNA-Dependent RNA Polymerase Assays.

Polymerizations were performed in 25-ml reaction volumes. The reaction mixture contained 20mM Tris, pH 7.5, 20mM $MnCl_2$, 6mM $MgCl_2$ 25mM NaCl, 25% glycerol, 0.5 lg/l bovine serum albumin, 2mM dithiothreitol. Our standard reactions used 600 ng of recombinant proteins, 100 mM GTP and 300 ng of template poly(rC) preannealed with 25 ng of primer (rG) with or without decreasing amounts of compound. Reactions were incubated at 30°C for 120 min then stopped by addiction of 2 ml of 200mM EDTA. 175ml of PicoGreen Quantitation Reagent (Molecular Probes), diluted 1/345 in TE, were added to each sample and incubated at room temperature, protected from ambient light, for 5 min. Fluorescence of samples in wells of a 96-well microtiter plate was determined in a fluorescence microplate reader.

2)- Computational

Structures of Receptors

The published X-ray crystal structures of BVDV RdRp (9) (PDB code 1S48; resolution 3.0 Å) and HCV RdRp (10) (PDB code 2BRK; resolution 2.3 Å) were used as starting structures for molecular simulations. BVDV and HCV RdRp consist of 588 and 532 amino acid residues respectively (Figure1). In HCV, residues Leu21 to Asn35 form a highly flexible loop, which is unresolved in the crystal structure. This loop ($_{21}$LPINALSNSLLAHHN$_{35}$) is supposed to play major role in protein-ligand interaction and hence was generated by using the Loopy module of Jackal modeling package (55). Simulations of both receptors in their uncomplexed form (vide infra) were performed in order to assess the presence of major conformational changes in the proteins and obtain starting structures for docking.

Structures of Complexes

Structures of the complexes 227G-BVDV and 227G-HCV were generated through molecular docking using the Autodock 4 package [56-57]The structures of the receptors used in docking runs were taken from equilibrium MD simulations of the apoenzymes, namely selecting the conformation with the lowest RMSD from the average structure extracted from the last 10ns of MD (i.e. at equilibrium, vide infra). In the absence of literature data indicating the presence of allosteric sites in BVDV RdRp, the C_α atom of Ile261 was chosen as the grid centre for all docking runs. The choice is guided by mutagenesis experiments performed on BVDV, showing that mutation Ile261Met produces resistance to 227G (see Ref. Coupled paper; Ile261Met mutation is resistent mutaent for phenyl-benzimidazole [58] and arylazoenamine .[59] For HCV RdRp, the C_α atom of Leu30 was taken to be the grid center as Leu30 lies in the experimentally known binding cavity of indole compounds in HCV RdRp. For both complexes grid maps were made of 30x40x40 points in x, y, z directions, distant 0.375 Å from each other.

The Lamarckian genetic algorithm was used ([57] with the following parameters: 10 million generations and energy evaluations; population size of 250; 100 runs (as recommended for blind docking) [60]; mutation rate of 0.02; crossover rate of 0.80; elitism value of 1. For local search, the so-called pseudo-Solis and Wets algorithms were applied by the default parameter . [61, 57]

Finally, two separate conformers for each system (227G-BVDV and 227G-HCV) were selected on the basis of 1) their binding energy, 2) cluster population and 3A) vicinity to the mutation in the case of BVDV-227G complex, 3b) orientation featuring key interaction with the protein as found in X-ray structures of complexes between HCV and compounds of the 227G family ,[59,62] in case of HCV-227G complex. These complexes were used as the starting structures for MD simulations.

MD Simulation protocol

Parametrization.

The AMBER99, TIP3P, and AMBER-modified Aaqvist [63,64] force fields were used to model proteins, water and ions respectively. Concerning 227G, since no experimental structure is available, the molecule was first drawn with ACD Chem. Sketch 11.0. The resulting molecular geometry was then optimized at the HF/6-31G(d) level up to a convergence in energy of 10^{-5} au using the Gaussian03 package [[65,66]. The *gaff* force field [Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. J. Comp. Chem. (2004) 25:1157–1174.][67] was used except for charges, calculated following the standard AMBER protocol: from the electrostatic potential map generated by the molecule (calculated with Gaussian03), the RESP [68] charges were derived using the *antechamber* module of AMBER.

Dynamics.

State-of-the-art all-atom MD simulations in the presence of explicit water and counter-ions were carried out with the ORAC package[69] for both receptors, in their apo forms and in complex with 227G. Details of the simulation are reported in Table 1. For all of the system the following procedure was used: First, geometry optimizations were carried out with a two-step protocol: (i) 10000 cycles (2000 steepest descent plus 8000 conjugate gradients) with harmonic restraint of k=10 kcal/(mol Å^2) on each heavy atom of the solute; (ii) 20000 conjugate gradients cycles with no restraints. Next, heating up to 300 K was achieved by linearly increasing the temperature in 100 ps of *NVT* MD, while imposing restraints of 1 kcal/(mol Å^2) on the solute. Restraints were then released for 100 ps, and as a last step preceding the productive dynamics, 1 ns of NPT MD was carried out in order to relax the simulation box. Finally, 20 and 10 ns long simulations were performed respectively for each apo-enzyme and adduct in explicit water solution under the NPT ensemble. Pressure and temperature were regulated at 1 atm and 310K using the isotropic Andersen–Parrinello–Rahman barostat [70,71]and the Nosé-Hoover thermostat [72,73], respectively.

Electrostatic interactions were evaluated using the soft Particle Mesh Ewald schemes with 1 Å grid spacing and a cut-off of 10 Å, the same used for Lennard-Jones interactions. A MTS-Respa algorithm with five shells (time steps of 0.5-1-2-4-12 fs) was used to integrate equations of motion, in conjunction with the SHAKE algorithm applied to bonds involving hydrogens.

Table :1

Receptor	Simulation length	Total atoms	Counter ions	Box dimension (side, angle)
$HCV_{complex}$	10ns	56092	14Cl	88.9 A, 109.47 dgrees
$BVDV_{complex}$	10ns	55696	17Cl	88.9x88.9x88.9 A;109x109x109
$BVDV_{wild}$	20ns	55705	17Cl	88.9x88.9x88.9 A;109x109x109
$BVDV_{mutant}$	20ns	55703	17Cl	88.9x88.9x88.9 A;109x109x109

Analysis of structures and dynamics.

For each analysis we took 10ns of production run. An inventory of the structural details and interactions was obtained by performing an extensive qualitative and quantitative analysis of the MD simulations in terms of hydrogen bonds (HB), hydrophobic contacts (HpH) and water-mediated interactions. The HB's between 227G and RdRp are counted using VMD scripts with the following threshold parameters: a distance of at most 3.2 Å and donor-hydrogen acceptor angle larger than 150°. HpHs are counted when non-polar atoms are separated by at most 3 Å. The durable HB is defined as the one with a lifetime equal to or higher than 20% of MD simulation time, and transient HB are the one with a lifetime shorter than 20% of simulation time [74]. For SB (Salt-Bridge), the cutoff distance is 3 Å. The criterion to determine a π-π interaction was that the SICD (Short inter atomic carbon-carbon distance) should be smaller than 4.8 Å .The RMSD residues wise was plotted taking reference average snapshot from the equilibrium dynamics to characterize the flexibility of the protein, in particular in the template entrance region, the behaviour of the minimum distances between key residues in flexible loops in this region, and the area of the entrance region along the whole equilibrium dynamics.. The energetic analysis was performed in the usual manner: the interaction energy was decomposed as the sum of individual components, each with a physical meaning and evaluated via the force field terms. In particular, we examined variations in the 227G-BVDV RdRp electrostatic and van der Waals interactions..[75]

Results

Experimental

Antiviral activity of 227G compound.

Recently, 21 compounds belonging to the chemical class of 2-phenyl benzimidazoles were evaluated in cell-based assays for their cytotoxicity and antiviral activity against a panel of RNA and DNA viruses, comprising, among the others, viruses representative of two of the three genera of the *Flaviviridae* family, i.e. Pestiviruses (Bovine Viral Diarrhoea Virus, BVDV) and Flaviviruses (Yellow Fever Virus, YFV). Hepaciviruses was not screened since it can hardly be used in routine cell-based assays.

As reported in Table 2, compound 227G was identified as the most potent and selective inhibitor of BVDV replication, with an EC_{50} value of 0.8 μM. 227G had no cytotoxic effect on uninfected MDBK cells (CC_{50} >100 μM), thus showing a selectivity index >125.

In the light of the interesting activity against BVDV and the lack of activity against YFV, we considered necessary to assess the 227G activity against the *Hepacivirus* genus of the *Flaviviridae* family, through the use of the Huh7 cell-based HCV replicon system. It turns our that compound 227G inhibited NS5B-directed viral replication, with an interesting EC_{50} of 1.1 μM (Table 2).

In vitro RdRp assays.

Since benzimidazoles are reported to be non-nucleoside inhibitors, the selected compound was evaluated for its activity on BVDV and HCV-1b recombinant RdRps. Historically RNA polymerase activity has been measured by the incorporation into RNA of radioisotopically labeled nucleotides; however, due to logistics reasons, we sought a non-isotopic detection method. We previously report the development of an assay for RNA polymerase activity based on the use of the RiboGreen dye to detect transcripts by fluorescence.

As reported in Table 2, 227G strongly inhibits BVDV RdRp activity in a dose-dependent manner, with an IC_{50} value of 0.002 μM; HCV1b RdRp was also inhibited, even though with an higher IC_{50} value of 0.4 μM. Thus, viral polymerases were confirmed to be the molecular target of 227G, which appears to be in our knowledge the first one belonging to this class which is active against both viruses.

For this reason, 227G was deeply investigated with the aid of bioinformatic tools, in order to exploit the molecular basis of its mechanism of action, which we supposed to be different in BVDV compared to HCV.

Table 2: Comparative evaluation of: i) antiviral activity of 227G against BVDV (cell-based assay) and against HCV-1b (Huh7 cell-based HCV replicon system assay); ii) inhibitory activity of 227G in enzymatic assays on recombinant purified RdRps of BVDV and HCV-1b.

	Cell-based and replicon HCV assays				*Enzymatic assays on RdRp*	
Compound	**BVDV**		**HCV-1b**		**BVDV**	**HCV-1b**
	CC_{50} [μM]	**EC_{50} [μM]**	**CC_{50} [μM]**	**EC_{50} [μM]**	**IC_{50} [μM]**	**IC_{50} [μM]**
227G	**>100**	**0.8**	**11.3**	**1.1**	**0.002[a]**	**0.4[a]**

[a] Data are the average of three independent experiments and standard deviations were less than 10%. % Residual activity was plotted versus increasing concentrations of compound. The curve was fit with Kaleidagraph (Synergy Software) to obtain IC_{50} values

Computational

Docking:

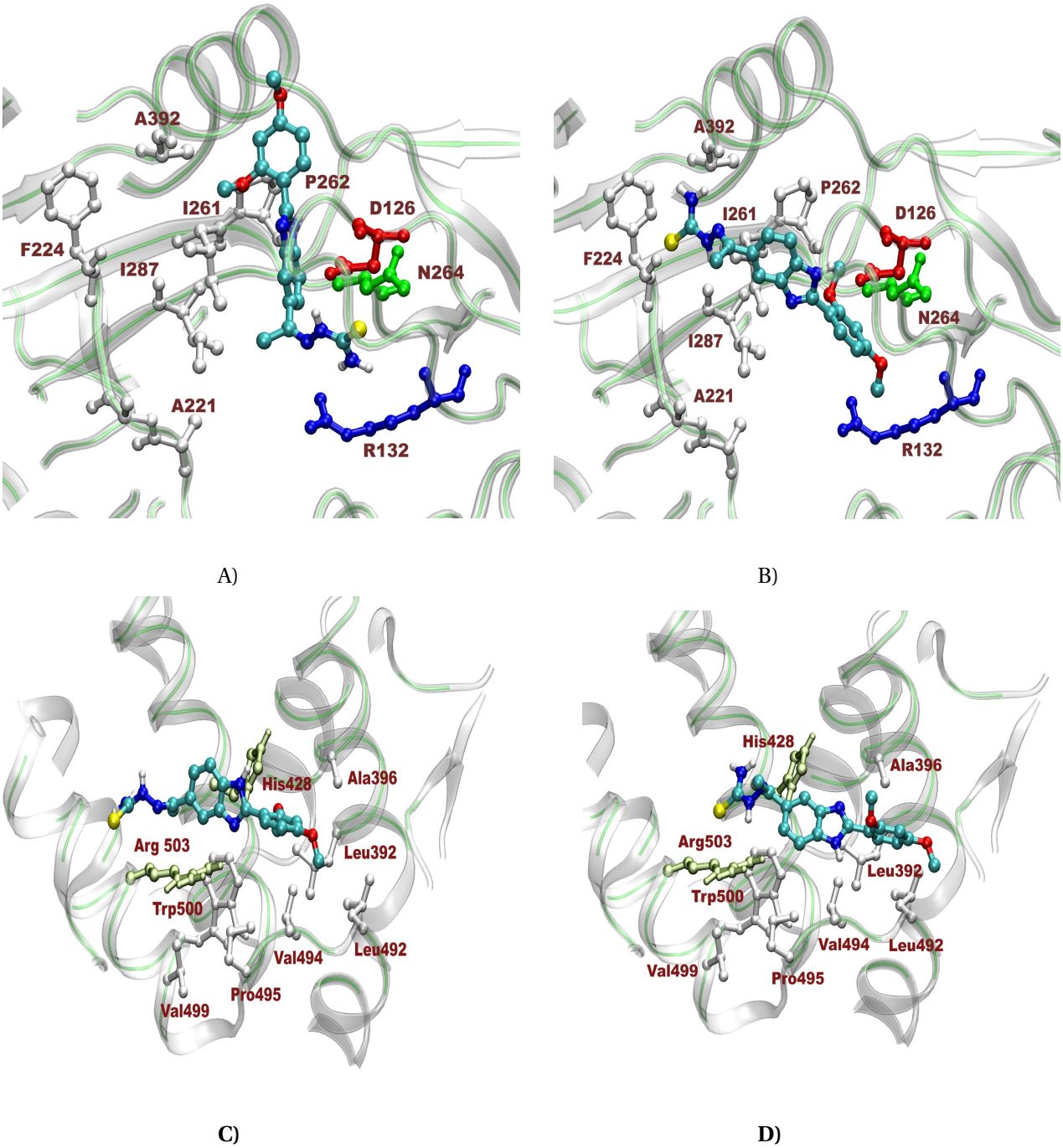
It has been suggested in literature that the binding mode of benzimidazole-containing compounds is similar to that of indole-based inhibitors [52]. Thus, as stated in Systems and Methods section, we used this previous experimental information to perform a guided docking of 227G onto HCV RdRp, and to compare with the binding mode of our compound.

In case of BVDV, due to the absence of crystallographic data of the complex indicating the binding of NNI in the RdRp, apo-enzyme (PDBID-1S48,)[76] was used for molecular modeling studies. Due to the lack of prior knowledge about putative binding cavity of NNI's, the proximity of the 227G to residue Ile261 was considered to be pivotal in choosing the starting structures [77] This residue has been shown by mutagenesis experiments [58,59] to be very important for the interaction of benzimidazole compounds with BVDV.

In each system two clusters have been found that satisfy all of the three criteria mentioned in Materials and Methods.

Representative conformers of each of them are shown in Figure 2. In case of 227G-HCV, the two lowest energy conformers were found to be in accordance with the expected mode of binding of benzimidazole compounds, as inferred from crystallographic studies .[52] (Fig.8)

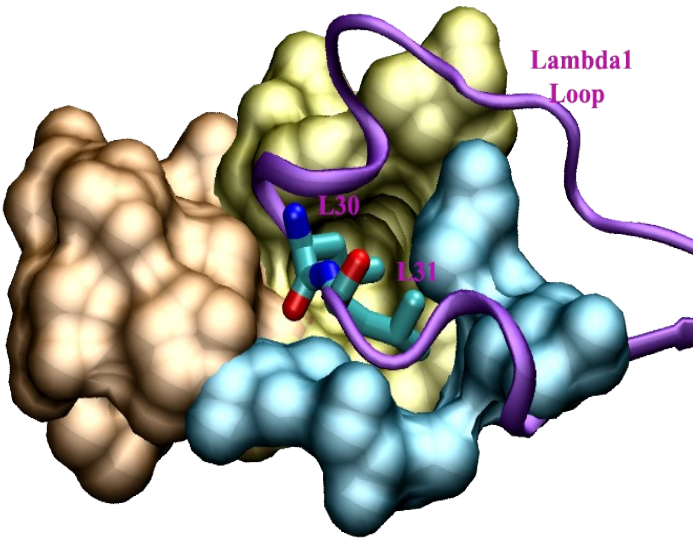
A comparative inspection revealed that in one conformer (Figure 2A) the part of 227G called arm1 occupies sub-pocket-1 (Figure 3B) of the binding cavity in contrast to what shown in Figure 2B, wherein the sub-pocket is empty. Furthermore, the methyl group of the tail region of 227G (Figure 1A) anchors to the extended hydrophobic region (Figure 3B) of the cavity indicating compliance to the experimental crystallographic studies. In BVDV, we obtained two clusters wherein 227G was found to be within attacking distance to Ile261, with opposite orientations. Since the two possible low-energy binding modes for the 227G in HCV and BVDV RdRp appear to be reasonable and to some extent supported by experimental evidence, both conformations were chosen for subsequent MD simulations.



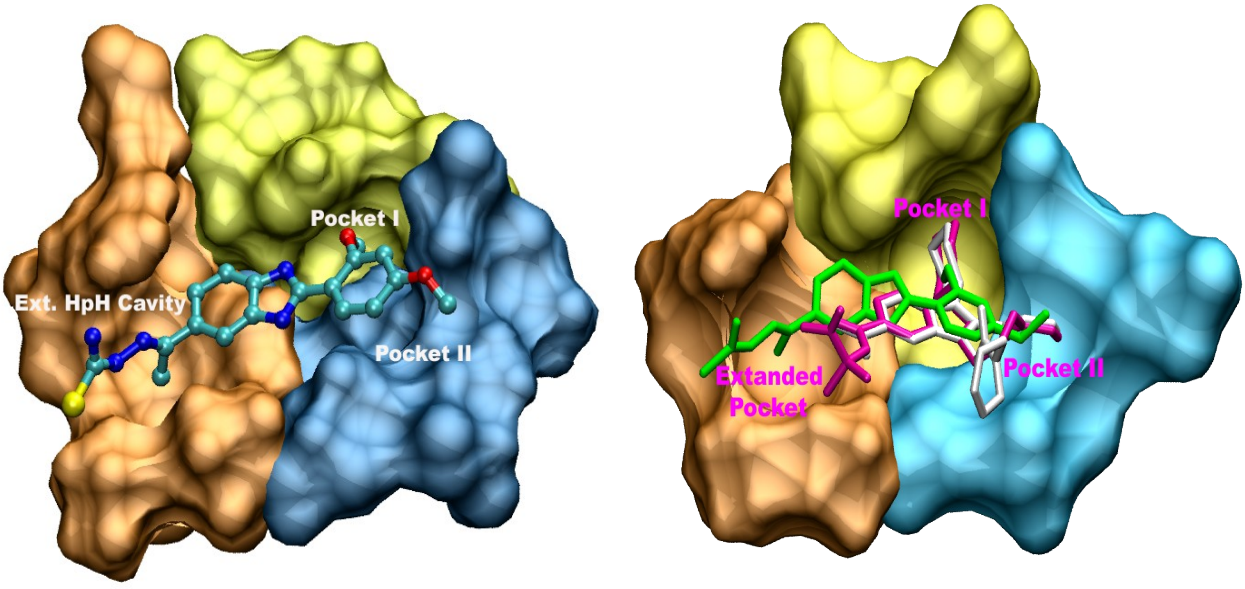
A) B)

C) **D)**

Figure2: Two different orientation of 227G in BVDV and HCV RdRp. Orientations are taken on the basis of the criteria mention in material and methods. During MD simulation we observed, one orientation of 227G is stable enough during whole length of trajectory while other one resulted into protein-227G dissociation in each case. (A) stable state of 227G in case of BVDV while B)- is unstable state. C)- stable state of 227G in HCV and D)- unstable state respectively. . The instability of 227G is because its failure to acquire its arm and tails in to their respective pockets. In both cases the key residues are same and their mode of interaction is almost same, but stability of only one pose clearly explained the specificity of 227G orientation.



A)-



B)- C)-

Figure3 : HCV RdRp, surface view representing the different pockets of cavity. Pocket1 and pocket2 are colored yellow and blue respectively while the extended hydrophobic cavity is represented in orange. A)depicts the active state of the protein wherein the Lys30 and Lys31 residues of the lambda1 loop sub optimally occupy the sub pocket1 and pocket2 of the cavity. B)227G accommodates itself in the HCV RdRp cavity. Arms of 227G occupy the sub pockets which in the apo protein is occupied by residues of lambda1 loop. C)- Binding cavity of HCV RdRp, indicating three different sub-pockets, I, II and extended pocket. Different arms of inhibitor occupying different pockets of the cavity. The CMF (white), 2DSX (pink) and 227G are green in color. The CMF and 2DSX correspond to their respective PDB ids.

MD Simulations

After approximately 3ns of MD simulation of the complexes represented in Figure 2B and 2D, the ligand spontaneously dissociated from the receptor and got completely solvated. The dissociation can be mainly due to the incorrect orientation of the compounds into the binding cleft. At opposite, simulations starting from configurations shown in Figure 2A and 2C did not exhibit any unbinding of the compounds along the whole MD run (20ns) and were then considered for detailed analysis of the binding features in BVDV and HCV, respectively. These are described in the following subsection for each complex.

BVDV.

COMPLEX

227G accommodate itself in a cavity encircled by four loops (L1, L2, L3 and L4). Loop L1 (Pro388-Ile398) and L2 (Ala221-Asn229) originate from the finger domain, loop L3 (Leu530-Gly537) belongs to the thumb domain, while loop L4 (Leu125-Arg132) is the part of the N-terminal domain. Further, the indole moiety of 227G stacks against the motif-I (Ile261-Lys266), which is located on the flexible finger-tip region Figure 4A. The residues lining the compound include mostly hydrophobic (Ala221, Ala222, Phe224, Ile261, Pro262, Ile287, Ala392) and polar (Thr160, Thr162 Asn217, Asn264 and Ser533) amino acids, although also three basic (Arg127, Arg130, Arg132) and two acidic (Asp126, Glu128) are present. This is consistent with the chemical properties of the compound, whose main body is made of aromatic rings and tail is polar in nature (Figure1C). Consistently, the major stabilizing factor of 227G is provided by the strong hydrophobic interaction with the receptor (-47.6Kcal/mol), with main contributions from residues Arg127, Glu128, Ile261 and Pro262. This feature is clearly identifiable in (Table4) and Figure 4B, where we collect the interaction energies and the (dynamic) interaction pattern related to HpH contacts, respectively.

In addition to hydrophobic interactions, a network of inter- and intra-molecular HBs stabilizes 227G in the pocket Figure5. In particular, a strong HB is established between N2 of 227G's indole moiety and the main chain of Asp126 (Average Dynamic length (ADL) = 2.3 Å). This interaction is present throughout all the simulation, has occupancy larger than 80%, and involves alternatively atoms OE1@Asp126 and OE2@ASP126 of the L4 loop. In addition to intermolecular bonds, a persistent intra-molecular HB exists during the entire simulation between side chain of Asp126 and HD@ASN264 of Motif-I. This interaction plays a significant role in stabilizing the former mentioned HB, which effectively locks the protein-ligand complex thereby securing the ligand stable inside the binding site.

Another source of stabilization of the complex comes from water-mediated interactions between Glu128 and N1@227G: although the mediating water molecule changes during the simulation, the interact persists for the entire simulation time.

This water mediated HB is stabilized by another set of intra-molecular HBs between residues Asp126 and Glu128 of the protein. The interaction between these two residues, weak at the onset of the simulation, is strengthened once the water-mediated interaction comes into picture. Another HB is formed between H18@227G and OG@Ser533 with ADL=2.5 Å and occupancy of 46%, which creates a linkage with loop L3 (Table3 & Figure10-suppli). We also found a transient HB formed between Ala221 and S1@227G (~occupancy 12%), further stabilizes 227G in the binding pocket. [Table 3 and Fig 5]

Interestingly, such intra-molecular bonds (either direct or water-mediated) were not observed in the simulation of the apo of the protein; thus, they are due to the presence of 227G and build a strong network that stabilizes ligand in the cavity.

HCV.

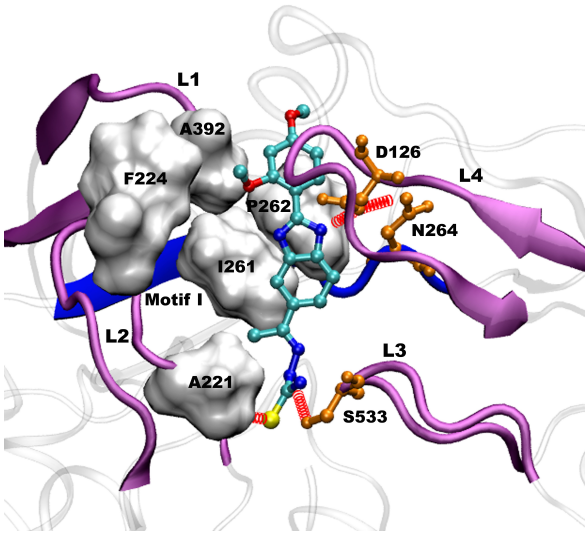
COMPLEX

227G sits very comfortably in the binding cleft present in the thumb domain of HCV RdRp which, in the WT, is occupied by the loop emanating from the finger domain. Primarily, the cavity is constituted by three alpha helices H1, H2 and H3. The residues outlining the cavities are Ile23, Asn24, Leu392, Ala395 Ala396, Ser423, Thr427, His428, Ser431, Leu492, Val494, Pro495, Val499, Trp500, His502 and Arg503 (Figure. 3C). As for HCV, the inhibitor-protein interaction is mainly hydrophobic in nature as the major contribution to the total interaction energy (-45.2 kcal/mol) is provided by the VdW term (-34.0 kcal/mol) (Table5). Val494 makes strong HpH interaction with the phenyl moiety of 227G with a percentage lifetime larger than 70% of the simulation time (Figure 3D). Additionally, Trp500 of helix H1 interacts with the indole moiety for ~40% of the simulation time and also with the methyl group of the tail region, which acts as an anchor to stabilize the tail of 227G into the cavity (see Discussion). Another stabilizing factor is the π-π interaction between Trp500 and ring2 of the 227G indole moiety. The phenyl ring along with the arms of 227G (Figure. 1c) occupies the cavity that in the apo-enzyme is occupied by Leu31 (Di Marco et al JBC, 2005). Residue Ala396 (H3) makes favorable hydrophobic contacts with the group of 227G. Further, Pro492 interacts with the ring I and thereby enhances the stability of 227G. Finally, an additional stability contribution comes from HpH contacts between 227G and the residue Pro495 and Leu392(H3). We did a comparative study of 227G with previously reported other indole compounds 2CMF and 2DSX, which binds in the same binding site. 227G binds exactly in the same fashion with a low RMSD value of 0.43 Å with CMF and 2DSX (Figure 3C)

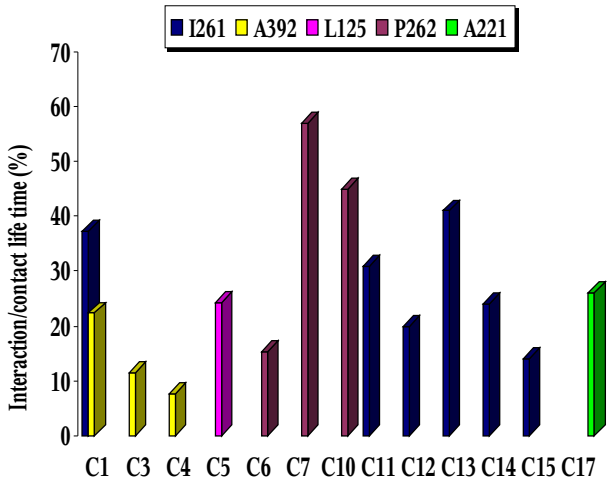
Table 3: Hydrogen Bond Interaction Network:

Inter-Hydrogen Bond (HB)			
Occupancy (%)	**Donar**	**Acceptor**	**ADL*(Á)**
80.0	227G-N2	Asp126-OD	2.3
45.0	227G-N5	Ser533-OG	2.5
12.0	Ala-221	227G-S1	2.8
Water-mediated Hydrogen bonds			
80	Glu128	WAT	2.2
79	227G-N1	WAT	2.3
Intra-Hydrogen Bond (HB)			
82	Asn273-ND2	Asp126:OD	2.1
84	Glu128-N	Asp126:OD	1.9

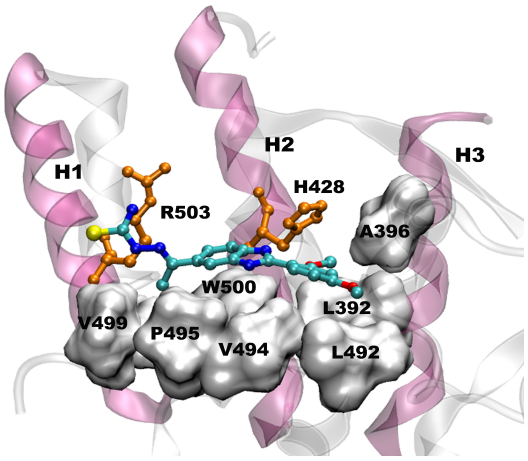
***-**Average Dynamic length



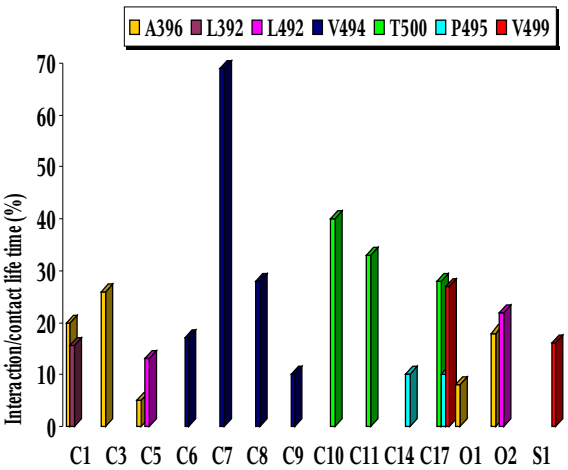
A)



B)



C)



D)

Figure 4: Interaction map of BVDV RdRp and HCV RdRp in complex with 227G. (A). The spiral rings indicate hydrogen bonds (HB) involving residues Asp126, Ser533 and Ala221, as well as a salt-bridge between Asp126 and Asn264. Four loops L_1, L_2, L_3 and L_4 are represented in pink while Motif-I is in blue. Residues contributing in HpH interactions are in represented by VdW surface and are coloured white. (C) the helices (H_1, H_2 and H_3) outlining the cavity are coloured pink. Residues outlining the cavity are shown in VdW surface view as well as indicated by ball and stick representation, (B & D)Statistics of HpH contacts formed between 227G-BVDV 227G-HCV complexes respectively. Labels on the X-axis refer to the atoms of 227G as in Fig. 1c while on the Y-axis the interaction lifetime is reported as fraction of the entire simulation time.

	Residue	Ele	VdW	Total		Residue	Ele	VdW	Total
BVDV	**Asp126**	-23.8(3.5)	-2.1(0.9)	-25.8(3.3)	**HCV**	**Val37**	0.1(0.1)	-0.3(0.2)	-0.2(0.2)
	Arg127	2.2(1.8)	-4.6(1.1)	-2.4(1.9)		**Leu392**	-1.3(0.6)	-1.3(0.6)	-2.6(1.1)
	Glu128	-3.7(2.7)	-4.4(1.3)	-8.5(3.1)		**Ala396**	0.4(0.7)	-1.8(0.6)	-1.4(1.0)
	Arg130	2.3(0.9)	-0.2(0.1)	2.0(0.9)		**Ile424**	-0.8(0.2)	-0.4(0.1)	-1.1(0.2)
	Arg132	8.1(1.7)	-2.2(0.7)	5.9(1.7)		**His428**	-4.1(1.4)	-3.7(0.8)	-7.8(1.5)
	Ala221	-0.8(1.0)	-1.7(0.6)	-2.5(1.1)		**Leu492**	-0.5(0.7)	-1.6(0.8)	-2.1(1.2)
	Phe224	-0.1(0.2)	-0.4(0.3)	-0.5(0.2)		**Val494**	-0.2(0.3)	-2.5(0.6)	-2.7(0.7)
	Ile261	-1.5(0.4)	-4.1(0.8)	-5.6(0.9)		**Pro495**	-0.2(0.4)	-2.3(0.5)	-2.5(0.6)
	Pro262	0.4(0.9)	-5.2(0.7)	-4.8(1.1)		**Val499**	1.2(0.8)	-2.4(0.5)	-1.3(0.9)
	Asn264	1.9(0.9)	-2.8(0.6)	-0.8(1.1)		**Trp500**	0.1(0.4)	-5.9(0.7)	-6.0(0.8)
	Ala392	-0.7(0.6)	-1.5(0.4)	-2.1(0.8)		**Arg503**	-2.6(3.1)	-4.9(0.9)	-7.5(3.5)
	Arg503	2.1(2.0)	-4.1(1.2)	-2.0(2.7)		**Tot**	-11.3(4.6)	-34.0(2.5)	-45.2(5.5)
	Tot	-34.0(6.0)	-47.6(3.5)	-81.7(7.3)					

Table: 4A Residues wise interaction energy of 227G A)- BVDV RdRp-227G B)- HCV RdRp-227G.

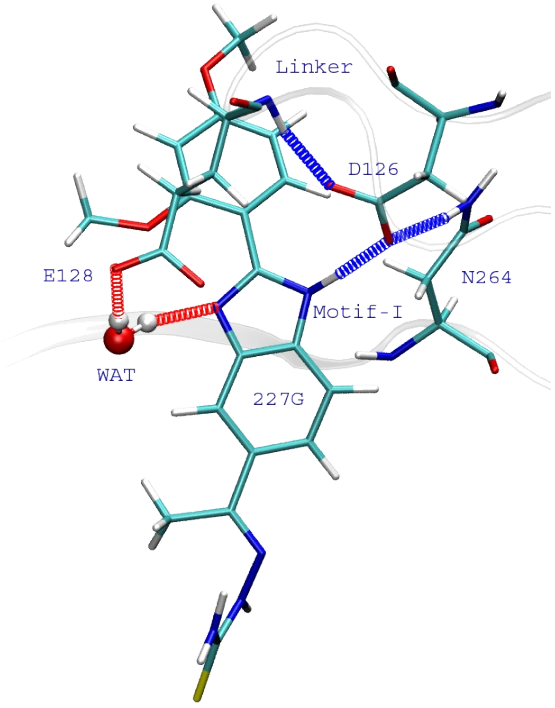


Figure5: Hydrogen Bond Interaction network: Inter (227G-Asp126), Intra (Glu128-Asp126) & (Asp126-Asn264), water mediated interaction (227G:N1-WAT) & (WAT:Glu128), as mentioned in Table3. This network is strong enough having occupancy more than 80 %.

Comparison With APO-Protein:

RMSD-Time wise

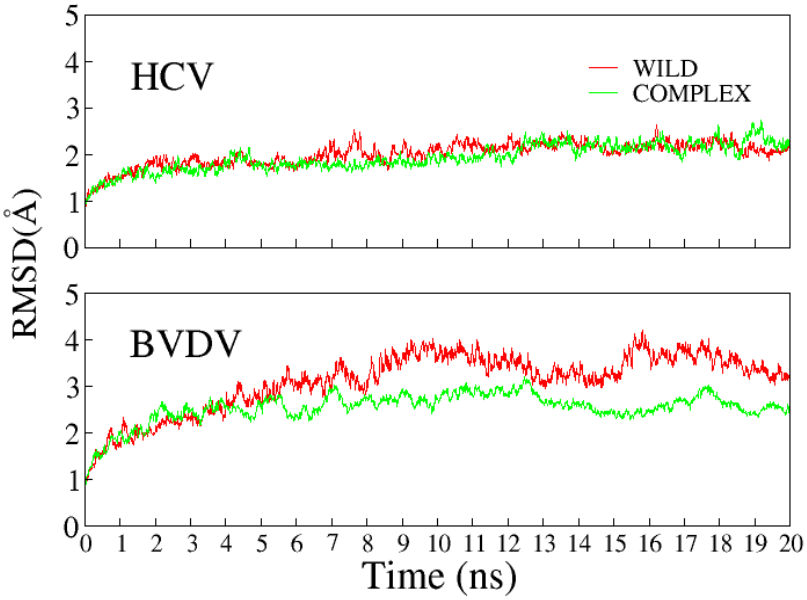


Figure6: RMSD of whole system as a function of time relative to the starting structures in both HCV and BVDV, APO and COMPLEX systems respectively.

We did RMSD analysis as we noticed a substantial fluctuation in BVDV apo enzyme, while the HCV system is quite stable(Fig 6).The root-mean-square deviations (RMSD) of the Cα atoms for the apo and complex of BVDV and HCV RdRp, as function of time are plotted in Figure 6. To take into account completely stabilized systems all the analysis has been carried out in the last 10ns of the simulations.

As depicted in Figure6 the backbone RMSD of Complex 227G-BVDV was found to stabilize with an average of 2.6Å. However, the apo protein showed a larger RMSD, thereby indicating a greater deviation when compared to the complex system. A qualitative examination of the RMSD trajectories from these simulations of the apo and complex system suggests that binding of 227G tends to provide conformational rigidity to the otherwise flexible protein. The dynamics of the apo protein shows a periodic behavior wherein, initially the system moves deviates to an RMSD of 4.0 Å upto 10ns, then it steeps down to 3.0 Å at 12ns . This behavior is repeated when the RMSD again jumps up to 4.0 Å at 16ns and then again coming back to 3.0 Å in the end of the simulation. Figure6 clearly showed that the two systems exhibit different dynamical behavior; with the apo-protein displaying conformational flexibility higher than that of the complex system.

On the contrary, the apo and complex system of HCV RdRp exhibited similar dynamical behavior. After an initial increase of the RMSD of 2.0 Å at around 2ns the Apo system stabilizes around an average value of 1.9 Å throughout the simulation, while the complex deviated around an average value of 1.7 Å exhibiting similarity in dynamical evolution.

RMSD-Residue-Wise

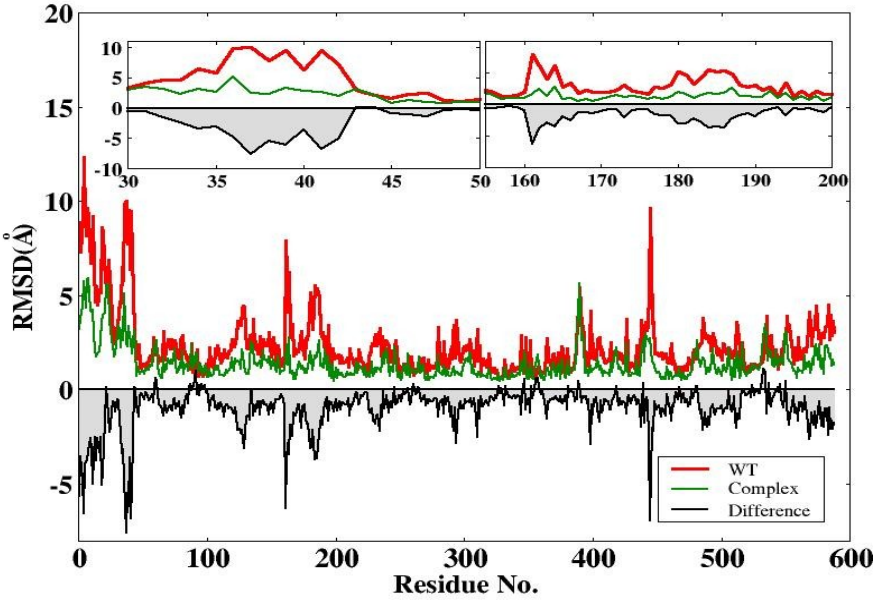


Figure7: Average RMSD during the simulation for each residue with respect to a reference snapshot. We observed huge fluctuations in the wild type as compared to the complex. Maximum peaks were obtained in the N-terminl region, residue surrounding motifI and motifII and loop L3 in Apo which got stabilised on 227G binding .

The relatively high values and the irregular but periodic profile of the Apo-system RMSD plot reflect structural changes of highly flexible protein regions. The root-mean-square deviation,(RMSD) per residue graph, (Figure7) highlight the regions of protein which show significantly larger flexibility in apo protein as compared complex system. Major region of interest; primarily because they make part of the 227G binding cavity; wherein the peaks were observed were the Loop L4 region; also known as the linker; and Loop L3. As already mentioned, Loop L4 and L3 along with L2 and L1 constitute the binding cavity. In the complex (227G-BVDV), 227G interact with the residues, in particular makes strong HB with Asp126, and another residue Glu128 involve in making water-mediated HB with 227G and intra HB with Asp126 and Asn264. This particular networking into the binding site doesn't let linker to move back in complex system thereby keeping it rigid, while in the Apo protein L4 is free to move back and forth. The larger peaks of L4, in the apo-protein can be explained as it is a part of the N-terminal domain, which as described in literature, has an innate flexibilty. This flexibility of the Nterminal domain along with fingertip region is important for the translocation of template during the polymerization. The same hold true for the L3 which interacts with 227G in the complex making HB with Ser533, providing rigidity to this loop, while it is free to move in the apo protein.

Area with respect to APO-protein:

We calculated and compared the area of the cavity mouth in the crystal structure of BVDV RdRp, in the structure of the Apo after relaxation and in the equilibrated configuration of the complex. In order to calculate this area, we considered the triangle defined by the lines connecting the terminal residues of the loops involved in making the cavity (Figure 8C). We were aware that the area calculated in this manner is an approximation and its absolute value has no quantitative meaning. Nevertheless, a comparison of the mouth areas provides a qualitative indication of the probability for a template to enter the polymerization site. We observed that there is complete closure of the template entrance channel due to presence of 227G in the system. A visual depiction of the cavity closure is depicted in Fig.8A & B. In-depth analysis of the area graph showed an enormous reduction in the area enclosed by the loops surrounding the 227G.(Figure.5C &D).

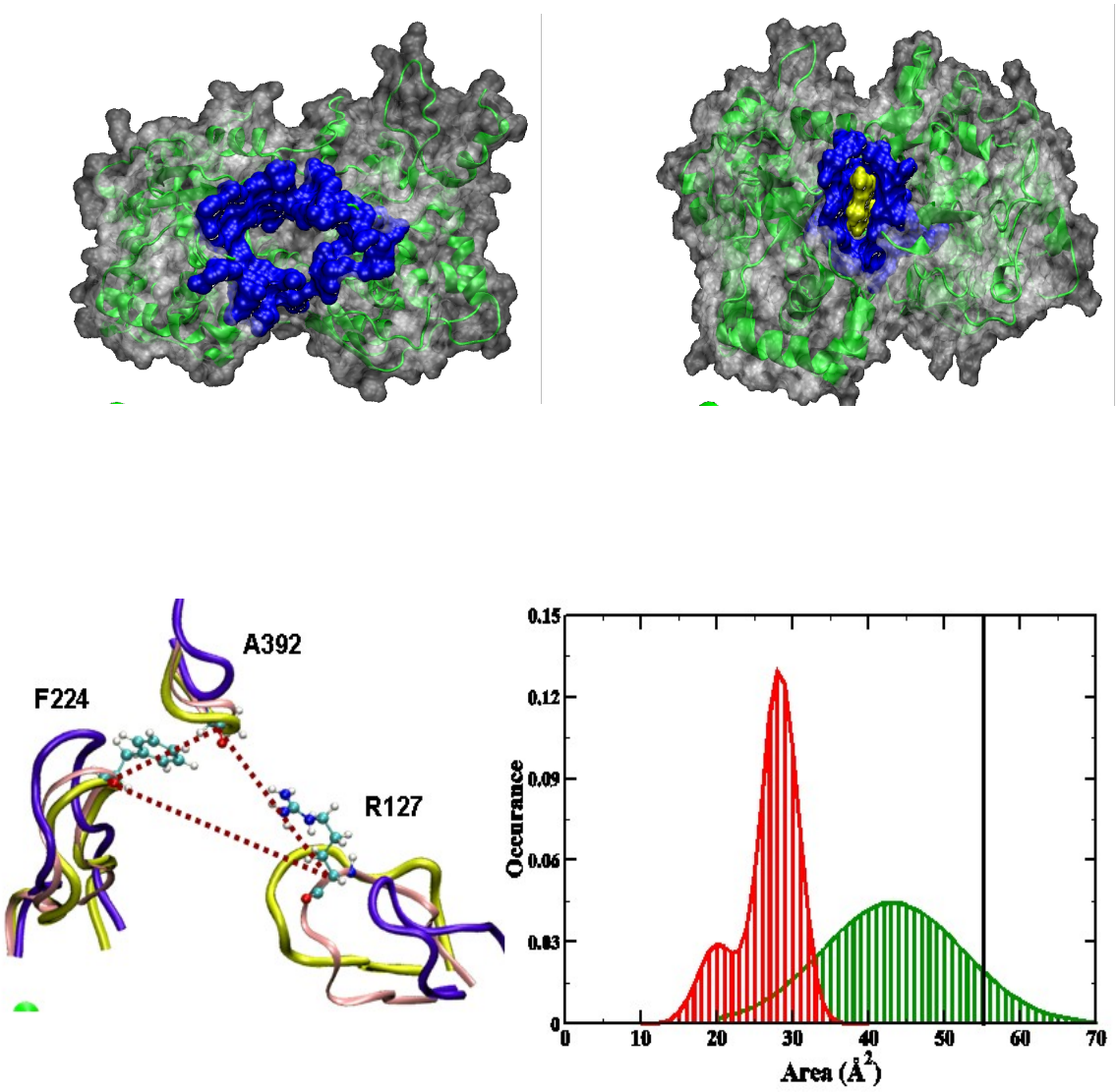


Figure8: Top view of the template entrance channel of BVDV RdRp is indicated by blue and the rest protein is in ribbon and green color. The putative binding cavity encircled by Loops L_1, L_2, L_3 and L_4 lye in the template entrance channel. (A)- The cavity view in crystal structure, indicates enough space for the entrance of template, while in (B)- 227G, indicated in yellow color, totally blocks the entrance to the channel. (C) Depicts the superposition of the averaged structures of the loops in complex (yellow) and Apo (violet) while the crystal is presented in red color. For the calculation of area, we considered L1, L2 and L4 which showed maximum variation between complex and Apo protein. The distances between residues R127(L_4-white), F224 (L_2-yellow) and A392 (L_1-orange) sampled during the simulations were used to calculate the area of the cavity in the complex and in apo-protein.(D) Distribution of the area during the simulations for the Apo-protein is shown by the green curve while the red indicates the complex system. The black line indicates the value extracted from the X-ray structure.

Discussion and Conclusion:

In the framework of our research for antiviral compounds of both synthetic and natural origin, a 2-phenylbenzimidazole derivative, 227G, emerged for its selective activity against a representative of the Pestivirus genus. In particular, being active against BVDV and HCV 1b recombinant RdRp in enzyme-based assays, 227G has RdRp as molecular target in both viruses.

When assessed for its activity in cell-based assays against a panel of DNA and RNA viruses, compound 227G showed an EC_{50} of 0.8 μM against BVDV, whereas proved to be inactive against a representative of the Flavivirus genus, YFV (vaccine strain 17D). Thus, to assess the activity range of the compound we measured its activity against HCV: 227G exhibited anti-HCV activity in the Huh-7 cell-based replicon system, with an EC_{50} value of 1.1 μM. These promising data lack of a microscopic interpretation necessary to possibly extend the activity spectrum of this class of compounds.

To accomplish this challenging task, a computational study represented a powerful tool as it helped us to identify binding site, binding mode, as well as to have a deeper dynamic insight into the mechanism of inhibition of both viruses. Additionally, the possibility of assembling the interaction inventory enables to identify hot-spots on both RdRp and inhibitor sides.

BVDV

The initial structure of BVDV complexed with 227G, which was extracted from the docking procedure, has the 227G inserted in a cavity surrounded by the residues of the fingertip region. As discussed in the results the cavity is made of two loops of finger domain, one from the thumb domain and a loop from the N-terminal domain. The part of N-terminal domain (residue Arg125 to Asn132), involved in interaction with 227G called linker . This N-terminal region is one of the most peculiar characteristics of BVDV RdRp, being absent in other polymerases .

The other significant region which take part in making cavity is finger-tip region, it originate from the finger domain and stack against thumb domain via several beta-strands. .[78-80] There is yet no clear evidence of the role of the N-terminal domain, but it is believed that the finger-tip region along with N-terminal (loop L4) forms the entrance of the template binding channel. The flexibility of the finger-tip region and N-terminus region is an intrinsic property of the polymerase, necessary for the binding and translocation of template during elongation.

Binding of 227G in this region may reduce this flexibility, hampering an efficient translocation, as it can be appreciated in Figure.6. Here we report the average RMSD during the simulation for each residue with respect to a reference snapshot. This reference snapshot is extracted from the MD simulations and refers to the structure with the lowest RMSD, with respect to the average configuration taken over the trajectories. The green and red curves refer to the complex and WT type systems, respectively.

Analysis of the residue wise RMSD indicate a vast difference in the conformational change in the finger-tip region of the WT and complex protein during the simulations. We observed a greater fluctuations of the residues in the finger-tip region of WT protein which stabilises in the complex. This indicates that the binding of 227G provides additional stability to the system.

Another significant fluctuation can be pinpointed in the L4 Linker region of the N-terminal domain. The figure 7 clearly indicates huge fluctuations in the WT protein as compared to the complex protein. With this we conclude that binding of 227G in the putative binding cavity clearly reduces the flexibility of the N-terminal domain as well as the finger-tip region which is pivotal for the concerted movement of the thumb and finger domain necessary for the translocation of template.

Analysing our MD trajectories, we found that the N-terminal domain contributes to stabilize 227G in the putative binding cavity. During MD simulation of the WT, the loops move closer to each other (coloured in pink in Figure 8C), if compared to the situation in the crystal structure (in blue, in Figure 8C). In the presence of the 227G, our MD results indicated that these loops are even closer, thereby resulting into a further closure

of the cavity. As mentioned in the result, in order to quantify this hypothesis we did area analysis of the mouth of the cavity in all three cases (WT, relaxed and complex), we observed there is wide reduction of occurrence of N-terminal in case of complex, while there is a wide range of occurrence in WT type. As a consequence, the binding of the 227G seems to block the entrance of the template channel. Thus, we hypothesised that binding of 227G results in closure of template channel resulting into hindrance to the entry of the template into the polymerization site, thereby preventing polymerization itself.

An alternative explanation of the inhibition action of 227G has been suggested, always based on the position of Ile261.The fingertip region contains polymerase motifs-I and II, which are involved in RNA template and NTP binding .[81,82] The residue I261 is highly conserved and is found to lie in motif-I (261IPKNEK266) (Figure 4A). The motif-I is located close to the NTP (i+1) binding site and plays a role for the binding of the incoming NTP's. Thus, binding of 227G near to Ile261 in the apo-enzyme may inhibit the binding of the incoming NTP's and thereby finally the polymerization. Loop L4, referred to as linker interacts with other viral polymerase thereby concentrating them in the replication complex which could be advantageous for more productive viral replication. Interaction of 227G with the L4 region could reduce the scope of interaction of L4 with other viral polymerase. This in turn may hinder the dimmer formation and affect the polymerisation.

HCV

The mode of inhibition of indole/benzimidazole compounds against HCV RdRp has been put forth by previous crystallographic studies [52], remarkably, it turns out to be substantially different from that found in BVDV. Indeed, accordingly to experimental evidence for others indole/benzimidazole inhibitors [83-84], we found that 227G binds onto the thumb domain of the enzyme in HCV RdRp while we report that 227G docks in the finger domain of BVDV RdRp. (Figure 1A &B)

The analysis of the interaction of 227G with HCV polymerase suggests that inhibition may occur following the same mechanism of inhibition as reported previously , [84]wherein binding of 227G might prevent the interaction between the fingertip loop lambda1 and the thumb domain (Figure 3B), forcing the enzyme in an opened “inactive” conformation .[85-87] This mechanism is substantially different from that hypothesised in the case of BVDV. The role of λ1 loop is to provide the conformational flexibility to the concerted movements of the thumb and fingers domain, which are the essential process in replication [88] . The tip of λ1 loop has an alpha-helix structure, which stacks against the thumb domain and has strong hydrophobic contacts with the surrounding residues, mainly with Val494, Pro495, and Trp500. The idea is that in order to reach the access of the otherwise buried cavity, the compound interacts with the residues Lys30 and Lys31 of the lambda1 loop forcing the loop to open and in turn accommodating its arms into the sub-pockets initially occupied by Lys30 and Lys31.

In an effort to identify features of the inhibition mechanism and interactions hotspots, we performed analysis of the interaction pattern between 227G and HCV. The hydrophobic cavity for the benzimidazole group of compound comprises two small closely spaced pockets (Pocket-1 and 2, Figure 3 [52] . The side chains of Leu392, Ala395, Ile424, Leu425, His428, and Phe429 line in Pocket1, and residues Val37, Leu392, Ala393, Ala396, Leu492, and Val494 form pocket 2 (Figure. 3). In the apo-enzyme residues Leu30 and Leu31 of the alpha-helix of lambda1 loop occupy pocket1 and pocket2, respectively. These interactions are disrupted by the presence of 227G in the complex. Indeed, after 10ns of MD simulation arm1 of 227G (see Figure.1C for notation) fitted comfortably in pocket2 of the protein. Likewise, the phenyl ring and arm2 fully occupy pocket1, which in the apo protein is only partially filled with the side chain of Leu31 (Figure.3A). The arm1 and arm2 contributed to the stability of 227G in the cavity by forming strong hydrophobic contacts. As shown in Figure 3, 227G accommodates very nicely into the cavity since the O-methyl groups of 227G protrude into the pocket more deeply than the residues located in the tip of the fingertip. Furthermore, there is another hydrophobic region in the binding cavity where the methyl group of 227G tail anchors the 227G by establishing strong interaction with Trp500 and Val499, thereby providing further stability. The benzimidazole moiety of 227G additionally contributes to the stabilization of 227G in the pocket by making strong hydrophobic contacts with the side chains of Val494, Trp500, and Pro495.The mode of interaction of 227G obtained in our study was found to be in accordance with that suggested on the basis of the crystal structures . Thus, it is possible to conclude that binding of 227G on the surface of the thumb domain, in a cavity that in the free enzyme is normally occupied by one of the fingertips, disrupts the fingertips-thumb interactions and forces the enzyme into a open, inactive conformation.

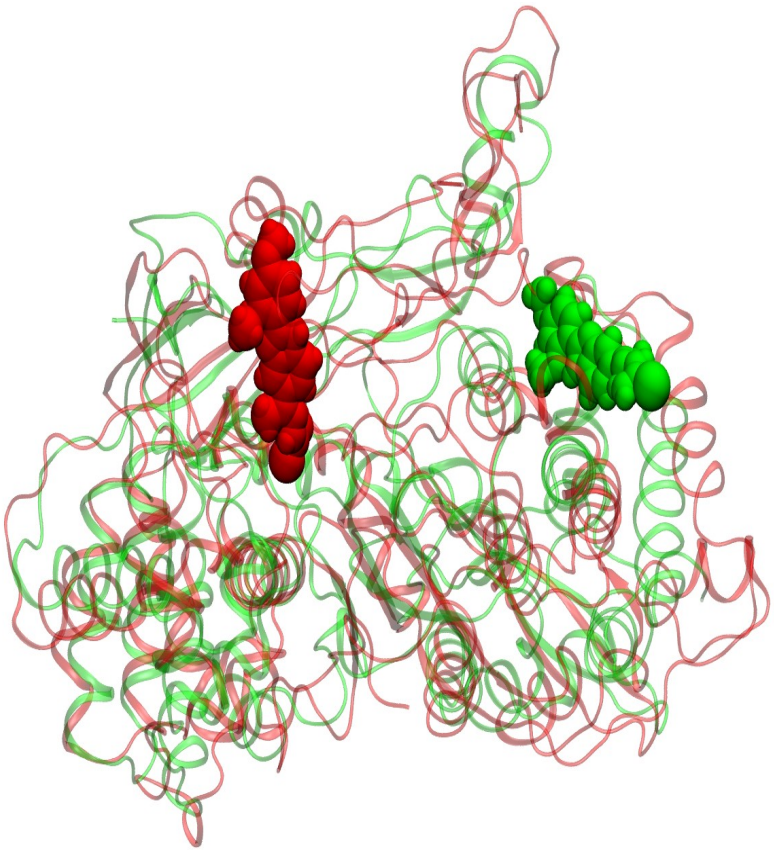


Figure 9: Superposition of HCV and BVDV RdRp complexes, taking their average pdb from their respective simulation in each cases. HCV RdRp is in green in color while BVDV RdRp is red in color. 227G in each case is rendered in VDW and lies in thumb and finger domain of HCV and BVDV system respectively.

Conclusion:

1)-The present study contributes to the understanding on how the polymerases of flaviviridae can be targeted by specific inhibitors. At the best of our knowledge, this is the first time that a benzimidazole derivative is reported to be active against both BVDV and HCV RdRp in the low micromolar range.

2)-Despite 227G could supplement the series of NNI's discovered by [89-90] endowed with the same dual antiviral activity, it is our opinion that BVDV could be considered a predictable surrogate model for HCV only in the case of NIs, since they bind to an evolutionarily conserved portion of the enzyme, which remained unchanged in the flaviviridae RdRp and in other viral polymerases.

3)-Here we defined the inventory of interactions established between 227G and HCV and BVDV RdRp. This was done with an aim to apply a structure based strategy to pinpoint possible changes in the compounds aiming at improving its potency.

4)-In this paper we proposed the probable mechanism of inhibition of 227G in BVDV RdRp , which are all different from the one reported for benzimidazole compounds against HCV RdRp. We showed that the mechanism of inhibition involves different domain in the two RdRps, in BVDV it bind in finger domain while in case of HCV in thumb domain (figure 9), thus weakening the validity of the use of BVDV as a surrogate model of HCV. Though sometimes activity against BVDV revealed a positive indication for HCV, as in the case if 2,5-disubstituted imidazo[4,5-c]pyridines, indeed, in our opinion it is not possible to exclude that compounds inactive or low active against BVDV can be surprisingly active against HCV. Despite the promising results obtained with our benzimidazole derivative, we could not disprove the concept of a limited spectrum of activity for the antiviral therapy.

All these issues are challenging and require a microscopically well-funded knowledge of the processes involved in the inhibition. About the validity of BVDV as a surrogate model, the strengths and weaknesses of its use are pinpointed. At the same time, rules to identify the range of validity of antiviral therapy might be traced back from our microscopic picture of the compound-target interaction. About the second issue, our hope is that drawing the complete inventory of the interactions and mapping such hot-spots on the structure of both targets and compound, a crucial information that together with experimental data contribute to the design of improved inhibitors. Hopefully, these forthcoming results will contribute to deciphering the molecular determinants of inhibitor activity and will provide the key to understand the molecular mechanism of inhibition. The final outcome will be a major understanding on how the Flaviviridae RdRp could be targeted by specific inhibitors, thus allowing a rational (in silico) drug design.

References

1. World Health Organization. 1998. WHO concerns on hepatitis C. Lancet 351:1415.
2. Lindberg A. (2003) Bovine viral diarrhoea virus infections and its control. A review.. Vet Q. 25(1):1-16..
3. Esteban JI, Sauleda S, Quer J. The changing epidemiology of hepatitis C virus infection in Europe. J Hepatol. 2008;48:148–162
4. Geiss BJ, Stahla H, Hannah AM, Gari HH, and Keenan SM Focus on Flaviviruses: Current and Future Drug Targets Future Medicinal Chemistry, May 2009, Vol. 1, No. 2, Pages 327-344.
5. QL Choo, G Kuo, AJ Weiner, LR Overby, DW Bradley, and M Houghton. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome.21 April 1989: 359-36221 April 1989: 359-362
6. Miyamura T, Saito I, Katayama T, Kikuchi S, Tateda A, Houghton M, Choo QL, Kuo G. Detection of antibody against antigen expressed by molecularly cloned hepatitis C virus cDNA: application to diagnosis and blood screening for posttransfusion hepatitis. Proc Natl Acad Sci U S A. 1990 Feb;87(3):983–987.
7. Cohen J . V. The scientific challenge of hepatitis C. Science. 1999 Jul 2;285(5424):26-30.
8. De Francesco R, Carfi A. Advances in the development of new therapeutic agents targeting the NS3-4A serine protease or the NS5B RNA-dependent RNA polymerase of the hepatitis C virus. Adv Drug Deliv Rev. 2007;59:1242–1262.
9. Thomas, D.L., Seeff, L.B.: Natural history of hepatitis C. Clin. Liver. Dis. 9 (2005), 383–398
10. Shi ST, et al. Preclinical characterization of PF-00868554, a potent nonnucleoside inhibitor of the hepatitis C virus RNA-dependent RNA polymerase. Antimicrob Agents Chemother 2009;53:2544-2552.Shi ST, et al. Preclinical characterization of PF-00868554, a potent nonnucleoside inhibitor of the hepatitis C virus RNA-dependent RNA polymerase. Antimicrob Agents Chemother 2009;53:2544-2552.
11. WORLD JOURNAL OF GASTROENTEROLOGY 13, 5673, 2007
12. Magdalena Sarasin-Filipowicz, Markus H Heim.Interferon-induced gene expression in chronic hepatitis C. Future Virology 2010 5:1, 25-31
13. Castet, V., C. Fournier, et al. (2002). "Alpha Interferon Inhibits Hepatitis C Virus Replication in Primary Human Hepatocytes Infected In Vitro." J. Virol. 76(16): 8189-8199.
14. Hoofnagle, J. H. and L. B. Seeff (2006). "Peginterferon and Ribavirin for Chronic Hepatitis C." New England Journal of Medicine 355(23): 2444-2451
15. Lauer, G. M., and B. D. Walker. 2001.N. Engl. J. Med. 345:41–52
16. Cao F. et. al, J. Viral. Hepat, 2010
17. Licia Tomei et al; Antiviral Chemistry & Chemotherapy 16:225–245, 2005.
18. Bernd et al, Best practice and Research clinical gastroenterology,2008, and annals of Hepatology; 2009, review.
19. AMERICAN JOURNAL OF GASTROENTEROLOGY Volume: 105 Issue: 5 Pages: 989-1004 Published: 2010
20. De Clercq, E. (2004). Non-nucleoside reverse transcriptase inhibitors (NNRTIs): past, present, and future. Chem. Biodivers. 1, 44–64.
21. Hilde Azijn,1 Ilse Tirry,1 Johan Vingerhoets,1 Marie-Pierre de Béthune,1 Guenter Kraus,1 Katia Boven,2 Dirk Jochmans,1 Elke Van Craenenbroeck,1 Gaston Picchio,2 and Laurence T. Rimsky; TMC278, a Next-Generation Nonnucleoside Reverse Transcriptase Inhibitor (NNRTI), Active against Wild-Type and NNRTI-Resistant HIV-1 ,Antimicrobial Agents and Chemotherapy, February 2010, p. 718-727, Vol. 54,
22. Lin, C. et al , 2005, J. Biol. Chem. 280: 36784–36791
23. Lin, C. et al, 2004. J. Biol. Chem. 279:17508–17514
24. De Francesco, R., Migliaccio, G., 2005. Challenges and successes in developing new therapies for hepatitis C. Nature 436, 953–960.
25. EXPERT OPINION ON INVESTIGATIONAL DRUGSVolume: 17Issue: 3Pages: 303-319 Published: MAR 2008
26. The Journal of Infectious Diseases 2010;202:65–74; BEST PRACTICE & RESEARCH IN CLINICAL GASTROENTEROLOGYVolume: 22 Issue: 6Pages: 1123-1136 Published: 2008
27. WORLD JOURNAL OF GASTROENTEROLOGY Volume: 13 Issue: 43 Pages: 5673-5681 Published: NOV 21 2007

28. Expert Opin Investig Drugs Volume: 19 Issue: 8 Pages: 963-75 Published: 2010
29. BEST PRACTICE & RESEARCH IN CLINICAL GASTROENTEROLOGY Volume: 22 Issue: 6 Pages: 1123-1136 Published: 2008; Bernd et al, Best practice and Research clinical gastroenterology,2008, and annals of Hepatology; 2009
30. Tonelli et al, Antiviral activity of benzimidazole derivatives. II. Antiviral activity of 2-phenylbenzimidazole derivatives Bioorganic & Medicinal Chemistry, Volume 18, Issue 8, April 2010, Pages 29
31. Akhtar, T et al, ACTA PHARMACEUTICA,58, 2,135-149,2008
32. Coutard,B et al. ANTIVIRAL RESEARCH,Volume:78,Issue:1,Pages:37-46,2008
33. Mazzei, M et al, BIOORGANIC & MEDICINAL CHEMISTRY,16,5,2591-2605,2008
34. Nathaniel A Brown, Expert Opinion on Investigational Drugs, June 2009, Vol. 18, No. 6 : Pages 709-725;
35. Expert Opin Investig Drugs Volume: 19 Issue: 8 Pages: 963-75 Published: 2010 Aug.
36. Wayne H. Miller et al. Biochemical Pharmacology Volume 31, Issue 23, 1 December 1982, Pages 3879-3884
37. Y C Cheng et. al., Antimicrob Agents Chemother. 1981 September; 20(3): 420–423
38. M C Starnes , Gao,W,Ting, R.Y.C., and Cheng, Y-C(1998), JBC,263, 5132-
39. Ferrer-Orta C, et al. Current Opinion in Structural Biology, 2006
40. Edwin H. Rydberg et al; JMB, 2009
41. Licia Tomei, J.Viro, Dec. 2003, p. 13225–13231
42. Hirashima S et al; 2006 J Med Chem. 2006 Jul 27;49(15):4721-36
43. Tonelli M et al; Bioorg Med Chem. 2010 Apr 15;18(8):2937-53.
44. Göker H et al, Farmaco. 1998 Jan;51(1):53-8.
45. Buckwold VE et al.Antiviral Research, 2003; 60: 1-15
46. Oriana Tabarrini et al, J. Med. Chem. 2006, 49, 2621-2627
47. Stephen D. Woodhouse et al; Antimicrob Agents Chemother. 2008 May;52(5):1820-8.
48. Wallingford CT.; Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA Jr, Vreven T, Kudin KN, Burant JC, Millam JM, GAUSSIAN. (2003) Pittsburgh PA: Gaussian, Inc]
49. Salim MT et al; Antivir Chem Chemother. 2010 Apr 14;20(5):193-200
50. Benzaria S, et al. Antiviral Chemistry & Chemotherapy, 2007; 18: 225-242
51. Neyts et al, A Novel, Highly Selective Inhibitor of Pestivirus Replication That Targets the Viral RNA-Dependent RNA Polymerase;2006,p. 149-160, Vol. 80, No. 1
52. Stefania Di Marco et al;Interdomain Communication in Hepatitis C Virus Polymerase Abolished by Small Molecule Inhibitors Bound to a Novel Allosteric Site 2005; JBC
53. Ikegashira, K., Oka, T., Hirashima, S., Noji, S., Yamanaka, H., Hara, Y., Adachi, T., Tsuruha, J., Doi, S., Hase, Y., Noguchi, T., Ando, I., Ogura, N., Ikeda, S., Hashimoto, H., 2006. Discovery of conformationally constrained tetracyclic compounds as potent hepatitis C virus NS5B RNA polymerase inhibitors. J. Med. Chem. 49, 6950–6953.
54. Edwin H. Rydberg et al;Structural Basis for Resistance of the Genotype 2b Hepatitis C Virus NS5B Polymerase to Site A Non-Nucleoside Inhibitors 2009; JMB
55. Loopy.(http://trantor.bioc.columbia.edu/programs/jackal)
56. R. Huey, G.M. Morris, A.J. Olson and D.S. Goodsell, J. Comput. Chem. 28 (2007), p. 1145.
57. Morris GM, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ
58. Tonelli, M.; Boido, V.; La Colla, P.; Loddo, R.; Posocco, P.; Paneni, M. S.;Fermeglia, M.; Pricl, S. Bioorg. Med. Chem. 2010, 18, 2304
59. Gabriele Gilibertia, Cristina Ibba, Esther Marongiua, Roberta Loddoa, Michele Tonellib, Vito Boidob, Erik Laurinic, Paola Posoccoc, Maurizio Fermegliac and Sabrina Priclc , Bioorg. Med. Chem. 2010
60. Hetenyi et al; 2002; protein science
61. F.J. Solis and R.J.-B. Wets. (1981) "Minimization by random search techniques", Mathematical Operations Research, 6, 19-30.
62. Carlsson et al; J. Med Chem; 2008
63. Aaqvist J. Ion-water interaction potentials derived from free energy perturbation simulations. J. Phys. Chem. (1990) 94:8021–8024;
64. Aaqvist J. et al. Phys. Chem. B, 2008, 112 (30), pp 9020–9041

65. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JJA, Vreven T, Kudin KN, Burant JC, et al, Revision C.02 ed. Gaussian, Inc. (2004)
66. Wallingford CT.; Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA Jr, Vreven T, Kudin KN, Burant JC, Millam JM, GAUSSIAN. (2003) Pittsburgh PA: Gaussian, Inc]
67. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. J. Comp. Chem. (2004) 25:1157–1174.
68. Bayly CI, Cieplak P, Cornell W, Kollman PA. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. J. Phys. Chem. (1993) 97:10269–10280
69. ORAC
70. Andersen HC. Molecular dynamics simulations at constant pressure and/or temperature. J. Chem. Phys. (1980) 72:2384–2393
71. Parrinello M, Rahman A. Polymorphic transitions in single crystals: a new molecular dynamics method. J. Appl. Phys. (1981) 52:7182–7190)
72. Nosé S. A molecular dynamics method for simulations in the canonical ensemble. Mol. Phys. (1984) 52:255–268
73. Hoover WG. Canonical dynamics: equilibrium phase-space distributions. Phys. Rev. A (1985) 31:1695–1697
74. Kumar et al; J Phys. Chem; 2010
75. Cheatham 3rd, T. E., P. Cieplak, and P. A. Kollman. 1999. A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. J. Biomol. Struct. Dyn. 16:845–862.
76. Kyung H. Choi et al; 2004 PNAS
77. [Paeshuyse 2006] Paeshuyse J, Leyssen P, Mabery E, Boddeker N, Vrancken R et al. (2006) A novel, highly selective inhibitor of pestivirus replication that targets the viral RNA-dependent RNA polymerase. J Virol 80: 149-60
78. Bar-Nahum G, Epshtein V, Ruckenstein AE, Rafikov R, Mustaev A, Nudler E. A ratchet mechanism of transcription elongation and its control. Cell. 120:183-93. 2005
79. Guo Q, Sousa R (2006) Translocation by T7 RNA polymerase: a sensitively poised Brownian ratchet. J Mol Biol
80. Yin YW, Steitz TA (2004) The structural mechanism of translocation and helicase activity in T7 RNA polymerase. Cell 116: 393–404
81. Koonin, E. V. (1991) J. Gen. Virol. 72, 2197–2206.
82. O'Farrell et al; Substrate complexes of hepatatis C virus RNA Polymerase(HC-J4)., 2003
83. Tomei, L., Altamura, S., Bartholomew, L., Biroccio, A., Ceccacci, A., Pacini, L., Narjes, F., Gennari, N., Bisbocci, M., Incitti, I., Orsatti, L., Harper, S., Stansfield, I., Rowley, M., De Francesco, R., Migliaccio, G., 2003. Mechanism of action and antiviral activity of benzimidazole-based allosteric inhibitors of the hepatitis C virus RNA-dependent RNA polymerase. J. Virol. 77, 13225–13231.
84. Ikegashira, K., Oka, T., Hirashima, S., Noji, S., Yamanaka, H., Hara, Y., Adachi, T., Tsuruha, J., Doi, S., Hase, Y., Noguchi, T., Ando, I., Ogura, N., Ikeda, S., Hashimoto, H., 2006. Discovery of conformationally constrained tetracyclic compounds as potent hepatitis C virus NS5B RNA polymerase inhibitors. J. Med. Chem. 49, 6950–6953.
85. Meitian Wang et al ; J. Biol Chem 2002.
86. Biswal, B.K., Wang, M., Cherney, M.M., Chan, L., Yannopoulos, C.G., Bilimoria, D.,Bedard, J., James, M.N., 2006. Non-nucleoside inhibitors binding to hepatitis C virus NS5B polymerase reveal a novel mechanism of inhibition. J. Mol. Biol. 361, 33–45.
87. George Kukolj et al ; Hepatitis C Virus Subgenomic Replicons in the Human Embryonic Kidney 293 Cell Line Journal of Virology, January 2004, p. 491-501, Vol. 78
88. Sreedhar Chinnaswamy et al , Conformations of the monomeric hepatitis C virus RNA-dependent RNA polymeraseJ. Biol Chem ; 2010:2 21–39
89. Neyts et al, The Predominant Mechanism by Which Ribavirin Exerts Its Antiviral Activity In Vitro against Flaviviruses and Paramyxoviruses Is Mediated by Inhibition of IMP Dehydrogenase J. Virol., February 1, 2005; 79(3): 1943 - 1947.; J.virol;2005
90. Neyts et al, A Novel, Highly Selective Inhibitor of Pestivirus Replication That Targets the Viral RNA-Dependent RNA Polymerase;2006,p. 149-160, Vol. 80, No. 1

Chapter 4

Quinoline Tricyclic Derivatives.

Design, Synthesis and Preliminary *In vitro* and *In silico* Antiviral Activity Against Flaviviridae Family of Three New Classes of

Virus-Encoded RNA-Dependent RNA Polymerase (RdRp) Inhibitors

Abstract

In this preliminary study three new classes of RNA-dependent RNA polymerase (RdRp) of Flaviviridae inhibitors, the linear *N*-tricyclic systems derived by condensation of the quinoline nucleus with 1,2,3-triazole, imidazole or pyrazine (obtaining triazolo[4,5-*g*]quinolines, imidazo[4,5-*g*]quinolines and pyrido[2,3-*g*]quinoxalines respectively), has been discovered. In particular the activity of the title compounds were evaluated in cell culture systems against YFV (as viruses representative of *Flaviviruses*) and BVDV (*Pestiviruses*) . Furthermore the cytotoxicity was evaluated in parallel cell-based assays. 1:1 Mixture of bis-triazoloquinolines (**1m**), imidazoquinolines (**2e**,**h**) and pyridoquinoxalines (**4h**,**j** and **5n**) showed anti-BVDV activity in the range 1-5 μM in *cell-based* assays. Mutation experiments versus both NS5b and NS3 enzymes of resistant strains of BVDV of two selected compounds (**2h** and **5m**) confirmed the inhibition of the enzyme target. Overall, by these experiments the imidazo[4,5-*g*]quinoline (**2h**) emerged as a potent BVDV polymerase inhibitor endowed with IC_{50} = of 0.06 μM. Herein, we have utilized several biophysical computational approaches, ranging from Docking, standard molecular dynamics and metadynamics , in order to predict the binding site, to identify key determinants of ligand binding, the energetics of unbinding as well as the escape mechanism of the lead compounds PS999 and PS1036 in BVDV RdRp. MM/PBSA techniques were used to calculate the binding free energies of 227G in both HCV and BVDV RdRp. Calculated binding free energies were then converted to computational IC50 which were then compared with the experimentally available IC_{50} values

Keywords: Triazolo[4,5-*g*]quinolines / Imidazo[4,5-*g*]quinolines / Pyrido[2,3-*g*]quinoxalines / Anti-viral activity / Flaviviridae / BVDV / YFV / HCV / Cytotoxicity / SAR / RNA-dependent RNA polymerase / *In silico* evaluation / Mutations in NS5b.

Introduction

The Flaviviridae family, containing viruses with single-stranded positive-sense RNA genomes (ssRNA$^+$), comprises three genera and several viruses that are currently unassigned to specific genera. The *Hepacivirus* genus includes the hepatitis C virus (HCV). Viruses such as GB virus-A and GB virus-A-like agents, GB virus-D and GBV-C or hepatitis G virus, while at present not formally classified within the Hepacivirus genus, are closely related to HCV and represent unassigned members of Flaviviridae. Within this family are also comprised the *Flavivirus* genus, with viruses such as Dengue Fever, Yellow Fever (YFV), West Nile, Japanese encephalitis and tick-borne encephalitis, and the *Pestivirus* genus, which includes bovine viral diarrhea (BVDV), border disease and classical swine fever viruses. Viruses within the Flaviviridae family cause significant diseases in human and animal populations.

HCV is a major cause of human hepatitis, globally.[1] The World Health Organization (WHO) estimates that over 170 million people worldwide are presently infected with this virus.[2,3] Most infections become persistent, and about 60% of cases progress towards chronic liver disease. Chronic HCV infection can lead to development of cirrhosis, hepatocellular carcinoma and liver failure.[4,5] Pegylated interferon in combination with ribavirin is used in the clinic for hepatitis due to HCV. Unfortunately this therapy has limited efficacy and is often associated with severe and adverse events.[6]

Flaviviruses are important human pathogens and are also prevalent throughout the world and cause a range of acute febrile illness and encephalitic and hemorrhagic diseases. Although an effective vaccine against YFV has been available since the late 1930s, utilization is incomplete in many areas.[7]

Pestivirus infections of domesticated livestock cause significant economic losses worldwide.These viruses cause a range of clinical manifestations including abortion, teratogenesis, respiratory problems, chronic wasting disease, immune system dysfunction and predisposition to secondary viral and bacterial infections. BVDV can also establish a persistent infection (PI) in animals that remain viremic throughout life and serve as continuous virus reservoirs.[8] Persistently infected animals often succumb to fatal mucosal disease. Furthermore, BVDV shows the ability to cross the placenta of susceptible animals causing a variety of fetal infections.[9]

In conclusion currently, with the exception of YFV, no vaccine exists against the various Flaviviridae members; furthermore there are no antiviral pharmaceuticals to prevent or treat Flavivirus or Pestivirus infections and the association of interferon and ribavirin has limited efficacy in HCV therapy. Therefore new drugs for therapy and chemoprophylaxis are clearly needed for infections and diseases caused by viruses belonging to the Flaviviridae family.

Recently we reported the synthesis of the [4,7]phenantroline nucleus, an angular *N*-tricyclic systems derived by expansion of the quinoline nucleus with pyridine, and of several related derivatives which emerged

as a new class of antiviral agents endowed with both *in silic*o and *in vitro* activity against viruses ssRNA^{+}.[10] Molecular modeling of the interactions between [4,7]phenantrolines and the RNA-dependent RNA polymerase (RdRp), termed NS5B in the case of the Hepaciviruses and Pestiviruses, confirmed the supposed activity against HCV NS5B of this angular *N*-tricyclic system. This research is inserted in a large program of design, synthesis and both *in silic*o and *in vitro* screening of new classes of non-nucleoside anti-Flaviviridae agents, in particular against HCV.[11]

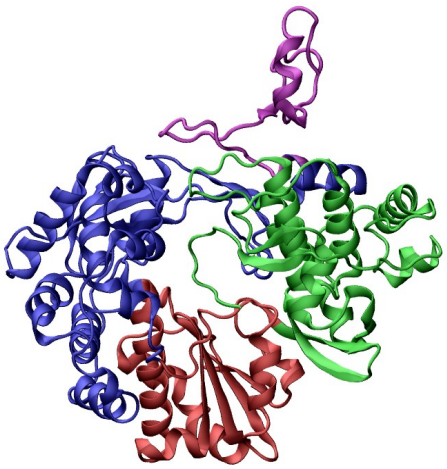


Figure1: Cartoon representation of BVDV RdRp. Different domains of RdRp viz thumb in green, finger in blue, palm in red and unique N-terminal is magenta in color.

Here we report the synthesis, *in vitro* antiviral screening, citotoxicity studies and molecular modeling of three series of linear *N*-tricyclic systems, derived from the quinoline ring, against representative Flaviviridae viruses. In particular we have studied triazolo[4,5-*g*]quinolines, imidazo[4,5-*g*]quinolines and pyrido[2,3-*g*]quinoxalines. Unsubstituted nucleuses of the three above quinoline derivatives were by some of us prepared since 2000,[12] and some their derivatives were subsequently synthesized and tested in anticancer, antifungal and antibacterial screenings.[13-14] In the aim to verify if also these linear *N*-tricyclic systems were able to protect the cells from the cytopathogenicity induced by Flaviviridae viruses we have re-synthesized and submitted the unsubstituted nucleuses and their known derivatives of Figures 1-3 (compounds **1a-j**, **2a-c**, **3a-d**, **4g-i** and **5k-l**) to antiflaviviridae screening. In particular we evaluated both *in silic*o and *in vitro* activity against YFV (as viruses representative of *Flaviviruses*) and BVDV (*Pestiviruses*). Due to the lack of efficient cell culture systems for the multiplication of HCV (*Hepacivirus*), title compounds were evaluated only *in silico*. Furthermore the citotoxicity was evaluated in parallel cell-based assays. On the light of the interesting results obtained we have designed the syntheses of new inhibitors of the RdRp of Flaviviridae bearing at this class of linear *N*-tricyclic systems **1k-m**, **2d-q**, **3d-f**, **4j** and **5m-o** of Figures 1-3. All the new compounds were submitted to the same assays above indicated. Furthermore two selected compounds (**2h** and **5m**) were submitted to mutation experiments versus NS5b (RdRp) and NS3 (helicase) of resistant strains of BVDV, in the aim to confirm the enzyme target and to verify the insurgence of resistance against these compounds.

Despite the interesting properties of these compounds, nothing is currently known about their molecular mechanism of interaction with virus polymerases. In an effort to shed light thereon, we performed here a thorough in silico study of the interaction of two lead compounds (namely PS999 of Imidazoquinoline and PS1036 of pyridoxoquinoline) with the BVDV RdRp.

This choice has been driven by the availability of data on the selection of BVDV RdRp resistant mutants for these two drugs. These data are not available at the moment for the third class of compounds, so we limit our study to the two compounds above. We used a thorough application of different computational approaches, ranging from docking, standard molecular simulations to characterize the structural and dynamical determinants of the interaction, meta dynamics to predict the energetics of unbinding of the compound PS999 and PS1036 from the putative binding and to estimate undocking pathway and PoissonÐBoltzmann surface area (MM/PBSA) calculations to provide an estimate for the free energy of binding (that can be compared with experiments to validate the binding site and therefore the models of the complexes).

NOTE: Results of only the computational part is presented in this chapter.

Material and Methods

MM-PBSA Calculation.

All energetic analyses were done using a single trajectory approach, where each large ribosomal subunit-inhibitor complex, large ribosomal subunit, and inhibitor snapshots were taken from the snapshot of the performed MD trajectory. According to the MM-PBSA method[15,16] binding free energy (ΔG_{bind}) of each system could be conceptually summarized as follows

$$\Delta G_{bind} = G_{com} - G_{rec} - G_{lig} \quad (1)$$

$$\Delta G = \Delta E_{MM} + \Delta G_{solv} - T\Delta S \quad (2)$$

In which

$$\Delta E_{MM} = \Delta E_{bond} + \Delta E_{angle} + \Delta E_{torsion} + \Delta E_{vdw} + \Delta E_{EEL} \quad (3)$$

$$\Delta G_{solv} = \Delta G_{PB} + \Delta G_{SA} \quad (4)$$

$$\Delta G_{SA} = \gamma SA + b \quad (5)$$

where G_{com}, G_{rec}, and G_{lig} are the free energy for the complex, receptor, and ligand (inhibitor), respectively. Each term is calculated by averaging the energy of molecular mechanics (ΔE_{MM}), the solvation free energy (ΔG_{solv}), and the vibrational entropy term ($T\Delta S$) as in (2). ΔE_{MM} (3) denotes the average molecular mechanics energy contributed by bonded (Ebond, Eangle, and Etorsion) and nonbonded (E_{vdw} and E_{EEL}) terms, and the individual nonbonded contribution of the binding pocket to inhibitor was further decomposed on a residue basis using MM-PBSA[17,18,19,20,21].

ΔG_{solv} (4) is the solvation free energy given by ΔG_{PB}, polar solvation free energy evaluated using the Poisson-Boltzmann equation, and ΔG_{SA}, nonpolar contribution to solvation free energy from the surface area.29 The electrostatic solvation free energy was calculated using DELPHI[22,23] software, with low dielectric medium for solute (ε=1) and high dielectric medium for solvent (ε=80). Atomic radii were taken from PARSE29 with an additional value of 1.90 Å for phosphorus,[24] and in order to be consistent with molecular mechanics energy calculation, the partial charges on solute were taken from the Amber 99 force field[21] and from our ab initio calculation (see above) for inhibitors. An 80% boxfill lattice with grid spacing 0.5 grid/Å was applied, and 10,000 linear iteration steps were required to obtain energy convergence. A nonpolar contribution to solvation free energy was determined from equation (23), where the surface area was calculated using molsurf from AMBER 8, and γ and b are 0.00542 kcal/mol Å2 and 0.92 kcal/mol, respectively, for use with PARSE atomic radii.29 The solvent probe radius was set to 1.4 Å. Residues within 25 Å from the mass center of ligand were used for PBSA calculation (this included the ligand with the exception of water molecules [25]

The solute entropy contribution (-TΔS) was estimated by normal-mode analysis18 using the NMODE module in AMBER 8.0. Each of the complex, receptor, and ligand systems were first minimized before switching over for normal-mode analysis, using a distance dependent dielectric constant of ε = 4r (r)interatomic distance), in the absence of solvent to mimic solvent screening33 until the convergence criteria of 0.0001 kcal/mol Å was achieved. Normal mode calculation was extremely time-consuming and computationally expensive for a large system; therefore, only residues within 10 Å from the mass center of inhibitor (excluding water molecules and ions) were used here. As discussed in refs 34 and 35 the differences of the calculated entropy value is quite small for different conformations, and normal-mode analysis calculation usually gave an inaccurate estimation for solute entropies, thus giving merely qualitative estimates of the solute entropy.36 Therefore, our calculation was only based on the average entropy value obtained from 25 snapshots taken from the final 2000 ps MD trajectories with a time interval of 20 ps.

Simulation Methodology (Metadynamics)

Initial relaxation of the system with the PS999 and PS1036 was performed for 20 ns at constant volume and temperature using the MD software package ORAC.[26] The process of undocking occurs on a timescale (~200 μs)21 which cannot be reached by standard MD simulations with an all-atom representation. To overcome this problem we used the metadynamics algorithm.[27] This algorithm, based on a history dependent biasing potential added in a subspace defined by a chosen set of reaction coordinates $s_\alpha(x)$ is aimed at reconstructing the multidimensional free energy of a given process. At time t the biasing potential V_G is given as the sum of repulsive Gaussian functions added with a frequency $1/\tau G$

$$V_G(s(x), t)) = W \sum_{\substack{t'=\tau G,2\tau G,3\tau G... \\ t'<t}} \exp(-(s(x) - sG(t'))^2/2(\delta s)^2) \qquad (1)$$

where W is the Gaussian height, δs is the Gaussian width. Due to this potential, the system is discouraged from revisiting the configurations already sampled. Metadynamics, not only allows the acceleration of rare events, but also the reconstruction of the free energy $F_G(s, t) = -V_G$, which is an approximation of F(s) in the region Σ(s) explored by $s(x_G(t))$ up to time t.22 The accuracy of free energy reconstruction is dependent upon the Gaussian parameters W, δs and τ_G. Details of the metadynamics algorithm have been previously described.[27]

Choice of Reaction Coordinates.

The choice of reaction coordinates is pivotal to obtain the best approximation of the free energy.

The CVs used here to describe PS999 and PS1036 dissociation are:

1. The distance d_{CMs} between the centers of mass of the ligand and the center of the mass of the protein (for PS999 and PS1036, respectively). A similar choice of CVs has been applied e.g. in Ref.
2. The number of hydrophobic contacts n_{hph} between non polar carbons on the ligand and on the bases that it covers in the starting structure, modeled as a coordination number:

$$n_{hph} = \sum_{ij} \frac{1 - (r_{ij} / r_0)^a}{1 - (r_{ij} / r_0)^b} \qquad (2)$$

The parameters a and b have values of 6 and 12, respectively, while r_0=6 Å accounts for the typical carbon-carbon distance (4/4.5 Å) and the thermal motions' amplitude (1.5/2 Å). A similar CV has been used in Refs. 44, 95.

The Gaussian parameters were w≈0.25 kcal/mol (0.1 Kj/mol), δs_{CMs}=0.5 Å, and δs_{hph}=6 in both the case of PS999 and PS1036. The time interval between two successive Gaussian depositions was set to 0.5 ps in all the simulations.

Free energies surfaces were calculated as a function of d_{CMs}, n_{hph}) . In addition,simulations were performed in which the three CVs were kept active. In every metadynamics run the dissociation event was seen after a few ns (see Results) and occurred following a very similar mechanism, which means that the relevant slow motions of the systems are captured by our CVs. The activation free energies associated to the detachments were extracted by stopping the summation over Gaussians just after the complete detachment of the drug .

For the sake of clarity, we show here only the free energy profiles as a function of d_{CMs} and n_{hph} in case of PS999•RdRp and PS1036•RdRp.
Furthermore, we used metadynamics as an exploratory method to identify the different minima along the the undocking pathway of the PS999 and PS1036. Then, we performed additional metadynamics simulations to calculate the free energy barriers connecting all of the minima. The "forward" energy barriers connecting each minimum (for example from Mini-I to Mini-II) was evaluated by counting the contour lines separating each minimum from "top to bottom" in the free energy surface (FES) resulting from the dissociation simulation.

To obtain the "backward" energy barriers connecting each minimum (for example from mini2 to mini1), we launched additional metadynamics simulations starting from each minimum identified and simulated the reverse process. From this new FES obtained, we then evaluated the "backward" energy barriers by counting the lines separating each minimum from "bottom to top". In the end we put together the "forward and backward" energy barriers in the 1D free energy profile (1D-FEP) that thus sums up the complete energetic of the translocation process. From the 1D-FEP we can evaluate the effective barrier of translocation ΔG as the difference between the energy at highest barrier and at the dissociation of the the inhibitors from the putative binding site.

From each identified minimum we performed additional standard MD simulations (up to 5 ns) to characterize the structural details and the interaction of PS999 and PS1036 with BVDV RdRp. The hydrogen bonds (Hbond) between RdRp and inhibitors are counted using VMD25 scripts with the following threshold parameters: a distance of at most 2.8 Å and donor-hydrogen-acceptor angle of at least 130°. Hydrophobic contacts (Hphobic) are counted when nonpolar atoms are separated by at most 3 Å. We define durable Hbonds as the ones that have a lifetime equal to or higher than 20% of simulation time.

The change in enthalpy energy of PS999 and PS1036 in the different minima identified was done by evaluating the non bonded interactions (van der Waals + electrostatics) between inhibitors and all other atoms (protein, ions, water molecules) which lie within a cuttoff of 10.0 Å from inhibitors, and for electrostatic interactions we adopt the same scheme (soft particle mesh Ewald schemes26) as for the simulations. The configurational entropy of inhibitors in the different minima was calculated from the covariance matrices of the atomic fluctuations as proposed by Andricioaei and Karplus.

Results:

We used computational protocol for the following mentioned objectives: 1) Use of Docking Protocol to provide a first guess of the putative binding cavity; 2) Use of MD simulation to relax the structure in aqueous solution 3) Use of Metadynamics to predict the binding poses of the compounds from their respective minima; 4) Use of meta dynamics to predict the energetics of unbinding of the compound PS999 and PS1036 from the putative binding cavity; 5) Use of MMPBSA for the calculation of binding free energies, computational IC_{50} values and comparison with the experimental IC_{50} values, of the poses identified by the metadynamics. Generally speaking, and in harmony with previous findings of our group, the procedure of comparison of experimental value with calculated value are very well validated (28, 29, 30,). Here, in the presence of mutagenic results along with the results of enzymatic assays (IC_{50}), we considered only the lead compound PS999 and PS1036 as a proof-of-principle . The above mentioned application have been applied to the discovery of interesting details about the binding modes of these drugs with the residues lining the putative binding site.

Docking:

Binding site of NNI in BVDV is yet not known for these classes of compounds, hence resistant mutations A392 and I261 of lead compound PS999 and PS1036 respectively were used as grid centers for the automated molecular docking runs. In the absence of knowledge regarding the binding sites of the these compounds, one of the major criteria for selecting the most biologically relevant docking pose in both compounds was, the proximity of the conformers to their respective resistant mutations.

MD Simulation

In order to validate the docking poses obtained in both compounds, first a relaxation of 10ns of standard MD simulations were performed for each complex system(PS999/PS1036-BVDVRdRp). From the residue wise electrostatic and VdW energy obtained from the MD-simulations, we were able to identify residues that were important for binding and which were nearly common for both PS999 and PS1036. The key residues which showed favorable interaction are Arg127, Glu128, Arg132, Ile261, Pro262, Leu263, Asn264 and Ala392 in both cases. The consensus in the residues interacting with both compounds and the highly favorable interaction energies obtained, strongly supported the possibility of binding of these compounds into a common site, present at the entrance to the template channel.

Interestingly, binding site identified for the PS999 and PS1036 was similar to the one identified for 227G, compound belonging to Benzimidazole class.(Refer chapter 1.) The residue I261 and A392 are located at the entrance to the template channel, in close vicinity to each other . Conglomeration of resistant mutations A392E(PS999) and I261M(227G), I261T(PS1036) as well as Identification of stable binding poses for each of the compounds belonging to different classes on the finger domain of BVDV RdRp indicate that domain could be a "***hot-spot***" for NNI binding in BVDV RdRp(Figure1).

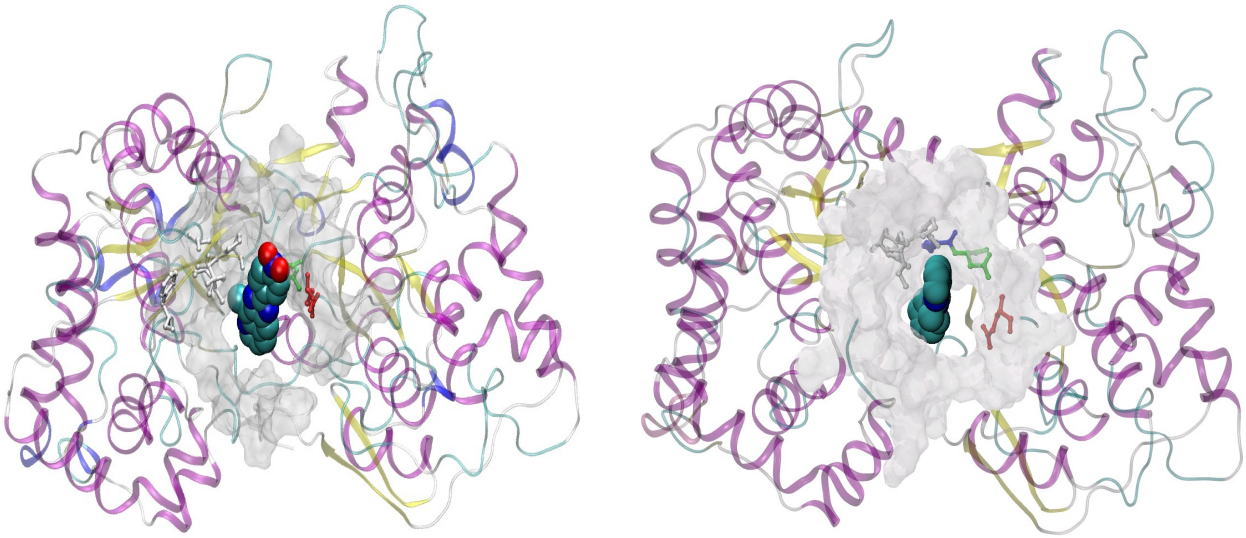


Figure1: Complete overview of the location of binding site of PS999 and PS1036 complex systems. The snapshot taken were the average structures for both the systems. The compounds are render in VdW, binding site residues are in ball and stick model, while binding cavity is in surface view. Whole protein rendered in ribbon, coloring in secondary structure form.

Putative Binding cavity

The putative binding cavity is made up of four loops (L1, L2, L3 and L4) (ref our previous findings). The Loop L4 (Pro388-Ile398) and L3 (Ala221-Asn229) originate from the finger domain, loop L2 (Leu530-Gly537) belongs to the thumb domain, while loop L1 (Leu125-Arg132) is the part of the N-terminal domain. Motif-I, which is located on the flexible fingertip region, was also found to flank the cavity. The residues lining the compounds include mostly hydrophobic (Ala221, Ala222, Phe224, Ile261, Pro262, Ile287, Ala392), this is consistent with the chemical properties of the compounds, whose main body is made of aromatic rings . Other residues of binding site are polar (Thr160, Thr162, Asn217, Asn264 and Ser533) amino acids, although also three basic (Arg127, Arg130, Arg132) and two acidic (Asp126, Glu128) are present.

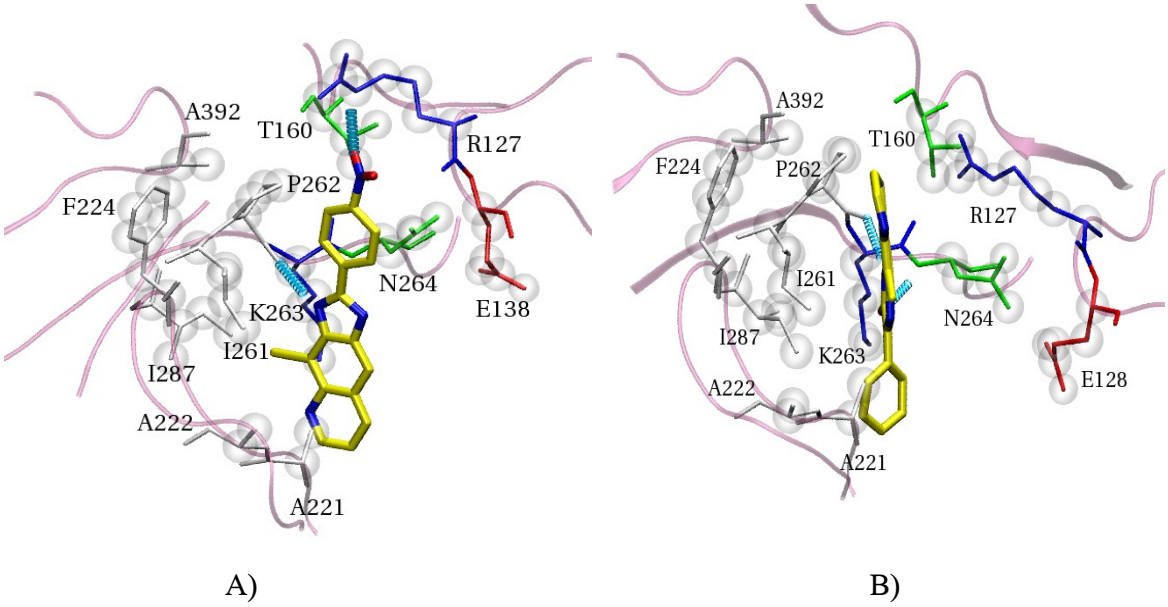


Figure2: Expanded view of interaction between key residues of binding cavity of PS999 (A)and PS1036(B). Snapshot taken from the deepest minima in both PS999 and PS1036. Binding site residues are shown here using ball-sticks representation(basic:blue, acidic:red, hydrophobic:white). In panel (A) Interaction map of PS999, residues involved in making HB with PS999 are Pro262 of motif-I and Arg127 of loop L1, while in panel (B) Interaction map of PS1036, Pro262 and Asn264, both from motif-I.

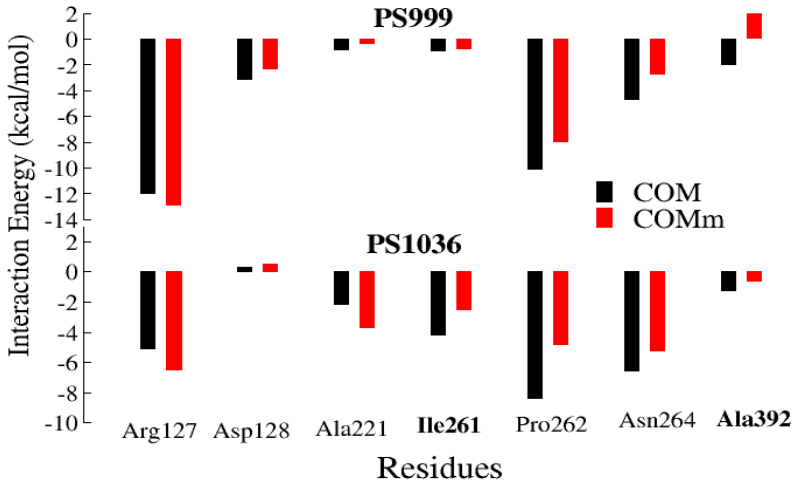


Figure3: Average ligand-residue interaction energies in kcal/mol over compound PS999 and PS1036 for the residues that contribute most to the ligand -surrounding. In the upper panel the interaction energy of WT and Mutant with PS999. In lower panel interaction energy of PS1036. The residues highlighted in bold are resistant mutant which reduces considerable in mutant (Red bar) with comparison to WT (black bar).

Residue	Ele	VdW	Total
Arg127	-10	-2.4	-12
Glu128	-1	-2.1	-3.1
Ala221	-0.1	-1	-1
Ile261	2	-3.4	-1.4
Pro262	-7.8	-2.3	-10.2
Asn264	-0.5	-4.2	-4.8
Ala392	-1.4	-1.3	-2.3

Table1: Interaction Energy of key residues in PS999-RdRp Complex (Kcal/mol)

Average structure obtained from the MD simulations for each system were then used as a starting structure for the metadynamics runs. Free energy calculations were used to explore the ligand binding space in an explicit solvent. These simulations were used to identify alternative stable binding pose, if present other than the initial docking structure, and to predict the energetics of unbinding. The free energy simulations were performed as a function of two collective variable that took into account the hydrophobic contacts and the distance between the center of mass of the inhibitor with the center of mass of the binding cavity. These variables have been already used to characterize ligand-target molecular recognition processes using the metadynamics approach.

Interestingly, we found that in case of PS1036,the deepest minima corresponded to the initial starting structures obtained by the docking procedures, while , in case of PS999 an alternative binding pose was identified as mini2. Under the action of metadynamics, the ligand leaves the starting position, and explores the whole binding site and finally takes its way out from the enzyme through the template entrance channel. A detailed description of the unbinding pathway, key interactions of the compound's in each minima along with the energetic required for the dissociation of the compounds from there respective binding pockets to the external environment, are discussed in following subsection.

PS999:
PS999 is the representative compound of Imidazoquinoline class of compounds. A deep inspection of the free energy profile of the unbinding process indicates that poses of PS999 clustered around three minima's, mini-1, mini-2 and mini-3, wherein mini-1 corresponds to the pose obtained from docking. However, the deepest minima, representing the most energetically stable pose of PS999 corresponded to mini2. It is important to note that standard docking programs AutoDock (31), which assume the protein to be rigid, fail to predict this second binding and more stable binding pose. Apart from performing a comparative analysis of the minima's obtained, with a view to determine the most stable and biologically relevant binding pose; our approach also helped us to have a guess of the putative undocking procedure. In order to accurately quantify the interaction pattern in each minima obtained, as well as to better characterize the undocking process, additional standard molecular simulations of 5ns were launched from each identified energy minimum.

Mini-1: The optimized location and the key interactions of the inhibitor within the putative binding site of BVDV RdRp are shown in figure 2(A). PS999 stabilizes in the cavity in such a way that the N3 atom of azole moiety makes a strong HB with the Pro262, (part of motif-I), exist throughout the simulation with occupancy more than 80%. A closer inspection of the free energy profile revealed that ~6 kcal/mol energy was required for the exit of PS999 from Mini1. (Figure 4A)

We did the decomposition of interaction energy of binding site residues to understand the contribution of each residue in the stability of PS999 into the binding pocket. The major contribution obtained inform of electrostatic came from residue Arg127 (-9.5kcal/mol) and then residue Pro262 (-7.7kcal/mol)(Figure 3, Table 1)Proceeding with metadynamics, the ligand leaves the Mini-1 and while exploring the catalytic site finds a second minima mini-2. Mini-2 was found to have the deepest basin and hence suggested an extremely good thermodynamic stability of PS999. As inferred from the free energy profile an unbinding energy of ~7 kcal was required by PS999 to exit from the binding pocket.

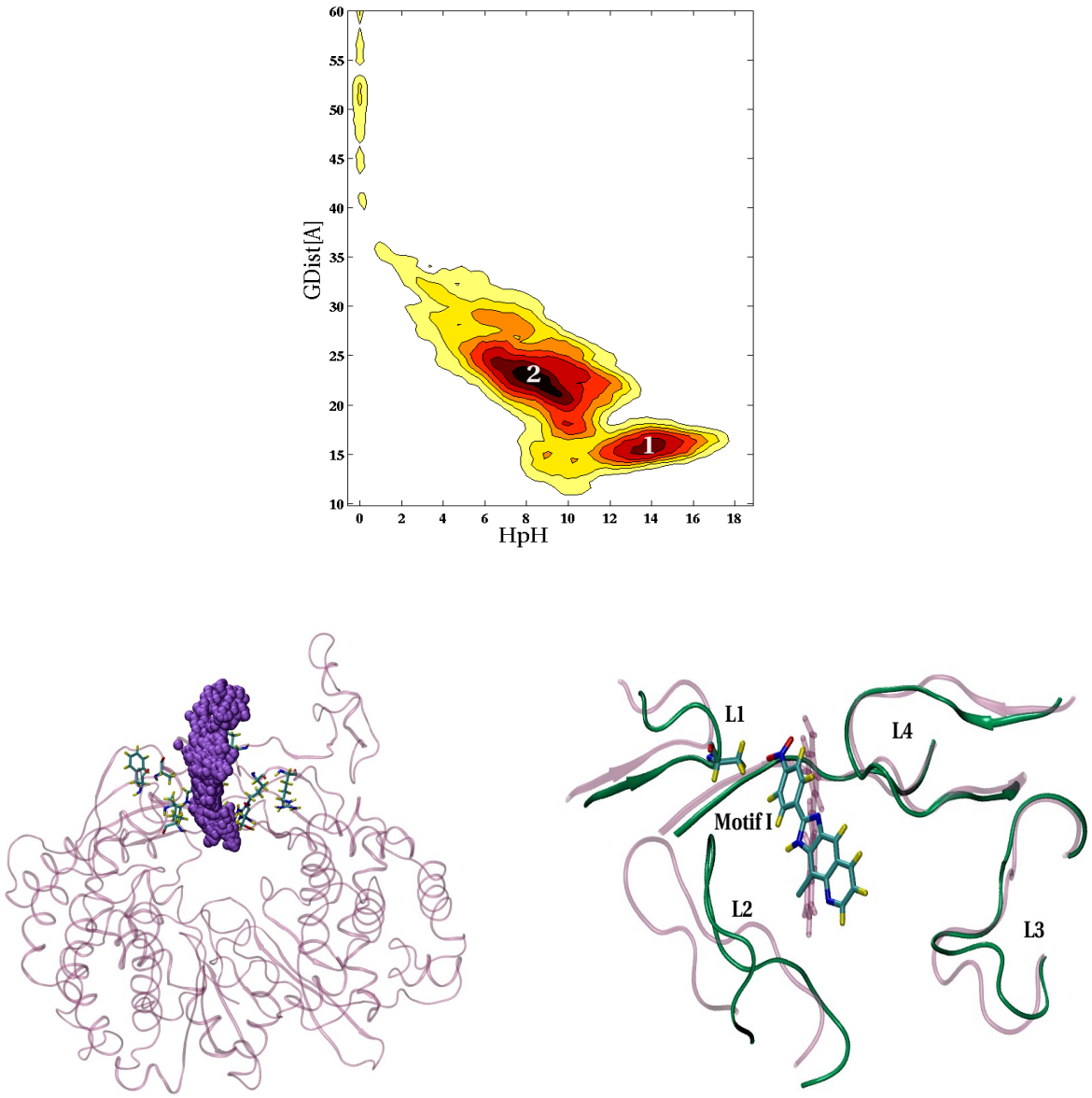


Figure4: A)-The two dimensional Free energy surface(FES) of PS999'S undocking process, as a function of *gdist* and *HpH* contacts. Each line correspond to 1kcal/mol. where each color and each line correspond to 1kcal/mol A)-this panel represent the FES of PS999, representation two deep minima while panel **B)**- Evolution of the center of mass of PS999 during the metadynamics simulations of the undocking process. The residues of the putative binding site are highlighted in licorice. C)- Pictorial representation of closing of binding site in mini2 (green in color), as compared to the mini1(purple and transparent).

Mini2. As already mentioned, in order to confirm the stability as well as to understand the dynamical and structural property of PS999 in Mini-2(Figure 4A), we extracted a snapshot from Mini-2 and carried out a 5ns standard MD run. The PS999 was found to be stable during the simulation. Major interactions contributing to the stability of PS999 in Mini-2 include a network of strong inter and intra molecular HBs. In particular a HB is formed between N atom of the guanidium group of Arg127 and polar oxygens of PS999 (ADL =2.8A). The interaction persists throughout the simulation with occupancy more than 65%. This interaction strengthen more in mini-2 as it is also present in mini-1 but with less occupancy (30%). Another strong HB exists between polar oxygen of PS999 with Thr160. Other source of stabilization of the complex comes from water-mediated interactions between Glu128 and N3@PS999 and Arg127 and O2@PS999: although the mediating water molecule changes during the simulation, this interaction persists for the entire simulation time. PS999 is further stabilized by strong HpH with the main contributions coming from residues Ala221, Phe224, Ile261, Pro262, Ile287 and Ala392. Interestingly, it was noticeable that the HpH contact of Ala392 was increased in mini2 more than 50% than in mini1.

The salvation analysis revealed that mimi2 is more solvent exposed than mini1. We did a 1st shell water analysis around the PS999 in both min1 and mini2 cases, and we observed in case of mini1, the no. of water molecule around PS999 were 6 while in case of mini2 they were 8. It seems more logical as in mini1, PS999 was more buried into the binding pocket and covered by three loops L1, L3 and L4 while in case of mini2 PS999 lye on the surface mainly interacting with loop L1 and L2 residues.

Undocking Mechanism: Meta-dynamics approach helped us to reveal the undocking mechanism of PS999 from the putative binding cavity into the outer environment. During movement from mini1 to mini2 a rotation of PS999 occurs around its major inertia axis. This leads into the first event observed in the undocking mechanism wherein breakage of HB between Pro262 and N2 atom PS999 occurs, thereby leading to the detachment of the indole moiety from the motif-I region of the BVDV RdRp. Due to this reorientation, the benzyl group of PS999 gets stacked against the guanidium group of Arg127. This stacking further strengthens the HB formed between the O2 atom of PS999 and Arg127, resulting into an occupancy of more than 50% during the standard MD simulation of mini2. Furthermore, N4 and N2 atom of the indole moiety of PS999 starts making water mediated interaction with Glu128. We also measured an rmsd of 1.3 Ang for the heavy atoms of binding site residues between Mini1 and Mini2. This value reflects the fact that there could be a local conformational change in the binding site residues of Mini1 and Mini2. In order to have a better understanding of the conformational changes taking place, we calculated minimum distances between the residues of loops L1, L2, L3 and L4 which describe the binding cavity. We found that the loops L2-L4 and L3-L4 sampled smaller distances in mini2 as compared to mini1. The histogram distribution in case of min1 was found to be centered around ~16.7 A in mimi2 and 18.4 A in mini1 for loop L2-L4 while 10.5 in mini2 and 11.3 in mini1 for loop L3-L4 respectively. This indicated a closure of the loops in mini2 as compared to mini1. (Figure4(C).

Escape: Due to the loss of HpH contacts with loop L3 residues , PS999 moves from the metastable state mini2 into the solvent. The electrostatic interaction between nitro groups of PS999 with the guanidine group of Arg127 was the last interaction to be lost and PS999 came out to be more solvent exposed region. Finally, all the interaction with the protein was lost and with the movement of linker residues and loop2, PS999 escaped out completely into the solvent. The escape mechanism is depicted in Figure 4 (B).

PS1036:
PS1036 is the representative compound of Pyridoxoquinoline class of compounds. Here, the deepest minima laso correspond to the initial docked posed of docking. The in depth speculation of FES of the unbinding process of PS1036 indicated that the different poses of PS1036 clustered around mainly two metastable state represent minimas, mini-1, mini-2 .Interestingly, the deepest minima, representing the most energetically stable pose of PS1036. PS1036 stabilize itself into the binding pocket all most in same orientation and with same key residues of the binding site as in case of PS999. The RMSD between PS999 and PS1036 in stable orientation is 1.03A. A respective study of each minima has been performed to insect the key interaction of PS1036 during its undocking mechanism. We performed additional MD-simulation to analyze dynamical and structural properties of PS1036 in their respective minima identified along the undocking path.

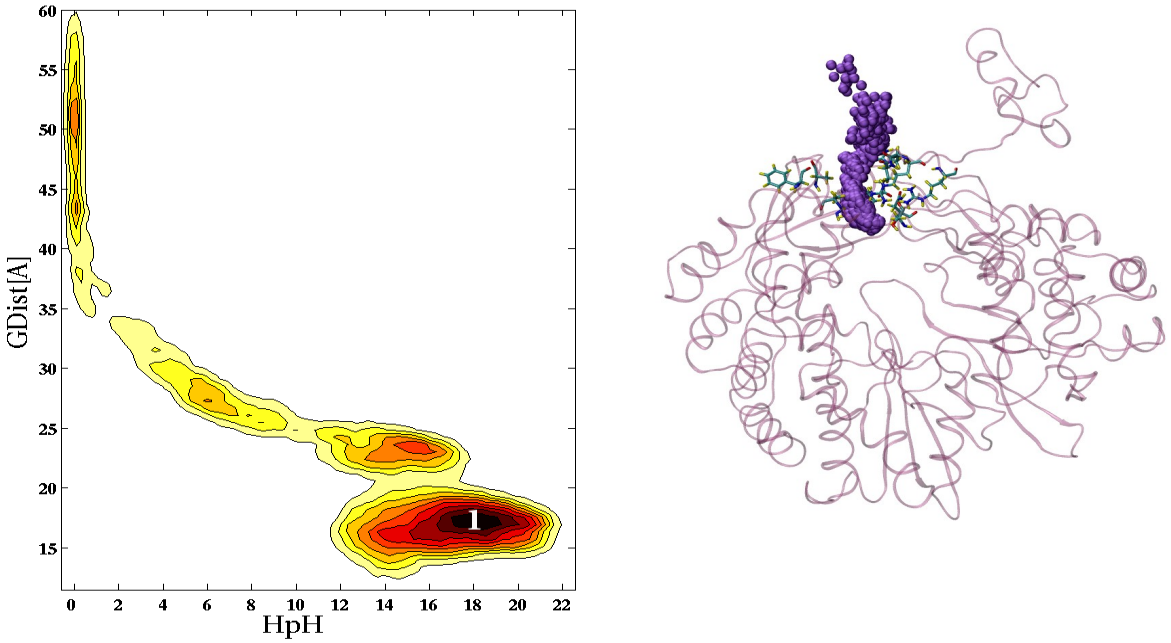


Figure5: A)-The two dimensional Free energy surface(FES) of PS1036'S undocking process, as a function of *gdist* and *HpH* contacts. Each line correspond to 1kcal/mol. where each color and each line correspond to 1kcal/mol A)-this panel represent the FES of PS1036, representation two deep minima while panel 2B)- represent the FES of PS1036 with only one minima. **B)**- Evolution of the center of mass of A)- PS1036 during the metadynamics simulations of the undocking process. The residues of the putative binding site are highlighted in licorice.

Mini1:
As we mentioned earlier that the deepest minima belong to the initial docked pose in the first minimum. Furthermore, we did a deep inspection to generate the binding map of PS1036. Here, PS1036 didn't go any major conformational change.

The important flexibility of PS1036 in mini1 is further evident of it's stability as average root-mean-square fluctuation (rmsf) was ~0.26 Å and the conformational entropy calculated to be TS=5.2kcal/mol. We observed the main stable factor of PS1036 was the two most durable HB formed with the N3 and O1 atom of the pyrido part of the PS1036 along with residue Pro262 and Asn264, N3:O@Pro262 (ADL= 1.9A) and O1:HN@N264 (2.2A) respectively with the life time greater than 80%. Another stabilizing factor was HpH contacts, mainly with residues Arg127 of loop L4, Ala392 of loop L3 ,Ile261 and P262 belonging to motif-I region with VdW energy

values -5.4, -1.6, -5.4 and -2.1kcal/mol respectively.(Table2, Figure 2B) In terms of electrostatic contribution, mainly provided by residues Ala221, Pro262 and Asn264 ~ -1.2; -6.4 and -4.7kal/mol respectively. This would imply a high barrier to go from mini1 to mini2. With the visual inspection of FES the free energy barrier obtained was ~9kcal/mol to go from mini1 to mini2.

Table2: Interaction Energy of key residues in PS1036-RdRp complex (Kcal/mol)

Residue	**Ele**	**VdW**	**Total**
Arg127	0.3	-5.4	-5.1
Ala221	-1.2	-0.9	-2.4
Ile261	1.3	-5.4	-4.2
Pro262	-6.4	-2.1	-8.4
Asn264	-4.7	-1.9	-7
Ala392	-0.1	-1.6	-1.2

Mini2:
Here, a HB formed from the same pyrido group(N3 atom) of PS1036 and with the backbone oxygen atom of Thr69 with ADL=2.25A at the simulation of 2.3ns with the occupancy at this period was more than 52%.The phenyl tail of the PS1036 was tilted around 30 degree and at this state, making interaction with loop L3 residues in form of HpH contacts mainly Ala221, Leu225 and Phe224. Here the major unstabilizing factor was the loss of HB with residue Asn264. Another major loss reported here, was in form of electrostatic contribution from Asn264 and Pro262 ~0.8 and 3.6 kcal/mol in mini2 with respect to mini1. Finally, as PS1036 moves from mini2 to out of the pocket it started important interaction with terminal residues of loop L1 and L4.

Undocking mechanism:
With metadynamics approach, we tried to understand the undocking process of PS1036 from its respective binding site. The major step of undocking process was the breaking of HB with Asn264. This loss put PS1036 into the detachment from the motif-I region and move it away from the loop L1; because of this loss, it started moving, around 30 degree from its original axis and reached close to loop L3. The instability caused by the loss of HB with Asn264, imposing PS1036 from mini1 to min2. Due to change its initial position, now the phenyl ring of PS1036 making HpH contacts with loop L3 residues. Apart from these HpH contacts it also has HB with Pro262 but with less occupancy than mini1. We calculated the root mean square deviation (rmsd) of heavy atoms ~ 1.23Å between the mini1 and mini2. While the rmsf of PS1036 in mini1 is 0.26Å, which increased in mini2 ~0.43Å. These values are very well evident of loss of stability of PS1036 from mini1 to mini2. Furthermore, we calculated the minimum distances between the loops as we did in case of PS999, to understand effect on loops during the detachment of PS1036. Interestingly we observed the distance between loop L1 and L3 and L1 and L2 increased considerably, which led the widening of binding site, similarly as in case of PS999.

Escape:
From the mini2 to exit of the PS1036 was mainly guided by the terminal residues of loop L2, L4 and finally with loop L1. Here, we observed the major interaction of PS1036 with residue Arg36 inform of π-cation reaction. With this interaction at the end, Arg36 stuck with PS1036 as an anchor and parallely moved the PS1036 towards the solvent environment. Finally, PS1036 lost its all interaction and reached to solvent. (Figure 5A and B)

During the undocking process we observed in case of PS999, the drug moved in between loop L1 and L2 while in case of PS1036 it moved out from Loop L1 and Loop L4. The undocking or escape of the inhibitor is

important to understand as it present some rational clues which could be important in the further modification of the inhibitor with respect to their key residues which unveil during the path analysis.

Binding free energies changes:

Only once the best pose has been identified by metadynamics and all eventual available criteria have been satisfied, a more accurate and quantitative method as the MM-PBSA analysis sampled by molecular dynamics simulations was used to predict binding free energy. Since, the minimum energy pose obtained from metadynamics corresponded to the initial docking position hence the MM-PBSA was performed from the snapshot taken from the 10ns relaxation of the initial docking pose.

By MM-PBSA analysis,32 the total free energy of binding into electrostatic, VdW, and solute-solvent interaction could be separated, gaining, thus, additional insights into the physics of the inhibitors-protein association process. . Several studies have shown good correlation with the experimental results when adapting MM-PBSA in studying the binding free energy of inhibitors to receptor, and consequently in-depth information about the binding mode of inhibitors to the receptors could be acquired. Nevertheless, there are number of issues related with MM-PBSA that needs careful consideration when one wants to employ this method, and one of them is to justify the necessity to evaluate snapshot of the complex, receptor and ligand from a single trajectory or running separate trajectories for all three term. The single trajectory method has an apparent advantage as only one trajectory of the complex is needed; However, assumptions have to be made as snapshots of receptor and ligands taken from the trajectory of the complex are of equivalent free energy to those that would be taken from separate trajectories. 33. As demonstrated by kuhn and coworkers,34 the use of MM-PBSA on a single, relaxed complex structure is adequate to provide a good binding affinity ranking and give a fair correlation with an experimental value which prompted us to use single trajectory approach in our work. For this energetic analysis, 100 equally spaced snapshots were taken at intervals of 8ns of each MD trajectory. The binding free energy and the energy components of the complexes are summarized in table?.

Generally speaking, and in harmony with out previous findings of our group (28,29,30,). According to this computational perspective, Table3, electrostatic (ΔG^{E}) and VdW (ΔG^{VdW}) terms in the gas phase provide the major favorable contributions to the inhibitor binding, whereas owing to the polar character of two compounds, the desolavtion penalty paid by these molecules upon binding is also quite substantial which is evident by unfavorable polar solvation energies. The non-polar solvation energies (ΔG^{NP}), which corresponds to the burial of SASA upon binding, barely contribute to the binding.

Further insight into the forces involved in complex formation can be obtained by analyzing the electrostatic (ΔG^{Etot}) and non-electrostatic (ΔG^{NPtot}) contributions in Table3 . As demonstrated by numerous studies, the electrostatic contribution generally disfavors the docking of ligand and receptor molecules because the unfavorable change in the electrostatic of solvation is mostly, but not fully, compensated by the favorable electrostatic within the resulting ligand-receptor complex. Indeed, from Table3, we can appreciate that, despite the favorable electrostatic energies in gas phase (ΔG^{E}), which are -4.5 and -4.6 in PS999 and PS1036 respectively, the contribution of the polar solvation energies to binding which are 24.3 and 22.9 in PS999 and PS1036 respectively, the sum of ΔG_{EL} and ΔG_{PB}, does not favor the binding. Table 3 also suggests that the net result of non electrostatic interactions (ΔG_{NP}), the sum of ΔG_{vdw} and ΔG_{NP}, is favorable for the formation of the complexes, and it should be noted that this behavior has been proposed previously as a general trend for the non-covalent ligand-receptor associations.35. From the above results, we conclude that the binding free energies obtained for these complexes are driven by more favorable non-polar interactions rather than by electrostatic interactions.

Another issue when dealing with the MM-PBSA is the approximation of solute entropy. The calculation of solute entropy term using normal-mode analysis is normally time-consuming, computationally expensive, and difficult to converge in the calculation, and most of the time it can be omitted [36,37,38], when calculating binding free energy. Our MD simulation showed that this trend can not be omitted if one wants to have a clearer picture on the energy terms. As shown in Table3, the initial ΔG_{bind} values (excluding solute entropy) for the PS999 and PS1036 were -28.6 and -26.6 respectively and one can see that these values were far from the final ΔG_{bind}. The solute entropy contribution calculated from the normal mode analysis was -18.4 for PS999 and -18.3 for PS1036, and it has significantly reduced the estimated binding free energy to -10.8 and -8.9 in

these two systems.

The opposing interplay between a more favorable enthalpic(ΔH) term and an unfavorable entropic term, as is also seen in our simulations is known as enthalpy/entropy compensation and is fundamental property of non-covalent interactions. This enthalpy/entropy compensation results in changes in ΔG values. More importantly, it is not confined to binding in aqueous solution, nor should it be ascribed to errors of measurement. it arises because bonding opposes motion and, also reciprocally, motion opposes bonding. The two effects can be traded off against each other because the strength of non-covalent bond is, at room temperature, comparable to thermal energies that oppose them.39

As was recently reviewed, allostery is a purely thermodynamic phenomenon in which a binding event leads to loss of freedom of motions of the binding partners, including their internal motions; Thus, it is entropy-unfavorable. Besides a disorder to order transition generally involves formation of a more cooperative set of interactions within the protein that replaces a less cooperative set of interactions within the protein that replaces a less cooperative set of interactions between the protein and the solvent. then, the added ligand-receptor interactions are enthalpy favorable because, during binding, the interactions get increasingly tighter. Concordantly, the above-mentioned trend is followed by complexes analyzed herein, as shown by the results listed in table3 The result obtained herein is highly encouraging as we got a good agreement between the value of calculated binding free energy and experimental data. The IC_{50} values thus calculated in both cases.

Drug Resistance

All of these results and observations led us to hypothesize the probable possibilities of mechanism of drug resistance. An Atomic-detailed consequences of point mutations in drug targets would be of great value in the design of novel compounds to inhibit receptors mutants. A detailed analysis In order to understand as to how the mutations A392E and I261T renders the drugs inactive against the enzyme, we introduced the mutations "*in-silico*" and relaxed the structure for 10ns.As we can see from the table1, each mutation causes a distinctive energetic change from the wild type protein. In the case of PS1036, I261Tdecreases the net electrostatic component greatly $\Delta E_{EL}+\Delta G_{PB}$=+5.7kcal/mol vs +3.9 kcal/mol for the wild type RdRp. For PS999, A392E causes a substantial decrease in both electrostatic and VdW energies. The free energy of binding results listed in table 3 are fully compatible with the indices of resistance level that were estimated from the experimental IC_{50} values.

We speculated the modeled systems PS999 with 392E and PS1036 with 261T, in details. We observed the major conformation changes in the binding site and the N-terminal region in form of Root-Mean-Square-Deviation(rmsd) in both cases. The average rmsd of binding site from the equilibrium trajectory with respect to the starting structure in the wild type complex is 1.38Å while in case of mutant it increased to 2.14Å. In case of PS1036 it is 1.06Å and in mutant complex it increased to 2.86Å. Another substantial changes, we observed in N-terminal region, where the rmsd is 0.78Å in PS999 and 1.03Å in PS1036. The change in N-terminal was notable as it is a crucial part in the polymerization and unique in BVDV RdRp(.

To investigate electrostatic effect of the $A392E^{PS999}$ and the $I261T^{PS1036}$ mutations, the electrostatic potential grids were computed by numerical solution of the linearized Poisson-Boltzman equations and mapped onto the solvent accessible surface of the binding pockets. (Figure 6). As shown in the figure A392E mutation was found to perturb the electrostatic potential inside the binding pocket. The electrostatic potential surface of the bound PS999 are fairly complementary with the wild type complex. The major mutational effect, observed was, the sever electrostatic repulsion between the oxygens of nitro group of PS999 and the carboxylic group of Glu392. As we can see from the figure 6 There is a strong electrostatic repulsion in case of mutant PS999 while in case of wild-PS999 it fitted comfortably into the binding pocket.

The computed interaction energies of binding between the PS999, PS1036 and their respective binding site residues are reported in Figure 2. Residue-based decomposition of the non-covalent interaction energies between WT and A392E and I261T mutant are in Figure 2. The decomposed interaction energies of binding predicted the A392E mutation A392E decreased the binding of PS999 to RdRp. The difference between the calculated **Δ**G of wild type and A392E is 3.6 kcal/mol, while the difference in value of ΔG in PS1036 WT and I261T is 2.3 kcal/mol. As shown in graph, A392E mutation perturbs the energy contribution of critical residues to the binding of PS999 to the RdRp.

The largest effect of the A392E mutation involves a considerable decrease on the intermolecular dispersive(ele vdw) interaction residues A221, A222, A224, I261 and P262 in comparison with WT complex. In particular, ΔG of WT complex and A392E are -2.8 and -1.4kcal/mol, respectively, suggesting A392E mutation significantly decrease the dispersive interaction between A392 and PS999. Another remarkable difference observed in WT and A392E is the high electrostatic energy of residue R127 in case of A392E. The residue R127 of loop L1 is very flexible in nature as it has very high B-factor. It making HB with PS999 in WT complex but with less occupancy. This occupancy enhanced asymptotically, as the flexible nature of R127 is hindered by E392 in case of mutant. The terminal carboxylic group of E392 making strong HB with guanidium group of R127, locks R127 in such a orientation which increases the interaction of it with nitro group PS999. However, the overall non covalent interaction energies of WT and A392E are similar.

Table3)-

System	ΔE_{VDW}	ΔE_{EL}	ΔG_{PB}	ΔG_{NP}	PB_{tot}	ΔG_{EL+PB}	ΔG_{vdw+NP}	$T\Delta S_{solute}$	ΔH	$\Delta G^a/\Delta G^b$	$IC_{50}{}^a/IC_{50}{}^b$
PS^{999}	-30.0(±2.7)	-18.6(±5.2)	24.3(±5.5)	-4.6(±0.7)	-28.8(±3.1)	5.7(±5.2)	-34.6(±2.8)	-18.4(±2.5)	-28.9	-10.5/-9.8	0.02/0.06
$PS^{999}{}_m$	-28.6(±2.9)	-17.9(±4.3)	21.8(±4.8)	-4.4(±1.2)	-25.2(±3.3)	5.9(±3.9)	-33.0(±1.8)	-18.3(±2.1)	-27.1	-6.9	1.6
PS^{1036}	-28.3(±3.2)	-20.2(±2.3)	22.9(±5.1)	-4.5(±0.19)	-26.6(±2.8)	2.9(±3.2)	-32.8(±1.2)	-18.3(±2.8)	-29.9	-8.9/-8.1	0.2/1.0
$PS^{1036}{}_m$	-27.4(±2.9)	-19.2(±3.1)	20.9(±5.3)	-4.2(±0.8)	-24.7(±3.5)	1.7(±3.8)	-31.6(±2.3)	-18.2(±1.9)	-29.2	-6.6	14.7

The experimental values of ΔG_{bind} were calculated from experimental IC_{50} value by fallowing relationship

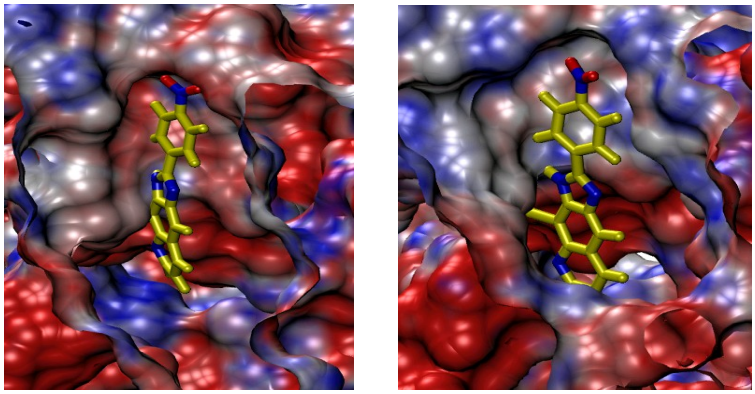
$\Delta G_{bind} = RT \ln IC_{50}$

(R is the universal gas constant, T = 298K)

a- Calculated value

b-Experimental value

All the energies are in kcal/mol and all IC_{50} values are in micro molar.



A B

Figure6: Electrostatic potential map of the PS999-binding pocket of BVDV RdRp. A)- PS999 in WT B) PS999 in A392E mutant (positive in blue; negative in red and neutral in white in color). The major notable difference on electrostatic map of PS999 is pointed with arrow in both cases, shown the effect of mutation A392E

In case of PS1036 the major noticeable effect of mutation I261T was the change of Isoleucine to threonine, as a change from HpH to polar amino acid. The I261 is the one of the key residue contributing -4.6kcal/mol in the stability of PS1036 into the binding pocket. In particular, there is a observable opening of binding site in case of I261T. Indeed, an opening of loop L1 is seen in I261T which is not observed in WT. Interestingly, we observed there is a considerable reduction in the VdW volume of T261 residue in Place of I261 in WT. Apart from

loss of HpH contacts with I261, there is other significant change observed in I261T in form of HB. In case of mutant I261T, PS1036 lost its both HB with residue P262 and N264, which was present in WT complex. The total energetic loss of ΔG in WT and I261T is 2.3 kcal/mol.

The loss of energies in mutant system is quite considerable, and the significant loss of overall stability of the system allows a conformational readjustment of the mutant protein binding site which, in turn, results in a loss of interaction between all residues lining the protein binding pocket and the compounds. But we didn't found any substantial change in entropy, as it would be because the backbone of both the compounds, PS999 and PS1036 is highly hydrophobic, which perfectly compatible with the nature of binding site, and they bind in the same region.

One of the most important benchmark in this study, is the correspondence between the estimated free energies of binding and the experimental measured IC_{50} values. Indeed, there is a good agreement between the trend exhibited by the IC_{50} values reported in table2 and the corresponding biological activity determined for these two compounds in BVDV infected cell line. Although we obviously can not directly compare the computed binding free energy with the IC_{50} values deriving from experiment, we can observe that, as in our previous experiences of our group, the rank of the compounds with respect to their activity towards its putative target, the RdRp of BVDV, is aptly maintained.

Conclusion:
Screening studies have come out with two new series of NNIs, Imidazoquinoline and Pyridozoquinolines as selective and potent inhibitors of BVDV RdRps.

Herein, we applied a combined experimental and molecular modeling strategy to explore the molecular basis of interaction of the PS999 and PS1036 with the BVDV RdRp. Based on the resistant mutations obtained for PS999 and PS1036, Binding site for these compounds, as also is in the case of 227G, were found to be located in the template entrance channel. Binding of different classes of compounds onto the same binding site made it a HOTSPOT for the binding of NNI of different classes in the BVDV RdRp.

Key residues playing pivotal role in binding of the NNIs were identified.

Metadynamics was used to identify alternate binding poses other than initial docking pose, as well as detailed undocking pathway was estimated in order to identify key events taking place during the dissociation process. Hypothesis regarding the mechanism of inhibition as well as the mechanism of resistance were put forth by performing extensive MD simulations.

Binding free energies of the NNI-BVDVRdRp complex were calculated which showed similar trend to the corresponding experimatal IC50 values for both the complexes. This was an encouraging performance and validate the model of the binding site, given that all molecular modeling studies were performed in the absence of any available crystal structure of the protein in complex with lead compounds. Also, the adopted procedure was able to correctly predict drug binding affinities in the presence of mutated protein residues involved in drug resistance.

Reference :

1. Hayashi, P. H.; Di Bisceglie, A. M. The progression of hepatitis B- and C-infections to chronic liver disease and hepatocellular carcinoma: epidemiology and pathogenesis. Med. Clin. North Am. 2005, 89, 371-389.
2. WHO. Global surveillance and control of hepatitis C. J. Viral Hepat. 1999, 6, 35-47.
3. Memon, M. I.; Memon, M. A. Hepatitis C: an epidemiological review. J. Viral Hepat. 2002, 9, 84-100.
4. Echevarria-Mayo, J. M. Etiology and pathogenesis of viral hepatitis. Enferm. Infecc. Microbiol Clin. 2006, 24, 4-56.
5. Bosch, F. X.; Ribes , J.; Cleries, R.; Diaz, M. Epidemiology of hepatocellular carcinoma. Clin. Liver Dis. 2005, 9, 191-211.
6. Fried, M. W.; Shiffman, M. L.; Reddy, K. R.; Smith, C.; Marinos, G.; Gonçales, F. L. Jr.; Häussinger, D.; Diago, M.; Carosi, G.; Dhumeax, D.; Craxi, A.; Lin, A.; Hoffman, J.; Yu, J. Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. N. Engl. J. Med. 2002, 347, 975-982.
7. Monath, T. P. Yellow fever as an endemic/epidemic disease and priorities for vaccination. Bull. Soc. Pathol. Exot. 2006, 99, 341-347.
8. O'Connor, A. M.; Sorden, S. D.; Apley, M. D. Association between the existence of calves persistently infected with bovine viral diarrhea virus and commingling on pen morbidity in feedlot cattle. Am. J. Vet. Res. 2005, 66, 2130-2134.
9. Chase, C. C.; Elmowalid, G.; Yousif, A. A. The immune response to bovine viral diarrhea virus: a constantly changing picture. Vet. Clin. North Am. Food Anim. Pract. 2004, 20, 95-114.
10. Carta, A.; Loriga, M.; Paglietti, G.; Ferrone, M.; Fermeglia, M.; Pricl, S.; Sanna, T.; Ibba, C.; La Colla, P.; Loddo, R. Design, synthesis and preliminary in vitro antiviral activity of 4,7-phenantrolines and 4,7-phenantrolin-1-ones against a single-stranded positive sense RNA genome viruses, Bioorg. Med. Chem. 2007, 15, 1914-1927.
11. Angusti, A.; Auzzas, L.; Boido, V.; Carta, A.; Ciliberti, N.; Ferrone, M.; Loddo, R.; La Colla, P.; Loriga, M.; Manfredini, S.; Mazzei, M.; Nieddu, E.; Paglietti, G.; Panemi, M. S.; Pricl, S.; Rassu, G.; Sparatore, F.; Tasso, B.; Vitale, G. Hep Dart (2005) Frontiers in Drug Development for Viral Hepatitis, December 11-15, 2005 Kohala Coast, Big Island, Hawai USA, Abstract 039. Global Antiviral Journal, 2005, 1, 37-38.
12. Sanna, P.; Carta, A.; Paglietti, G. Synthesis of two novel tricyclic rings: triazolo[4,5-g]quinolines and pyrido[2,3-g]quinoxalines derived from 6,7-diaminoquinolines, Heterocycles 2000, 53, 423-432.
13. Carta, A.; Sanna, P.; Gherardini, L.; Usai, D.; Zanetti, S. Novel functionalized pyrido[2,3-g]quinoxalines as antibacterial, antifungal and anticancer agents, Il Farmaco 2001, 56, 933-938.
14. Carta, A.; Boatto, G.; Paglietti, G.; Poni, G.; Setzu, M. G. Caredda, P. Synthesis and biological evaluation of triazolo[4,5-g]quinolines, imidazo[4,5-g]quinolines and pyrido[2,3-g]quinoxaline. Part II, Heterocycles 2003, 60, 833-842.
15. Carta, A.; Palomba, M.; Corona,. P. Synthesis of substituted aminoquinolines as useful intermediates for preparation of aromatic N-tricyclic systems, Heterocycles, 2006, 68, 1715-1722.
16. Carta, A.; Paglietti, G. A new synthesis of triazolo[4,5-g]quinolines and unexpeted ring reduced products by treatment with hydrazine hydrate, Arkivoc 2004, 5, 66-75.
17. Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.;Srinivasan, J.; Case, D. A.; Cheatham, T. E. 3rd. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc. Chem. Res. 2000, 33 (12), 889–897
18. Kuhn, B.; Donini, O.; Huo, S.; Wang, J. M.; Kollman, P. A. MM-PBSA applied to computer-assisted ligand design. In Free energy calculations in rational drug design; Rami Reddy, M.; Erion, M. D.,Eds.; Kluwer Academic/Plenum Publishers: New York, 2001; pp 243-251.
19. Sitkoff, D.; Sharp, K. A.; Honig, B. Accurate Calculation of Hydration Free-Energies Using Macroscopic Solvent Models. J. Phys. Chem. 1994, 98 (7), 1978–1988.
20. Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.;Srinivasan, J.; Case, D. A.; Cheatham, T. E. 3rd. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc. Chem. Res. 2000, 33 (12), 889–897
21. Kuhn, B.; Donini, O.; Huo, S.; Wang, J. M.; Kollman, P. A. MM-PBSA applied to computer-assisted ligand

design. In Free energy calculations in rational drug design; Rami Reddy, M.; Erion, M. D.,Eds.; Kluwer Academic/Plenum Publishers: New York, 2001; pp 243-251.
22. Sitkoff, D.; Sharp, K. A.; Honig, B. Accurate Calculation of Hydration Free-Energies Using Macroscopic Solvent Models. J. Phys. Chem. 1994, 98 (7), 1978–1988.
23. Rocchia, W.; Sridharan, S.; Nicholls, A.; Alexov, E.; Chiabrera, A.; Honig, B. Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: Applications to the molecular systems and geometric objects. J. Com- put. Chem. 2002, 23 (1), 128–137.
24. Honig, B.; Sharp, K.; Yang, A. S. Macroscopic Models of Aqueous- Solutions - Biological and Chemical Applications. J. Phys. Chem. 1993, 97 (6), 1101–1109.
25. Rashin, A. A. Buried Surface-Area, Conformational Entropy, and Protein Stability. Biopolymers 1984, 23 (8), 1605–1620.
26. Procacci, P.; Darden, T. A.; Paci, E.; Marchi, M. J. Comput. Chem.1997, 18, 1848
27. Laio, A.; Parrinello, M. Proc. Natl. Acad. Sci. U.S.A. 2002, 99,12562.
28. 28. Tonelli, M.; Vazzana, I.; Tasso, B.; Boido, V.; Sparatore, F.; Fermeglia, M.; Paneni M. S.; Posocco, P.; Pricl, S.; La Colla, P.; Ibba, C.; Secci, B.; Collu, G.; Loddo, R.Bioorg. Med. Chem. 2009, 17, 4425.
29. Carta, A.; Loriga, M.; Paglietti, G.; Ferrone, M.; Fermeglia, M.; Pricl, S.; Sanna, T.;Ibba, C.; La Colla, P.; Loddo, R. Bioorg. Med. Chem. 2007, 15, 1914.
30. For recent applications of our MDSA protocol see, for instance: (a) Zampieri, D.;Mamolo, M. G.; Laurini, E.; Fermeglia, M.; Posocco, P.; Pricl, S.; Banfi, E.;Scialino, G.; Vio, L. Bioorg. Med. Chem. 2009, 17, 4693; (b) Mazzei, M.; Nieddu,E.; Miele, M.; Balbi, A.; Ferrone, M.; Fermeglia, M.; Mazzei, M. T.; Pricl, S.; La Colla, P.; Marongiu, F.; Ibba, C.; Loddo, R. Bioorg. Med. Chem. 2008, 16, 2591; © Zampieri, D.; Mamolo, M. G.; Vio, L.; Banfi, E.; Scialino, G.; Fermeglia, M.; Ferrone, M.; Pricl, S. Bioorg. Med. Chem. 2007, 15, 7444; (d) Carta, A.; Loriga, M.; Piras, S.; Paglietti, G.; Ferrone, M.; Fermeglia, M.; Pricl, S.; La Colla, P.; Collu, G.;Sanna, T.; Loddo, R. Med. Chem. 2007, 3, 520. and references cited therein.
31. Morris, G. et al. Automated Docking Using a Lamarckian Genetic Algorithm and an Empirical Binding Free Energy Function. Journal of Computational Chemistry 19, 1639–1662 (1998)
32. Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Srivasan, J.; Case, D. A.; Cheatam, T. E., III. Acc. Chem. Res. 2000, 33, 889–897
33. Page, C. S.; Bates, P. A. Can MM-PBSA calculations predict the specificities of protein kinase inhibitors. J. Comput. Chem. 2006, 27 (16), 1990–2007.
34. Kuhn, B.; Gerber, P.; Schulz-Gasch, T.; Stahl, M. Validation and use of the MM-PBSA approach for drug discovery. J. Med. Chem. 2005, 48 (12), 4040–4048.
35. Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Srivasan, J.; Case, D. A.; Cheatam, T. E., III. Acc. Chem. Res. 2000, 33, 889–897
36. Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. 3rd. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc. Chem. Res. 2000, 33 (12), 889–897
37. Wang, J.; Morin, P.; Wang, W.; Kollman, P. A. Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA. J. Am. Chem. Soc. 2001, 123 (22), 5221–5230.
38. Otyepka, M.; Kriz, Z.; Koca, J. Dynamics and binding modes of free cdk2 and its two complexes with inhibitors studied by computer simulations. J. Biomol. Struct. Dyn. 2002, 20 (2), 141–154.
39. Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Srivasan, J.; Case,

Chapter5

Identification of binding cavity for novel Non-Nucleodside inhibitors of HCV RdRp: A molecular docking study

Key-Abstract

The virus encoded RdRp has emerged as a prime target in the search for specific HCV antivirals. Identification and successes of NNI, which are non-competitive diverse small molecules, against other polymerase, have encouraged them to be used as effective antivirals against HCV RdRp. Screening studies have came out with compounds belonging to Pyridoquinoxaline(PS1097) Phenanthroline(PS1101) and Imidazoquinolines (PS1126), effective against HCV RdRp at a very low micromolar range. Information regarding the binding cavity of these compounds are still lacking. Herein, we have utilized docking procedure to investigate binding sites, binding modes as well as binding affinity of these compounds in HCV RdRp.To rule out any bias, each of the three reported HCV NS5B NNI binding site represented by (A) Benzofurans (B) N,N-disubstituted phenylalanine and (C) benzothiadiazine inhibitors was examined for the binding of the compounds. We then used all atom standard molecular dynamics (MD) simulations to investigate the most probable binding site for the compounds as in both the complexes, there was a dissociation in two out of the three cavities under investigation. Therefore, Identifying the binding cavities and orientation of our compounds can help us to build a microscopically well-funded picture to elucidate the mode of action of compounds and thereby providing clues for rational drug design.

Introduction

Hepatitis C virus (HCV) infection is a major cause of chronic liver disease worldwide[1,2]. Chronic hepatitis C (CHC) may lead to liver cirrhosis and hepatocellular carcinoma. Hepatitis C Virus infection is estimated to be four to five times more prevalent than HIV-1, with over 200 million cases globally, of which $4.1 million infections exist in the United States alone. Despite its large medical and economical impact, there are no vaccines or efficient therapies without major side effects. The problems faced in HCV research are manifolds. HCV undergoes rapid genetic evolution during replication, resulting in a vast mix of variants, thereby presenting additional challenge towards eradication of the virus in infected patients. Development of vaccines or broadly effective therapy for all genotypes of HCV have been severely hampered by the lack of appropriate experimental systems. The only animal that can reliably be infected with HCV is the chimpanzee, but ethical reasons, high costs, and the inherent difficulties in working with large animals severely limit their experimental utility. However, the greatest limitation for HCV research has been the lack of a cell culture system that supports the efficient and reliable propagation of the virus. Among limited options the most effective therapy which include Type I interferons (IFNs), IFNα and IFNβ, are crucial and potent components of the early host response against virus infection[3] and recombinant (pegylated) IFNα2a and IFNα2b are widely used for the treatment of CHC and chronic hepatitis B. The current standard treatment of CHC with pegylated IFNα (pegIFNα) and ribavirin leads to cure in about 50% of patients with side effects such as headache, fever, severe depression, myalgia, arthralgia and hemolytic anaemia. [4,5] The cause of treatment failure in the remaining half of the patients is poorly understood. Thus, there is an urgent need to develop improved therapeutic options to combat HCV infections.

The RNA-dependent RNA polymerase encoded by HCV, which is strictly required for viral replication, has been the focus of intense drug discovery activity as they are considered as significant targets for therapeutic intervention. Several reasons that makes HCV RdRp an excellent target for antiviral investigation include: 1)-Successes stories in targeting the polymerases from other viral systems such as human immunodeficientciency virus (HIV); cytomegalovirus (CMV); herpes simplex virus (HSV);hepatitis B virus (HBV). ,coupled with recent advances in experimental systems for studying the HCV polymerase.: 2)-The most promising antivirals target viral proteins or processes that are not endogenous to host cells. Close structural homologs of the HCV RNA-dependent RNA polymerase (RdRp) do not exist within the uninfected host cell and hence it makes it a interesting target for fighting HCV infection; 3)- extensive functional and structural characterization (The high-resolution three-dimensional crystal structures of the NS5B polymerase is available), not only make it lucrative target for combating HCV but also for further computer aided drug designing efforts.

In common with other polymerases, the structure of HCV RdRp is assimilated to that of right hand, consisting of palm, thumb and a finger domain. Two loops emanate from the thumb domain and stack against the finger domain. From a structural point of view, the de novo replication implies that the enzyme adopts at least two different conformations. The conformations more frequently seen in crystal structure correspond to the “closed” or the “active” state where the the loops extending from the finger domain are in contact with the thumb. Disruption of the fingertip-thumb domain forces the protein into a more “open” inactive conformation. Many of the polymerase inhibitor works by breaking the fingertip-thumb connection. From a biochemical point of view, the polymerase works in three stages: 1) initiation, 2) transition and 3)elongation phase. In theory, polymerase inhibitors could potentially target all these steps either by acting at one or several of them.

Polymerase inhibitors falls into two categories: Nucleoside and non Nucleoside inhibiors. Nucleoside analogue inhibitors (NI), discovered through the rational search for substrate analogues, target the active site of polymerase, either competing with natural NTP substrates and/or acting as “chain terminators”, or causing a mutational “error catastrophe” by being incorporated into the elongating nascent RNA molecule. NIs are often endowed with a broad-spectrum antiviral activity, due to the fact that, during evolution, RNA viruses have remained unchanged certain sequences of their RdRp, by virtue of the key role in the process of transcription and replication of viral genomes. This is the case, for example, of the 2’-methyl nucleosides capable of inhibiting the multiplication of RNA viruses belonging to different families. Various nucleoside analog inhibitors are widely used for the treatment of human immunodeficiency virus (HIV), hepatitis B virus (HBV), and herpes simplex virus (HSV). A nucleoside analog drug ribavirin (1-beta-D-ribofuranosyl-1, 2, 4-triazole-3-carboxamide), a purine-analog with broad-spectrum antiviral activity,[6] is used for the treatment

of HCV infected patients. Recently, some nucleoside and non-nucleoside analogs have been tested as the potent candidates of anti-HCV agents. However, the effectiveness of nucleoside analogues has been reported to be compromised due to the generation of resistant mutants and the adverse side-effect, necessitating the development of new inhibition targets and inhibitors of NS5B. Accordingly, the non-nucleoside inhibitors binding to allosteric sites distinct from the active site have begun to be identified through high throughput screening (HTS) and crystallographic analysis of the inhibitor–NS5B complex. NNI target the alloenzyme free of substrate and un-complexed with any other non-structural replicative proteins. They are inactive when the enzyme has entered the processive RNA elongation phase[7] , which suggests that they target preferentially the initiation phase, by blocking the enzyme, thereby preventing conformational change needed for the polymerization.

Till date, four well documented NNI binding sites have been reported [8,9,10]. The derivatives of benzofuran and benzothiadazine have been reported to bind to two out the four allosteric binding sites present in the palm domain (fig 1, site C and site D)[11,12]. Benzofuran compound HCV-796 has been shown to have significant antiviral effects in patients chronically infected with HCV[13,14]. In addition two series of compounds based on the thiophene and benzimidazole (fig 1) scaffolds have been reported to inhibit NS5B by binding to two different binding pockets in the thumb domain of NS5B(site A and site B) [15,16]. NNI's have there own set backs including adverse side effects rapid selection of resistant virus mutant, or naturally occurring polymorphism in site “B” and “C”.[17,18]. This leads to an urgent need of applying several approaches like structure based drug designing, and chemical modifications[19-25] to increase the potency as well as to come up with new antivirals.

Here, we used enzyme based screening efforts to identify two compounds PS1097,PS1101 and PS1126 belonging to pyridoxoquinoxalines, phenanthrolines and imidazoquinolones class respectively, with high antiviral potency and selectivity against replication of HCV RdRp. Imidazoquinolines were first identified as compounds endowed with anti-viral activity in guinea pigs infected with HSV; they are now used as vaccine adjuvants to elicit effective anti-viral immune response. As far as we know, it is the first time that a direct antiviral activity of compounds belonging to the imidazoquinoline class against HCV RdRp is reported. Between these PS1097 seems to be highly selective for HCV. Most importantly PS1097 effectively reduced viral RNA with an IC_{50} ~ 7.5μM in the cell-based HCV replicon system suggesting that it was able to access the perinuclear membrane and inhibit polymerase activity in the context of the HCV replicase complex. Having said this, in absence of knowledge regarding the binding cavities as well the mechanism of inhibition, extensive molecular modelling studies, ranging from rigorous molecular docking to long molecular dynamics simulations, were performed. In order to identify probable binding site of the compounds,in order to rule out any bias, extensive molecular docking studies were performed on all the four NNI binding sites known so far. Further , several biophysical techniques were applied to investigate the binding affinities of the biologically relevant poses obtained from docking. Finally, hypothesis on the mechanism of inhibition for each compound was put forth.

Experimental Data:

Comp	Anti-BVDV Activity (μM)			IC_{50}	
	CC_{50}	EC_{50}	SI	BVDV	HCV
PS1126	>100	5.8	>17	97	2
PS1097	>100	6	>16	40	4
PS1101	63	9	7	0.8	1

Computational Recipe:

Receptors
Since we were clueless regarding the binding site of the three classes of inhibitors viz, PS1097(pyridoxoquinoxalines), PS1101(Phenanthrolines) and PS1126(Imidazoquinolines), we started off by making a catalog of all the available x ray structures of nearly all HCV RdRo-NNI complexes present in Protein Data Bank[26]. Our idea was to categorize the existing NNIs on the basis of the four known NNI-binding site(fig 1). This was done in order to identify key determinants viz, residue involved in Hydrogen Bonding(HB) with the inhibitors and those making favorable HpH interaction , for each of the four site.

Compounds belonging to N,N-disubstituted phenylalanine,[27,28], thiophene,[29,30], thiazol-one [31,32] dihydropyrones,[33,34]and acyl pyrrolidine [35,36] molecular scaffolds occupy a common binding site in the thumb subdomain (~35Å from the polymerase active site), denoted by site 'B' in fig1. Likewise, compounds belonging to benzothiadiazine analogues,[37,38] and acyl pyrrolidine analogues[39,40], benzylidene analogues [41,42] acrylic acid derivatives,[44] and proline sulfonamide analogues[43] bind in a cavity 10 Å away from the catalytic site in the palm domain.

A third class of compounds consists of indole-derived NNIs [45,46] that were reported to bind to an allosteric pocket in the thumb domain, close in space but clearly distinct (Figure 1) from site B. this allosteric site has been depicted as site A in fig1. Lastly, compounds belonging to class benzofuran have an affinity for a cavity present in the palm domain depicted as site D in fig.1.

With the extracted information from the reported PDB's of crystallized ligands corresponding to the key residues of each binding site, we performed subsequent molecular docking simulations, and ligand based alignments to locate the proper position of our newly reported compounds in each of the four binding site.
The X ray crystal structure used in this study is 1NB4.

Complexes.

Docking. Autodock 4 was used for all docking calculations[48]. The AutoDockTools (ADT) package was employed to generate the docking input files and analyze the docking results. The binding site of these three class of compounds (Imidazoquinoline; Pyridoxoquinoline, Phenanthroline) have not been reported. In the absence of literature data for the binding of these class of compounds, we performed molecular docking in all four binding sites taking carbon α atom of the respective resistant mutant P495, 423, 414, and C316 for binding site A, B, C and D respectively as a grid center, for each compounds. The interaction pattern and key residues of crystallized ligand (table1), which correspond to the binding site A, B, C, and D were used for the identification of proper orientation of our compounds into the four different binding site. Two different grid boxes, one for each binding site, were centered on the average mass center of the ligands. Thus, a grid box of 91 × 91 × 86 points and a grid spacing of 0.375Å was set in order to accommodate the NNIs that bind on to the site A and B of NS5B polymerase. The second grid box (spacing, 0.375Å) of 62 X 75 X 75 points was implemented for more buried cavity, in such a way to accommodate the NNIs that bind the site C and D of NS5B. The grid maps were generated for each atom probe, describing its interactions with the compounds. Autogrid 4, as implemented in the Autodock software package, was used to generate grid maps. The Lamarckian genetic algorithm (LGA)[48] was employed to generate orientations or conformations of the ligands within the binding site. The global optimization started with a population of 150 randomly positioned individuals, a maximum of 2.5×10^8 energy evaluations, and a maximum of 27,0000 generations, as mentioned the protocol for blind docking. A total of 200 runs was performed, while all the remaining run parameters were maintained at their default settings. A cluster analysis was carried out using 2 Å as the root-mean-square deviation tolerance. Docking experiments were also tried using a single grid box comprising the two binding sites. Although Autodock performed quite well in selectively positioning the ligands in the right pockets, the results in term of root-mean-square-deviation (rmsd) values were not fully satisfying. This is likely due to the fact that with only 100 runs such a large region is not sufficiently explored.

Assessment of Docking:
Redocking. Further refinement of docking was performed, aiming to check out the more realistically binding mode conformations of compounds with no experimental data. The choice of the best conformation was based on the assumption that, only the first ranked docked conformation should be considered (that is, the conformer characterized by the lowest estimated free energy of binding)[49,50] in other cases, the lowest energy conformer of the most populated cluster should also be taken into account. We applied the second protocol for investigation of binding mode of newly identified potent lead of different class of compounds .
MD Simulation: The protocol to rum MD simualtionis the same as mentioned in chapter1.

MM-PBSA:
The single trajectory approach [62,63] was used on 200 snapshots extracted from the MD dynamics of the RdRp-227G complexes were extracted to perform binding free energy calculation.
According to the MM-PBSA method[64,65] binding free energy (ΔG_{bind}) of each system could be conceptually summarized as follows:

$$\Delta G_{bind} = G_{com} - G_{rec} - G_{lig} \quad (1)$$

$$\Delta G = \Delta E_{MM} + \Delta G_{solv} - T\Delta S \quad (2)$$

in which:

$$\Delta E_{MM} = \Delta E_{bond} + \Delta E_{angle} + \Delta E_{torsion} + \Delta E_{vdw} + \Delta E_{EEL} \quad (3)$$

$$\Delta G_{solv} = \Delta G_{PB} + \Delta G_{SA} \quad (4)$$

$$\Delta G_{SA} = \gamma SA + b \quad (5)$$

G_{com}, G_{rec}, and G_{lig} are the free energy for the complex, receptor, and ligand (inhibitor), respectively. Each term is calculated by averaging the energy of molecular mechanics (ΔE_{MM}), the solvation free energy (ΔG_{solv}), and the vibrational entropy term ($T\Delta S$) as in (2). ΔE_{MM} (3) denotes the average molecular mechanics energy contributed by bonded (E_{bond}, E_{angle}, and $E_{torsion}$) and nonbonded (E_{vdw} and E_{EEL}) terms. The nonbonded contribution of the binding pockets were further evaluated on a residue basis using MM-PBSA.

ΔG_{solv} (4) is the solvation free energy calculated by summing polar (ΔG_{PB}, evaluated using the Poisson-Boltzmann equation) and nonpolar (ΔG_{SA}, proportional to the surface area, 3) terms. ΔG_{PB} was calculated using the DELPHI[65,66] package, with low dielectric medium for solute (ε=1) and high dielectric medium for solvent (ε=80). Atomic radii were taken from PARSE [64] with an additional value of 1.90 Å for phosphorus[67], while partial charges on solute were taken from the Amber 9 force field, the same used to perform MD simulations. An 80% boxfill lattice with grid spacing 0.5 grid/Å was applied, and 10,000 linear iteration steps were required to obtain energy convergence. The nonpolar contribution to solvation free energy was determined from equation [66], where the surface area was calculated using molsurf, and γ and b are 0.00542 kcal/mol Å^2 and 0.92 kcal/mol, respectively [64]. The solvent probe radius was set to 1.4 Å. Residues within 25 Å from the mass center of ligand were used for PBSA calculation (this included the ligand and water molecules (5 A from the drug)[69]

The solute entropy contribution (-TΔS) was estimated by normal-mode analysis [1] using the *nmode* module of AMBER 9.0. Before normal-mode analysis each of the complex, receptor, and ligand structures were optimized in the absence of explicit solvent and using a distance dependent dielectric constant of $\varepsilon = 4r$ (where r is the interatomic distance), to mimic solvent screening [68] until the convergence criteria of 0.0001 kcal/mol Å was achieved. Normal mode calculation is extremely time-consuming and computationally expensive for a large system; therefore, only residues within 10 Å from the mass center of inhibitor (including water molecules and ions) were used here [69]. As discussed in [70] and [71] the differences in the calculated entropy value is quite small for different conformations, and normal-mode analysis calculation usually gave an inaccurate estimation for solute entropies, thus giving merely qualitative estimates of the solute entropy.[72] Therefore, our calculation was only based on the average entropy value obtained from 100 snapshots taken from the MD trajectories of the extracted minima

Results

Docking analysis:

As already discussed in M&M, in order to rule out any bias, extensive docking protocol was done on all the Allosteric sites for NNI binding site A indole and benzimidazole[45], site B(dihydropyranone[33], thiophene carboxylic acids[27,20], cyclopentylethyl-phenoxy acrylic acids[42], site C benzothiadiazines[38], reversible arylsulfonyl rhodanines[39] and Site D[13,14,15] known till date [Fig-1]. In the absence of knowledge regarding the binding cavities of our compounds PS1097, PS1101 and PS1126, a carefully constructed multi-point strategy was considered, in order to select the most biologically relevant binding pose, for all the four binding cavities, which were to be considered as starting structures for further MD runs. The selection of the most probable binding pose also depended on the interactions made with the resistant mutations of the specific sites as well as the key determinants we identified on analysis of various crystal structures and which were used to define each cavity..

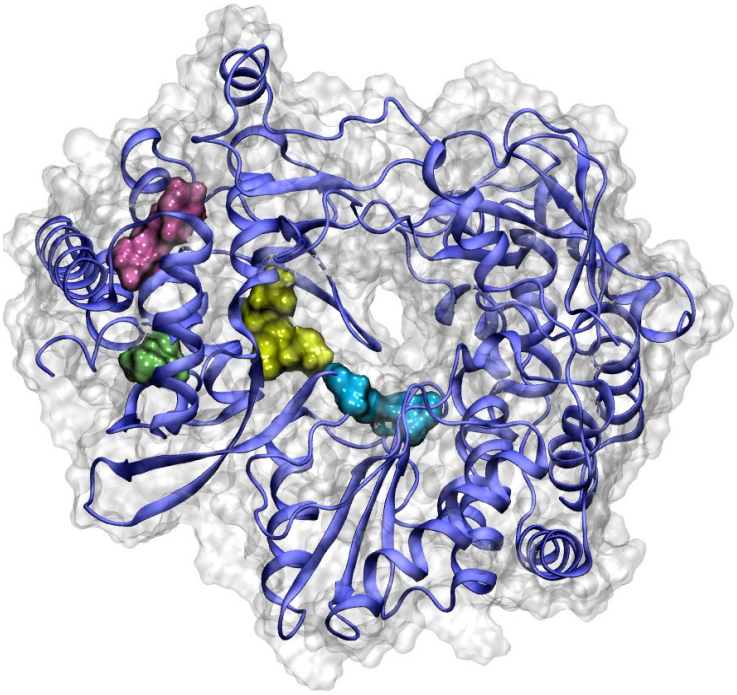


Figure1. Ribbon representation of the overall structure of HCV NS5B polymerase in blue along with transparent surface view in white. Solid surface representations of four allosteric binding sites are also shown. The dark red colored surface corresponds to the thumb allosteric binding pocket of indole group "siteA"; the dark green colored surface corresponds to the 2nd thumb allosteric binding site "site B", the dark yellow colored surface corresponds to the allosteric binding site situated in the palm domain "site C" and the dark skyblue colored surface correspond to 2nd palm domain binding site "site D".

Each of the four cavity was defined with reported resistant mutations. It has been reported that for majority of the compounds binding to the allosteric 'A' site, P495L act as a resistant[52] (Dutartre et al., 2005), for site B the resistant mutation identified is M423T mutant for the compounds [18](Pauwels et al., 2007), for site C M414T and Y448H demonstrated a clear and universal resistance to the inhibitors binding to this site, while for site D C316Y is reported to inhibit the activity of the compounds.

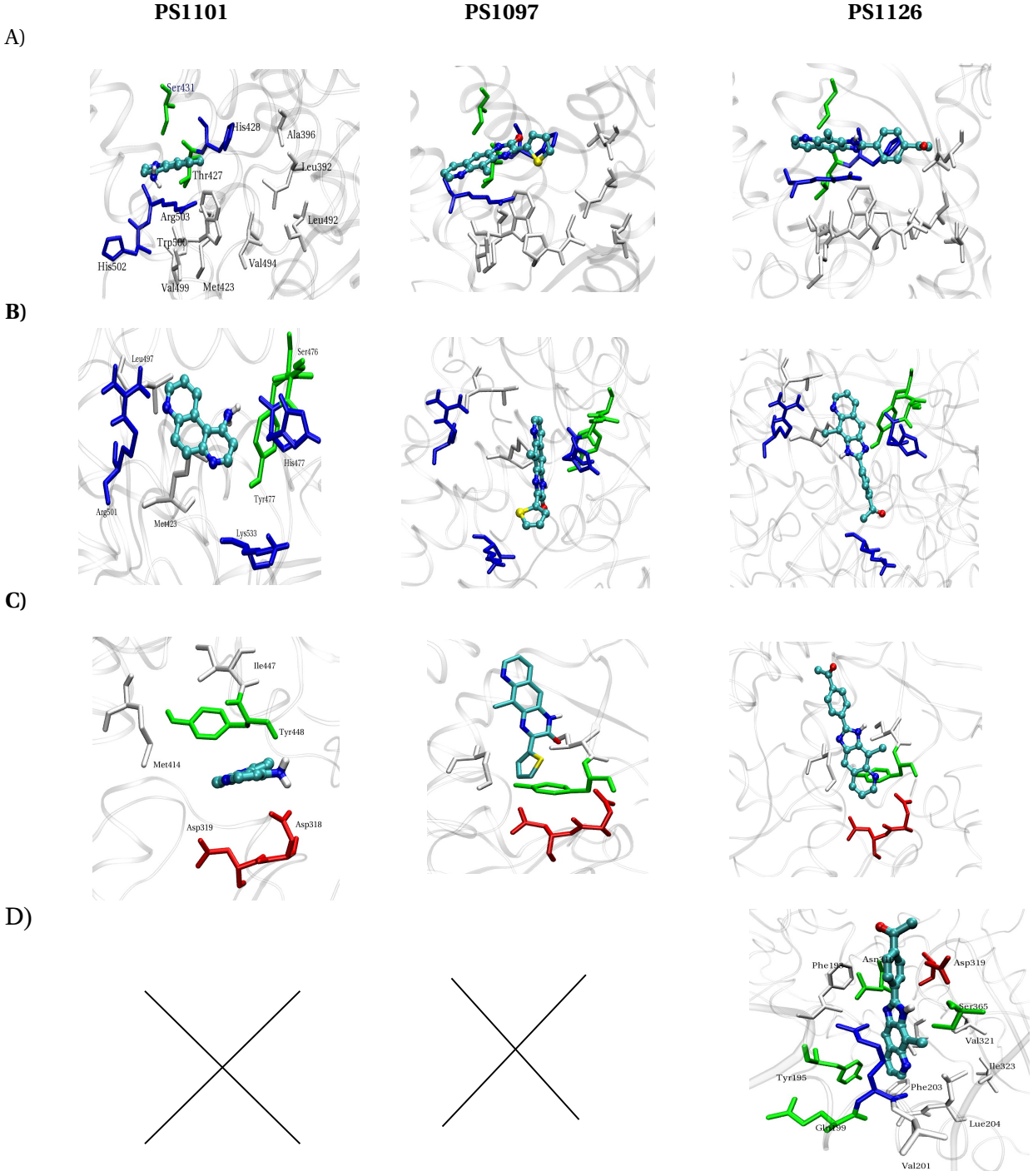


Figure 2: Depicts the final docking conformations taken for each compound in each of the four binding site. Residues defining the cavity, as obatined by our analysis of X ray structures are presented as licorice. It should be noted here that while in case of PS1126 we were able to get good clusters in all the four site, on the contrary , there docking was not succefull for site D in case of PS1101 and PS1097.

Apart from following a normal procedure of picking up orientations on the basis of high population of conformers and low energy, we also did a thorough investigation of each conformer from the clusters obtained on the basis of their vicinity to the resistance mutations as well as the interactions made by the compounds with the key structural determinants of the binding cavity. Applying this strategy we came out with one conformer, for PS1126, in each site, While Docking failed to produce good clusters in case of site D for PS1101 and PS1097(Fig.2). Finally we started with three starting conformers each for compounds PS1101 and PS1126 in site A, B and C while 4 conformers for PS1126, which were found to be within attacking distance with the reported resistant mutations for respective sites.

MD Analysis:
For all three compounds, we launched 10MD simulation for cavity A, B, C and D. For compound PS1097, there was an early dissociation from the cavities A, B, and D. while the conformer we obtained in site C was found to be stable throughout the 20ns. Same way the compound PS1101 was also found stable up to 20ns in cavity C. While compound PS1126, starting structure of site D was stable while in the rest, there was an unbinding and movement of the compound away from the binding cavity. The survived launches were then considered for detailed analysis of the binding features, which are described in the following section.

Compound1(PS1126)

PS1126 was found to bind in a cavity reported for binding of benzofuran (also known as "Hinge" inhibitor) compounds [11,12]. In contrast to the broad open surface exposed in the palm1 binding pocket is deep, narrow, and enclosed descending from the palm C binding site. NNI binding site in palm2 has unique structural features. The deep enclosure within the core of the conserved and highly structured palm domain, featuring narrow entry and exit portals.The cavity is predominantly hydrophobic in nature. Its an elongated and highly enclosed pocket present in the palm domain of HCV RdRp. It lies between primer grip motif (residue I363-N369) and beta sheets (residue P214-Y219, D319-G325).[15]. According to the data obtained by the x-ray crystallographic structure FQK[15], the pocket is formed by residues Gly557, Gly449, Gly554, Tyr448, Gln199, Val201, Leu204, Ile323, Leu314, Ser365, Phe203, Asp319, Asn316, Tyr195, Phe193, Val321, Arg200, Pro197, Leu360 and Ile363. Total interaction energy between PS1126 and the binding pocket was calculated as described in material and method. The total energy came out to be -32.6 kcal/mol. Fig. 3b summarizes the Total interaction energy contributions (favorable hydrophobic as well electrostatics interactions) between PS1126 and the Binding pocket. Major contribution to the stability of PS1126 in the pocket is provided by the VdW interactions (-25.1kcal/mol) while the total electrostatic contributions are -7.1kcal/mol.

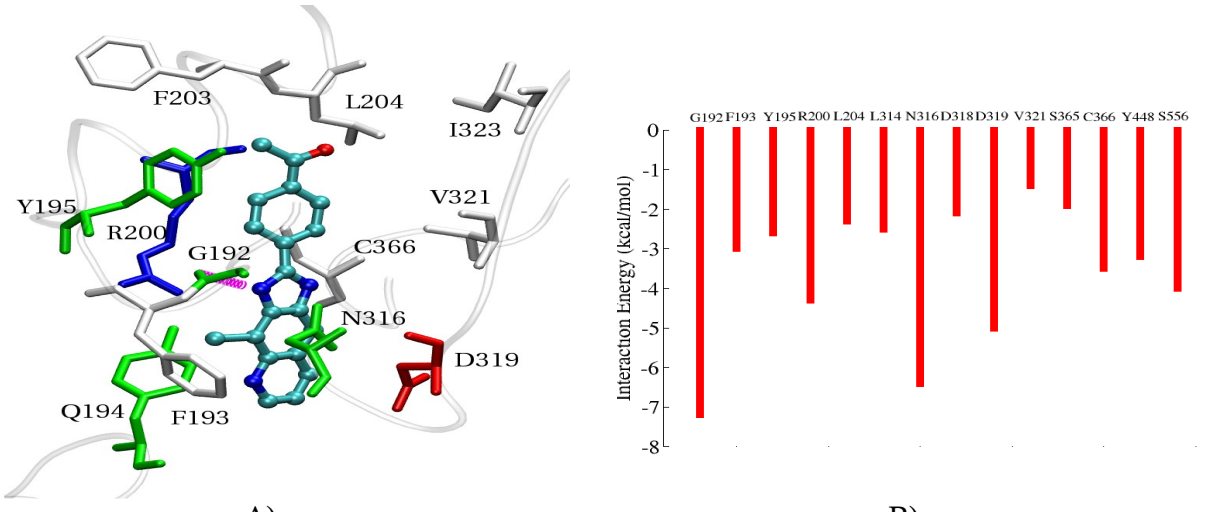


Figure3: Expanded view of interaction between key residues of binding cavity of PS1126. Snapshot taken from the average structure from the equilibrium of 20ns MD simulation. The Binding site residues are shown here using sticks representation(basic:blue, acidic:red, hydrophobic:white) while compounds are in ball and stick in cyan color. In panel (A) and (B)Interaction map and interaction energy of PS1126 Average ligand-residue interaction energies in kcal/mol over compound PS1126 for the residues that contribute most to the ligand -surrounding.

Table:1 Interaction Energy of key residues in PS1126-RdRp complex (Kcal/mol)

Residue	Ele	VdW	Total
Phe193	-0.5(±0.3)	-2.4(±0.8)	-2.9(±0.9)
Arg200	1.5(±2.1)	-6.0(±1.0)	-4.3(±2.4)
Leu204	-3.4(±0.4)	-2.0(±0.5)	-2.3(±0.6)
Leu314	0.7(±0.2)	-3.3(±0.5)	-2.6(±0.6)
Asn316	-2.4(±1.2)	-4.1(±0.7)	-6.4(±1.2)
Asp318	-2.0(±0.5)	0.2(±0.0)	-2.1(±0.7)
Asp319	-3.1(±1.4)	-2.0(±0.6)	-5.0(±1.5)
Val321	-0.3(±0.2)	-1.1(±0.4)	-1.4(±0.5)
Ser365	-0.8(±2.2)	-2.0(±0.5)	-2.0(±2.1)
Ser366	-0.7(±0.6)	-2.8(±0.7)	-3.5(±1.0)
Tyr448	-0.17(±0.1)	-3.1(±0.7)	-3.2(±1.1)

A strong network of HB was observed which stabilized PS1126 into the binding cavity. PS1126 accommodates itself in such way that a HB with occupancy more than 70.4% exist throughout the simulation of 20ns with G192@N1. Likewise, another HB is formed between Ser556 and N2 atom of PS1126. Furthermore, another HB between Cys366 and N3 atom with occupancy more than 25% helps in locking the PS1126 into the binding cavity. The comparative analysis of the different crystal [15] structures of inhibitors binding to the binding cavity D showed that the residues Asn316 and Arg200 played crucial role in stabilizing the inhibitors. Analyses of the Residue wise contribution to the interaction energy showed that in our case , residue Asn316(-6.4 Kcal/mol) and Arg200(-4.3Kcal/mol), both present at entrance to the binding cavity, were major contributors to the total interaction energy. Further residues Ser556, Gly192, Asp319 and Tyr448 provide stability to PS1126 into the cavity by providing favorable HpH and electrostatic contacts as is listed in table 1.

Arg200 as a trigger to the switch for opening of the otherwise buried cavity

Site D inhibitors are often called as 'Hinge inhibitor' [8,22,52] as they bind to the palm domain and induce a hinge movement of the active site Arg200 side chain to generate a relatively small pocket where the inhibitor is accommodated. Based on this hypothesis, we performed a more in depth analysis of the local conformational changes taking place at the mouth of the cavity, which may lead to the opening and eventually creation of a pocket wherein PS1126 gets accommodated.

Deep investigation of the MD simulation of the Apo system revealed that three residues Arg200 (conserved core of palm domain), Cys366(primer grip) and Asn316(beta sheet of GDD motif) lying at the mouth of the cavity, interact with each other and act as a "*gate*" for the opening of the buried cavity. Arg200 interact with the Asn316 making strong HB with occupancy more than 80%(ADL 2.0 A), thereby closing the entrance to the cavity, as can be seen in Fig??. The residues undergo conformational changes, including breaking of HB between Arg200 and Asn316, allowing passage for entry of PS1126 into the cavity. (FIG).
In order to further explore events which lead to the opening of the buried cavity C, we superimposed average structures of the Apo, complex system and crystal structure. Interestingly, our computational simulations indicated that Arg200 and Cys366, underwent significant conformational changes upon PS1126 binding. We observed that side chain of Arg200 in crystal structure completely hindered the entry to the binding pocket while the average structure of complex showed a complete deviation of the side chain away from the cavity.

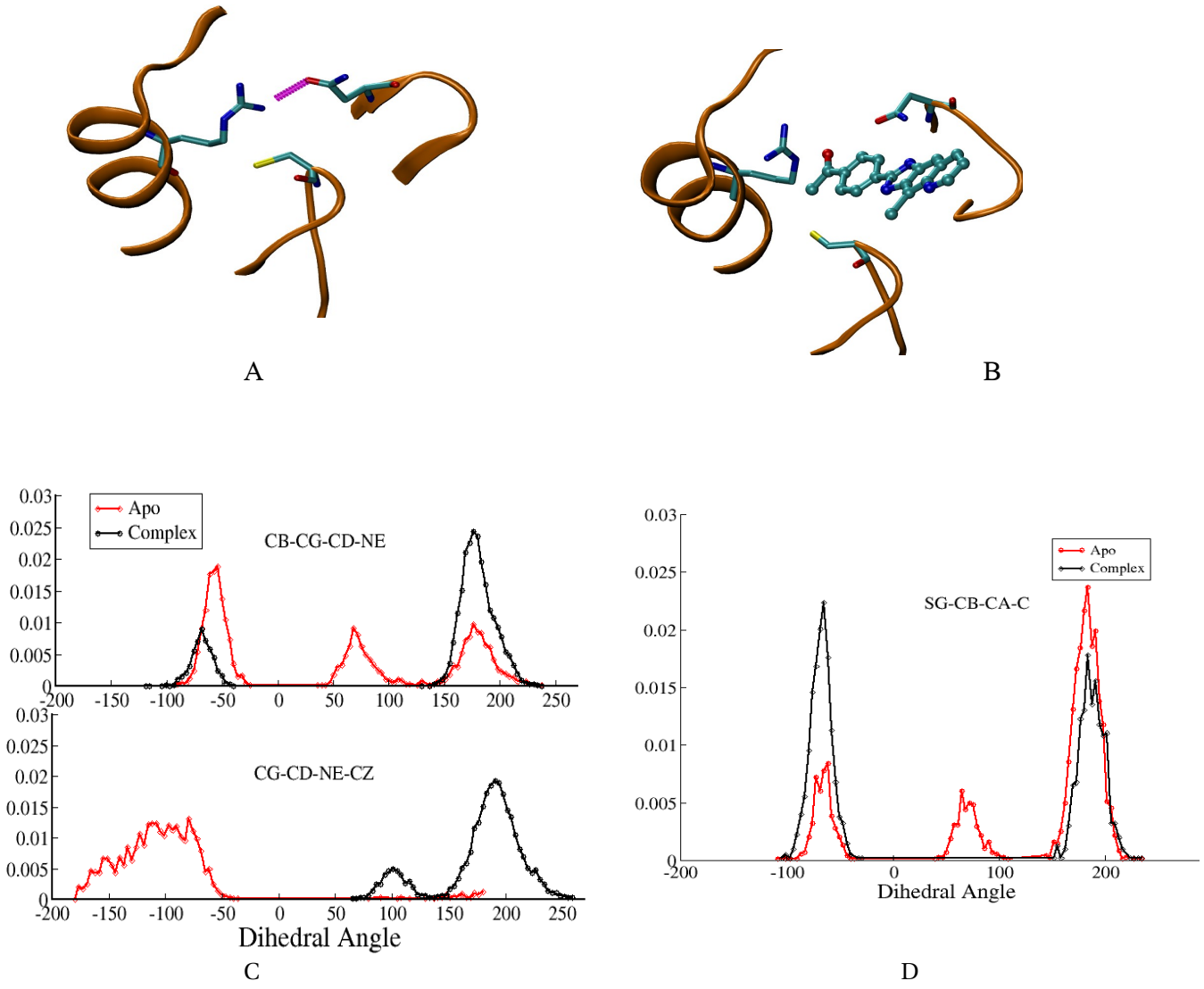


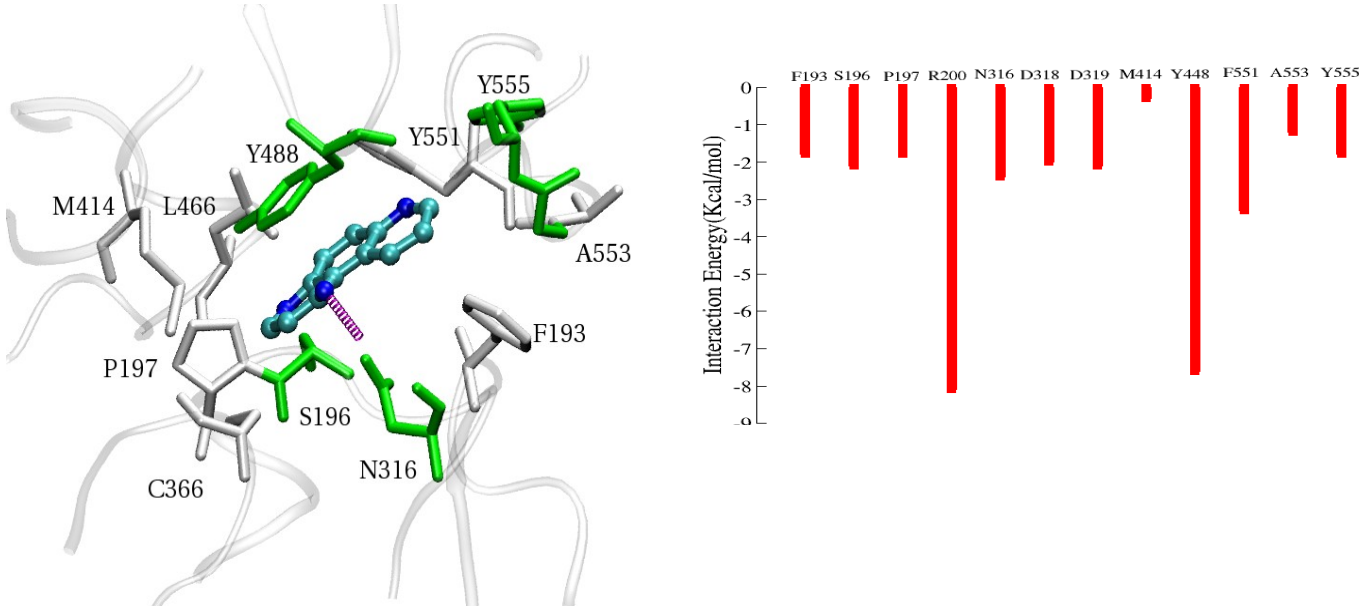
Figure4: In the panel one; the three key residues Arg200, Asn316 and Cys366 shown in stick. The arrangement of these residues, " *hinge*", with the spiral ring represent the HB between Arg200 and Asn316, which opens in presence of PS1126. In the panel two, the possible population states for torsions of residue Arg200. The atoms forming ramachandran torsions (Φ, Ψ) are : Φ torsion are:CB-CG-CD-NE; and Ψ torsion are : CG-CD-NE-CZ. The Ramachandran dihedral illustrating the change in the side chain dihedral angles upon ligand binding. In panel (C)-for Φ torsion(Apo) [-80:-25], [30:120] and [150:225] and for complex [-100:-40], [130:225]. In panel (D)- Ψ torsion are [-180:-50] for Apo and for complex [150:250]. The panel C is the population states of Cys366 in presence of ligand. The distribution for Apo are [-100: -40], [50:110] and [150:215], while for complex [-100:-40] and [160:230].

With the above analysis, we inferred that the residue Arg200 and Cys366 act as gate for the entry of PS1126 into the otherwise buried cavity. Interaction of these residues with PS1126 throws them into an open confirmation providing a clear passage of the inhibitor into the binding cavity.

further, we also observed a different dihedral transition between the Apo and the complex structure as shown in fig. In the Apo system dihedral CB-CG-CD-NE of the Arg200 side chain showed three phase distribution, exploring one conformation more than the other two, indicating more freedom of movement in the APO state. However, on binding of PS1126 the side chain was induced into a more favorable orientation to accommodate the subsequent binding of PS1126. The difference in the conformations explored by the side chain of Arg200 was more pronounced in dihedral CG-CD-NE-CZ, wherein , the side chain in case of Apo occupied an altogether different conformation space as compared to the complex systems. All these, study supported the view that entry of PS1126 induced a reorientation of the side chain of Arg200 into a more open state. Furthermore, visual inspection of the trajectory indicated that another major reorientation took place in the Cys366 residue of the primer grip. The side chain of Cys366 exhibited three kinds of orientation during the MD simulation for the free enzyme while a biphasic behavior was exhibited in the complex indicating a more restricted movement favorable for the binding of PS1126.

PS1101

PS1101 was found to be stable in Site C, which is located in the core of the polymerase in proximity to the enzyme active site and at the interface between thumb and palm domains. PS1101 is surrounded by several sequentially disparate and important features of the polymerase . The cavity is flanked by primer grip(363 to 369)or the motif 'E', that links the palm and the thumb domains, the C-terminal tail that protrude towards and into the catalytic core, the β flap (442 to 452)and the active site GDD (317, 318, 319) motif.(ref).The cavity is primarily hydrophobic in nature. Residues lining the cavity include Phe193, Ser196, Pro197, Arg200, Asn316, Asp318, Asp319, Met414, Tyr448, Phe551, Ala553 and Tyr555.



A) B)

Figure5: Expanded view of interaction between key residues of binding cavity of PS1097, PS1101 and PS1126. Snapshot taken from the average structure from the equilibrium of 20ns MD simulation. The Binding site residues are shown here using sticks representation(basic:blue, acidic:red, hydrophobic:white) while compounds are in ball and stick in cyan color. In panel (A) and (B)Interaction map and interaction energy of PS1097; panel (C) and (D) Interaction map and interaction energy of PS1101 and panel (E) and (F) are Interaction map and interaction energy of PS1126. Average ligand-residue interaction energies in kcal/mol over compound PS1097, PS1101 and PS1126 for the residues that contribute most to the ligand -surrounding.

Residue	Ele	VdW	Total
Phe193	1.9(±0.4)	-4.0(±0.6)	-2.0(±0.6)
Ser196	-1.0(±1.0)	-1.3(±0.5)	-2.1(±1.0)
Pro197	-0.2(±0.2)	-2.0(±0.4)	-2.0(±0.4)
Arg200	-4.5(±2.4)	-3.5(±0.7)	-8.1(±2.6)
Asn316	-1.4(±1.2)	-1.5(±0.4)	-3.0(±1.2)
Asp318	-1.9(±0.6)	-0.1(±0.1)	-2.0(±0.7)
Asp319	-2.0(±0.7)	-0.1(±0.0)	-2.1(±0.7)
Tyr448	0.3(±0.8)	-8.0(±0.8)	-8.0(±1.2)
Phe551	-0.6(±0.5)	-3.0(±0.7)	-3.3(±0.7)
Ala553	-1.0(±0.4)	-0.4(±0.4)	-1.4(±)
Tyr555	-0.3(±0.3)	-1.5(±0.4)	-2.0(±0.5)

Table2: Interaction Energy of key residues in PS1101-RdRp complex (Kcal/mol)

PS1101 stabilities itself into the cavity by making a strong hydrogen bond with the N3@ PS1101 and HD@Cys316. The HB persist throughout the 20 ns simulation with an occupancy of more than 65% with average dynamic length ~2.8A . Owing to the highly hydrophobic nature of the cavity the major contribution to the total interaction energy (-34Kcal/mol) is provided by the hydrophobic interactions (-25.0 kcal/mol). The residue wise contribution to the energy is listed in table2. Residue Tyr448 belonging to the β hairpin loop makes maximum hydrophobic interaction (-7.9 kcal/mol) with PS1101. In depth analysis of the trajectory revealed that a strong π-π stacking exist between the phenyl ring of Tyr448 and PS1101 throughout the simulation thereby providing additional stability to PS1101 into the binding cavity. Additionally, Arg200 provides further support to PS1101 into the binding cavity by contributing favorably to the TOT interaction energy (-8.1kcal/mol). Other residues which contribute to stabilize PS1101 into the cavity include Cys366 (primer grip), Phe193 and Pro197 (palm domain) .Fig(5).

PS1097

PS1097 accommodated itself in cavity which can be characterized as an extended Site C. The wall of the binding cavity include residue from the palm and the thumb domain. The cavity is overlapping but more buried with the one of PS1101. Predominantly, the major interactions occur between PS1097 and the the residues of Beta hairpin loop, inner core of the thumb domain and the primer grip.

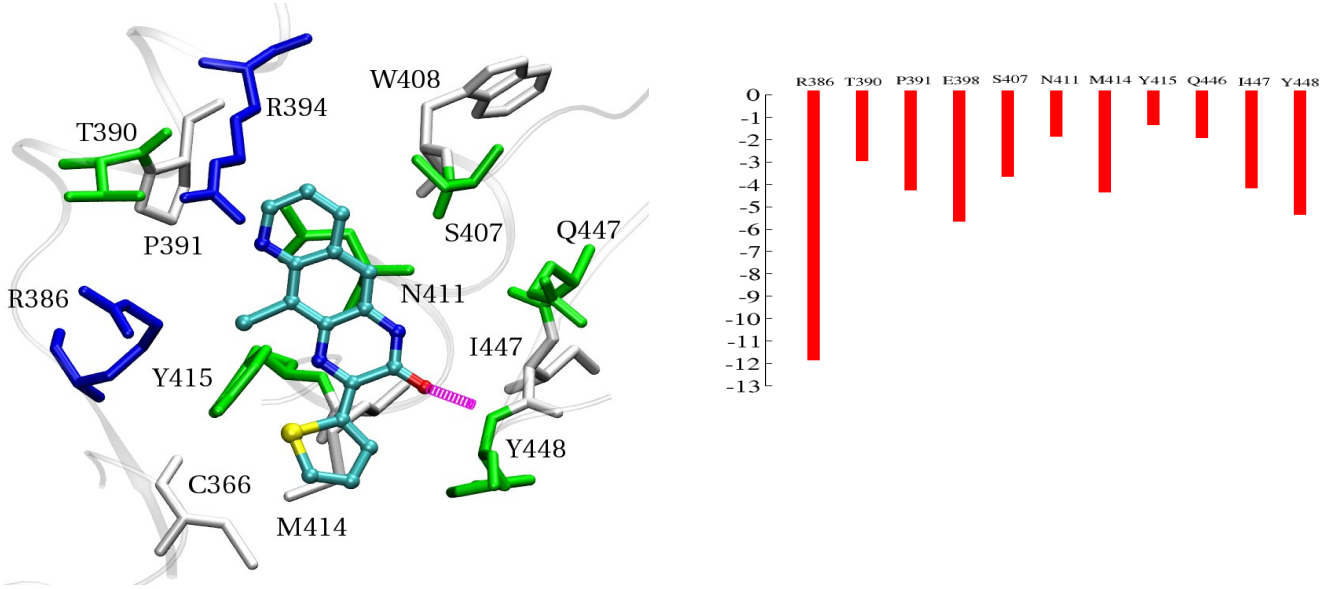


Figure6: Expanded view of interaction between key residues of binding cavity of PS1097, PS1101 and PS1126. Snapshot taken from the average structure from the equilibrium of 20ns MD simulation. The Binding site residues are shown here using sticks representation(basic:blue, acidic:red, hydrophobic:white) while compounds are in ball and stick in cyan color. In panel (A) and (B)Interaction map and interaction energy of PS1097; panel (C) and (D) Interaction map and interaction energy of PS1101 and panel (E) and (F) are Interaction map and interaction energy of PS1126. Average ligand-residue interaction energies in kcal/mol over compound PS1097, PS1101 and PS1126 for the residues that contribute most to the ligand -surrounding.

Residue	**Ele**	**VdW**	**Total**
Arg386	-9.9(±1.5)	-1.7(±0.4)	-11.7(±1.7)
Thr390	-1.7(±0.6)	-1.1(±0.3)	-2.8(±0.7)
Pro391	-1.6(±0.4)	-2.3(±0.5)	-4.1(±0.7)
Glu398	-5.4(±1.0)	-0.2(±0.1)	-5.5(±1.1)
Ser407	-2.0(±0.8)	-1.6(±0.4)	-3.5(±1.3)
Asn411	2.7(±1.2)	-4.4(±0.8)	-2.0(±1.3)
Met414	-0.1(±0.3)	-4.1(±0.6)	-4.2(±0.6)
Tyr415	1.6(±0.3)	-3.0(±0.4)	-1.2(±0.4)
Gln446	-1.2(±0.8)	-0.5(±0.1)	-2.0(±0.8)
Ile447	-2.4(±0.5)	-2.0(±0.3)	-4.0(±0.8)
Tyr448	-2.9(±1.0)	-2.3(±0.7)	-5.2(±1.0)

Table3 : Interaction Energy of key residues in PS1097-RdRp complex (Kcal/mol)

PS1097 is located in the cavity in such a way that a strong HB(Hydrogen Bond) is formed between O atom of PS1097 and Tyr448 of the β hair pin. The HB persist throughout the 20ns of simulation with an occupancy of more than 80%. Detailed study on crystal structures (2GIQ,3CO9,2QE2) of different classes of compounds binding to the the cavity C exhibited that in majority of the cases HB between the inhibitors with Tyr448 was very common. This showed that, in the absence of the experimental data indicating the binding cavity of the pyridoquinoxaline compounds, the binding orientation, we obtained from our computational methodologies were coinciding well with the already published crystal structures. To quantify the energetic contribution of binding site residues we performed interaction energy analysis. we observed the total interaction energy between PS1097 and the binding pocket of PS1097, came out to be 48 kcal/mol. In PS1097-HCV RdRp system, both electrostatic(-23.3 kcal/mol) and VdW(-24.5 kcal/mol) contributed all most equally to the total interaction energy. Major contributors of the hydrophobic contacts(HpH) are from residues Asn411, Met414, Pro391, Tyr448 and Tyr415 while residues Arg386, Tyr448, Ile447 and Glu398 provided strong electrostatic interaction crucial for the stability of the PS1097 in the binding pocket. (Table3, Fig 6)

1)-Mechanism of Inhibition:
In case of HCV RdRp the active site is encircled by a β hairpin loop(A442-Y452) and the C-terminus(L545-H562) which fold into the cavity.[53,54]. Due to the strategic positioning of the β loop as well as the C-terminus region at the Exit site, it is believed that in inactive state, both the β loop and the C-terminus blocks the channel required for the double stranded RNA product to exit the polymerase. [55](Butcher, S. J.;Nature 2001, 410, 235-240.) literature study have indicated that these regions may be involved in template binding or they may act as a scaffold for the initiation complex. In order to give passage to the newly formed double stranded RNA product, it is mandatory for the beta hairpin as well as C-terminus region to undergo local conformational changes giving way for the exit to the polymerization product.

1.1) **Mechanism of Inhibition of PS1101:**
PS1101 binds in the cavity in such a way that it makes strong interactions with the residues of most important features of HCV RdRp viz: the primer grip, beta hairpin loop, C-terminal, and the GDD motif. This opens up myraid ways of mechanism of inhibition of HCV Polymerase by PS1101. Ps1101 strongly interacts with the Beta hairpin loop as well the Cterminal residue, thereby it may restrict the movement of the Beta loop as well Cterminal which are necessary to give passage for the exit of the newly formed RNA double strand. We observed that binding of PS1101 blocked the exit channel of the HCV RdRp which can be clearly seen in fig?.. . Hence it could meddle with the exit of the polymerization product.

furthermore, since PS1101 interacts with GDD motif it is more likely that it could interfere with the active site or may compete with the nucleotide binding which may eventually result into freezing of the enzyme into its RNA synthesis initiation conformation when the first phophodiester bond formation occurs.[56](stephane betzi et al;2009;antiviral research)

1.2)- **Mechanism of Inhibition of PS1097:**
As is the case with some of the inhibitors of different classes[18](2AX1) binding to site C PS1097 was found to bind approx 10 Ang away from the GDD motif , making it more unlikely to interfere with the active site or compete with the nucleotide binding.
PS1097 was found to be sandwiched between the β hairpin loop and residues belonging to the inner core of thumb domain. PS1101 is stabilised in the cavity by making favorable interaction with residues Glu446 and Met447(fig?) and also a strong HB with Tyr 448. All the former mentioned residues are located in the beta hairpin loop, which in turn is located next to the active site and is thought to be involved in *de-novo* initiation[25]
We propose a two point hypothesis regarding the mechanism of inhibition of HCV RdRp due to PS1097.
1)-Strong interactions between the PS1097 and β-hairpin loop tends to restrict the latter. This could freeze any conformational change in the thumb domain which is required to bind to the template or form the platform for the initiation complex.
2)-PS1097 binds in the channel needed for the exit of the neo-synthesized dsRNA. Thus binding of PS1097 not only blocks the exit but also hampers the local conformational change required in β hairpin loop by limiting its flexibility .

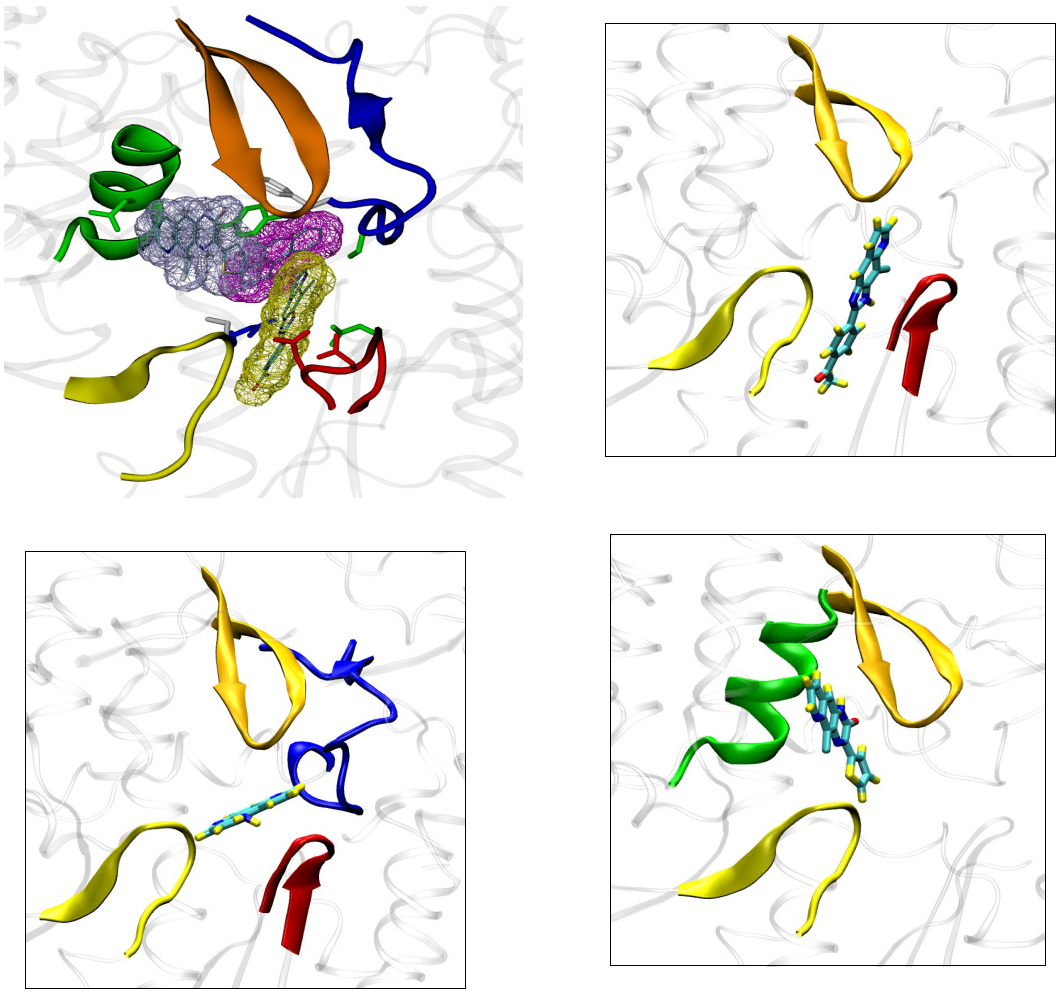


Figure7: Key structure playing crucial role in polymerization. These five structurally and functionally important region shown as β hairpin in brown, C-terminal in blue, inner core of thumb domain in green , primer-grip in yellow and catalytic triad, red in color. The compounds PS1097(sky blue), PS1101 (mauve), and PS1126(yellow)in wireframe into their specific binding mode in their binding site. The resistance mutation of that corresponding residues are shown in stick (res-type).

1.3)-Mechanism of Inhibition of PS1126:
A detailed analysis of the interaction pattern exhibited by PS1126 showed that the inhibitory effect may have many origins. literature studies have put forth that site C inhibitors which interact with the β hairpin loop may reduce the flexibility of the β hairpin and eventually of the thumb domain which is crucial for template binding.[25]. In our case PS1126 accommodates itself in such a way that it makes favorable interaction with Tyr448 belonging to the beta hairpin loop. The interactions with the beta hairpin loop may results into loss of conformational changes of thumb domain needed for a more concerted movement between the thumb and the finger domain essential for template translocation.

Furthermore, PS1126 interacts with the primer grip by making favorable interactions with Ser365 region. This makes it more likely that the binding of the primer grip could effect the binding of the substarte at the active site or primer binding via stearic hinderance.
literature study have indicated that compound HCV-796 may inhibit the HCV RdRp either by preventing incorporation of nucleotides or may impair the ability of the enzyme to discriminate between ssRNA and dsRNA[57] . PS1126 also interacts with residue Asp319 belonging to the GDD motif of the palm domain. this indicate that the inhibitor works by either hampering the incorporation of nucleotide or the formation of ternary complex itself.

Table:4 Comparison of Calculated and Experimental Binding Affinities.

	PS1097	**PS1101**	**PS1126**
ΔE_{EL}	-36.0(±1.9)	-42.9(±2.7)	-19.8
ΔE_{VDW}	-52.17(±2.9)	-39.9(±1.8)	-25.9
ΔG_{PB}	24.1(±2.7)	13.7(±3.2)	1.5
ΔG_{NP}	-5.8(±1.2)	-5.6(±0.9)	-4.7
PB_{tot}	-33.9(±3.1)	-29.8(±3.4)	-30.1
ΔG_{EL+PB}	-11.9(±3.9)	-29.2(±2.6)	-18.3
ΔG_{vdw+NP}	-57.9(±4.7)	-45.6(±3.8)	-30.6
$T\Delta S_{solute}$	-18.9(±2.3)	-17.6(±1.8)	-21.6(±2.1)
ΔH	-69.8	-74.8	0.3
ΔG^{a}	-11.6	-10.2	-8.5
ΔG^{b}	-8.8	-8.1	-7.7
$IC_{50}{}^{d}$	0.04	1.0	2.0

Table4: The experimental values of ΔG_{bind} were calculated from experimental IC_{50} value by fallowing relationship

$$\Delta G_{bind} = RT \ln IC_{50}$$

(R is the universal gas constant, T = 298K)
a- Calculated value
b-Experimental value
c-not available
d-experimental IC_{50} value in μmolar
All the energies are in kcal/mol and all IC_{50} values are in micro molar.

Slight deviation in the binding affinity were reasonable given the approximations that were made in the simulations as the conformational change in the ligand and/or receptor binding were neglected because of our use of a single trajectory in the bound and the unbound state.

Table 4 also displays binding free energy components that help us understand the dominant ligand structural determinants of the magnitude dependence of binding. The 'gas phase' van der waal energies were found to be the most important in complex formation in all the three inhibitor-complex systems and their values are -37.5 kcal/mol, -35.6 kcal/mol and -39.6kcal mole in case of PS1097, PS1101 and PS1126 respectively. The 'gas phase' electrostatic energies are -20.5 kcal/mol, -9.5 kcal/mol and -24.3kcal/mol for the binding of PS1097, PS1101 and PS1126 respectively. The non-polar solvation free energy favors the complex formation, which contributes -3.5 kcal/mol, -2.9kcal/mol and -3.8kcal/mol for the case of PS1097, PS1101 and PS1126 respectively.

The stability of the structure in the simulations and the close reproduction of the experimental binding mode for all the three compounds PS1101, PS1097 and PS1126 suggest that it is reasonable to use the simulation data for further analysis. The binding free energy calculated were compared using the MMPBSA approach are compared with the experimental free energies of binding (absolute affinities, Table 1;). The calculations predict that binding affinities of PS1101, PS1097 and PS1126 are -12.2 kcal/mol, -11.6 kcal/mol and -11.5 kcal/mol respectively. It is noteworthy that the MMPBSA predicted binding free energies were found to be in good agreement with the experimental data where the experimental binding affinity for the PS1101, PS1097 and PS1126 were found to be -8.1 kcal/mol, -8.8 kcal/mol and -7.7 kcal/mol respectively.

Table 4 reports contribution of the hydrophobic/apolar and electrostatic interactions to the susbtrate-binding free energy. We found that the predicted binding free energy is dominated by the magnitude of the apolar contribution(delGvdw + delGnp). Due to the result of highly unfavorable electrostatic solvation energy in Table 1(delGpb, 26.1, 24.2 and 33.6 kcal/mol), the contribution of electrostatic interactions to the solvation free energy (delGsolv) is unfavorable for three inhibitor-protein complexes(Table2). Table 2 examine the total binding free energy contribution in terms of components dominate by distinct apolar/hydrophobic and electrostatic terms in the inhibitor-substrate interactions and solvation terms. The two component extracted in Table 2 underscore the fact that the polar/hydrophobic contribution dominate the binding free energy of our three inhibitors in their respective binding pockets. The present binding free energy results confirm that their binding affinities are determined by apolar/hydrophobic terms rather than polar interactions with the binding-site residues.

Conclusion:

Here, we have reported , for the first time, three compounds PS1097 PS1101 and PS1126 belonging to three different classes viz; Phenonthrolines, pyridoxoquinoxalines and imidazoquinilines active agianst HCV RdRp at a very low micrmolar range.

Extensive Insilico study helped us to identify putative Binding cavity of the compounds belonging to Pyridoxoquinoxaline , Imidazoquinoxaline and phenonthrolines scaffolds. This was an encouraging performance, given that all molecular modeling studies were performed in the absence of any available crystal structure of the protein in complex with these classes of inhibitors.

Key binding pattern of the compounds in the cavity were identified with an aim of further optimisation of the lead compounds in order to increase their potency.

Also, the adopted procedure was used to predict drug binding affinities between the inhibitors and HCV RdRp. The Calculated binding free energies were found to be in agreement with the experimental data. Furthermore, we were successful in throwing light on the probable mechanism of inhibition of the three compounds reported in this paper.

In conclusion, our hierarchical molecular modeling procedure achieved a high reliability,thus constituting an attractive strategy in drug lead exploration. From this point of view, the presented results might provide useful hints in guiding the rational design more selective and potent compounds against HCV RdRp.

References:

1. WHO (January 2006, posting date). Initiative for Vaccine Research (IVR): hepatitis C [Online]. http://www.who.int/vaccine_research/diseases/viral_cancers/en/ index2.html.;
2. Feld, J. J. & Hoofnagle, J. H. (2005). Mechanism of action of interferon and ribavirin in treatment of hepatitis C. Nature, 436, 967–972.
3. Biron CA, Sen GC (2001) Interferons and Other Cytokines. In: Knipe DM, Howley PM, editors. Fundamental Virology. 4th ed. Philadelphia: Lippincott Williams & Wilkins. pp. 321–351
4. Fried MW, Shiffman ML, Reddy KR, Smith C, Marinos G, et al. (2002) Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. N Engl J Med 347: 975–982.
5. Manns MP, McHutchison JG, Gordon SC, Rustgi VK, Shiffman M, et al. (2001) Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomised trial. Lancet 358: 958–965.
6. Potter CW, Phair JP, Vodinelich L, Fenton R,Jennings R.Antiviral, immunosuppressive and antitumour effects of ribavirin.Nature1976; 259:496-7.
7. Beaulieu, 2 Ma et al., 2005
8. De Francesco, R., Carfi, A., 2007. Advances in the development of new therapeutic agents targeting the NS3-4A serine protease or the NS5B RNA-dependent RNA polymerase of the hepatitis C virus. Adv. Drug Deliv. Rev. 59, 1242–1262.
9. De Francesco, R., Migliaccio, G., 2005. Challenges and successes in developing new therapies for hepatitis C. Nature 436, 953–960.
10. Betzi, S., Suhre, K., Chétrit, B., Guerlesquin, F., Morelli, X., 2006. GFscore: a general nonlinear consensus scoring function for high-throughput docking. J. Chem. Inf. Model. 46, 1704–1712. Biswal, B.K., Cherney, M.M., Wang, M., Chan, L., Yannopoulos, C.G., Bilimoria, D., Nicolas, O., Bedard, J., James, M.N., 2005. Crystal structures of the RNA-dependent RNA polymerase genotype 2a of hepatitis C virus reveal two conformations and suggest mechanisms of inhibition by non-nucleoside inhibitors. J. Biol. Chem.280, 18202–18210.
11. Dhanak, D.; Duffy, K. J.; Johnston, V. K.; Lin-Goerke, J.; Darcy, M.; Shaw, A. N.; Gu, B.; Silverman, C.; Gates, A. T.; Nonnemacher, M. R.; Earnshaw, D. L.; Casper, D. J.; Kaura, A.; Baker, A.;Greenwood, C.; Gutshall, L. L.; Maley, D.; DelVecchio, A.;Macarron, R.; Hofmann, G. A.; Alnoah, Z.; Cheng, H. Y.; Chan,G.; Khandekar, S.; Keenan, R. M.; Sarisky, R. T. Identification and biological characterization of heterocyclic inhibitors of the hepatitis C virus RNA-dependent RNA polymerase. J. Biol. Chem. 2002, 277, 38322–38327.
12. Howe, A. Y., Cheng, H., Johann, S., Mullen, S., Chunduru, S. K., Young, D. C., Bard, J., Chopra, R., Krishnamurthy, G., Mansour, T., andO'Connell, J. (2008) Antimicrob. Agents Chemother. 52, 3327–3338
13. Villano, S., Howe, A. Y., Raible, D., Harper, D., Speth, J., and Bichier, G. (2006) Hepatology 44, 607A
14. Villano, S., Raible, D., Harper, D., Speth, J., Chandra, P., Shaw, P., and Bichier, G. (2007) Hepatology 46, S24
15. Hang, J.Q., Yang, Y.,Harris, S.F., Leveque, V., Whittington, H.J., Rajyaguru, S., Ao-Ieong, G., McCown, M.F., Wong, A., Giannetti, A.M., Le Pogam, S., Talamas, F., Cammack, N., Najera, I., Klumpp, K. Slow binding inhibition and mechanism of resistance of non-nucleoside polymerase inhibitors of hepatitis C virus. (2009) J.Biol.Chem. 284: 15517-15529
16. Chan, L., Pereira, O., Reddy, T. J., Das, S. K., Poisson, C., Courchesne, M.,Proulx, M., Siddiqui, A., Yannopoulos, C. G., 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., and Bedard, J. (2004) Bioorg. Med. Chem. Lett. 14, 797– 800
17. Kukolj, G., McGibbon, G. A., McKercher, G., Marquis, M., Lefebvre, S.,Thauvette, L., Gauthier, J., Goulet, S., Poupart, M. A., and Beaulieu, P. L.(2005) J. Biol. Chem. 280, 39260 –39267
18. Ludmerer, S.W., Graham, D.J., Boots, E., Murray, E.M., Simcoe, A., Markel, E.J., Grobler, J.A., Flores, O.A., Olsen, D.B., Hazuda, D.J., LaFemina, R.L., 2005. Replication fitness and NS5B drug sensitivity of diverse hepatitis C virus isolates characterized by using a transient replication assay. Antimicrob. Agents Chemother. 49,2059–2069.
19. Pauwels, F., Mostmans, W., Quirynen, L.M., van der Helm, L., Boutton, C.W., Rueff, A.S., Cleiren, E., Raboisson, P., Surleraux, D., Nyanguile, O., Simmen, K.A., 2007. Binding-site identification and genotypic profiling of hepatitis C virus polymerase inhibitors. J. Virol. 81, 6909–6919.

20. Beaulieu, P.L., 2007. Non-nucleoside inhibitors of the HCV NS5B polymerase: progress in the discovery and development of novel agents for the treatment of HCV infections. Curr. Opin. Investig. Drugs 8, 614–634.
21. Biswal, B.K., Wang, M., Cherney, M.M., Chan, L., Yannopoulos, C.G., Bilimoria, D.,Bedard, J., James, M.N., 2006. Non-nucleoside inhibitors binding to hepatitis C virus NS5B polymerase reveal a novel mechanism of inhibition. J. Mol. Biol. 361, 33–45.
22. Ikegashira, K., Oka, T., Hirashima, S., Noji, S., Yamanaka, H., Hara, Y., Adachi, T., Tsuruha, J., Doi, S., Hase, Y., Noguchi, T., Ando, I., Ogura, N., Ikeda, S., Hashimoto, H., 2006. Discovery of conformationally constrained tetracyclic compounds as potent hepatitis C virus NS5B RNA polymerase inhibitors. J. Med. Chem. 49, 6950–6953.
23. Koch, U., Attenni, B., Malancona, S., Colarusso, S., Conte, I., Di Filippo, M., Harper, S., Pacini, B., Giomini, C., Thomas, S., Incitti, I., Tomei, L., De Francesco, R., Altamura, S., Matassa, V.G., Narjes, F., 2006. 2-(2-Thienyl)-5,6-dihydroxy-4-carboxypyrimidines as inhibitors of the hepatitis C virus NS5B polymerase:discovery, SAR, modeling, and mutagenesis. J. Med. Chem. 49, 1693–1705.
24. 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. Identification and structure-based optimization of novel dihydropyrones as potent HCV RNA polymerase inhibitors. Bioorg. Med. Chem. Lett. 2006, 16, 4834–4838.
25. Lee, G.; Piper, D. E.; Wang, Z.; Anzola, J.; Powers, J.; Walker, N.;Li, Y. Novel inhibitors of hepatitis C virus RNA-dependent RNA polymerases. J. Mol. Biol. 2006, 357, 1051–1057.
26. Powers, J. P.; Piper, D. E.; Li, Y.; Mayorga, V.; Anzola, J.; Chen, J. M.; Jaen, J. C.; Lee, G.; Liu, J.; Peterson, M. G.; Tonn, G. R.; Ye, Q.; Walker, N. P.; Wang, Z. SAR and mode of action of novel non-nucleoside inhibitors of hepatitis C NS5b RNA polymerase. J. Med.Chem. 2006, 49, 1034–1046.
27. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.;Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. Nucleic Acids Res. 2000, 28, 235–242.
28. Wang, M.; Ng, K. K.; Cherney, M. M.; Chan, L.; Yannopoulos, C. G.;Bedard, J.; Morin, N.; Nguyen-Ba, N.; Alaoui-Ismaili, M. H.; Bethell, R. C.; James, M. N. Non-nucleoside analogue inhibitors bind to an allosteric site on HCV NS5B polymerase. Crystal structures and mechanism of inhibition. J. Biol. Chem. 2003, 278, 9489–9495.
29. Chan, L.; Reddy, T. J.; Proulx, M.; Das, S. K.; Pereira, O.; Wang, W.; Siddiqui, A.; Yannopoulos, C. G.; Poisson, C.; Turcotte, N.;Drouin, A.; Alaoui-Ismaili, M. H.; Bethell, R.; Hamel, M.; L'Heureux, L.; Bilimoria, D.; Nguyen-Ba, N. Identification of N, N-disubstituted phenylalanines as a novel class of inhibitors of hepatitis C NS5B polymerase. J. Med. Chem. 2003, 46, 1283–1285.
30. Chan, L.; Das, S. K.; Reddy, T. J.; Poisson, C.; Proulx, M.; Pereira,O.; Courchesne, M.; Roy, C.; Wang, W.; Siddiqui, A.; Yannopoulos, C. G.; 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. 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. 2004, 14, 793–796.
31. Chan, L.; Pereira, O.; Reddy, T. J.; Das, S. K.; Poisson, C.; Courchesne,M.; Proulx, M.; Siddiqui, A.; Yannopoulos, C. G.; 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. 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.2004, 14, 797–800.)
32. Yan, S.; Appleby, T.; Larson, G.; Wu, J. Z.; Hamatake, R.; Hong, Z.;Yao, N. Structure-based design of a novel thiazolone scaffold as HCV NS5B polymerase allosteric inhibitors. Bioorg. Med. Chem. Lett. 2006, 16, 5888–5891.
33. Yan, S.; Larson, G.; Wu, J. Z.; Appleby, T.; Ding, Y.; Hamatake, R.; Hong, Z.; Yao, N. Novel thiazolones as HCV NS5B polymerase allosteric inhibitors: Further designs, SAR, and X-ray complex structure. Bioorg. Med. Chem. Lett. 2007, 17, 63–67.
34. Love, R. A.; Parge, H. E.; Yu, X.; Hickey, M. J.; Diehl, W.; Gao, J.;Wriggers, H.; Ekker, A.; Wang, L.; Thomson, J. A.; Dragovich, P. S.;Fuhrman, S. A. Crystallographic identification of a noncompetitive inhibitor binding site on the hepatitis C virus NS5B RNA polymerase enzyme. J. Virol. 2003, 77, 7575–7581.
35. 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.

Identification and structure-based optimization of novel dihydropyrones as potent HCV RNA polymerase inhibitors. Bioorg. Med. Chem. Lett. 2006, 16, 4834–4838
36. Burton, G.; Ku, T. W.; Carr, T. J.; Kiesow, T.; Sarisky, R. T.; Lin-Goerke, J.; Baker, A.; Earnshaw, D. L.; Hofmann, G. A.; Keenan,R. M.; Dhanak, D. Identification of small molecule inhibitors of the hepatitis C virus RNA-dependent RNA polymerase from a pyrrolidine combinatorial mixture. Bioorg. Med. Chem. Lett. 2005, 15, 1553–1556.
37. Slater, M. J.; Amphlett, E. M.; Andrews, D. M.; Bravi, G.; Burton,G.; Cheasty, A. G.; Corfield, J. A.; Ellis, M. R.; Fenwick, R. H.; Fernandes, S.; Guidetti, R.; Haigh, D.; Hartley, C. D.; Howes, P. D.; Jackson, D. L.; Jarvest, R. L.; Lovegrove, V. L.; Medhurst, K. J.; Parry, N. R.; Price, H.; Shah, P.; Singh, O. M.; Stocker, R.; Thommes
38. Dhanak, D.; Duffy, K. J.; Johnston, V. K.; Lin-Goerke, J.; Darcy, M.; Shaw, A. N.; Gu, B.; Silverman, C.; Gates, A. T.; Nonnemacher, M. R.; Earnshaw, D. L.; Casper, D. J.; Kaura, A.; Baker, A.;Greenwood, C.; Gutshall, L. L.; Maley, D.; DelVecchio, A.;Macarron, R.; Hofmann, G. A.; Alnoah, Z.; Cheng, H. Y.; Chan,G.; Khandekar, S.; Keenan, R. M.; Sarisky, R. T. Identification and biological characterization of heterocyclic inhibitors of the hepatitis C virus RNA-dependent RNA polymerase. J. Biol. Chem. 2002, 277, 38322–38327
39. Tedesco, R.; Shaw, A. N.; Bambal, R.; Chai, D.; Concha, N. O.; Darcy,M. G.; Dhanak, D.; Fitch, D. M.; Gates, A.; Gerhardt, W. G.;Halegoua, D. L.; Han, C.; Hofmann, G. A.; Johnston, V. K.; Kaura, A. C.; Liu, N.; Keenan, R. M.; Lin-Goerke, J.; Sarisky, R. T.; Wiggall, K. J.; Zimmerman, M. N.; Duffy, K. J. 3-(1,1-dioxo-2H-(1,2,4)-benzothiadiazin-3-yl)-4-hydroxy-2(1H)-quinolinones potent inhibitors of hepatitis C virus RNA-dependent RNA polymerase. J. Med. Chem.2006, 49, 971–983
40. Lee, G.; Piper, D. E.; Wang, Z.; Anzola, J.; Powers, J.; Walker, N.;Li, Y. Novel inhibitors of hepatitis C virus RNA-dependent RNA polymerases. J. Mol. Biol. 2006, 357, 1051–1057.
41. Powers, J. P.; Piper, D. E.; Li, Y.; Mayorga, V.; Anzola, J.; Chen, J. M.; Jaen, J. C.; Lee, G.; Liu, J.; Peterson, M. G.; Tonn, G. R.; Ye, Q.; Walker, N. P.; Wang, Z. SAR and mode of action of novel non-nucleoside inhibitors of hepatitis C NS5b RNA polymerase. J. Med.Chem. 2006, 49, 1034–1046.
42. Pfefferkorn, J. A.; Greene, M. L.; Nugent, R. A.; Gross, R. J.; Mitchell,M. A.; Finzel, B. C.; Harris, M. S.; Wells, P. A.; Shelly, J. A.; Anstadt,R. A.; Kilkuskie, R. E.; Kopta, L. A.; Schwende, F. J. Inhibitors of HCV NS5B polymerase. Part 1: Evaluation of the southern region of(2Z)-2-(benzoylamino)-3-(5-phenyl-2-furyl)acrylic acid. Bioorg. Med.Chem. Lett. 2005, 15, 2481–2486.
43. Pfefferkorn, J. A.; Nugent, R.; Gross, R. J.; Greene, M.; Mitchell, M. A.; Reding, M. T.; Funk, L. A.; Anderson, R.; Wells, P. A.; Shelly, J. A.; Anstadt, R.; Finzel, B. C.; Harris, M. S.; Kilkuskie, R. E.; Kopta,L. A.; Schwende, F. J. Inhibitors of HCV NS5B polymerase. Part 2:Evaluation of the northern region of (2Z)-2-benzoylamino-3-(4- phenoxy-phenyl)-acrylic acid. Bioorg. Med. Chem. Lett. 2005, 15,2812–2818.
44. Gopalsamy, A.; Chopra, R.; Lim, K.; Ciszewski, G.; Shi, M.; Curran,K. J.; Sukits, S. F.; Svenson, K.; Bard, J.; Ellingboe, J. W.; Agarwal,A.; Krishnamurthy, G.; Howe, A. Y.; Orlowski, M.; Feld, B.;O'Connell, J.; Mansour, T. S. Discovery of proline sulfonamides as potent and selective hepatitis C virus NS5b polymerase inhibitors.
45. Evidence for a new NS5b polymerase binding site. J. Med. Chem.2006, 49, 3052–3055.
46. Di Marco, S.; Volpari, C.; Tomei, L.; Altamura, S.; Harper, S.; Narjes, F.; Koch, U.; Rowley, M.; De Francesco, R.; Migliaccio, G.; Carfi, A. Interdomain communication in hepatitis C virus polymerase abolished by small molecule inhibitors bound to a novel allosteric site. J. Biol. Chem. 2005, 280, 29765–29770.
47. Ikegashira, K.; Oka, T.; Hirashima, S.; Noji, S.; Yamanaka, H.; Hara, Y.; Adachi, T.; Tsuruha, J.; Doi, S.; Hase, Y.; Noguchi, T.; Ando, I.;Ogura, N.; Ikeda, S.; Hashimoto, H. Discovery of conformationally constrained tetracyclic compounds as potent hepatitis C virus NS5B RNA polymerase inhibitors. J. Med. Chem. 2006, 49, 6950–6953.
48. loopy (http://trantor.bioc.columbia.edu/programs/jackal)
49. Morris, G.M., Goodsell, D.S., Huey, R., Olson, A.J., 1996. Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4. J.Comput. Aided Mol. Des. 10, 293–304.
50. Ren, J.; Esnouf, R.; Garman, E.; Somers, D.; Ross, C.; Kirby, I.;Keeling, J.; Darby, G.; Jones, Y.; Stuart, D.; et al. High resolution structures of HIV-1 RT from four RT-inhibitor complexes. Nat. Struct. Biol. 1995, 2, 293–302.) in other cases, the lowest energy conformer of the most populated cluster should also be taken into account
51. Ragno, R.; Frasca, S.; Manetti, F.; Brizzi, A.; Massa, S. HIV-reverse transcriptase inhibition: inclusion of ligand-induced fit by cross-docking studies. J. Med. Chem. 2005, 48, 200–212.)

52. Dutartre, H., Boretto, J., Guillemot, J.C., Canard, B., 2005. A relaxed discrimination of 2 -O-methyl-GTP relative to GTP between de novo and elongative RNA synthesis by the hepatitis C RNA-dependent RNA polymerase NS5B. J. Biol. Chem. 280, 6359–6368.
53. Dutartre, H., Bussetta, C., Boretto, J., Canard, B., 2006. General catalytic deficiency of hepatitis C virus RNA polymerase with an S282T mutation and mutually exclusive resistance towards 2 -modified nucleotide analogues. Antimicrob. Agents Chemother. 50, 4161–4169.
54. P. Thommes, M. Convery, E. Williams, D. Mesogiti, M. Slater, S. Bains, F. Gray, D. Haigh, The R200-hinge site, a novel inhibitor binding site on the HCV NS5B RNA-dependent RNA polymerase, 13th International Meeting on Hepatitis C Virus and Related Viruses, Cairns, Queensland, Australia, August 27-31, 2006, vol. 426
55. Ago, H., Adachi, T., Yoshida, A., Yamamoto, M., Habuka, N., Yatsunami, K. & Miyano, M. (1999) Structure (Cambridge, U.K.) 7, 1417–1426.
56. Lesburg, C. A., Cable, M. B., Ferrari, E., Hong, Z., Mannarino, A. & Weber, P. C. (1999) Nat. Struct. Biol. 6, 937–942.
57. Butcher, S.J., Grimes, J.M., Makeyev, E.V., Bamford, D.H., Stuart, D.I., 2001. A mechanism for initiating RNA-dependent RNA polymerisation. Nature 410, 235–240. Clark, A.M., Labute, P., 2007. 2D depiction of protein–ligand complexes. J. Chem. Inf. Model. 47, 1933–1944.
58. Stéphane Betzi , Cécilia Eydoux , Cécile Bussetta, Marilyne Blemont, Pieter Leyssen, Claire Debarnot, Mohamed Ben-Rahou, Jacques Haiech, Marcel Hibert,Francoise Gueritte, David S. Grierson g, Jean-Louis Romette, Jean-Claude Guillemot, Johan Neyts, Karine Alvarez, Xavier Morelli, Hélène Dutartre , Bruno Canard; Identification of allosteric inhibitors blocking the hepatitis C virus polymeraseNS5B in the RNA synthesis initiation step; Antiviral Research 84 (2009) 48–59
59. Stefan Reich,Ralph Peter Golbik, René Geissler, Hauke Lilie and Sven-Erik Behrens ; Mechanisms of Activity and Inhibition of the Hepatitis C Virus RNA-dependent RNA Polymerase; April 30, 2010 285, 13685-13693.
60. Tomei, L., Altamura, S., Bartholomew, L., Biroccio, A., Ceccacci, A., Pacini, L., Narjes, F., Gennari, N., Bisbocci, M., Incitti, I., Orsatti, L., Harper, S., Stansfield, I., Rowley, M., De Francesco, R., Migliaccio, G., 2003. Mechanism of action and antiviral activity of benzimidazole-based allosteric inhibitors of the hepatitis C virus RNA-dependent RNA polymerase. J. Virol. 77, 13225–13231.
61. Tomei, L., Altamura, S., Bartholomew, L., Bisbocci, M., Bailey, C., Bosserman, M., Cellucci, A., Forte, E., Incitti, I., Orsatti, L., Koch, U., De Francesco, R., Olsen, D.B., Carroll, S.S., Migliaccio, G., 2004. Characterization of the inhibition of hepatitis C virus RNA replication by nonnucleosides. J. Virol. 78, 938–946.
62. Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.;Srinivasan, J.; Case, D. A.; Cheatham, T. E. 3rd. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc. Chem. Res. 2000, 33 (12), 889–897
63. Kuhn, B.; Donini, O.; Huo, S.; Wang, J. M.; Kollman, P. A. MM-PBSA applied to computer-assisted ligand design. In Free energy calculations in rational drug design; Rami Reddy, M.; Erion, M. D.,Eds.; Kluwer Academic/Plenum Publishers: New York, 2001; pp 243-251.
64. Sitkoff, D.; Sharp, K. A.; Honig, B. Accurate Calculation of Hydration Free-Energies Using Macroscopic Solvent Models. J. Phys. Chem. 1994, 98 (7), 1978–1988.
65. Rocchia, W.; Sridharan, S.; Nicholls, A.; Alexov, E.; Chiabrera, A.; Honig, B. Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: Applications to the molecular systems and geometric objects. J. Com- put. Chem. 2002, 23 (1), 128–137.
66. Honig, B.; Sharp, K.; Yang, A. S. Macroscopic Models of Aqueous- Solutions - Biological and Chemical Applications. J. Phys. Chem. 1993, 97 (6), 1101–1109.
67. Rashin, A. A. Buried Surface-Area, Conformational Entropy, and Protein Stability. Biopolymers 1984, 23 (8), 1605–1620.
68. Chong, L. T.; Duan, Y.; Wang, L.; Massova, I.; Kollman, P. A.Molecular dynamics and free-energy calculations applied to affinity maturation in antibody 48G7. Proc. Natl. Acad. Sci. U. S. A. 1999, 96 (25), 14330–14335.
69. Jacob Kongsted Æ Ulf Ryde, An improved method to predict the entropy term with the MM/PBSA approach, J Comput Aided Mol Des (2009) 23:63–71
70. Reyes, C. M.; Kollman, P. A. Structure and thermodynamics of RNA-protein binding: using

molecular dynamics and free energy analyses to calculate the free energies of binding and conformational change. J. Mol. Biol. 2000, 297 (5), 1145–1158.
71. Wang, J.; Morin, P.; Wang, W.; Kollman, P. A. Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA. J. Am. Chem. Soc. 2001, 123 (22), 5221–5230.
72. Cheatham, T. E., 3rd; Srinivasan, J.; Case, D. A.; Kollman, P. A. Molecular dynamics and continuum solvent studies of the stability of polyG-polyC and polyA-polyT DNA duplexes in solution. J. Biomol. Struct. Dyn. 1998, 16 (2), 265–280.

VI
CONCLUSION

Aim of this work is to use and exploit the advances being made in biophysical techniques relating to de novo drug designing in order to understand the molecular recognition process with respect to protein ligand system. In a nut shell, identification of novel binding site, estimation of protein-ligand binding affinities, exploring the changes in protein dynamics on ligand binding, throwing light on probable mechanism of inhibition and resistance, identifying principal interactions in a protein ligand system; all this with an aim to provide clues for further optimization of affinity and for the design of inhibitors that are less susceptible to the mutations; have been the major outcome of this thesis.

At the heart of this thesis lies the basic problem of development of new, effective antiviral compounds for combating debilitating human (HCV) and animal pathogen (BVDV) which is of paramount importance due to the huge life and economic loses caused by these viruses. The major part of my thesis focuses on NNIs (newly synthesized and biologically potent) which bind to BVDV and HCV RdRp. The compounds belonging to classes Benzimidazole, Imidazoquinoline, Pyridoxoquinoline and Phenathroline are used as a ligands against BVDV and HCV RdRp to understand protein-ligand recognition.

Herein, we have employed an arsenal of computational techniques to understand the molecular basis of interaction of NNIs against BVDV and HCV RdRp.

Project1)-(HCV-vs-BVDV)

To the best of our knowledge, we, for the first time identified a benzimidazole derivative active at low micromolar range against both HCV and BVDV RdRp. Benzimidazole class of compounds were known to bind in the thumb domain of HCV RdRp, but no information regarding the binding of same class of compounds was available in case of BVDV RdRp. We for the first time identified a novel binding site for the binding of Benzimidazole class of compounds in the finger domain of BVDV RdRp. The binding site was found to lie in the entrance to the template channel in BVDV RdRp.

We defined the inventory of interactions established between 227G and HCV and BVDV RdRp. This was done with an aim to apply a structure based strategy to pinpoint possible changes in the compounds aiming at improving its potency.

In this thesis we proposed the probable mechanism of inhibition of 227G in BVDV RdRp, which are all different from the one reported for benzimidazole compounds against HCV RdRp. We showed that the mechanism of inhibition involves different domain in the two RdRps, in BVDV it bind in finger domain while in case of HCV in thumb domain. This lead to our opinion that BVDV could be considered as a predictable surrogate model for HCV only in the case of NIs, since they bind to an evolutionarily conserved portion of the enzyme, which remained unchanged in the flaviviridae RdRp and in other viral polymerases, but despite the promising results obtained with our Benzimidazole, we could not disprove the concept of a limited spectrum of activity for the antiviral therapy.

Project2): BVDV

Screening studies have come out with two new series of NNIs, Imidazoquinoline and Pyridozoquinolines as selective and potent inhibitors of BVDV RdRps.
Herein, we applied a combined experimental and molecular modeling strategy to explore the molecular basis of interaction of the PS999 and PS1036 with the BVDV RdRp. Based on the resistant mutations obtained for PS999 and PS1036, Binding site for these compounds, as also is in the case of 227G, were found to be located in the template entrance channel. Binding of different classes of compounds onto the same binding site made it a HOTSPOT for the binding of NNI of different classes in the BVDV RdRp.
Key residues playing pivotal role in binding of the NNIs were identified.
Metadynamics was used to identify alternate binding poses other than initial docking pose, as well as detailed undocking pathway was estimated in order to identify key events taking place during the dissociation process.
Hypothesis regarding the mechanism of inhibition as well as the mechanism of resistance were put forth by

performing extensive MD simulations.
Binding free energies of the NNI-BVDVRdRp complex were calculated which showed similar trend to the corresponding experimatal IC50 values for both the complexes. This was an encouraging performance and validate the model of the binding site, given that all molecular modeling studies were performed in the absence of any available crystal structure of the protein in complex with lead compounds. Also, the adopted procedure was able to correctly predict drug binding affinities in the presence of mutated protein residues involved in drug resistance.

Project: 3(HCV)

We employed extensive molecular docking and MD simualtions in order to identified the putative binding pocket of three new class of compounds (Imidazoquinoline, Pyridoxiquinoline and Phenanthrolines) which are active against HCV RdRp at low micromolar range. Key binding pattern of the compounds in the cavity were identified with an aim of further optimisation of the lead compounds in order to increase their potency. Drug-protein binding affinities for all the three NNI-HCV RdRp system were calculated which were found to be in accordance with the experimental data.

Furthermore, we were also successful in throwing light on the probable mechanism of inhibition of the three compounds reported in this chapter. The results obtained might provide useful hints in guiding the rational design for more selective and potent compounds against HCV RdRp.

Hopefully, the result we have obtained so far will contribute to deciphering the molecular determinants of inhibitor activity and will provide the key to understand the molecular mechanism of inhibition. The final outcome will be a major understanding on how the HCV and BVDV RdRp could be targeted by specific inhibitors, thus allowing a rational (*in silico*) drug design.