DRUG RESISTANCE MECHANISM OF NS3/4A PROTEASE AND IDENTIFICATION OF
POTENT COMPOUNDS INHIBITING REVERSE TRANSCRIPTASE OF HEPATITIS VIRUS

Mr. Jirayu Kammarabutr

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Biochemistry and Molecular Biology
Department of Biochemistry
Faculty of Science
Chulalongkorn University
Academic Year 2018
Copyright of Chulalongkorn University

กลไกการดื้อยาของโปรตีเอสเอ็นเอส 3/4 เอ และการระบุสารยับยั้งรีเวิร์สแทรนสคริปเทสของเชื้อไวรัสตับอักเสบ

นายจิรายุ กัมมาระบุตร

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาชีวเคมีและชีววิทยาโมเลกุล ภาควิชาชีวเคมี
คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2561
ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

Thesis Title	DRUG RESISTANCE MECHANISM OF NS3/4A PROTEASE AND IDENTIFICATION OF POTENT COMPOUNDS INHIBITING REVERSE TRANSCRIPTASE OF HEPATITIS VIRUS
By	Mr. Jirayu Kammarabutr
Field of Study	Biochemistry and Molecular Biology
Thesis Advisor	Assistant Professor THANYADA RUNGROTMONGKOL, Ph.D.

Accepted by the Faculty of Science, Chulalongkorn University in Partial Fulfillment of the Requirement for the Master of Science

-- Dean of the Faculty of Science

(Professor POLKIT SANGVANICH, Ph.D.)

THESIS COMMITTEE

-- Chairman

(Assistant Professor KANOKTIP PACKDIBAMRUNG, Ph.D.)

-- Thesis Advisor

(Assistant Professor THANYADA RUNGROTMONGKOL, Ph.D.)

-- Examiner

(Professor SUPOT HANNONGBUA, Ph.D.)

-- Examiner

(Assistant Professor Rath Pichyangkura, Ph.D.)

-- External Examiner

(Associate Professor Kiattawee Choowongkomon, Ph.D.)

จิรายุ กัมมาระบุตร : กลไกการดื้อยาของโปรติเอสเอ็นเอส 3/4 เอ และการระบุสารยับยั้งรีเวิร์สแทรนสคริปเทสของเชื้อไวรัสตับ
อักเสบ. (DRUG RESISTANCE MECHANISM OF NS3/4A PROTEASE AND IDENTIFICATION OF POTENT COMPOUNDS
INHIBITING REVERSE TRANSCRIPTASE OF HEPATITIS VIRUS) อ.ที่ปรึกษาหลัก : ผศ. ดร.ธัญญดา รุ่งโรจน์มงคล

ไวรัสตับอักเสบนับเป็นปัญหาด้านสุขภาพที่สำคัญ ซึ่งมีผู้เสียชีวิตจากเชื้อไวรัสชนิดนี้ทั่วโลกมากถึง 1.34 ล้านคนในปี ค.ศ. 2015 โดยสาเหตุหลักของการเสียชีวิตเกิดจากภาวะตับอักเสบชนิดเรื้อรังจากไวรัสตับอักเสบชนิดบีและซีซึ่งคิดเป็น 95 เปอร์เซ็นต์ ผู้ติดเชื้อไวรัสทั้งสอง ชนิดนี้มีมากถึง 325 ล้านคนทั่วโลก ซึ่งมากกว่าจำนวนของผู้ป่วยติดเชื้อไวรัสเอชไอวีถึง 10 เท่า ไวรัสตับอักเสบถือเป็นสาเหตุหลักของการเกิด โรคไวรัสตับอักเสบชนิดซี ซึ่งไวรัสชิดนี้มีอัตราการกลายพันธุ์สูงเนื่องจากไวรัสดังกล่าวขาดกระบวนการตรวจสอบความถูกต้องของเอนไซม์ อาร์เอ็นเอพอลิเมอเรส โดยหนึ่งในโปรตีนเป้าหมายที่สำคัญสำหรับการออกแบบและพัฒนายาต้านไวรัสตับอักเสบซีคือเอนไซม์เอ็นเอส3/4เอ โปรทีเอส ขณะที่ไวรัสตับอักเสบบีสามารถก่อให้เกิดภาวะตับอักเสบชนิดเฉียบพลันและเรื้อรังได้ ปัจจุบันการรักษาการติดเชื้อไวรัสตับอักเสบบี ทำได้โดยการให้ยาในกลุ่มนิวคลีโอไทด์แอนะล็อก โดยยาในกลุ่มนี้สามารถเข้ายับยั้งแบบแข่งขันที่บริเวณเร่งปฏิกิริยาของเอนไซม์ รีเวิร์สแทรนสคริปเทสของไวรัสชนิดนี้ได้ ดังนั้นในงานวิจัยนี้จึงมุ่งศึกษาผลของการลดลงในประสิทธิภาพของสารยับยั้งอะซูนาพรีเวียร์ต่อโปรตีน เป้าหมายเอ็นเอส3/4เอ โปรทีเอสของไวรัสตับอักเสบซีสายพันธุ์ที่ 1 สายพันธุ์กลายที่ตำแหน่ง R155K และ D168A โดยอาศัยเทคนิคการจำลอง พลวัตเชิงโมเลกุล จากผลการวิเคราะห์องค์ประกอบหลักทางโครงสร้างของเอนไซม์สายพันธุ์กลาย แสดงให้เห็นว่าการกลายพันธุ์ของทั้งสอง ตำแหน่งสามารถส่งผลต่อการสูญเสียของโครงข่ายพันธะไฮโดรเจนระหว่างกรดอะมิโนที่ตำแหน่ง R123 R155 และ D168 ได้ นอกจากนี้จากการ คำนวณพลังงานยึดเหนี่ยวอิสระพบว่าประสิทธิภาพการยับยั้งของอะซูนาพรีเวียร์ต่อเอนไซม์เอ็นเอส3/4เอ โปรทีเอสของไวรัสตับอักเสบซี สายพันธุ์กลายที่ตำแหน่ง R155K และ D168A ลดลงอย่างมีนัยสำคัญเมื่อเทียบกับสายพันธุ์ดั้งเดิม ส่วนสำหรับไวรัสตับอักเสบบีได้ทำการศึกษา อันตรกิริยาที่เกิดขึ้นและประสิทธิภาพในการเข้าจับระหว่างยาต้านไวรัสเอชไอวี ได้แก่ สตาวูดีน ไดดาโนซีน และซิโดวูดีน และยาต้านไวรัสตับ อักเสบบี ลามิวูดีน ต่อโปรตีนเป้าหมายรีเวิร์สแทรนสคริปเทส โดยอาศัยเทคนิคการจำลองพลวัตเชิงโมเลกุล จากผลการคำนวณพลังงานยึด เหนี่ยวอิสระพบว่าสตาวูดีนมีประสิทธิภาพในการยึดจับกับเอนไซม์รีเวิร์สแทรนสคริปเทสได้ใกล้เคียงกับซิโดวูดีน และคาดว่าดีกว่าสารตั้งต้นและ ลามิวูดีน จากนั้นจึงทำการศึกษาลักษณะโครงสร้างของยาในรูปสารเชิงซ้อนและในรูปอิสระในน้ำเพื่อหาโครงสร้างของยาต้านไวรัสเอชไอวี ที่เสถียรที่สุด โดยอาศัยวิธีการจำลองพลวัตเชิงโมเลกุลแบบ replica exchange ซึ่งผลการคำนวณพบว่าโครงสร้างของยาที่จับกับโปรตีน เป้าหมายเป็นโครงสร้างที่เสถียรที่สุด นอกจากนั้นยังทำการค้นหายากลุ่มอื่นๆ ที่คาดว่าจะมีประสิทธิภาพการยับยั้งโปรตีนเป้าหมายนี้โดยใช้ รูปแบบสามมิติของคุณสมบัติทางเภสัชวิทยาของสารประกอบ ซึ่งได้ทำการคัดเลือกและนำไปทดสอบประสิทธิภาพการยับยั้งต่อเชื้อไวรัสตับ อักเสบบีต่อไป จากการค้นหายาโดยใช้รูปแบบสามมิติของคุณสมบัติทางเภสัชวิทยาของสารประกอบและความคล้ายคลึงของโครงสร้าง พบว่า ยากลุ่มอื่นๆ เช่น ฟลอกซูริดีน ไตรฟลูริดีน และโซฟอสบูเวียร์ น่าจะสามารถยับยั้งการทำงานของเอนไซม์รีเวิร์สแทรนสคริปเทสได้ และเมื่อทำ การทดลองพบว่าฟลอกซูริดีนมีความเป็นไปได้ในการยับยั้งไวรัสตับอักเสบบี แต่อย่างไรก็ตามยาตัวนี้ยังคงมีความเป็นพิษต่อเซลล์ตับ จากงานทั้ง สองส่วนนี้จะเป็นประโยชน์และมีความสำคัญอย่างยิ่งต่อการพัฒนายาที่มีประสิทธิภาพสูงยิ่งขึ้นในอนาคต

สาขาวิชา	ชีวเคมีและชีววิทยาโมเลกุล	ลายมือชื่อนิสิต ....
ปีการศึกษา	2561	ลายมือชื่อ อ.ที่ปรึกษาหลัก ....

5971923123 : MAJOR BIOCHEMISTRY AND MOLECULAR BIOLOGY

KEYWORD: HBV reverse transcriptase, Stavudine, Didanosine, Zidovudine, Molecular dynamics simulation, Pharmacophore-based screening, Replica exchange molecular dynamics, HCV NS3/4A protease, Asunaprevir, Principal component analysis

Jirayu Kammarabutr : DRUG RESISTANCE MECHANISM OF NS3/4A PROTEASE AND IDENTIFICATION OF POTENT COMPOUNDS INHIBITING REVERSE TRANSCRIPTASE OF HEPATITIS VIRUS. Advisor: Asst. Prof. THANYADA RUNGROTMONGKOL, Ph.D.

Viral Hepatitis has become a public health concern which caused 1.34 million deaths in 2015 and 95% of sickness and untimely deaths come from chronic hepatitis B and C virus infections. The epidemic of viral hepatitis B and C affects around 325 million people worldwide and is 10 times more than the global HIV epidemic. Hepatitis C virus (HCV), which is the main cause of hepatitis C, has a high mutation rate due to the lack of proofreading activity of the RNA polymerase enzyme. One of an attractive target for anti-HCV drug design and development is the NS3/4A serine protease. Whereas hepatitis B virus (HBV) infection can result in both acute and chronic diseases in the human liver. Current therapeutic drugs for HBV treatment are the nucleos(t)ide analogues (NAs) acting as competitive inhibitor against HBV reverse transcriptase (RT). In the present study, all-atom molecular dynamics (MD) simulations were performed to investigate the influence of the single point mutations (R155K and D168A) in the HCV genotype 1 NS3/4A protease on the structural dynamics, molecular interactions and susceptibility of asunaprevir (ASV). Principal component analysis (PCA) indicated that these two mutations caused a loss of the hydrogen bond network of residues R123--R155--D168. In addition, the free energy calculations based on different semiempirical QM/MM-GBSA methods revealed that the binding affinity of ASV with the two mutant forms was significantly decreased in the order of wild-type < R155K < D168A. In the case of HBV, MD simulations were performed on the HBV-RT in complex with anti-HIV drugs (stavudine, didanosine and zidovudine) and anti-HBV drug (lamivudine) to determine the ligand-protein interactions and binding efficiency towards the HBV-RT. The predicted binding free energies showed that stavudine had the binding affinity against HBV-RT similar to zidovudine and relatively better than the substrate and lamivudine. Afterward, the conformations of anti-HIV drugs in bound state with HBV-RT were investigated with their unbound form to elucidate the most stable conformation of anti-HIV drugs using replica exchange molecular dynamic. The results suggested that the conformations of all anti-HIV drugs in bound form were the most stable conformation. Moreover, the exploration of new potent drugs against HBV-RT, which were screened from the three-dimensional pharmacophore modeling was evaluated their anti-HBV activity. Based on pharmacophore-based screening and similarity search for 50% of nucleoside-based scaffold, the three commercial drugs, including floxuridine, trifluridine and sofosbuvir were selected. The *in vitro* study showed that floxuridine possibly reduced the HBV DNA but still had toxicity with hepatocytes. This research provided the useful structural information regarding the atomistic understanding of acquired drug resistance and a rational guidance for antiviral drug design and development.

Field of Study:	Biochemistry and Molecular Biology	Student's Signature ....
Academic Year:	2018	Advisor's Signature ....

ACKNOWLEDGEMENTS

Jirayu Kammarabutr

TABLE OF CONTENTS

Table

จุฬาลงกรณ์มหาวิทยาลัย
CHULALONGKORN UNIVERSITY

Figure

LIST OF ABBREVIATION

aa	amino acids
ASV	asunaprevir
atm	atmosphere
˚C	degree Celsius
EC_{50}	half maximal effective concentration
eIF	eukaryotic initiation factor
ER	endoplasmic reticulum
FBS	Fetal bovine serum
HBV	hepatitis B virus
HCV	hepatitis C virus
IRES	internal ribosomal entry site
K	Kelvin
kDa	kilodalton
kp	kilobase
MD	molecular dynamics
ml	milliliter
mM	millimolar
NCBI	national center for biotechnology information
ns	nanosecond
nm	nanometer
ORF	open reading frame
PDB	Protein Data Bank
ps	picosecond

RdRp	RNA-dependent RNA polymerase
REMD	replica-exchange molecular dynamics
RMSD	root-mean-square displacement
SASA	solvent accessible surface area
SPP	signal peptide peptidase
µM	micromolar
UTRs	untranslated regions
vdW	van der Waal
kcal/mol	kilocalorie per mole
MM/GBSA	molecular mechanics/generalized Born surface area
MM/PBSA	molecular mechanics/Poisson Boltzmann surface area
QM/MM-GBSA	quantum mechanics/molecular mechanics generalized Born surface area

CHAPTER I

INTRODUCTION

Viral Hepatitis has become a public health challenge of global proportion which caused 1.34 million deaths in 2015, indicating a number comparable to deaths caused by tuberculosis and higher than those caused by HIV. However, the number of deaths due to viral hepatitis is increasing over time, while mortality caused by tuberculosis and HIV is declining. Most viral hepatitis deaths in 2015 were due to chronic liver disease with 720,000 deaths by cirrhosis and primary liver cancer with 470,000 deaths due to hepatocellular carcinoma. Although there are 5 distinct types of viral hepatitis (A-E), chronic hepatitis B and C cause 95% hepatitis related sickness and untimely deaths. Hepatitis D is less common and occurs only in association with hepatitis B. The other viruses (A & E) are spread via contaminated food, water and cause acute infections and outbreak in areas of poor sanitation and inadequate waste disposal [1]. Acute infections are often short-lived and resolve within a few weeks. The epidemic caused by hepatitis B virus (HBV) affects mostly the African Region and the Western Pacific Region. The epidemic caused by hepatitis C virus (HCV) affects all regions, with major differences between and within countries. The epidemic of both viral hepatitis B and C affects 325 million people globally and is 10 times larger than the global HIV epidemic. In Africa, chronic viral hepatitis affects over 70 million Africans (60 million with hepatitis B and 10 million with hepatitis C). Antiviral medicines can cure more than 95% of persons with hepatitis C infection, thereby reducing the risk of death from liver cancer and cirrhosis, but access to diagnosis and treatment is low. In addition, there is currently no vaccine for hepatitis C. For HBV, a vaccine has been available since 1982. The vaccine is 95% effective in preventing infection and the development of chronic disease and liver cancer due to hepatitis B. Viral hepatitis can now be treated with new simple oral medications which cure hepatitis C infections within 12 weeks and effectively control chronic Hepatitis B infection with longer term treatment. However, there were more new HBV and HCV infections than patients who were started on treatment. By the concept of target inhibition networks of antiretroviral (ARV)

regimen for a treatment-naive patient, two from three of drugs were nucleoside reverse transcriptase inhibitors (NRTIs) and protease inhibitor (PI) which targeted reverse transcriptase and protease of virus, respectively. These two enzyme also occurred in both HCV and HBV, suggesting that reverse transcriptase and protease are the major target for inhibiting viral hepatitis.

1.1 Hepatitis C virus

1.1.1 Genomic organization

HCV is a small (50-80 nm in size) enveloped RNA virus belonging to the *Hepacivirus* genus of the *Flaviviridae* family [2]. The HCV particle consists of a nucleocapsid, which contains the positive single-stranded RNA genome of approximately 9.6 kb, which contains an open reading frame (ORF) encoding a polyprotein precursor of approximately 3000 residues flanked by untranslated regions (UTRs) at both ends covered by a host cell-derived lipid envelope [3]. The precursor is cleaved into at least 10 different proteins: the structural proteins Core, E1, E2 and p7, as well as the non-structural proteins NS2, NS3, NS4A, NS4B, NS5A and NS5B (**Fig. 1**). The E1 and E2 regions are the most variable, while the 5'UTR and terminal segment of the 3'UTR have the highest degree of sequence conservation among various isolates. The 5'UTR, which is ∼341 nucleotide in length, contains an internal ribosomal entry site (IRES), which is essential for cap-independent translation of viral RNA, from which four highly structured domains are produced. The 3'UTR varies between 200 and 235 nucleotide in length, including a short variable region, a poly(U/UC) tract with an average length of 80 nucleotide, and a virtually invariant 98 nucleotide X-tail region. The X region forms three stable stem–loop structures that are highly conserved among all genotypes. It appears that the 3'X region and the 52 nucleotide upstream of the poly(U/C) region, are crucial for RNA replication, while the remainder of the 3'UTR plays a role in enhancement of replication.
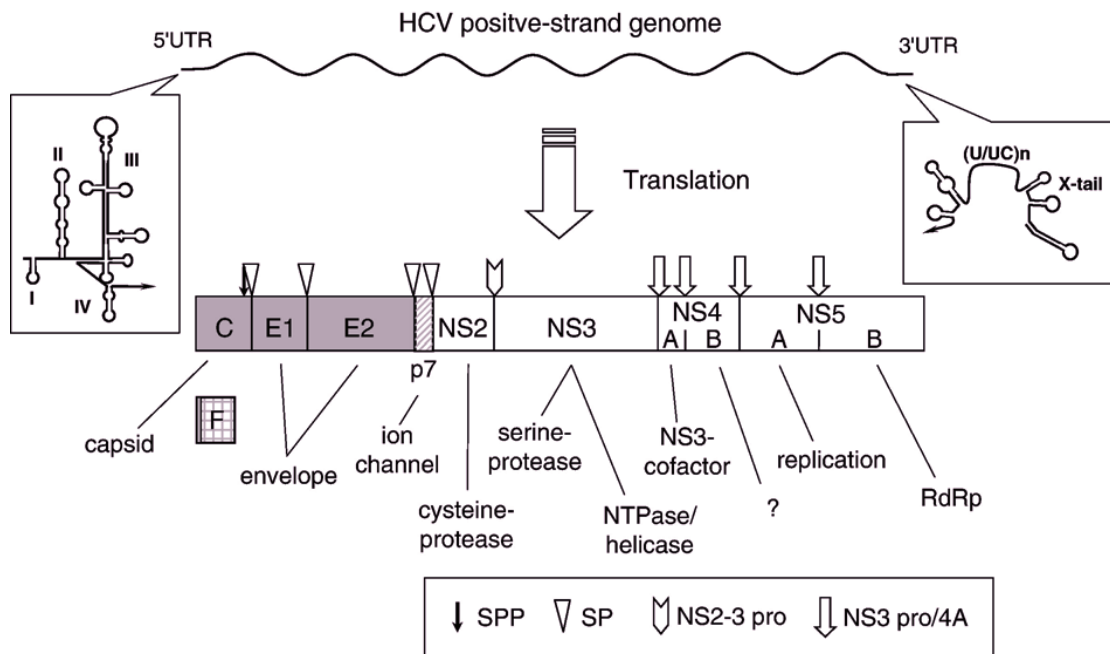
Figure 1 HCV genome organization and polyprotein processing

1.1.2 NS3/4A protease complex

NS3 is a multi-functional viral protein containing a serine protease domain in its N-terminal third and a helicase/NTPase domain in its C-terminal two-thirds. NS4A is a cofactor of NS3 protease activity. NS3/4A is a complex bifunctional molecule essential for viral polyprotein processing and RNA replication [2, 4]. The catalytic triad is formed by residues His 57, Asp 81 and Ser 139 of NS3. It catalyzes HCV polyprotein cleavage at the NS3/NS4A, NS4A/NS4B, NS4B/NS5A and NS5A/NS5B junctions. The central portion of NS4A is important for efficient processing of the nonstructural proteins by NS3. The central region of NS4A acts as a cofactor of NS3 serine protease activity, allowing its stabilization, localization at the ER membrane as well as cleaveage-dependent activation, particularly at the NS4B/NS5A junction.

1.1.3 Global prevalence and genotype distribution of hepatitis C virus infection

Currently, HCV infects over 185 million people and long-term HCV-infected patients are under an increased risk of developing liver diseases, such as cirrhosis and hepatocellular carcinoma. HCV strains are classified into seven recognized genotypes (1-7) on the basis of phylogenetic and sequence analyses of whole viral genomes. Genotype 1 is the most common HCV genotype worldwide, accounting for 83.4 million (46.2%) with over one-third of genotype 1 cases located in East Asia. HCV genotype 3 is the next most common and is estimated to account for 54.3 million (30.1%) cases globally, approximately three-quarters of which occur in south Asia. Genotypes 2, 4, and 6 are responsible for the majority of the remaining cases of HCV worldwide, with an estimated 16.5 million (9.1%), 15.0 million (8.3%), and 9.8 million (5.4%) cases, respectively. East Asia accounts for the greatest numbers of genotype 2 and genotype 6 HCV cases, while North Africa and the Middle East have the largest number of genotype 4 cases. It has been reported that genotype 1 is the most common genotype in 85 of the 117 countries and is highly prevalent worldwide [5]. Genotype 2 dominated in West Africa, genotype 3 in south Asia and parts of Scandinavia, genotype 4 in Central and North Africa, Genotype 5 in South Africa, and genotype 6 in SE Asia [5]. The diversity of genotypes is high in China and many Southeast Asian countries

1.1.4 HCV life cycle

1.1.4.1 Attachment and entry

Attachment of the virus to a cell followed by viral entry is the first step in the virus life cycle. In order to enter the host cell, the virus must first bind to a receptor on the cell surface. After attachment, the nucleocapsid of enveloped viruses is released into the cell cytoplasm as a result of a fusion process between viral and cellular membranes. The entry process is controlled by viral surface glycoproteins that

trigger the changes required for mediating fusion. By analogy, HCV envelope glycoproteins are believed to belong to class II fusion proteins. However, HCV envelope glycoproteins do not appear to require cellular protease cleavage during their transport through the secretory pathway. HCV entry into cells is pH-dependent and endocytosis-dependent but the identity of the HCV fusion peptide remains unclear.

1.1.4.2 Translation and post-translational processing

Decapsidation of viral nucleocapsids releases free positive-strand genomic RNAs into the cell cytoplasm, where they serve, together with newly synthesized RNAs, as messenger RNAs for synthesis of the HCV polyprotein. HCV genome translation is under the control of the IRES of the 5'UTR and the first nucleotides of the core-coding region. The HCV IRES has the capacity to form a stable pre-initiation complex by directly binding the 40S ribosomal subunit without the need of canonical translation initiation factors. This binary complex then binds to eukaryotic initiation factor 3 [6], as well as the ternary complex eIF-2, GTP, to form a 48S-like complex. A second GTP hydrolysis step involving initiation factor eIF5B then enables the 60S ribosomal subunit to associate, forming a functional 80S ribosome that initiates viral protein synthesis.

1.1.4.3 Polyprotein processing

The main translation product of the HCV genome is a large precursor polyprotein that is subsequently processed by cellular and viral proteases into mature structural and nonstructural proteins (**Fig. 1**). From the hydrophobicity profile and dependence on microsomal membranes, junctions at core/E1, E1/ E2, E2/p7, p7/NS2 are processed by host signal peptidases. Cleavage by the host signal peptidase yields the immature form of the core protein. The signal peptide is further processed by a host signal peptide peptidase (SPP) to yield the mature form of the core protein. HCV nonstructural proteins are processed by two viral proteases: cleavage between NS2

and NS3 is a rapid intramolecular reaction mediated by a NS2-3 protease spanning NS2 and the N-terminal domain of NS3. In contrast, the remaining four junctions are processed by the NS3 protein. The NS3-NS5B region is presumably processed by sequential cleavage: NS3/4A>NS5A/5B>NS4A/4B>NS4B/5A.

1.1.4.4 RNA replication

HCV replication is assumed to start with synthesis of a complementary negative-strand RNA using the positive-strand genome RNA serves as a template. In The second step, negative-strand RNA serves as a template to produce numerous strands of positive polarity that will subsequently be used for polyprotein translation, synthesis of new intermediates of replication or packaging into new virus particles, both steps of which are catalyzed by the NS5B RdRp. The positive-strand RNA progeny is transcribed in a five to ten fold in excess compared to negative-strand RNA. NS5B RpRd was initially thought to catalyze primer-dependent initiation of RNA synthesis, either through elongation of a primer hybridized to the RNA template or through a copyback mechanism.

1.1.4.5 Viral assembly and release

As with related viruses, the mature HCV virion likely consists of a nucleocapsid and outer envelope composed of a lipid membrane and envelope proteins. The N-terminal half of the core protein is important for nucleocapsid assembly. Nucleocapsid assembly generally involves oligomerization of the capsid protein and encapsidation of genomic. This process is thought to occur upon interaction of the core protein with viral RNA, and the core-RNA interaction may be critical for switching from RNA replication to packaging. Once a HCV nucleocapsid is formed in the cytoplasm, it acquires an envelope as it buds through an intracellular membrane. Interactions

between the core and E1/E2 proteins are thought to determine viral morphology. HCV particles are released from the cell through the secretory pathway [2, 4].

1.1.5 Treatment of HCV infection

The currently used anti-HCV agents are targeted inhibitors against the important NS3/NS4A protease, NS5B polymerase and NS5A viral proteins [7]. For the treatment of HCV patients, the combination of PEGylated-interferon-α and ribavirin is extremely expensive but only effective for ~50% of HCV patients [8, 9]. Moreover, this therapy causes several adverse side effects, including alopecia, rash/itching, nausea, thrombocytopenia and anaemia. Consequentially the NS3/4A protease has become one of the attractive targets for anti-HCV drug design and development [4]. Importantly, the NS3 proteolytic activity and the acceleration of the cleavage rate are associated with the NS4A cofactor [10]. The HCV therapy has been improved by the use of boceprevir [11-13] and telaprevir [14, 15], which specifically inhibit the NS3/4A protease enzyme. Although this treatment leads to an increased rate of sustained virologic response, such therapeutic drugs have a low success rate [16]. Recent drugs, such as simeprevir [17, 18] and vaniprevir [19, 20], were approved for the treatment of HCV-infected patients in 2013 and 2014, respectively. In addition, other inhibitors are currently used in clinical studies such as danoprevir [21, 22], grazoprevir [23], faldaprevir [24], vedroprevir [8] and asunaprevir (ASV; [25, 26]). Because HCV lacks a proofreading activity of its RNA polymerase enzyme, high mutation rates in response to the treatments have been widely reported, resulting in resistance to drugs/inhibitors [27-30]. ASV is a second-generation NS3/4A protease inhibitor that was approved for use in combination with daclatasvir in Japan for the treatment of HCV genotype 1-infected patients [31]. However, ASV has not yet been approved in the USA and Europe, where it is currently in phase III clinical trials [32]. This compound shows a strong antiviral activity against genotypes 1 and 4 of HCV [25, 26]. From *in vitro* studies, the

acquired resistance against ASV caused by the two single point mutations of R155K and D168A in the NS3/4A protease results in a lower susceptibility to ASV, with an ~21- and 23-fold reduced half maximal effective concentration (EC_{50}), respectively [29]. These two mutated residues are located near the protease active site (**Fig. 2**) and were also found to reduce the binding affinities of some other anti-HCV inhibitors, such as danoprevir [21, 22], simeprevir [17, 18] and grazoprevir [23], by changing the conformation of the active site [33-35]. For instance, the absence of a Nε, which is the nitrogen atom at epsilon position, of residue K155 prevented the interaction between residues 155 and 168, which disrupted the interaction of ASV with the target protein [36].
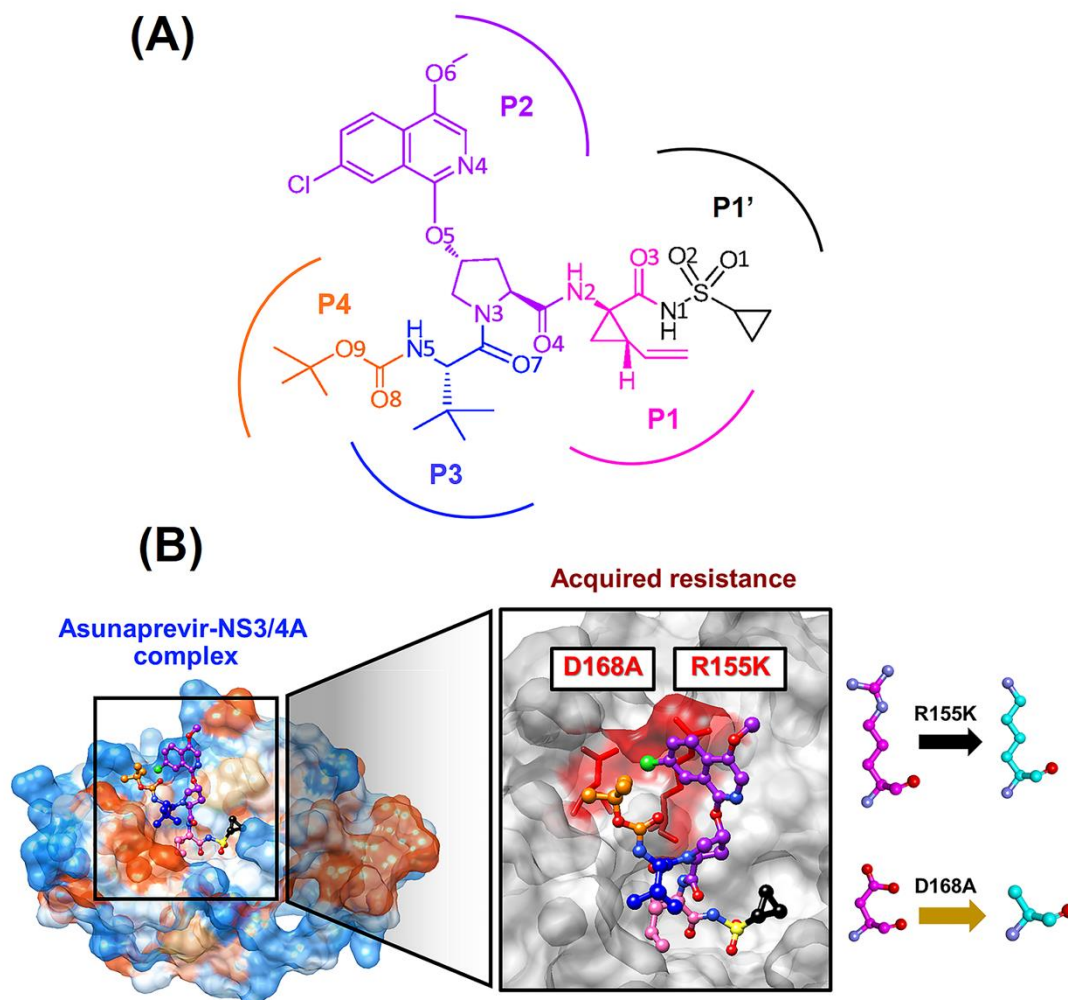
Figure 2 **(A)** Chemical structure of ASV, in which the P1', P1, P2, P3 and P4 substituents are shaded by black, pink, purple, blue and orange, respectively. **(B)** Three-dimensional structure of the ASV-NS3/4A complex, where the acquired resistance caused by the R155K and D168A point mutations towards the NS3/4A protease is also shown.

1.2 Hepatitis B virus

1.2.1 Genomic organization

HBV is a small enveloped DNA virus that classify to the family Hepadnaviridae, genus Orthohepadnavirus. Hepadnaviruses are para-retroviruses which reverse transcription occurs in the virus-producing cell from pregenomic RNA (pgRNA) [37]. HBV contains an open circular DNA genome, which is partially double-stranded and contains overlapping protein reading frames over much of its length, corresponding to a high information density (**Fig. 3**). The partially double-stranded DNA is represented in both the minus- and plus-strands, with the variably extended region of the plus-strand shown dashed, the direct repeats marked and the RNA primer shown as a zig-zag. The ribosome binding sites are shown as grey arrows. The overlap over most of the genome and are labelled C (and preC) for the core (and pre-core), P for the polymerase, E for the envelope proteins (with the large, medium and small surface proteins shown outside as dotted lines, to show the preS1 and preS2 at the N-termini) and X for protein X [37, 38].

1.2.2 HBV life cycle

HBV infects only hepatocytes and binds specifically to an unknown receptor on the pre-S1 region of the surface protein which is attached on the liver cell membrane. Presently, the precise mechanism of viral entry has not been reported yet. However, the mechanism of viral entry that the viral particle passes through cellular plasma membrane of liver cell by fusion of the viral membrane and cellular plasma membrane have been proposed as potential pathway [38]. The virus binds to the sodium taurocholate co-transporting polypeptide (NTCP) receptor on the surface of hepatocytes and is endocytosed, releasing its DNA-containing nucleocapsid into the cytoplasm. After core is released into the cytoplasm and transport relaxed circular DNA (rcDNA) containing nucleocapsid to the host nucleus, the rcDNA is repaired at the nick

in the minus-strand and the gap in the plus-strand by both viral and cellular enzymes to form covalently closed circular DNA (cccDNA) as shown in **Fig. 4**. This process aims to produce the cccDNA in the nucleus [39].
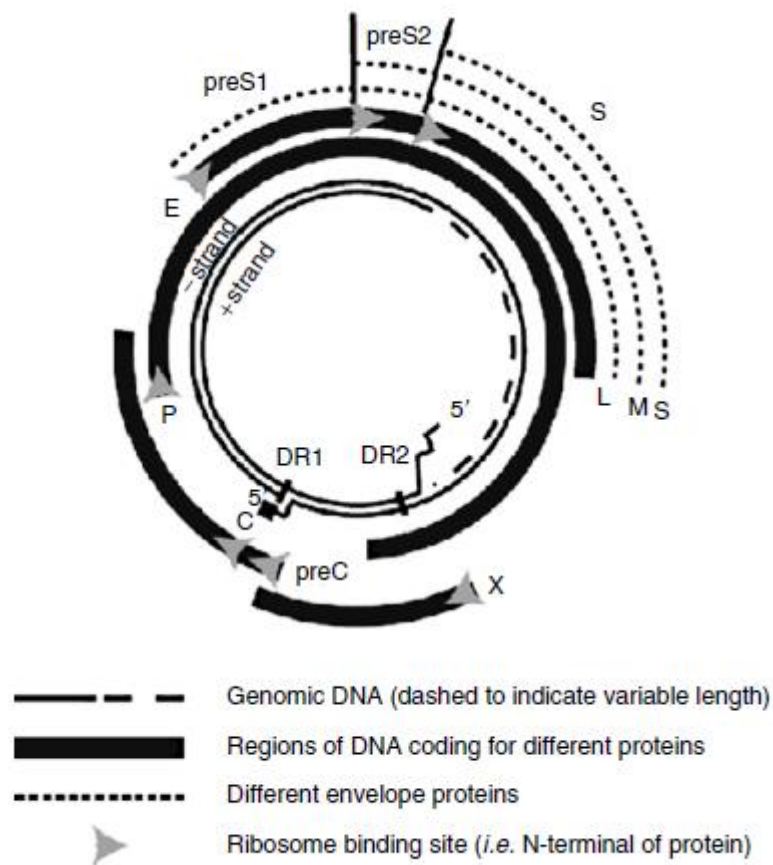
Figure 3 Genomic map for hepatitis B virus [37]

Integration of HBV DNA into the host genome also takes place. The cccDNA functions as a mini-chromosome and template for transcription of viral RNA. Due to the long half-life of hepatocytes, cccDNA will persist indefinitely within the host hepatocyte nucleus, thereby serving as a reservoir for reactivation of viral replication. The cccDNA consists of the precore/core gene, the polymerase gene, the PreS1/L-, PreS2/M- and Surface/S-gene and the X gene. The precore/core gene is coded for the nucleocapsid protein and for the secreted, non-structural, precore protein, the HBV

envelope antigen (HBeAg). Other three genes, the PreS1/L-, PreS2/M- and Surface/S-gene, are coded for the three envelope proteins.
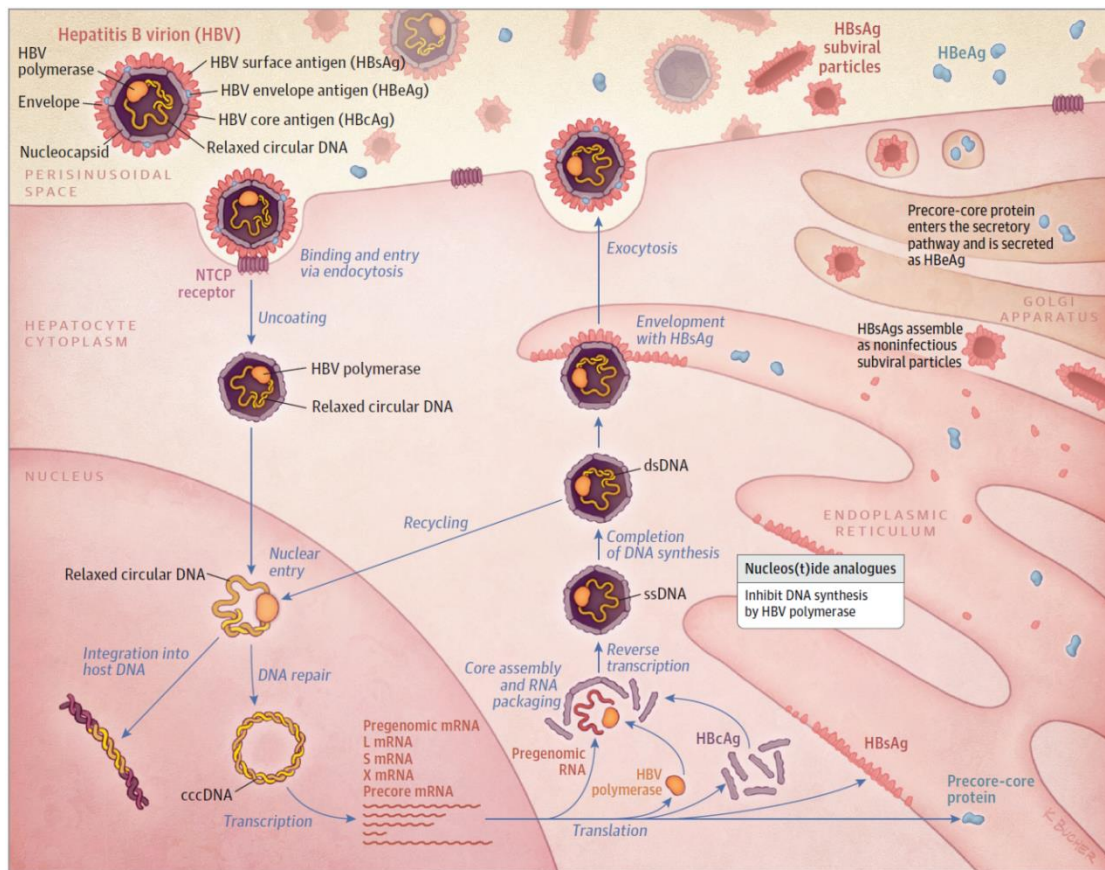
Figure 4 The hepatitis B virus and replication cycle [39]

The enveloped virions bind to the host cell by an unknown PreS1 region. The protein X has been suggested to be involved in many control processes in the infected cell, resulting in increased synthesis of viral proteins and also possible inhibition of apoptosis of the infected cell. Unfortunately, the mechanism of X protein is still the least understood. The cccDNA is transcribed into subgenomic RNA (sgRNA) and pregenomic RNA (pgRNA). The self-assembly of the RNA-containing viral nucleocapsid occurred via complex formation of the pgRNA with core protein and polymerase (RT domain). The assembly starts when the RT binds to a short hairpin element located

on the pgRNA. The RT in complex with RNA acts as the nucleus for assembly of core protein. This enzyme also consists of a terminal protein (TP) domain containing a tyrosine residue which acts as the primer for addition of the first deoxynucleotide triphosphate during synthesis of the minus-strand of the viral DNA. As the RT bound to pgRNA at a specific site (ϵ region) near 5'-terminus in Core assembly process, DNA synthesis starts copying specific sequence in ϵ region by receiving first deoxynucleotide triphosphate onto a tyrosine residue in the active site of the RT [37, 38]. The synthesis of minus-strand DNA proceeds through the 5'-direct repeat (DR1) to the 5'-terminus of pgRNA where it keeps for any further template. At this process, the primer DNA shifts from the 5'-DR1 onto the 3'-DR1, allowing further elongation of the primer. The RNase H activity starts to degrade the pgRNA (template) as the minus-strand DNA grows which means that an RNA/DNA double helix is the substrate for RNase H. After the RT reaches the 5'-terminus of the pgRNA for the second time, it will have a second minus-strand DNA with terminal repeats in the same orientation and a short RNA fragment, degrading by RNase H from the template. The short RNA fragment will hybridize to the 3'-DR1 region of the minus-strand DNA which it acts as the initial primer for the plus-strand synthesis. The synthesis of the DNA plus-strand begins when the end of the single strand DNA are close together which allow the primer (RNA fragment) to move from 3'-DR1 onto DR2 of minus-strand. After the 3'-end of primer bind to the 3'-DR1 of the minus-strand, the plus-strand is now ready for elongation process catalyzed by the RT function. These RNA-containing cores are still immature which are not ready to envelop and do not normally interact with the surface protein, until double-stranded DNA has been formed. The nucleocapsids are not only re-imported to the nucleus but also enveloped and released via the endoplasmic reticulum (ER). In case of import to the nucleus, the pgRNA is reverse transcribed into negative-strand DNA by the RT inside the nucleocapsid (as describe above). Afterwards, rcDNA is created by plus-strand synthesis from the negative-strand DNA for cccDNA amplification in the nucleus. On

the other hand, in cytoplasm, RNA-containing nucleocapsids undergo a maturation process to DNA containing nucleocapsids by reverse transcription from pgRNA to pgDNA. After import into the ER lumen, the envelope proteins are secreted by the cell either as small, non-infectious subviral spherical (SVRs) or as infectious virions. Especially, the non-infectious SVPs are produced in a thousand to million fold excess over virions [37, 38].

1.2.3 Reverse transcriptase (RT)

The RT is an enzyme used to synthesis the complementary ploy-DNA by utilizing both poly-RNA and poly-DNA as template, a process termed reverse transcription [40]. These enzymatic activities have been referred to as RNA-dependent DNA polymerase and DNA-dependent DNA polymerase activity. The two activities are appeared to share a common active site and have very similar heat inactivation profiles [40, 41]. As like known DNA polymerase, the RT also requires a primer to initiate DNA synthesis which both ribo and deoxyribo-oligomers can used as primer. The efficiency of transcription differs with the source of reverse transcriptase and the divalent cations used in the reaction. However, the RT lacks proofreading activity due to the absence of 3'-5' exonuclease activity [42]. For instance, RT from avian myeloblastosis virus [43] shows the large number of incorrectly paired bases when copying ribopolynucleotide and deoxyribopolynucleotide templates. Importantly, the frequency of error of RT from AMV is 1 in 600 when transcribing homopolymer templates and 1 in 6000 when copying alternating copolymer templates. Moreover, the RT can specifically degrade the RNA moiety of an RNA-DNA hybrid by RNase H activity. The degradation is independent on the synthesis of complementary DNA since preformed hybrids are susceptible to degradation [44, 45]. This enzyme acts as an exoribonuclease differ from cellular RNase H which is an endoribonuclease [46, 47]. A processive nuclease is defined as an enzyme which once bound to the polynucleotide chain substrate

completely degrades the substrate before being released. In addition, the digested RNA product has a phosphate-5' and 3'-OH, suggesting that the degraded RNA template can act as primer for the synthesis of the second strand of DNA. From all available data indicate that DNA polymerase and RNase H activities reside on the same polypeptide which has different functional sites [40, 41, 47, 48].

1.2.4 Diagnosis

To diagnose the HBV infection, commercially available serological tests are used to detect HBV surface antigen (HBsAg), HBeAg, HBV surface antibody (anti-HBs), HBV core antibody (anti-HBc), HBV envelope antibody (anti-HBe), and HBV DNA [39]. Note that Chronic HBV infection is defined as detection of HBsAg on 2 times measured at least 6 months apart. Diagnosis of acute and chronic HBV infection can classify into 4 cases: HBV infection, resolved infection, immunity and isolated core. Individuals who present HBsAg for more than 6 months are defined chronic infection. Individuals with chronic HBV infection are typically asymptomatic and are diagnosed during routine health maintenance or screening. In case of acute infection, anti-HBc will be detected in the form of immunoglobulin M (IgM). Only 5 to 10% of individuals with acute HBV infection will progress to chronic HBV infection. Interestingly, only one-third of individuals develop symptoms such as fever, fatigue, malaise, abdominal pain and jaundice during an acute HBV infection. Among individuals who recover from HBV infection, 80% will develop anti-HBs and all will develop anti-HBc which is presented in form of immunoglobulin G (IgG) [39]. Besides, HBsAg is now no longer detected in this case. Thus, testing for anti-HBc is important for identifying individuals who may have been previously infected. Essentially, individuals younger than 19 years should be vaccinated, resulting in gaining immunity. Only anti-HBs will be detected in immunity case. Note that individuals with anti-HBs levels less than 10 IU/L are considered not immune.

1.2.5 Phases of HBV Infection

Individuals with chronic HBV infection can transition between periods of immunologic activity and inactivity several times during a lifetime. Chronic HBV infection is divided into 4 phases: HBeAg-positive immunotolerance disease, HBeAg-positive immunoactive disease, HBeAg-negative inactive disease (inactive chronic HBV or low replicative), and HBeAg-negative immunoreactive disease. There are 2 immunologically active phases, including the HBeAg-positive immunoactive phase and HBeAg-negative immunoreactive phase, and 2 immunologically inactive phases, including the HBeAg-negative immunotolerance phase and the HBeAg-negative inactive phase. Individuals in the immunotolerance and HBeAg-negative inactive disease phases are asymptomatic while individuals in the HBeAg-positive immunoactive and HBeAg-negative immunoreactive disease phases can transition from asymptomatic to liver failure. To determine the phase of chronic HBV infection, the serological markers are used to detect the presence or absence of HBeAg and anti-HBe, the quantity of HBV DNA (HBV DNA level), level of alanine aminotransferase (ALT; known as a sensitive marker of liver inflammation) and the presence or absence of intrahepatic necroinflammation and fibrosis. The use of these biomarkers shows the presence of intrahepatic immunologic activity which provides a damage level of liver cell. More importantly, combining these biomarkers is important in determining the need for treatment initiation [39]. Individuals with acute HBV infection are positive for HBeAg. With the progress from acute HBV infection to chronic HBV infection, individuals can become negative for HBeAg (loss of HBeAg). The transition from positive to negative of HBeAg is associated with a decrease in HBV DNA level (<2,000 IU/mL) and suspension of liver inflammation and injury which is in the HBeAg-negative inactive disease phase. However, the duration of each phase is not well known, but most individuals with chronic HBV infection who infected perinatally have undergone HBeAg loss and only 6 to 10% of adults older than 40 years remain positive for HBeAg. During the course of

chronic HBV infection, the loss of HBsAg with or without seroconversion to anti-HBs and an undetectable HBV DNA level in the peripheral blood may occur which is associated with improved survival and reduced risk of liver failure and hepatocellular carcinoma.

1.2.6 Global prevalence and genotype distribution of hepatitis B virus infection

Hepatitis B virus (HBV) is a major health public problem, infecting over 240-257 million people worldwide. Chronic hepatitis B infection is defined as the detection of serum hepatitis B surface antigen (HBsAg; the viral glycoprotein) after 6 months of infection. By the infected with chronic hepatitis B virus, it can progress to low-grade liver inflammation, with episodes of transienthigh-grade liver inflammation and activation of fibrogenic processes, leading to liver fibrosis and cirrhosis, which may culminate in decompensated (symptomatic) liver disease and/or the development of hepatocellular carcinoma (HCC) in 25–40% of HBV carriers [39, 49]. In addition, HBV infection has further oncogenic potential after HBV integration into the host genome, which is an additional pathway contributing to HCC. Hepatocellular carcinoma and cirrhosis represent the fifth and the ninth most common cancer in males and females, causing 88,700 deaths in 2015 [50]. The prevalence of HBV infection in the general population depends on the region and ranges from 0.01% in northern European countries to 22.3% in the Pacific region (3–5). In Europe, the average HBV prevalence has been estimated at 1.5%, with 15 million people infected [1, 51].

There is 10 different HBV genotypes (A-J), according to the difference in overall genomic sequences or in sequences of gene coding surface antigen, with a distinct geographic distribution have been discovered which show the highest prevalence of HBV in Central and east Asia, sub-Saharan Africa and the Pacific regions [39]. Genotype A is most common found in northern Europe, North America, India, and Africa, while

genotype B and C are most prevalent in southeast Asia [51]. Genotype D is most prevalent in southern Europe, the Middle East, and India [43, 52]. In Europe, genotypes A and D were found to be most prevalent [53]. Genotype A is further divided in subtypes A1–A4. Subtypes A1, A3, and A4 are mainly found in Africa, whereas A2 prevails in northern and central Europe and in North America [52]. In addition, there are nine known subtypes of HBV genotype D [54]. HBV subtype D1 is predominant in the Middle East and North Africa, subtype D2 has been found in Albania, Turkey, Brazil, western India, Lebanon, and Serbia, and subtype D3 in Serbia, western India, and Indonesia. Subtype D4 has been rare and has been reported from Haiti, Russia, the Baltic countries, Brazil, Kenya, Morocco, and Rwanda. Subtype D5 is common in eastern India, subtype D6 in Indonesia, Kenya, Russia, and the Baltic region, and subtype D7 in Morocco and Tunisia. Subtypes D8 and D9 have recently been discovered and reported from West Africa and India [54]. For other genotypes, there were present in Europe sporadically, with most cases being imported to Europe either as a consequence of immigration from the Middle East, Asia, and Africa in past decades, or imported by European travelers that were infected abroad. Currently, HBV infected people still suffer from chronic hepatitis B even though the HBV vaccine shows 95% effective because most of the current HBV infected patients were born before the vaccine was widely accessible [50].

1.2.7 Treatment of HBV infection

There are 7 antiviral treatments for chronic HBV infection available in the United States which can be classify into 2 groups: interferons and nucleos(t)ides analogues. Interferons (alfa, beta, and gamma) are cytokines that are endogenously produced by immune system cells in response to viral infections and was approved in 1991 [39, 50]. Afterwards, it will be substituted by its pegylated from (Peg-IFN). The mechanism of Peg-IFN therapy related on the induction of long-term immune control

in response to viral infections (host immunomodulation effects) including boosting host immune response against infected hepatocytes, facilitating viral clearance. Nevertheless, the exact mechanism of an antiviral effect is not clear yet, but it is also believed to have direct antiviral including degradation of cccDNA and viral mRNA and inhibition of viral DNA. Even though all type of interferons have antiviral and immunomodulatory effects, but only alfa and beta interferons show more potent antiviral actions. Regardless, this antiviral treatment still has significant drawbacks such as an adverse safety profile and high response variableness which is ineligible and unsuitable for many patients. The adverse effects include flu-like symptoms: asthenia, pyrexia, myalgia and headache; thyroid dysfunction; gastrointestinal symptoms: nausea, diarrhea and decreased appetite; psychiatric disorders: depression and insomnia; dermatologic disorders: alopecia and pruritus; neutropenia; thrombocytopenia [55]. The best treatment response of Peg-IFN occurs in patients with HBV genotype A infection. In randomized clinical trials, Peg-IFN achieved higher rates of HBeAg loss, compared with lamivudine. Long-term treatment with interferon has demonstrated an association between therapeutic response (with HBeAg loss and sustained HBV DNA suppression) and higher incidence of HBsAg loss, improved liver histology and a reduction in cirrhosis and hepatocellular carcinoma compared with untreated controls. However, overall responses remain suboptimal in that only approximately one-third of patients achieve HBeAg loss and fewer achieve HBsAg loss [50].

NAs are widely used as anti-HBV drug agents against HBV and also HIV infections. The current NAs for HBV treatment include lamivudine (3TC) [56], telbivudine (TBV) [6], adefovir dipivoxil (ADV) [57], tenofovir disoproxil fumarate (TDF) [58], tenofovir alafenamide [59] and entecavir (ETV) [60]. The NAs are phosphorylated by intracellular kinases to form their diphosphate or triphosphate active forms. The active form of NAs prevents additional dNTPs from being incorporated by replacing themselves into HBV-

RT active site instead of dNTPs in elongating DNA strand [50, 61]. As the HBV-RT is homologous to the human immunodeficiency virus type 1 reverse transcriptase (HIV-1 RT), it could be reasonable to use HIV-1 RT drugs against HBV-RT [62]. The approved HIV drugs have been discovered, including stavudine (d4T) [63], didanosine [64], and zidovudine (ZDV) [65]. 3TC is the first oral NA approved by the United States Food and Drug Administration (FDA) for the treatment of chronic hepatitis B infection in 1998 [66] and has been the most experienced oral antiviral in chronic hepatitis B infection patients [67, 68]. However, long-term treatment of 3TC is associated with high rates of drug resistance, which lead to virological relapse and biochemical flare [66, 68, 69]. Therefore, 3TC is recommended as a second-line therapy for the treatment of chronic hepatitis B infection [66, 69]. d4T, known as synthetic analogue of thymidine, is a highly potent inhibitor of HIV-1 replication in vitro. About 40 % of d4T is renally eliminated by glomerular filtration andtubular secretion [63, 70]. Peripheral neuropathy, observed in about 12 % of patients, is the most important adverse effect of d4T [70, 71]. DDI is an analogue of deoxyadenosine and renally eliminated like d4T [64, 70]. The most significant toxicities associated with the use of DDI are pancreatitis, peripheral neuropathy, hepatic steatosis, and lactic acidosis, which are manifestations of mitochondrial toxicity. Peripheral neuropathy, observed in approximately 20 % of individuals, while pancreatitis may occur in 1 to 7 % of patients, resulting in rarely used because of toxicity [70, 72]. Next, ZDV is a pyrimidine analogue, has an azido group substituting for the 3'hydroxyl group on the ribose ring [65, 70]. Clinical use of this drug has evolved from use as monotherapy to use in combination with other antiretroviral drugs (ARVs). With long-term use, the predominant adverse effect of ZDV is myelosuppression, including neutro-penia and anemia, which have been shown to occur in 16 and 24 % of patients, respectively [70, 73]. This drug also is associated with myopathy, lactic acidosis, and hepatic steatosis, resulting from mitochondrial toxicity. Importantly, ZDV overdose has been observed in neonates, receiving postnatal HIV

prophylaxis, which there was a self-limited lactic acidosis and neutropenia in most cases [74, 75]. From *in vitro* studies, these compounds demonstrated well tolerate, excellent oral bioavailability, significant susceptibility to HIV [63, 65, 76] and has also been reported to inhibit HBV-RT activity. For instance, ZDV shows inhibition activity against HBV-RT by 21 to 41 percentage of inhibition of activity by two repeated experiments [65, 77].
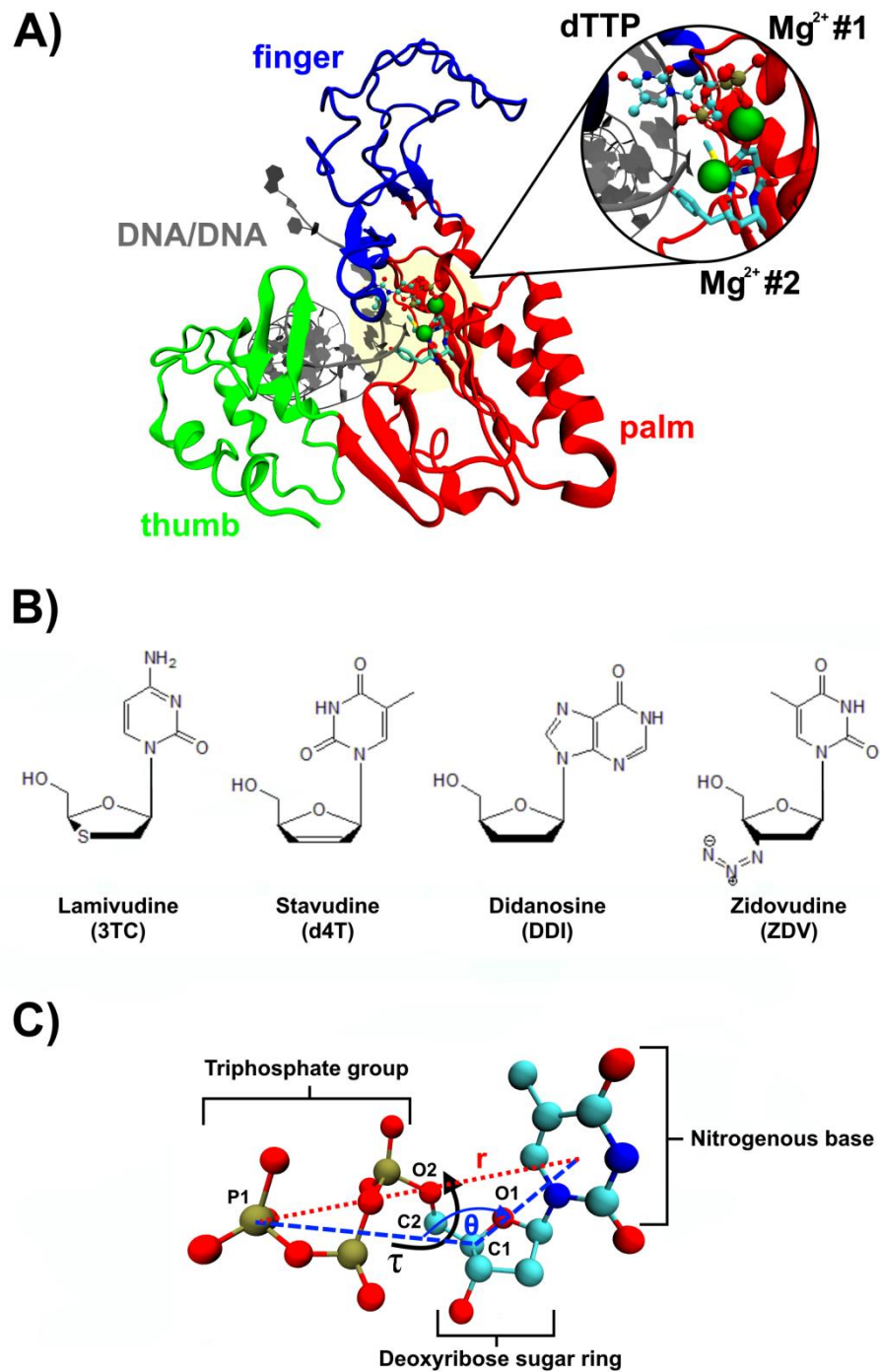
Figure 5 **A)** The 3D structure of HBV-RT composing of three subdomains: finger (blue), palm (red) and thumb (green) in complex with double-stranded DNA (grey), two Mg^{2+} (dark green) and dTTP substrate. **B)** The 2D chemical structures of anti-HBV and anti-HIV drugs. **C)** The topological angle (θ), torsional angle (τ),

distance (r) and the name of linked atoms were labeled and illustrated in nucleotides containing pyrimidine base.

1.3 Objectives

i)   To investigate the effect of such mutations on the ASV binding to NS3/4A, in terms of the inhibitor-protein interactions and binding efficiency towards the HCV NS3/4A protease

ii) To understand structural properties and interaction of HIV drugs including stavudine, didanosine and zidovudine against HBV-RT

iii) To study the various possible conformations change of anti-HIV drugs in unbound form and bound form with HBV-RT

iv) To screen for drugs that are potentially effective as anti-HBV and anti-HIV drugs against HBV-RT

v)   To investigate the inhibitory activity of anti-HIV drug and other drugs from screening against HBV

CHAPTER II

THEORIES

2.1 Homology modeling

Homology modeling is the computational approaches for protein three-dimensional structure modeling and prediction. Homology modeling builds an atomic model based on experimentally determined known structures that have sequence homology of more than 40%. It is also known as comparative modeling. If two proteins share a high sequence similarity, they are likely to have very similar three-dimensional structures. If one of the protein sequences has a known structure, then the structure can be copied to the unknown protein with a high degree of confidence. Homology modeling is moderately accurate for the positions of alpha carbons and inaccurate for side chain positions and loops. The others approaches are threading for <40% similarity and *ab initio* prediction for no homolog. Homology modeling divides into 3 types: basic modeling, advanced modeling and iterative modeling [78].

2.1.1 The overall homology modeling procedure

2.1.1.1 Template selection

The template selection involves searching the Protein Data Bank (PDB) for homologous proteins with determined structures. The search can be performed using a heuristic pairwise alignment search program like BLAST or FASTA. A database protein should have at least 40% sequence identity, highest resolution and the most appropriate cofactors for it to be a template sequence.

2.1.1.2 Sequence alignment

Once the template is identified, the full-length sequences of the template and target proteins need to be realigned using refined alignment algorithms to obtain optimal alignment.

2.1.1.3 Backbone model building

Once optimal alignment is achieved, the corresponding coordinates residues of the template proteins selected can be simply copied onto the target protein. If the two aligned residues are identical, coordinates of the side chain atoms are copied along with the main chain atoms. The average coordinate values of the templates are used in case of multiple templates.

2.1.1.4 Loop modeling

After the sequence alignment, there are often regions caused by insertions and deletions leads to gaps in sequence alignment. The gaps are modeled by loop modeling. The loops are generated by random many loops and searches for the one that has reasonably low energy and φ and ψ angles in the allowable regions in the Ramachandran plot.

2.1.1.5 Side chain refinement

The positions of side chains will be determined after the main chain atoms are built. A side chain can be built by searching every possible conformation by every torsion angle of the side chain to select the one that has the lowest interaction energy with neighboring atoms.

2.1.1.6 Model refinement and model evaluation

This step includes the energy minimization procedure on the entire model, which moves the atoms in such a way that the overall conformation has the lowest energy potential. The goal of energy minimization is to relieve steric collisions without altering the overall structure. In these loop modeling and side chain modeling steps, potential energy calculations are applied to improve the model. The final model has to be evaluated for checking the $\varphi-\psi$ angles, chirality, bond lengths, close contacts and also the stereochemical properties.

2.2 Molecular dynamics (MD) simulation

The study of the molecular structure is a key point in the understanding of biology. Biological function is based on molecular interactions, and these are a consequence of macromolecular structures. To observe molecular interactions by experimental processes is difficult, expensive, and takes a long time. MD simulation is used to overcome the limitations. MD simulation has advanced from simulating several hundreds of atoms to systems with biological relevance, including entire proteins in solution with explicit solvent representations, membrane embedded proteins, or large macromolecular complexes like nucleosomes or ribosomes. These simulations can capture a wide variety of important biomolecular processes, including conformational change, ligand binding, and protein folding, revealing the positions of all of the atoms at femtosecond temporal resolution. Importantly, the simulations can also predict the biological response from perturbations such as mutation, phosphorylation, protonation, or the addition or removal of a ligand at an atomic level. Moreover, MD simulations are often used in combination with a wide variety of experimental structural biology techniques, including X-ray crystallography, cryoelectron microscopy (cryo-EM), nuclear magnetic resonance (NMR) and electron paramagnetic resonance [79]. The choice of a proper potential is the most importance in obtaining accurate MD

simulations [80-82]. An approximate potential may be calculated from quantum mechanics and from the Born-Oppenheimer approximation in which only the positions of the atomic nucleus bonding are considered. The potentials may be divided into bonding potentials and long-range potentials. The bonding potentials involve interaction with two atoms (bond stretching), three atoms (bond bending), and four atoms (dihedral). Long-range interactions are associated with the Lennard-Jones potential (van der Waal) and the Columbic potential [83, 84]. The MD simulation method is based on Newton's second law or the equation of motion, $F_i = m_i a_i$, where F is the force exerted on the i particle, m_i is mass and a_i is acceleration of the i particle. From knowledge of the force on each atom, the acceleration can be calculated by the first-order differential of velocity respect to time or second-order differential of the position respect to time. Nowadays, MD simulation is achieved on popular programs such as AMBER, CHARMM, ACEMD, GROMACS, or NAMD.

2.3 The molecular mechanics/Poisson–Boltzmann or generalized Born surface area (MM/PBSA and MM/GBSA) methods

Computational methods that combine molecular mechanics energy and implicit solvation models, such as Molecular Mechanics/Poisson Boltzmann Surface Area (MM/PBSA) and Molecular Mechanics/Generalized Born Surface Area (MM/GBSA), have been widely exploited in free energy calculations [85]. Both of MM/PBSA and MM/GBSA allow for rigorous free energy decomposition into contributions originating from different groups of atoms or types of interaction [86, 87]. In MM/PBSA or MM/GBSA, binding free energy (ΔG_{bind}) between a ligand (L) and a receptor (R) to form a complex (RL) is calculated as:

$$\Delta G_{bind} = \Delta H - T\Delta S \approx \Delta E_{MM} + \Delta G_{sol} - T\Delta S \qquad (1)$$

$$\Delta E_{MM} = \Delta E_{internal} + \Delta E_{electrostatic} + \Delta E_{vdW} \qquad (2)$$

$$\Delta E_{internal} = \Delta E_{bond} + \Delta E_{angle} + \Delta E_{dihedral} \qquad (3)$$

$$\Delta G_{sol} = \Delta G_{PB/GB} + \Delta G_{SASA} \tag{4}$$

$$\Delta G_{SASA} = \gamma \cdot SASA + \beta \tag{5}$$

where ΔE_{MM}, ΔG_{sol} and $-T\Delta S$ are the changes of the gas phase MM energy, the solvation free energy and the conformational entropy upon binding, respectively. ΔE_{MM} includes $\Delta E_{internal}$ (bond, angle and dihedral energies), electrostatic for $\Delta E_{electrostatic}$ and van der Waals energies for ΔE_{vdW}. ΔG_{sol} is the sum of electrostatic solvation energy (polar contribution), $\Delta G_{PB/GB}$, and non-electrostatic solvation component (non-polar contribution), ΔG_{SASA}. The polar contribution is calculated using either GB or PB model, while the non-polar energy is estimated by solvent accessible surface area (SASA). The γ and β with the values of 0.00542 kcal/mol·Å^2 and 0.92 kcal/mol, respectively. The conformational entropy change $-T\Delta S$ is usually computed by normal-mode analysis on a set of conformational snapshots taken from MD simulations. MM/GBSA or MM/PBSA has been successfully applied to various protein-ligand or protein-protein/peptide complexes [88, 89]. In addition, MM/PBSA or MM/GBSA is sensitive to simulation protocols, such as sampling strategy of generating snapshots and entropy calculation methods as well as other parameters such as charge models, force fields, solute dielectric constant and radius parameters in continuum solvent models [90].

2.4 The quantum mechanics/molecular mechanics generalized Born surface (QM/MM-GBSA) method

The quantum mechanics/molecular mechanics (QM/MM) approach is a molecular simulation method that combines the strengths of the QM (accuracy) and MM (speed) approaches, achieving the prediction of ligand binding affinity and substrate specificity. The QM/MM-GBSA approach was calculated in the same manner as the MM/GBSA but the H term was computed by a hybrid of quantum mechanics (QM) and molecular mechanics (MM) [91, 92]:

$$E_{total} = E_{QM} + E_{MM} + E_{QM/MM} \tag{6}$$

where E_{QM} is energy of the QM region, E_{MM} is the energy of the MM region and $E_{QM/MM}$ is the interaction between QM and MM region.

2.5 The Replica-exchange molecular dynamics (REMD) method

It is usually difficult to obtain accurate canonical distributions at low temperatures by conventional simulation methods because simulations at low temperatures tend to get trapped in one of a huge number of local minimum-energy states. One way to overcome this multiple-minima problem is to perform a simulation based on non-Boltzmann probability weight factors so that a random walk in energy space may be realized. Random walks allow the simulation to pass any energy barrier and to sample a much wider phase space than by conventional methods. One recently approach to increase conformational sampling is parallel tempering. The parallel tempering, or the replica exchange MD (REMD) method, has emerged as one of the most widely used methods to enhance conformational sampling [93-95]. Several replicas of the system are simulated independently in parallel, under different conditions. At predefined intervals, neighboring pairs of replicas are exchanged with a specific transition probability. In most REMD studies, the temperature is used as the parameter that changes among the replicas. Exchanging temperatures allows conformations that are trapped in local minima at a low temperature to escape by passing to a higher temperature condition. Temperature based REM (T-REM) combines simulated tempering and multicopy simulations in a thermodynamic correct manner. Using T-REM, N independent copies (replicas) of the systems are propagated simultaneously at different fixed temperatures. At regular intervals pairs of replicas at successive temperatures are exchanged according to a Metropolis criterion allowing individual replicas to sample a range of temperatures. The exchange probability between two replicas (i and j) is calculated as:

$$P(i,j) = \exp(-\Delta) \qquad\qquad (9)$$

$$\Delta \equiv [\beta_i - \beta_j] \, (E_i - E_j) \tag{10}$$

$$\beta = 1/k_B T \tag{11}$$

where E is the potential energy of the system, T is the absolute temperature, and kB is Boltzmann's constant. At higher temperatures the increased thermal energy facilitates the exploration of conformational space and allows the system to cross barriers not readily crossed at the temperature of interest on a given time scale.

2.6 The pharmacophore-based virtual screening

The concept of pharmacophores is based on the assumption that the molecular recognition of a biological target shared by a family of compounds can be described by a set of common features that interact with a set of complementary sites on the biological target. The features are quite general such as hydrogen-bond donors, hydrogen-bond acceptors, positively and negatively charged or polarizable, hydrophobic regions, or metal-ion interactions. In addition, pharmacophore key elements might be a group of atoms and a part of the volume of the molecule together with geometrical constraint like distance, angles and dihedral angles [96]. The feature-based pharmacophore models have found extensive use in medicinal chemistry for hit and lead identification and during the subsequent lead to candidate optimization. A 3D pharmacophore is mainly based on the availability of the 3D structure of the binding site of the target. The model can be generated directly from the complex structure of the ligand and the target, according to the 3D structure of the target and a certain number of ligands. Unfortunately, pharmacophore representations cannot explain everything about binding of ligand to the biological target [97]. However, the pharmacophore approach still has proven to be successful, allowing the perception and understanding of key interactions between a target and a ligand.

CHAPTER III

MATERIALS AND METHODS

3.1 Software

AMBER14

AMBER16

Discovery Studio 2.5 Client

Discovery Studio 3.5 Client

Chimera 1.11

GaussView 3.09

Gaussian 09

DecoyFinder 2.0

MobaXterm 9.4

Pymol 1.1

SSH Secure Shell Client

VMD 1.9.2

OriginPro 8

VideoMach 5.9.0

KNIME Analytics Platform

LigandScout 4.2.6

National Center for Biotechnology Information (NCBI)

The European Bioinformatics Institute (EMBL-EBI)

NGL Viewer 2.0.0 server

Protein Data Bank (PDB)

PDB2PQR 2.0.0 Server

RAMPAGE server

SWISS-MODEL server
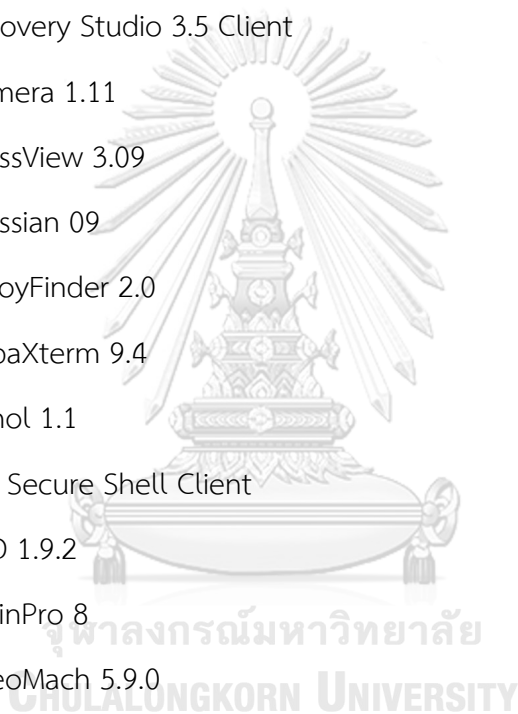
3.2 Computational methods

3.2.1 HCV NS3/4A protease

3.2.1.1 System preparation and ligand optimization

The three-dimensional structure of ASV in complex with the WT (PDB: 4WF8 and R155K mutant (PDB: 4WH6)) NS3/4A forms [98] was obtained from the protein data bank. The enzyme variant containing the D168A mutation was built by changing residue 168 from aspartate (D) to alanine (A) using the Discovery Studio 2.5 Accelrys Inc. [99]. The apo proteins were generated in all variants by removing the ligand from the complex structure. The protonation state of all ionizable residues was characterized using PROPKA3.1 [100] at pH 7.0. Note that, the cysteine (C) residue that is coordinated with Zn^{2+} was set as deprotonated state (CYM). The missing hydrogen atoms were added using the LeaP module [101] in AMBER14 [102]. The electrostatic potential charges of the ligand were computed at the HF/6-31(d) level of theory, while the restrained ESP charges and corresponding parameters of the ligands were generated using the antechamber and parmchk modules, respectively, in AMBER14 [79, 103, 104]. The AMBER ff03.r1 and GAFF force fields were applied for protein and ligand, respectively. Chloride ions were added for neutralizing the simulated systems. Each complex was solvated in a 80.9 X 69.5 X 77.9 $Å^3$ rectangular box of TIP3P water [105]. The distance between protein and solvation box edge was 10 Å. Prior to perform the MDSs, the added hydrogen atoms and water molecules were minimized with 1500 steps of steepest descents and conjugated gradient. Lastly, the whole system was minimized using the same procedures.

3.2.1.2 Molecular dynamics (MD) simulation

The MD simulations of all systems were performed under a periodic boundary condition in accord with the standard procedures [102]. All covalent bonds involving hydrogen atoms were fixed using the SHAKE algorithm [106]. The cutoff of 10 Å was

employed for nonbonded interactions, while the particle mesh Ewald summation method [107] was applied for calculating the long-range electrostatic interactions. A Langevin algorithm [108] was used to control the temperature with a collision frequency of 5.0 ps^{-1} for the first 3.5 ns and subsequently changed to 2.0 ps^{-1}. Each system was heated up to 298 K for 200 ps. Afterwards, all simulations, including the enzyme in APO and in CPX, were treated with the NPT ensemble at this temperature and a pressure of 1 atm using the PMEMD module. All systems were simulated for 100 ns at a time step of 0.002 ps [79]. In each system, 500 snapshots were collected every 1 ns of MD simulation. The binding free energy of all systems was performed using the semi-empirical QM/MM-GBSA method compared with MM/GBSA and MM/PBSA [109, 110]. The inhibitor and the two focused residues (155 and 168) were quantum mechanically treated, while the rest of system was performed using molecular mechanics. The per-residue decomposition free energy ($\Delta G_{bind}^{residue}$) based on the MM/PBSA method was performed to investigate the mutated amino acids (residues 155 and 168) associated with inhibitor binding. Additionally, the important motion of the protein in the WT and mutant (R155K and D168A) systems was characterized by PCA using CPPTRAJ module implemented in AMBER14 [102].

3.2.2 HBV reverse transcriptase

3.2.2.1 Homology modeling, system preparation and ligand optimization

Since the 3D structure of HBV-RT has not been available, the HBV-RT domain sequence (positions 349-693 based on NCBI sequence database (AY233278)) was submitted to SWISS-MODEL web-server [111] using a homologous HIV-RT structure (PDB: 4PQU) [112] as a template. Even though sequence alignment of both enzymes (HIV-RT and HBV-RT) showed only 20.9% of identity, the residues in the active site were conserved for these two enzymes. The Ramachandran plot (**Fig. 13**) showed the number of residues in favored and allowed regions of 78.4% and 20.1%, respectively.

Deoxythymidine triphosphate (dTTP), double-stranded DNA and two magnesium (Mg^{2+}) ions were remained in HBV-RT model according to their original orientation in HIV-RT crystal structure (PDB: 1RTD) [113]. The functional substituents of all HIV and HBV drugs were modified from dTTP substrate (PDB: 1RTD) using Discovery Studio 2.5 [114]. The protonation state of all ionizable residues was assigned using PROPKA3.1 [113] at pH 7.0. The missing hydrogen atoms were added using the LeaP module [101] implemented in AMBER16 [115]. The electrostatic potential (ESP) charges of ligand were calculated at the HF/6-31(d) level of theory, while the restrained ESP (RESP) charges and corresponding parameters of ligands were computed respectively using antechamber and parmchk modules in AMBER16 [103, 104]. The AMBER force fields ff14S, GAFF2 and OL15 were applied for protein, ligand and DNA, respectively. Sodium ions were added for neutralizing the simulated systems. Each complex was solvated using TIP3P water model [105] in a rectangular box (92.9 x 118.9 x 103.3 $Å^3$). The distance between protein and solvation box edge was 10 Å. with the spacing distance of 10 Å. Before performing MD simulations, the added hydrogen atoms and water molecules were minimized using 1,500 steps of steepest descents and conjugated gradient. Finally, the same procedures were used to minimize the whole systems.

3.2.2.2 Molecular dynamics (MD) simulations

MD simulations of all systems were performed under periodic boundary condition in according to the standard procedures [115]. SHAKE algorithm [106] was used to fix all covalent bonds involving hydrogen atoms. The cutoff of 10 Å was applied for non-bonded interactions, while the particle mesh Ewald summation method [107] was used for calculating the long-range electrostatic interactions. Langevin algorithm [108] was employed to control temperature with a collision frequency of 5.0 ps^{-1} for the first 3.5 ns and subsequently changed to 2.0 ps^{-1}. Each system was heated up to 310 K for 1 ns. After that, all systems were treated with *NPT* ensemble at this

temperature and pressure of 1 atm using the Langevin thermostat. All models were simulated until the simulation time reached 200 ns with time step of 0.002 ps [79]. The 2,500 snapshots of each system were collected in every 10 ns along the simulations. The binding free energy was predicted using the semi-empirical QM/MM-GBSA method and classical MM/GB(PB)SA [109, 110]. The inhibitor was quantum mechanically treated, while the rest of system was performed using molecular mechanics without considering entropy. The per-residue contribution free energy ($\Delta G_{bind}^{residue}$) based on the MM/GBSA method was performed to investigate the key binding residue associated with inhibitor binding.

3.2.2.3 Temperature replica exchange molecular dynamics (T-REMD)

All unbound forms of ligand in explicit TIP3P solvent were obtained after minimizing step of the whole system using the same methods as mentioned above. The number of non-interacting copies (replicas) was performed at different temperatures depending on the temperature range and the total number of particles in the system [95, 116]. Only pairs of replicas corresponding to neighboring temperatures were exchanged according to a Metropolis criterion [95, 116]. A pair of replicas (*i* and *j*) at neighboring temperatures (*n* and *m*) were exchanged with the probability: $P(i,j) = \exp(-\Delta)$, where $\Delta \equiv [\beta_n - \beta_m](E(q[i]) - E(q[j]))$ and $\beta = 1/k_B T$, q is the coordinate vector, E is the potential energy of the system, T is the absolute temperature, and k_B is the Boltzmann's constant. To obtain a suitable temperature of each replica and a uniform exchange probability, the temperatures were selected following to this equation: $T_i = T_0 \exp(ic)$ [95], where T_0 is the temperature of the first replica and c is constant value. We have used $T_0 = 298$ K and $c = 0.0345$, then generated the following sixteen temperatures (16 replicas): 298.0, 308.5, 319.3, 330.5, 342.1, 354.1, 366.5, 379.4, 392.7, 406.5, 420.8, 435.5, 450.8, 466.7, 483.0 and 500.0 K. Before starting the T-REMD simulations, each replica was performed a short simulation

for 2 ns to equilibrate the systems with an individual temperature. The T-REMD simulations were performed in the *NVT* ensemble. The pairs of replicas were exchanged every 5 ps with 40,000 times along the simulation 100 ns. Finally, the trajectories obtained from the 100 ns of each temperature were assigned as the production phase for further characterization of the structural parameters of each ligand.

To determine the probability of the conformations changes of anti-HIV drugs during the simulation, the trajectories were clustered based on structural analyses including the torsional angle (τ), distance (r) and topological angle (θ) between three subunits of compound (a nitrogenous base, deoxyribose sugar ring and triphosphate group). The conformation within a specified cutoff was categorized into the same cluster. The potential energy overlapping of all neighboring replicas was performed to verify the sufficient numbers of replica exchanges between pairs of replicas. The free energy surfaces (FESs) of compound were calculated from the ratio between the τ, r and θ in all conformations of each temperature as following: $\Delta G^T = -RT\log(P^T_{torsion}/ P^T_{angle})$ and $\Delta G^T = -RT\log(P^T_{torsion}/ P^T_{distance})$, where $P^T_{torsion}$, P^T_{angle} and $P^T_{distance}$ are the probabilities of the compound conformations, measured from the structural parameters of the τ, r and θ of three subunits, respectively, at the temperature *T*.

3.2.2.4 3D Pharmacophore-based screening

7,500 structures of 3TC and ZDV in complex with HBV-RT extracted from last 30-ns MD trajectories were used as a template for 3D pharmacophore-based screening (3D-PBS) using LigandScout 4.2.6 [96, 97]. Prior to perform 3D-PBS, the library of database was generated from DrugBank [117] containing 9,671 drug-like molecules. Note that only approved drugs were focused. The pharmacophore models were built from every two MD trajectories based on chemical diversity containing essential

features such as hydrogen bond (H-bond) donor, H-bond acceptor, hydrophobic and aromatic ring [118-120]. The same pharmacophore models were clustered and one model from each cluster was used as an agent model for screening. Subsequently, the hits identified from DrugBank database were filtered by applying similarity search for 50% of nucleoside-based scaffold. Finally, the obtained screening result was verified by conducting a receiver operating characteristic (ROC) plot. A set of inactive compounds (decoys) was created by mimicking the structures of active molecules obtained from PubChem database [121]. The ROC values were plotted and interpreted the term of sensitivity (true positive rate) and specificity (false positive rate) of screening results.

3.3 Experiment materials

3.3.1 Equipments

Centrifuge: Universal 320 R, Hettich, Germany

CO_2 Incubator: Forma™ Series 3 Water Jacketed, Thermo Fisher Scientific Inc., USA

Biosafety cabinets: Purifier Logic+ Class II, Type A2, Labconco, USA

Micropipette: Nichipet Premium LT, Nichiryo, Japan

Micropipette: Nichipet EXII, Nichiryo, Japan

MiniSpinPlus, Eppendorf Co., Inc., Germany

Thermo shaker incubator: MTH-100, China

Microcentrifuge, Heraeus Instrument, Germany

Real-Time PCR System: StepOnePlus™, Thermo Fisher Scientific Inc., USA

Spectrophotometer, Thermo Fisher Scientific Oy, Finland

Spectrophotometer: DS-11 Series, Denovix®, USA

Tissue culture plate: 24-well plates, Falcon®, USA

Tissue culture plate: 96-well plates, Falcon®, USA

Cell culture flask: T75 Flask, Falcon®, USA

Microscope: DMi1, Leica, China

Electronic pipetting: Easypet® 3, Eppendorf Co., Inc., Germany

Polypropylene (PP) centrifuge tubes: 15 mL, Axygen®, Mexico

Polypropylene (PP) centrifuge tubes: 50 mL, Axygen®, Mexico

Disposable serological pipettes: 5 mL, HCS®,China

Disposable serological pipettes: 10 mL, HCS®,China

Microtubes, Axygen®, China

PCR tubes, Kisker Biotech GmbH & Co. KG, Germany

Muse Cell Analyzer: Muse™, Merck, Germany

3.3.2 Chamicals

DMEM/High glucose: Hyclone™, GE Healthcare Life Sciences, USA

Fetal bovine serum (FBS), Hyclone™, GE Healthcare Life Sciences, USA

Non-essential amino acids (NEAA) 100X: Gibco, Thermo Fisher Scientific Inc., USA

0.25% Trypsin-EDTA (1X): Gibco, Thermo Fisher Scientific Inc., USA

3-(4,5-Dimethylthiazol-2-yl)-2,5-Diphenyltetrazolium Bromide (MTT): Invitrogen™, Thermo Fisher Scientific Inc., USA

Tenofovir disoproxil fumarate: the Center of Excellence in Hepatitis and Liver Cancer, the Faculty of Medicine, Chulalongkorn University, Thailand

Lamivudine, Glentham Life Sciences, UK

Stavudine, Glentham Life Sciences, UK

Trifluridine, Glentham Life Sciences, UK

Sofosbuvir, Glentham Life Sciences, UK

Floxuridine, Glentham Life Sciences, UK

3.3.3 Kits

GenUP™ gDNA Kit, Biotechrabbit, Germany

Muse® count & viability Kit: Muse™, Merck, Germany

3.3.4 Cell lines

The actively HBV-producing HepG2.2.15 human hepatoblastoma cell line was obtained from the Center of Excellence in Hepatitis and Liver Cancer, the Faculty of Medicine, Chulalongkorn University.

3.4 Experimental methods

3.4.1 HBV reverse transcriptase

3.4.1.1 Chemical reagents, cell line and culture

The HBV-producing HepG2.2.15 human hepatoblastoma cell line and TDF were obtained from the Center of Excellence in Hepatitis and Liver Cancer, Faculty of Medicine, Chulalongkorn University. The 3TC (#GP6029), d4T (#GP1591), fluxuridine (FdU, #GP4098), trifluridine (TFT, #GE1913) and sofosbuvir (SOF, #GP2923) were purchased from Glentham Life Sciences (UK). The HepG2.2.15 cells were maintained at 37°C under 5% CO_2 in Dulbecco's modified Eagle medium (DMEM) supplement with 10% (v/v) fetal calf serum, 50 mg/ml G418, non-essential amino acid (NEAA) and antibiotics (amphotericin B, penicillin, and streptomycin). The cultured medium was changed every three days.

3.4.1.2 Drug susceptibility assays

Drug stock solutions were prepared freshly in 100% DMSO as 200X stock and three additional 10-fold dilutions were prepared from these stocks. HepG2.2.15 cells were seeded into a 24-well plate at a density of 5 x 10^5 cells/well. Cells were treated

with different concentrations of drug(s) at days 1 and 3. Finally, cells were harvested on day 7 for real-time PCR (RT-PCR). Experiments were performed in duplicate.

3.4.1.3 Quantification of HBV DNA using real-time PCR and cytotoxicity assay

For RT-PCR, cellular DNA in HepG2.2.15 cells was extracted using GenUP™ gDNA Kit from biotechrabbit according to the manufacturer's instructions. Briefly, 1 ml of HBV DNA samples were amplified using 2X SYBR green PCR master mix (biotechrabbit) with the following primers for detecting intrahepatic HBV DNA: 5'-TCCTCCAAYTTGTCCTGGTYATC-3' and 5'-AGATGAGGCATAGCAGCAGGAT-3' and for HBsAg gene and *(ii)* 5'-GTGCACCTGACTCCTGAGGAGA-3' and 5'-CCTTGATACCAACCTGCCCAG-3' for β-globin gene. The PCR preparations of intrahepatic HBV DNA were incubated at 95°C for 3 minutes, followed by 50 cycles with 95°C for 15 seconds, 55°C for 20 seconds and 72°C for 30 seconds, and 60°C for one minute. Note that for β-globin, the annealing step was set as 58°C for 20 seconds instead. The data were analyzed using StepOne™ Software Version 2.1 (Applied Biosystems) and threshold cycle (C$_T$) values were obtained. To calculate copy numbers of HBsAg gene out of C$_T$ values, a standard curve was generated with 7-fold serial dilutions of an HBV plasmid (pHBV^{WT}; 10^3 to 10^9 copies per reaction) and 7-fold serial dilutions of a β-globin plasmid (10^3 to 10^9 copies per reaction), and the C$_T$ values for each sample were converted into HBV copy numbers. The HBV DNA inhibition activity was calculated as:

$$\text{The HBV DNA inhibition activity (\%)} = \frac{\text{Copy number}_{\text{HBsAg}}}{\text{Copy number}_{\beta-\text{globin}} \times S} \times 100$$

where *S* is the ratio of the copy number between HBsAg and β-globin of vehicle control (0.5% DMSO). Cell viability was assessed using the MTT assay. HepG2.2.15 cells were

plated in a 96-well plate at a density of 4.5×10^4 cells/well. After overnight incubation, cells were treated with various concentrations of a compound (1 µM to 100 µM) at days 1 and 3. After 7 days of incubation, the MTT solution (0.5 mg/mL) was added and incubated for another 1 hour. The medium was removed and 150 µL of DMSO was added to each well. The cell viability was detected by measuring the absorbance at a wavelength of 570 nm using spectrophotometer (Thermo Fisher Scientific Oy, Finland). The cell viability was calculated according to the following equation:

$$\text{Cell viability (\%)} = \frac{\text{Absorbance}_{570}(drug)}{\text{Absorbance}_{570}(0.5\% \, DMSO)} \times 100$$
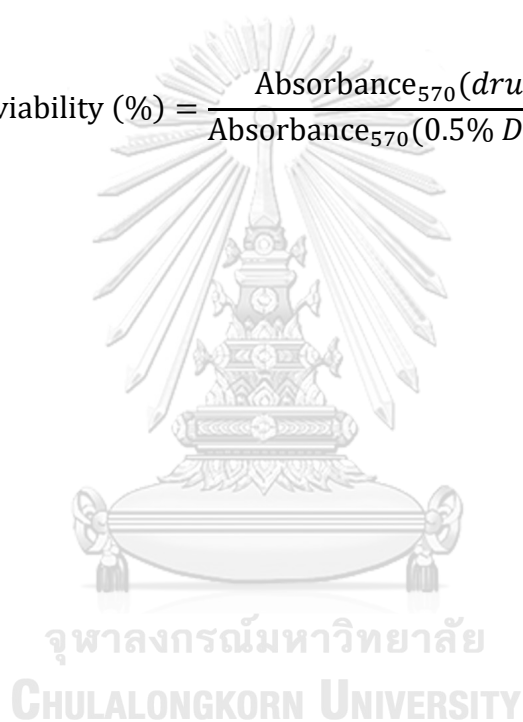
CHAPTER IV

RESULTS AND DISCUSSION

4.1 HCV NS3/4A protease

4.1.1 Stability of the global structure of each simulated model

The all-atom root-mean-square displacement (RMSD) was calculated in order to estimate the stability of each simulated model (WT and the R155K and D168A single point mutations in APO and in CPX with ASV) along the simulation time. As shown in **Fig. 6**, the RMSD value of each system rapidly increased during the first 10 ns and then fluctuated at ~2.2-2.7 Å in all the studied models until the end of the simulations. Altogether, the MD simulation systems tended to be stable after 60 ns. In this study, the equilibrated MD simulation trajectories extracted from the last 15 ns were used for further analysis of the *(i)* electrostatic network at the mutation site, *(ii)* protein conformation, *(iii)* inhibitor-protein hydrogen bond (H-bond) interaction and *(iv)* effect of mutations on the inhibitor binding affinity.
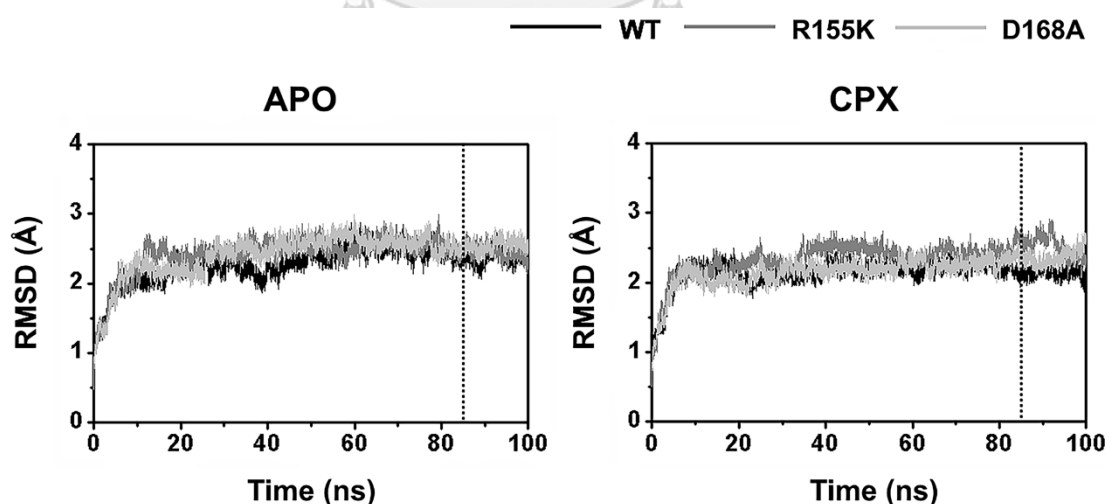

Figure 6 RMSD plots of the WT (black), R155K (grey) and D168A (light grey) HCV NS3/4A protease without (APO) and with (CPX) ASV binding for 100 ns of MD simulation

4.1.2 Electrostatic network at the mutation site

It has been reported that the electrostatic network formed at residues R123---R155---D168 plays an important role in the inhibitor binding [122]. In order to calculate this interaction network throughout the production phase of the simulation, we used the geometric criteria of: *(i)* a distance between the H-bond donor and acceptor of ≤3.5 Å and *(ii)* a H-bond angle of ≥120°. As shown in **Fig. 7**, a high percentage of H-bond occupation (%HB_{oc}) was detected in the WT system in both the APO and CPX. As expected, the R155K and D168A point mutations importantly decreased the electrostatic network formation compared to that of the WT NS3/4A, resulting in a lower binding affinity of ASV towards R155K and D168A NS3/4A. The point mutations decreased not only the electrostatic attractions but also the van der Waal (vdW) interactions, indicating that vdW forces also played a role in the drug binding (**Fig. 8**). The obtained results strongly agreed with previous reports [100, 105] as well as our QM/MM-GBSA free energy calculation (as discussed later and **Fig. 12**)

4.1.3 Protein conformation

To understand the effect of the R155K and D168A mutations on the protein conformational changes, PCA was performed [123]. These motions correspond to the correlated vibration or collective motion of groups of atoms. The scatter plots between the first (PC1) and second (PC2) principal components indicated the different distribution of protein conformations.

The PCA scatter plots (**Fig. 9A**) revealed that the conformational distributions of the holo (CPX) form of both mutations were dramatically different from those of the WT, implying that the R155K and D168A mutations in NS3/4A promoted protein conformational changes in a contrasting manner, which can significantly affect the protein-ligand binding affinity (**Fig. 12**, as discussed later). The first 15 PC modes of all studied models in the CPX and APO forms showed a percentage of accumulated

variance of >50%, indicating that these modes represent the important motions of protein (**Fig. 9B**).

Hydrogen bond	APO			CPX		
	WT	R155K	D168A	WT	R155K	D168A
O(D168)...N(R155)	99	99	97	99	99	98
OD1(D168)...NE(R155)	67	0	0	72	0	0
OD1(D168)...NH1(R155)	65	0	0	56	0	0
OD2(D168)...NE(R155)	82	0	0	66	0	0
OD2(D168)...NH1(R155)	58	0	0	72	0	0
OD1(D168)...NE(R123)	7	38	0	15	28	0
OD1(D168)...NH1(R123)	5	9	0	0	4	0
OD1(D168)...NH2(R123)	6	44	0	8	20	0
OD2(D168)...NE(R123)	3	45	0	14	24	0
OD2(D168)...NH1(R123)	1	5	0	0	10	0
OD2(D168)...NH2(R123)	3	40	0	7	23	0
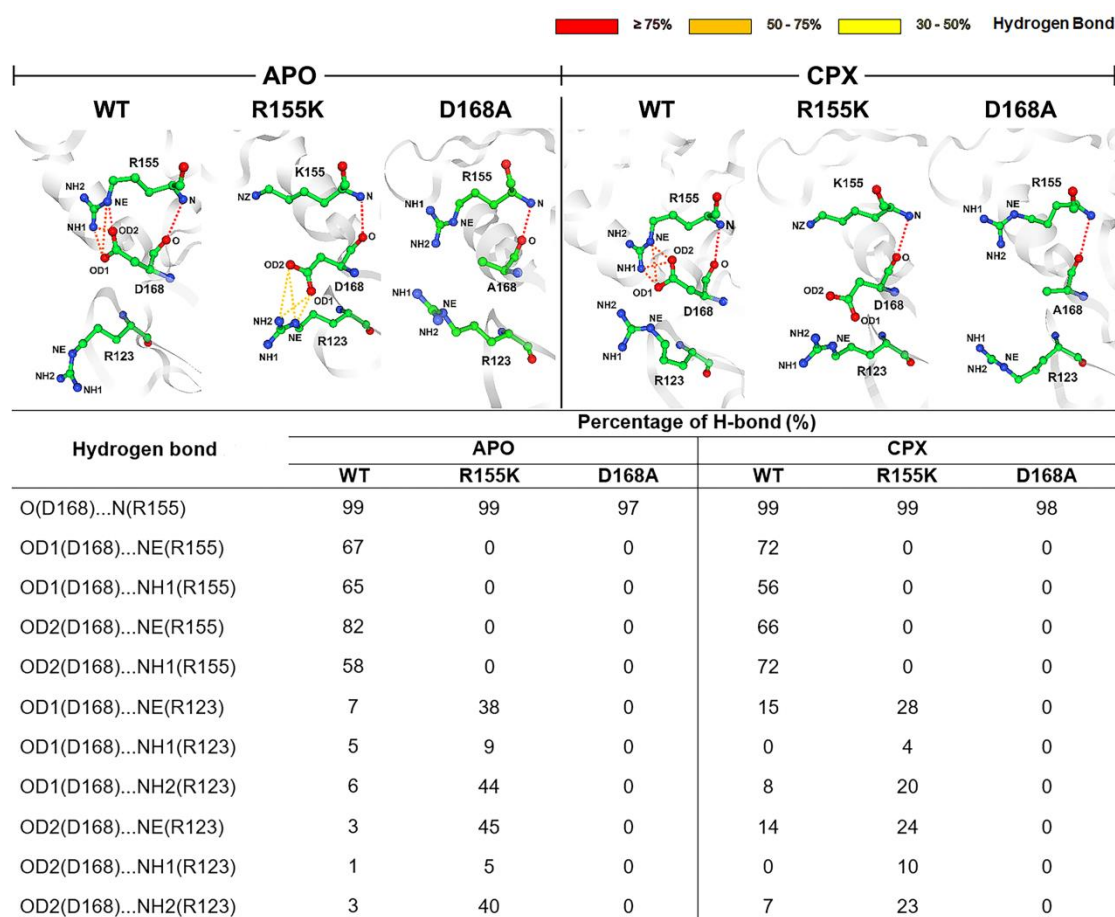
Figure 7 H-bond network of the three important NS3 residues (R123, R155 and D168) in the WT, R155K and D168A systems in both APO and CPX forms.

According to the porcupine plot of the WT system (**Fig. 10**, top), the direction of motion of residues R123, R155 and D168 moved towards the ligand binding, in a different manner from that in the APO form, suggesting that ligand binding induces the closed form of the protein, in which these three residues are important for ASV recognition. Additionally, the ligand interacted firmly in the pocket site due to the long side chains of two arginine residues. As expected, both R155K and D168A mutations caused the key binding residues 123,155 and 168 to significantly move outwards from

ASV, resulting in a loss of electrostatic attractions and H-bond interactions. In summary, these structural transformations could be the reason why ASV showed a low susceptibility towards the R155Kand D168A forms of the NS3/4A protease.
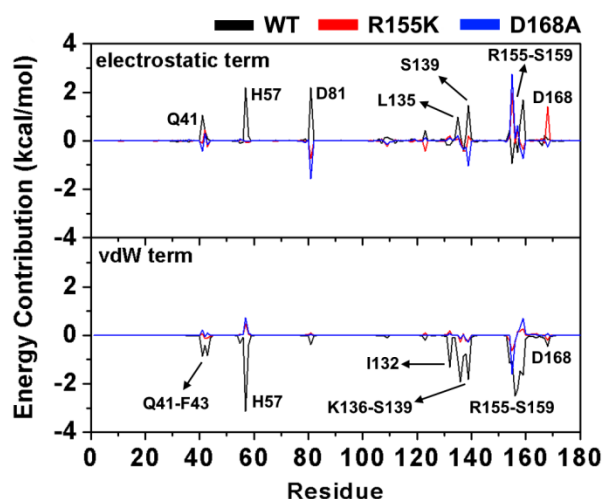
Figure 8 Averaged electrostatic and vdW energy contribution from each residue of NS3 protease domain in WT and mutant forms.

4.1.4 Inhibitor-protein H-bond interaction

One of the important factors involved in the stabilization of protein-ligand complexes is H-bonds. To measure such interactions throughout the overall simulation and equilibration phase (the last 15 ns), H-bond formation was calculated using the two geometric criteria described above, and the results are shown in **Fig. 11**. Among the five sites of ASV (P1' and P1-P4, **Fig.2A**), there were no H-bond interactions between the NS3/4A protease and ASV at the P2 and P4 moieties, in accord with the previous report on faldaprevir and danoprevir recognition [124]. For the WT NS3/4A (black), ASV showed 4 strong H-bonds (%HB$_{oc}$>70), one each with G137 and R155 at the P1' and P1 site, respectively, and two with A157 at the P3 site. With two firm H-bonds detected at the P3 site, the A157 residue played an important role as recognition site for anti-HCV protease inhibitors (**Fig. 11A**) [79]. Notably, the number of H-bonds formed in the mutated NS3/4A, especially R155K, was lower than those in the WT, suggesting that

these mutations attenuate the H-bonds, which is consistent with the loss of electrostatic network formation detected in the R155K and D168A mutations (**Fig. 11B**).
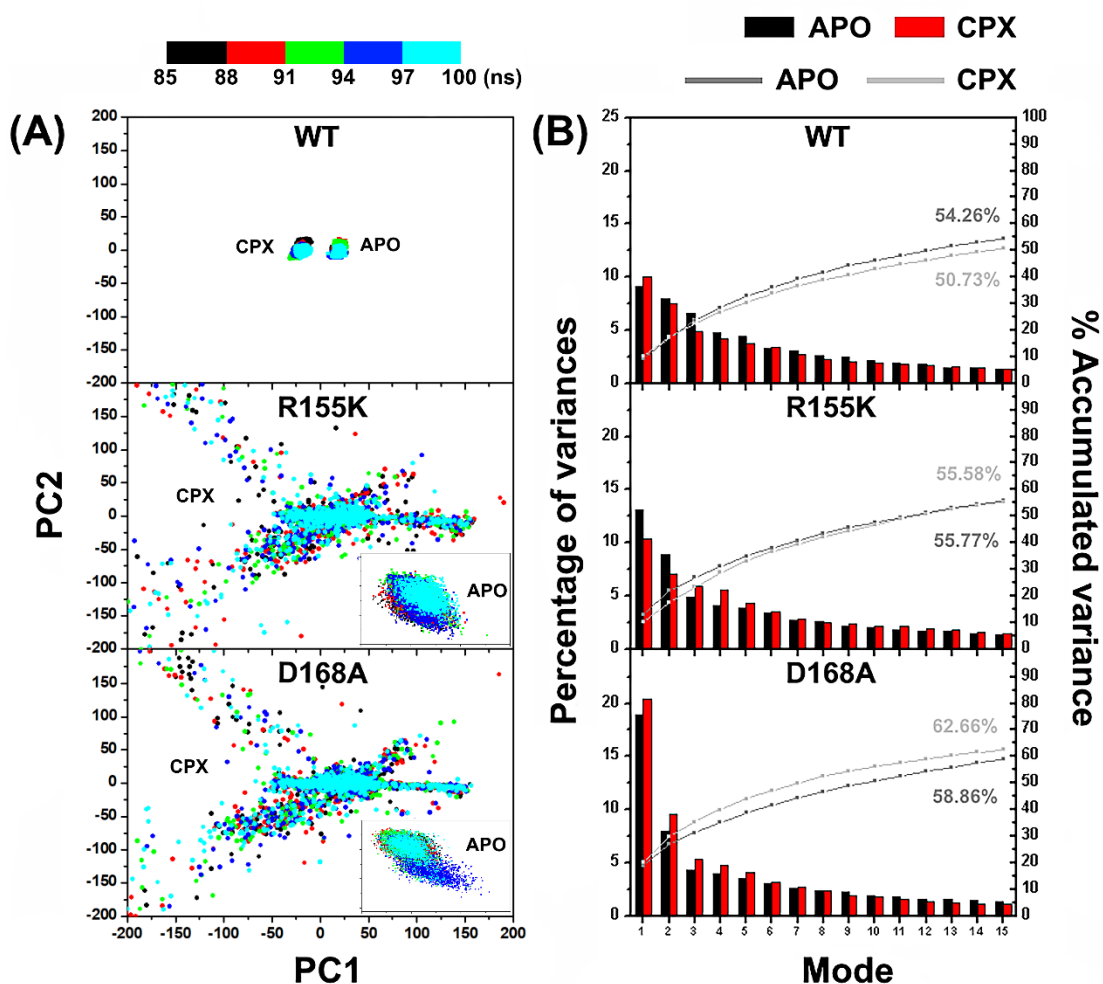
Figure 9 **(A)** The two-dimensional scatter plot of MD simulation trajectories between the first (PC1) and second (PC2) principal components. **(B)** The PCA scree plot of quantitative characters of the first 15 modes.

4.1.5 Effect of the R155K and D168A mutations in NS3/4A on the ASV binding affinity

To evaluate and compare the binding affinity of ASV against the WT and the R155K and D168A mutant strains of the NS3/4A protease, the QM/MM-GBSA binding free energy calculations were performed on 100 snapshots extracted from the last 15

ns of each MD simulation. Based on this approach, the ASV inhibitor and the NS3/4A residues 155 and 168 were treated by the semiempirical methods (AM1, RM1, PM3 and PM6) compared with MM/GBSA and MM/PBSA, while the remainder was described at the MM level. The results of the averaged binding free energies in each system are compared in **Fig. 12**, which demonstrated that the binding free energy ($\Delta G_{bind}^{residue}$) values obtained from all four levels of QM theory are in the same order of WT > R155K > D168A. The R155K and D168A single point mutations resulted in a significant reduction in the binding free energy by ~6 and ~9 kcal/mol, respectively, compared to the WT complex, in reasonable agreement with the experimentally observed 21- and 23-fold changes in the EC_{50} values, respectively. Unlike QM/MM-GBSA method, the free energy values between WT and both mutations obtained from MM/GBSA and MM/PBSA calculations were not significantly different. In addition, it should be noted that the calculated binding free energy does not give us an absolutely accurate binding free energy value in comparison with experimental data, but it evidently shows a relative binding affinity between WT and mutant strains toward ASV. Moreover, EC_{50} values which refer to concentration of a drug that causes the half maximal effect of an observed effect cannot directly compare to the binding efficiency of drug toward targeted protein.
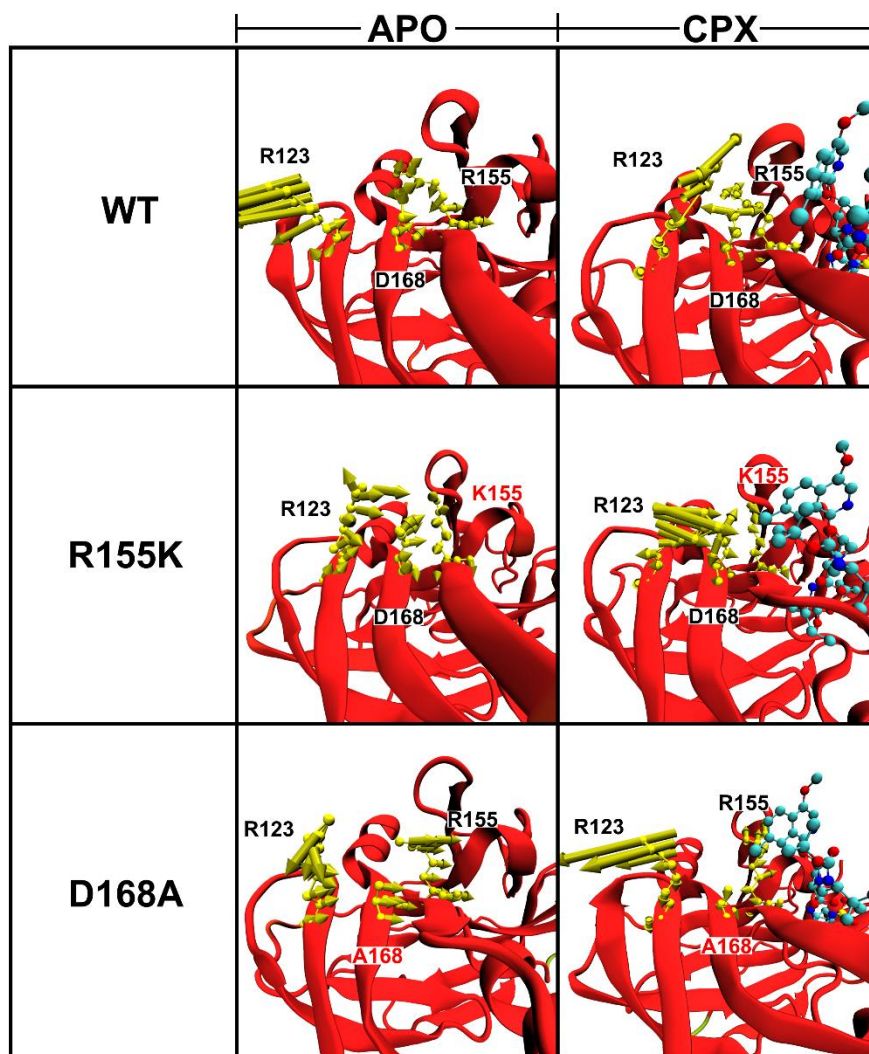
Figure 10 Porcupine plot of the PC1 of the studied systems. The arrow head and its length indicate the direction and amplitude of motions, respectively.

Moreover, the QM/MM-GBSA binding free energy calculations results were similar to previous computational studies, where the binding affinities of narlaprevir and vaniprevir towards the NS3/4AR155K mutant were significantly lower than that of the WT strain [125, 126], while the D168 A/V mutants of NS3/4A considerably decreased the susceptibilities of faldaprevir, danoprevir and vaniprevir [124, 125]. In summary, our QM/MM-GBSA binding free energy calculations not only supported the role the R155K and D168A mutations towards the weaker binding affinity of ASV, but

also strongly predicted the susceptibility of this inhibitor in same manner as the experimental data.
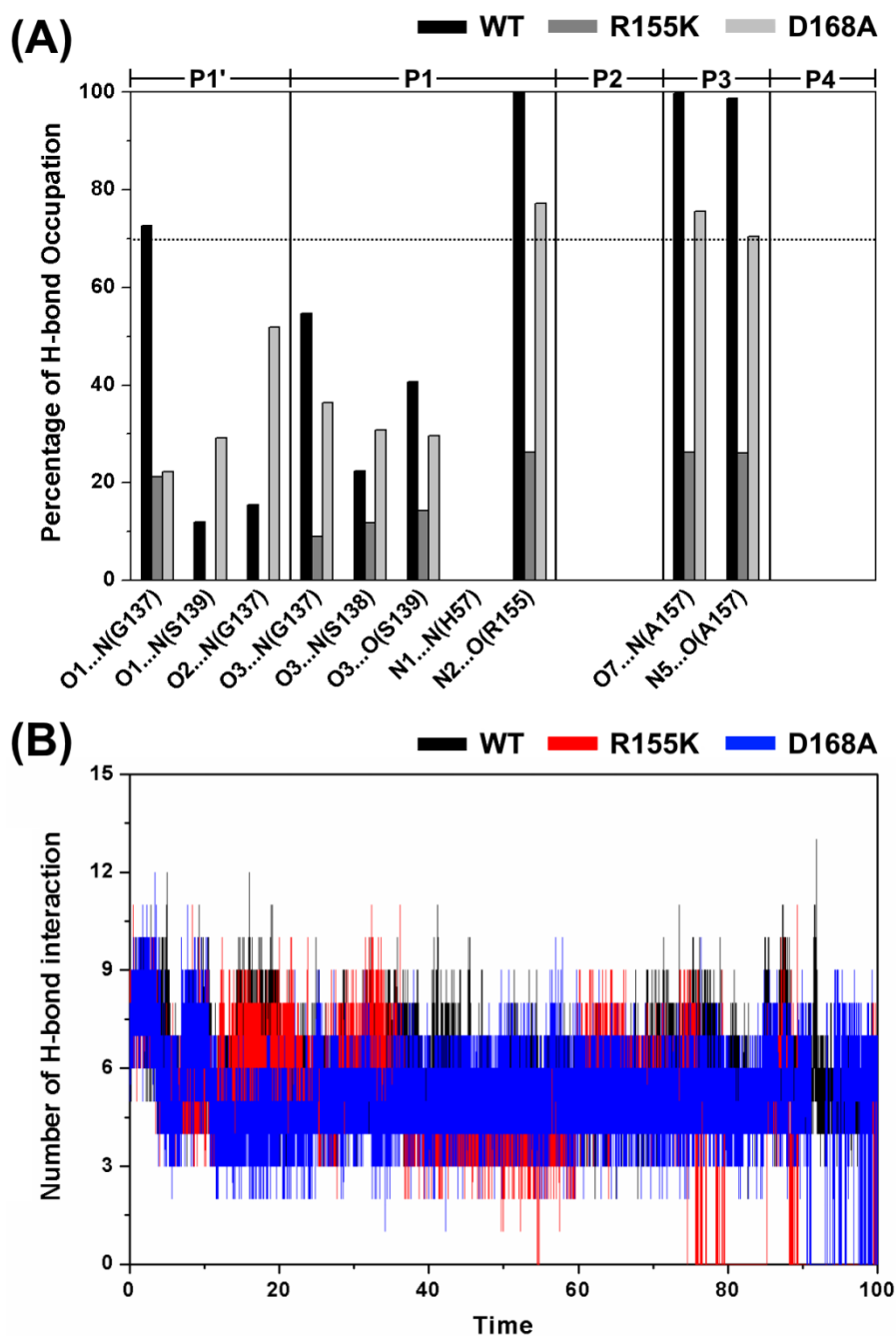
Figure 11 **A)** Percentage of H-bond occupation over the last 15 ns, and **B)** the number of H-bond (versus simulation time) of NS3 residues formed with ASV in the WT and the R155K and D168A mutant forms of NS3/4A.
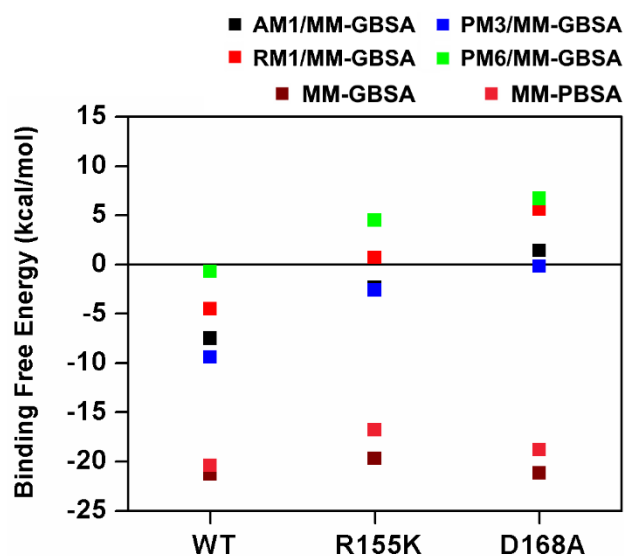
Figure 12 Binding free energy calculated by the MM/GBSA, MM/PBSA and the QM/MM-GBSA method, at the AM1 (black), RM1 (red), PM3 (blue) and PM6 (green) levels of QM theory for the WT and R155K/D168V mutant systems.

4.2 HBV reverse transcriptase

4.2.1 System stability of global structure

The root-mean-square displacement (RMSD) of residues around 5 Å of ligand in the active site was calculated to estimate the stability of all simulated models along the simulation times. As shown in **Fig. 14**, the RMSD value of protein residues in the active site of each system rapidly increased during the first 10 ns and then fluctuated at ~2.0-3.5 Å for all studied models until the end of the simulations. In contrast, the RMSD fluctuations of ligand were much less than those of protein by ~0.5-2.0 Å (except 3TC system). In the present study, the equilibrated MD simulation trajectories extracted from the last 30 ns were extracted for further analysis in terms of: the (*i*) key binding and drug-protein interactions, (*ii*) ligand topology, (*iii*) binding affinity of anti-HIV drugs against HBV-RT and (*iv*) virtual screening based on 3D-PBS.
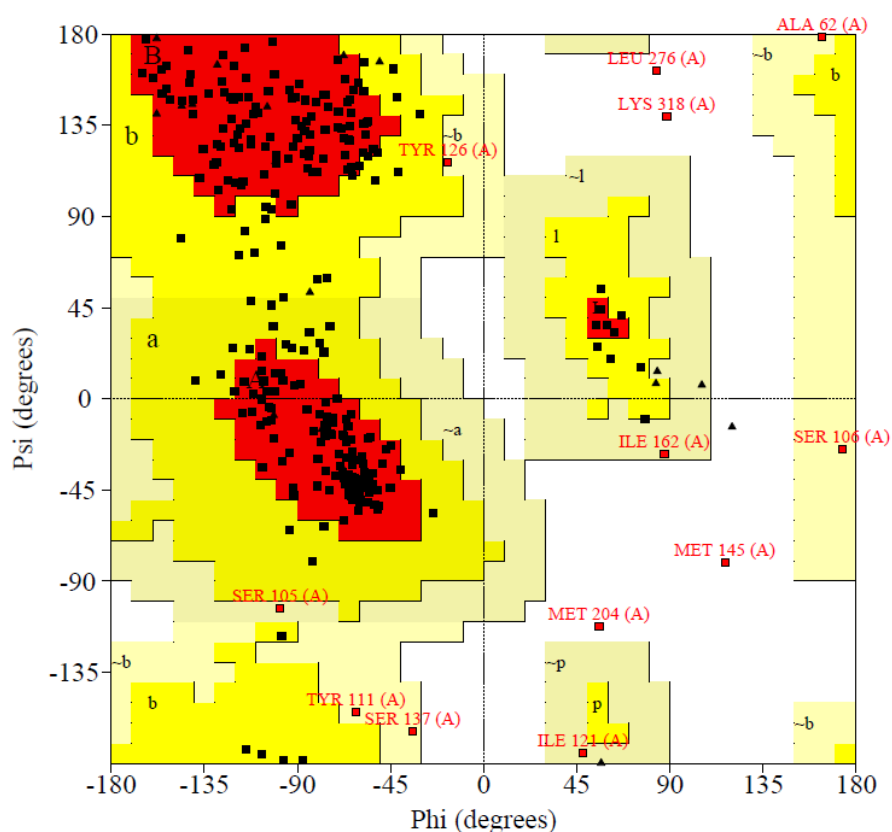
Figure 13 The Ramachandran plot of HBV-RT model which was modeled from HIV-RT.

4.2.2 Key binding and drug-protein interactions

The important amino acid residues involved in drug binding were characterized using $\Delta G_{bind}^{residue}$ calculation based on MM/GBSA method (**Fig. 15**). Note that the negative and positive values of the energy contribution correspond to the energy stabilization and destabilization, respectively. From the fingerprint of all systems with energy stabilization of ≤ -2 kcal/mol, the results demonstrated that there were seven polar residues (K32, R41, D83, S85, D205, N236 and K239) and three hydrophobic moieties (A86, A87 and F88) associated with drug binding. These results suggested that electrostatic attraction is the main force inducing protein-ligand complexation, which is consistent well with MM calculations (**Table 1**). The energy destabilization at D83

and D205 residues was occurred by the repulsion of negative charge between aspartic acid residues of HBV-RT and the phosphate groups of ligands.
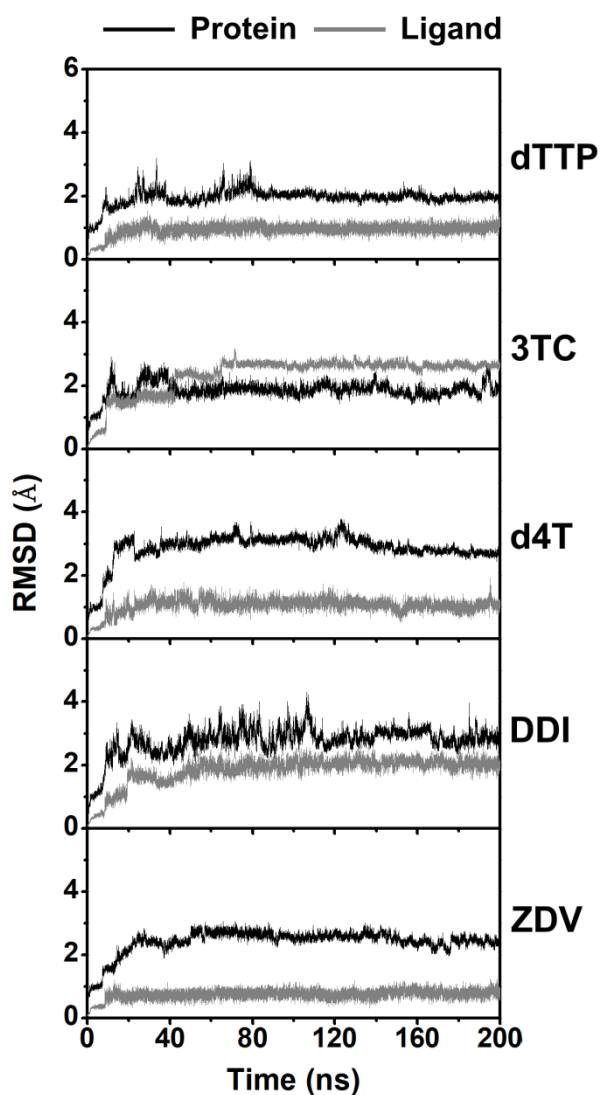
Figure 14 RMSD plots of residues around 5 Å from ligand in active site of HBV-RT (black) and ligand (grey), anti-HBV and anti-HIV drugs, systems for 200 ns of MD simulations

In order to support the role of Mg^{2+} as a cofactor, the $\Delta G_{bind}^{residue}$ of HBV-RT toward Mg^{2+} ions was calculated, and the obtained results (**Fig. 16**) revealed that two Mg^{2+} played an important role in stabilizing HBV-RT at D83 and D205 residues. Apart from amino acids residues, the double-stranded DNA, especially at 3'OH of nitrogenous

base, also contributed against ligand binding with the $\Delta G_{\text{bind}}^{\text{residue}}$ of ~-1 to -4 kcal/mol (**Fig. 17**). One of the important factors involving in the stabilization of protein-ligand complexes is H-bonds. To measure such interactions throughout the equilibration phase, H-bond formation was calculated using the two geometric criteria: (*i*) the distance between the H-bond donor and H-bond acceptor of ≤ 3.5 Å and (*ii*) the H-bond angle of ≥ 120° (weak H-bond, **Fig. 18A**) and ≥ 150° (strong H-bond, **Fig. 18B**). The obtained results showed that the H-bond formations were highly detected in ZDV system, especially R41 and S85 residues (**Figs 18A-B**). On the other hand, the DDI showed the lowest H-bonds (for both weak and strong H-bonds criteria) as compared with the other anti-HIV drugs. In the case of d4T, there were two firmly strong H-bonds detected at A86 and A87 residues, which are somewhat similar to the 3TC. Interestingly, the HBV-RT residues K32, A86, A87 and N236 could form several H-bond interactions with ligands rather than other amino acids, suggesting that these four residues are important recognized sites for ligand binding, which is in accord with the $\Delta G_{\text{bind}}^{\text{residue}}$ calculations (**Fig. 15**) and H-bond analysis from previous report [127].

Altogether, the obtained drug-protein interactions suggested that (*i*) the energy stabilization for drug binding not only came from protein target but also derived from the adjacent DNA molecules and (*ii*) electrostatic attractions and H-bond formations were the crucial interactions associated with drug binding

Table 1 MM calculations of anti-HBV and anti-HIV drugs for binding with HBV-RT based on MM/GBSA method.

Energy (kcal/mol)	Simulated system				
	dTTP	3TC	d4T	DDI	ZDV
ΔE_{vdW}	-23.13 ± 0.45	-18.89 ± 0.53	-17.92 ± 0.62	-15.73 ± 0.57	-21.07 ± 0.61
ΔE_{ele}	-622.46 ± 4.95	-394.17 ± 4.65	-670.69 ± 4.26	-360.51 ± 4.57	-659.71 ± 5.93
ΔE_{MM}	-645.59 ± 4.96	-413.07 ± 4.59	-688.61 ± 4.18	-376.24 ± 4.63	-680.78 ± 5.92
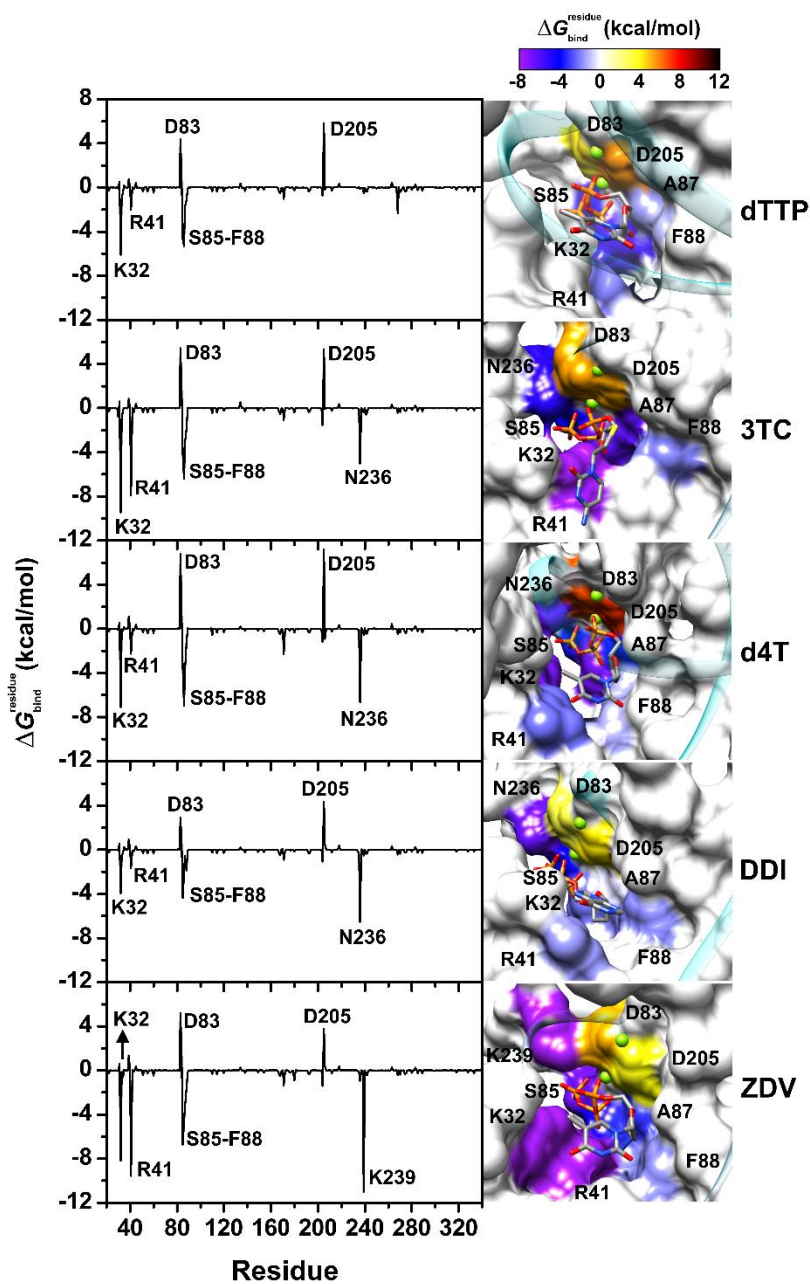
Figure 15 The per-residue contribution free energy of HBV-RT for dTTP, 3TC, d4T, DDI and ZDV in which energy contribution values are labeled from purple to black with -8 kcal/mol to 12.0 kcal/mol.
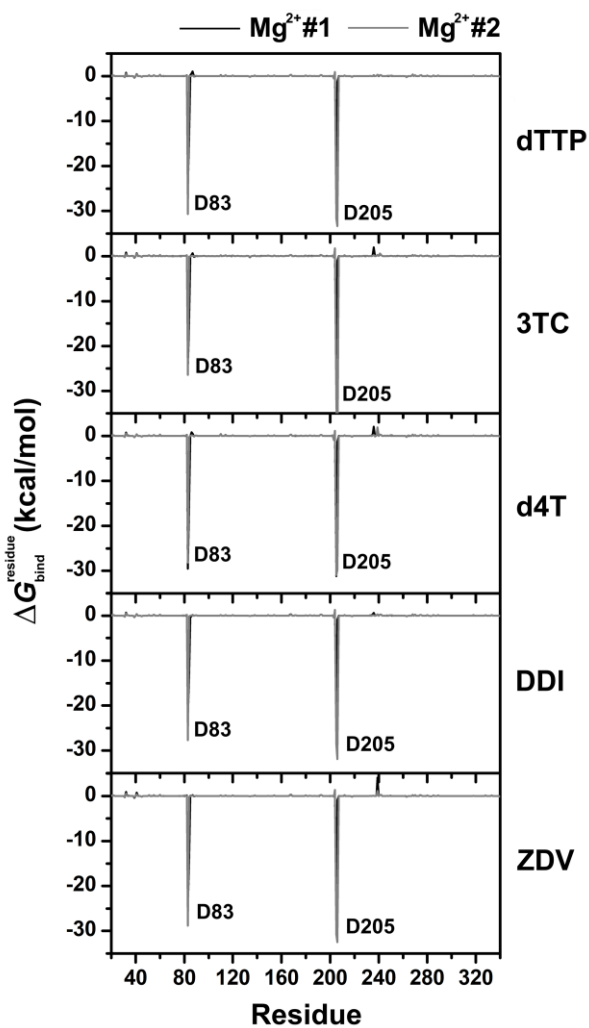
Figure 16 Averaged total energy from each residue of HBV-RT domain with Mg^{2+} in all systems.

4.2.3 Ligand topology

To characterize the conformation of HIV (d4T, DDI and ZDV) and HBV (3TC) drugs as well as HBV substrate (dTTP) in complex with and without HBV-RT, the structural parameters τ, r and θ (**Fig. 5C)** were calculated and plotted as the probability distribution over the production phase at 310 K. In addition, the representative snapshots extracted from T-REMD simulation are illustrated in **Fig 19B**. According to the canonical probability distributions (**Fig. 20**), there were enough overlaps between

all neighboring pairs of distributions in all systems, indicating the sufficient numbers of replica exchanges between pairs of replicas.
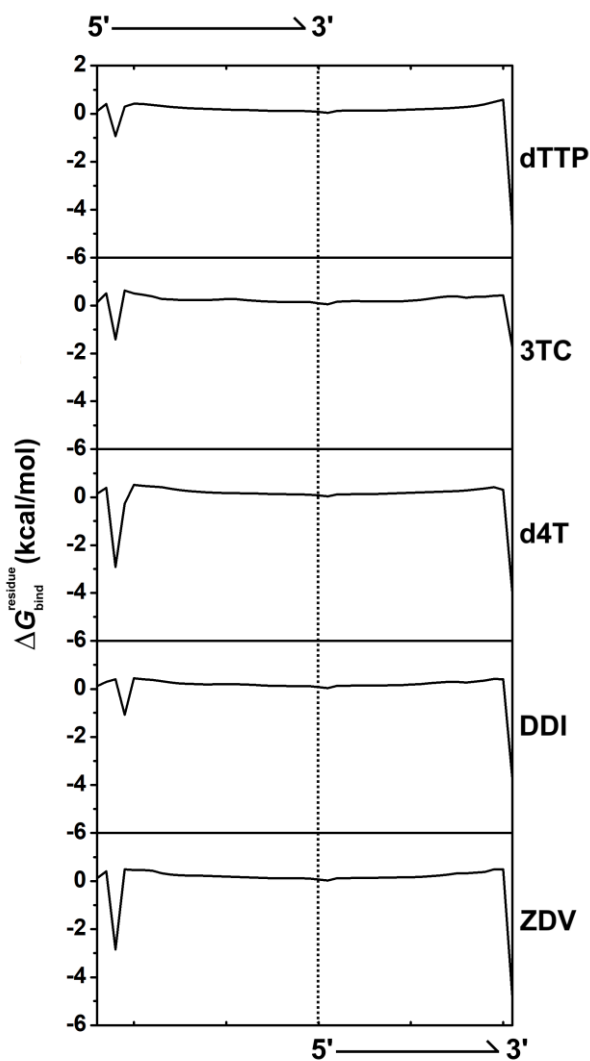
Figure 17 The per-residue decomposition free energy of double-strand DNA which in complex with HBV-RT for dTTP, 3TC, d4T, DDI and ZDV.

The τ and θ structural parameters were chosen to construct the FES profile (**Fig. 19**). In the case of unbound state, the distribution of τ of all ligands showed similar patterns (except d4T) in which the populations of τ and θ were mostly found at -180°, 60° and 180° and at 60° to 180°, respectively. By considering d4T system, the local

minima detected at τ of ~60° and at θ of ~60° to 120° was similar to that of ZDV. In addition, the DDI showed an almost identical conformational pattern with 3TC at the θ of -180°, 60° and 180°. As shown in **Fig. 21**, the distribution of r showed the similar trend with that of θ for both bound- and unbound states, in which only one population was detected (~5 to 13 Å for r (**Fig. 21**) and ~60° to 180° for θ (**Fig. 19**)). As expected, the FESs of ligands conformations in bound state were more negative than those of unbound ones, as evidenced by the obviously increased blue color, indicating that the protein-ligand complexation promoted the stable conformation of ligands.
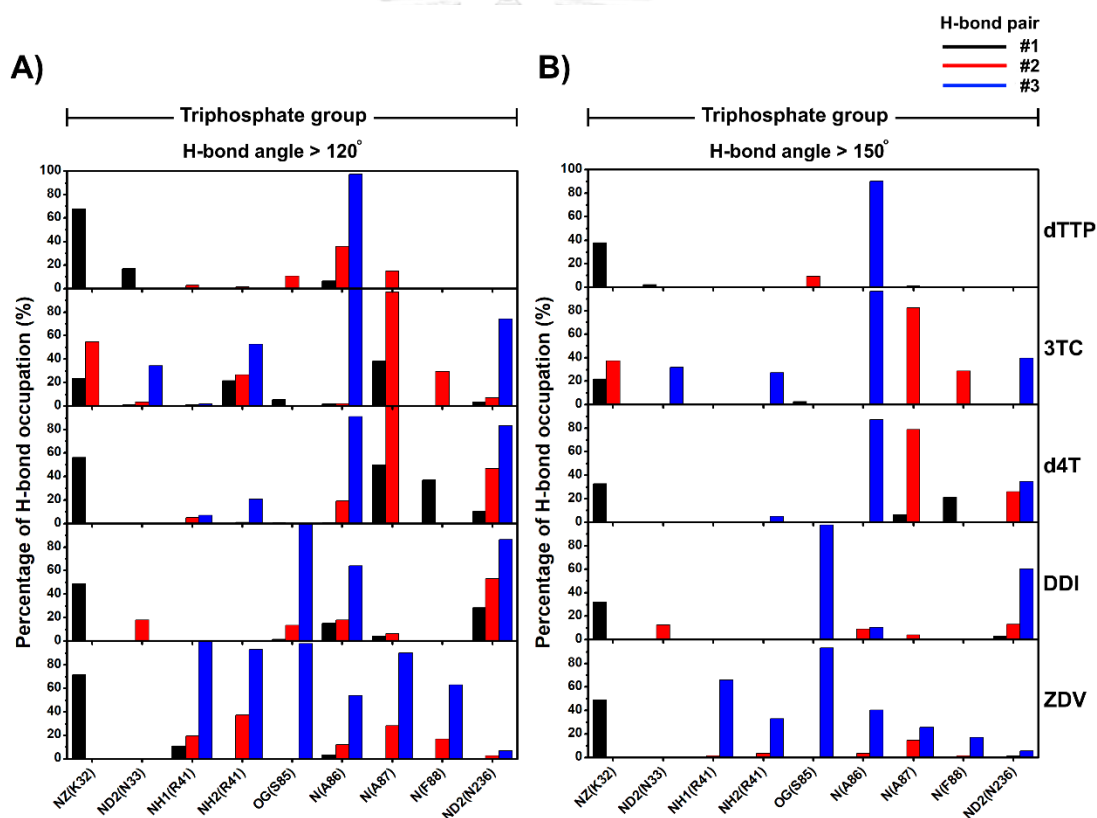
Figure 18 The H-bond occupation of HBV-RT formed with five ligands over 30 ns. **A)** Weak H-bond formation was calculated using angle ≥ 120°. **B)** Strong H-bond calculated using ≥ 150°. Each detected H-bond interaction was represented in black, red and blue.
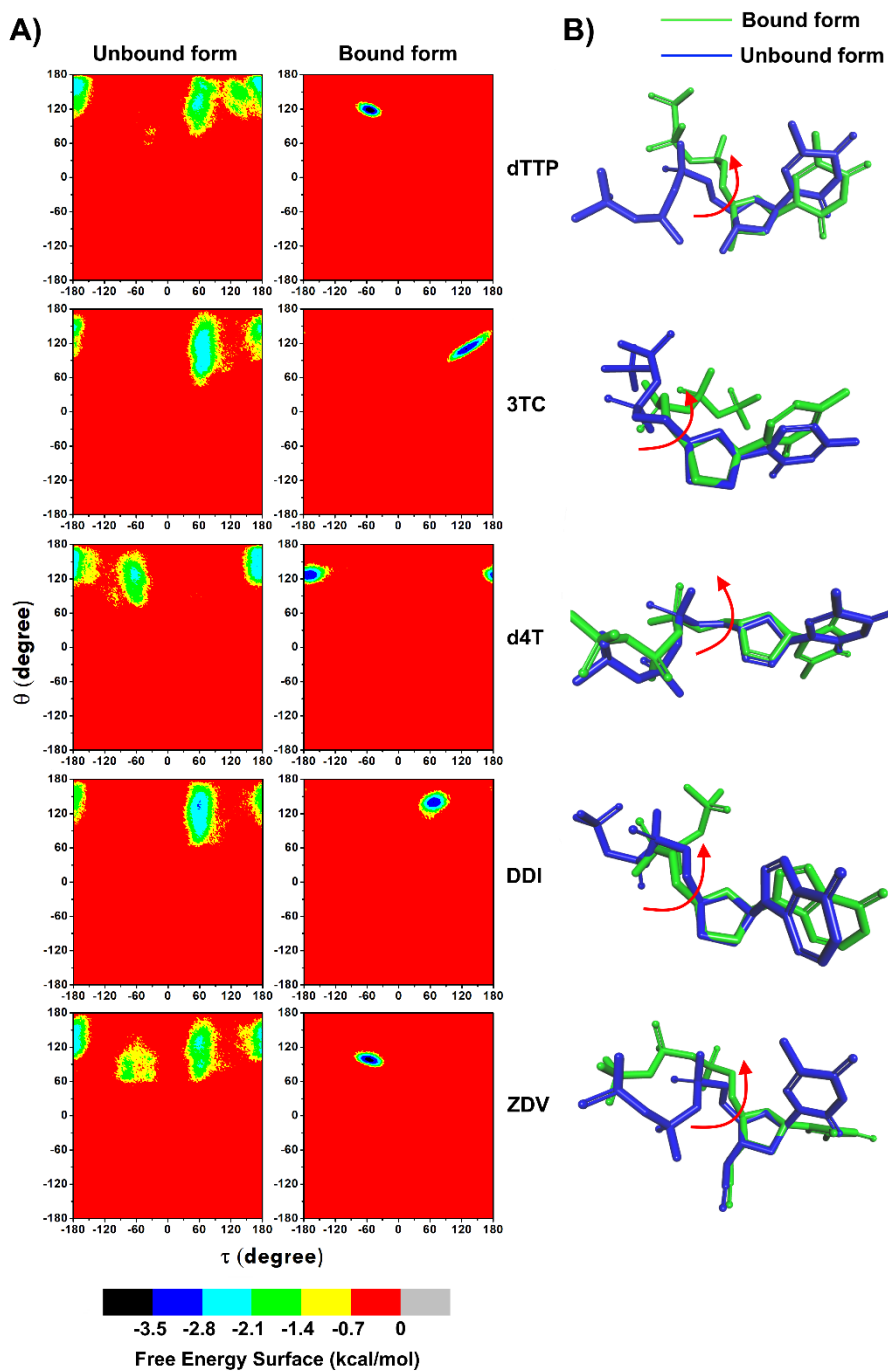
Figure 19 **A)** FES map of dTTP, 3TC, d4T, DDI and ZDV showing the distribution of τ and θ in unbound and bound forms at 310.0 K. The FESs (kcal/mol) were ranged from 0 (grey) to -3.5 kcal/mol (black). **B)** The representative conformation of ligand in bound (green) and unbound (blue) states.
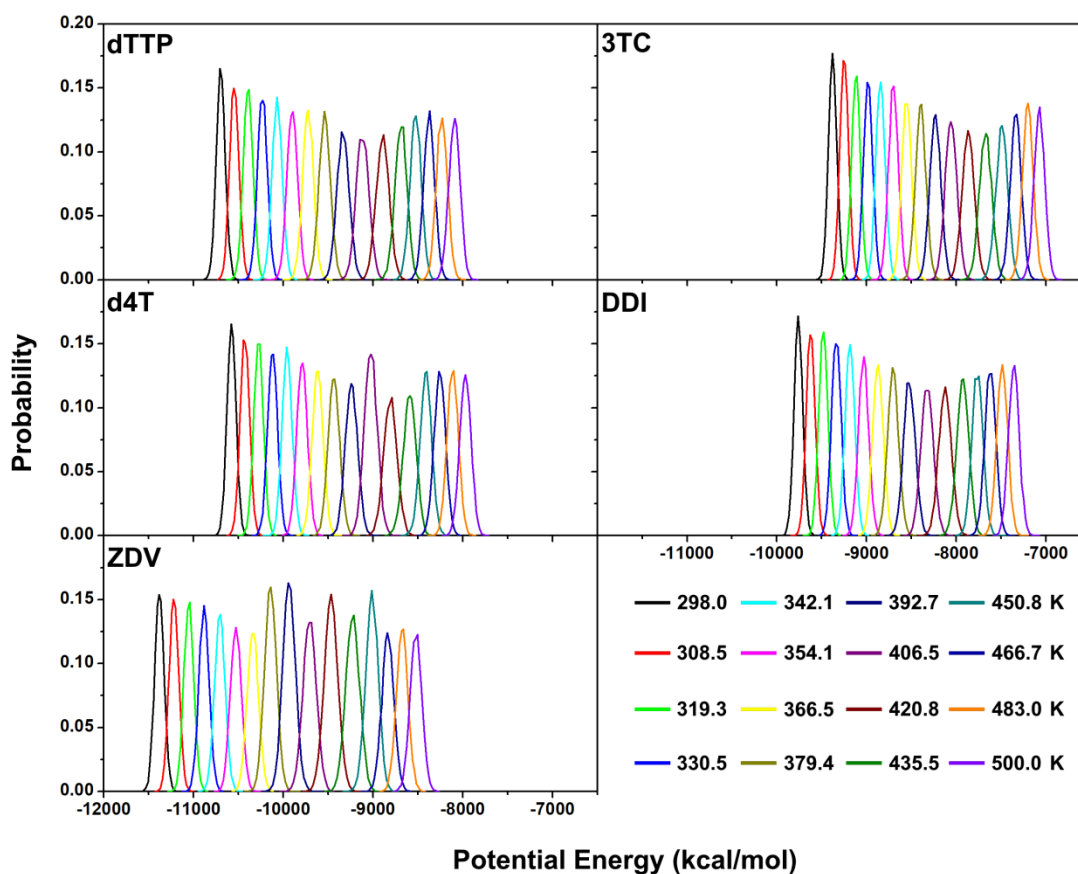
Figure 20 The probability distributions of the total potential energy of all ligands in all temperatures (replicas). Each temperature is represented in different colors.

4.2.4 Binding affinity of anti-HIV drugs against HBV-RT

The susceptibility of anti-HIV drugs against HBV-RT was predicted using the QM/MM-GBSA and MM/GB(PB)SA binding free energy (ΔG_{bind}) calculations on 100 snapshots extracted from the last 30 ns of each MD simulation. Based on QM/MM-GBSA, all drugs were treated by the semiempirical methods (AM1 and PM6), while the remainder was described at the MM level. Note that (*i*) the binding free energies were calculated without considering entropy and (*ii*) The QM/MM-GBSA and MM/GB(PB)SA calculations provided the same trend.
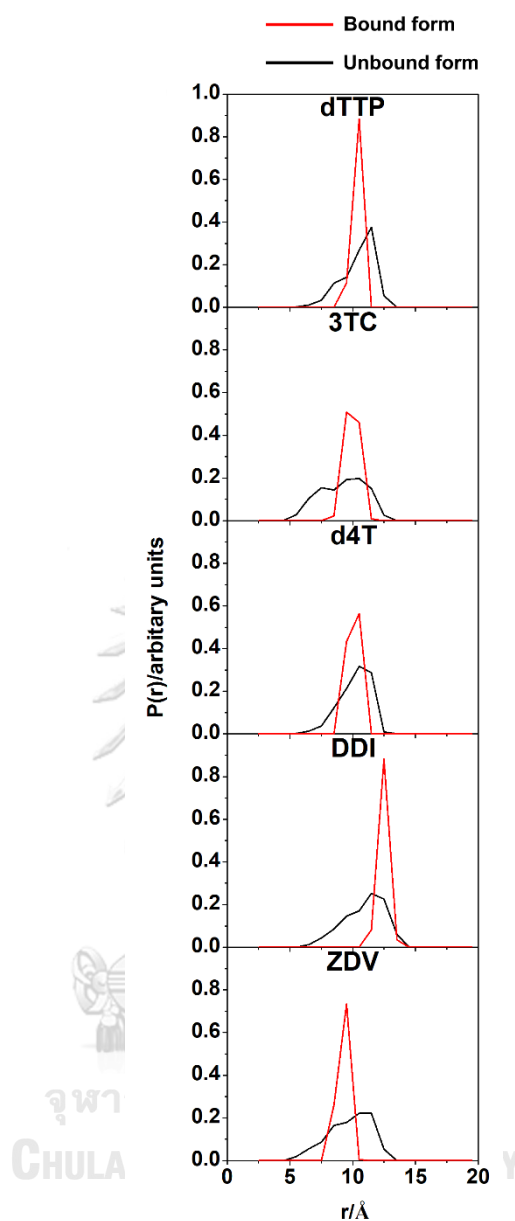
Figure 21 Distribution of distance (r) of dTTP, 3TC, d4T, DDI and ZDV in unbound and
bound form with HBV-RT at 310.0 K.

The obtained results (**Fig. 22**) demonstrated that the ΔG_{bind} values of anti-HIV
drugs obtained from all methods were in the same order of ZDV < d4T << DDI. The
binding affinity of d4T and ZDV drugs was higher than those of dTTP substrate, which
is in good agreement with previous study demonstrating that the binding strength of
anti-HBV drug tenofovir towards the HBV-RT was higher than that of dATP substrate

[61]. Although ZDV showed the highest affinity against HBV-RT (**Fig. 5**) and could potentially inhibit HBV-RT activity with the IC_{50} of 0.3 µM, the *in vivo* and clinical study results indicated that ZDV remains ineffective. Hence, d4T was chosen instead for further in vitro studies.
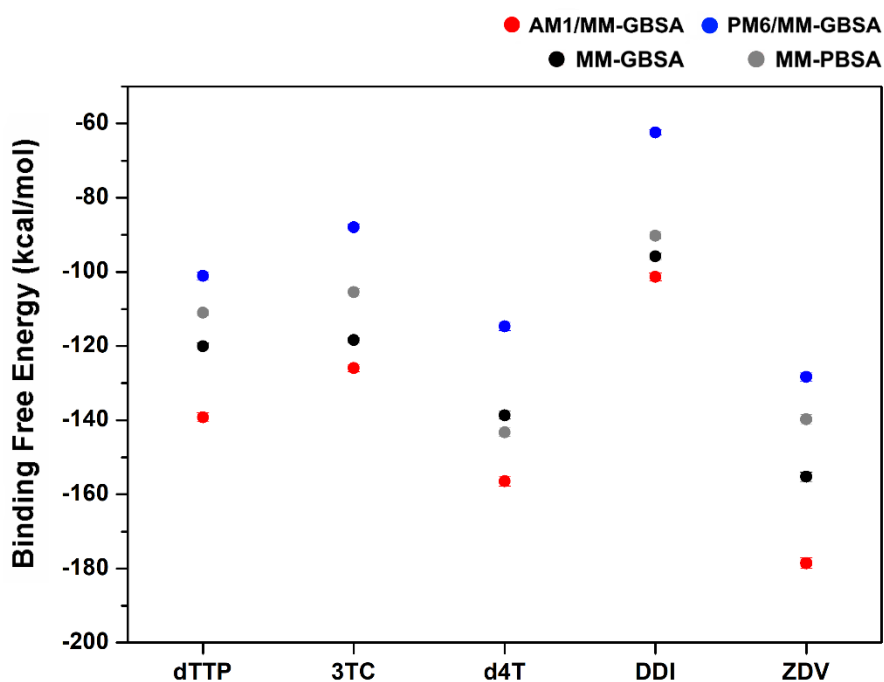
Figure 22 The predicted binding free energies of anti-HIV drugs compared with the substrate and anti-HBV drug to HBV-RT using QM/MM-GBSA and MM/GB(PB)SA calculations.

4.2.5 Virtual screening based on 3D pharmacophore modeling

To identify the new potent candidates against HBV-RT, the 3D-PBS was applied. First, only 2,329 approved drugs were selected from DrugBank database. After that, these compounds were filtered and collected by applying the respectively 3D-PBS (using pharmacophore features: H-bond donor (green arrow), H-bond acceptor (red arrow) and hydrophobic interaction (yellow sphere) (**Fig. 23**)) and similarity search for 50% of nucleoside-based scaffold; as a result, only 16 screened compounds were obtained. Subsequently, the anti-HBV, anti-HIV and other drugs, which already tested with HBV (e.g., zidovudine, telbivudine, stavudine, didanosine, tegafur and Isoprinosine)

[128-130], were removed from such 16 compounds, remaining only 10 drugs left as shown in **Table 2**. Finally, three of them, including floxuridine (FdU, an anti-cancer drug) [131], trifluridine (TFT, an anti-herpesvirus drug) [132], sofosbuvir (SOF, a high molecular weight drug) [133] were selected as representative compound for *in vitro* HBV DNA inhibition assay.

The performance of 3D-PBS and the ability of identifying the structural similarity between (*i*) 4,583 decoys and 10 active compounds for 3TC and (*ii*) 3,583 decoys and 10 active compounds for ZDV were determined using the receiver operating characteristic (ROC) plot, where the area under the curve (AUC) indicates the quality of ROC plot. The AUC value of > 0.50 indicated that the active molecules are reliable. As shown in **Fig. 24**, the AUC values are 0.50 (1%), 0.90 (5%), 0.95 (10%), and 0.91 (100%) for the 3TC and 1.00 (1%), 1.00 (5%), 1.00 (10%), and 0.77 (100%) for the ZDV; thus, the hit compounds derived from 3D-PBS are acceptable. Altogether, FdU, TFT and SOF, obtained from 3D-PBS, and the d4T, acquired from free energy calculations, were then examined for their anti-HBV activity *in vitro*.
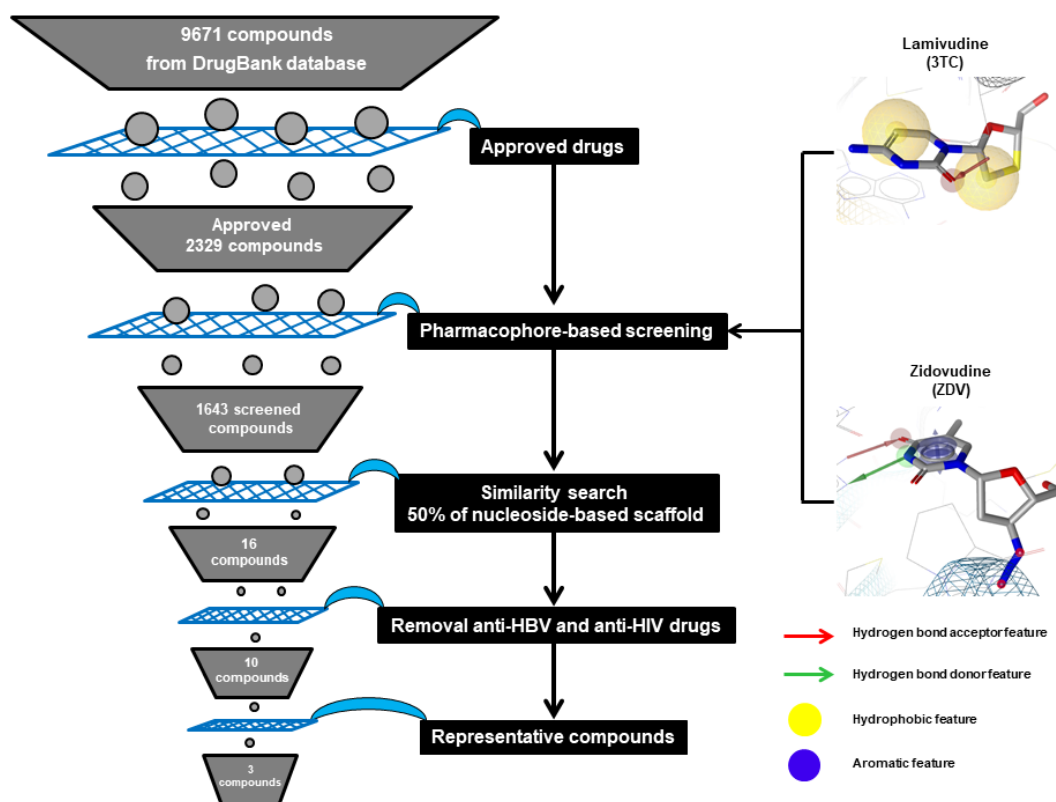
Figure 23 The screening workflow based on 3D-PBS and similarity search for 50% of
nucleoside-based scaffold.

4.2.6 *In vitro* Inhibition of HBV replication

To evaluate the anti-HBV activity of anti-HIV drug (d4T) and the screened
compounds (FdU, TFT and SOF) in comparison with anti-HBV drugs (TDF and 3TC),
HepG2.2.15 cells expressing HBV were treated with all focused compounds at different
concentrations of 1 μM, 10 μM and 100 μM for 7 days. Note that (*i*) the HepG2.2.15
cells intracellularly contain HBV DNA in both forms: integrated- and episomal DNAs
and (*ii*) the amount of integrated DNA was not affected by the addition of anti-HBV
agents, while the episomal DNA would be decreased if the HBV-RT is inhibited.
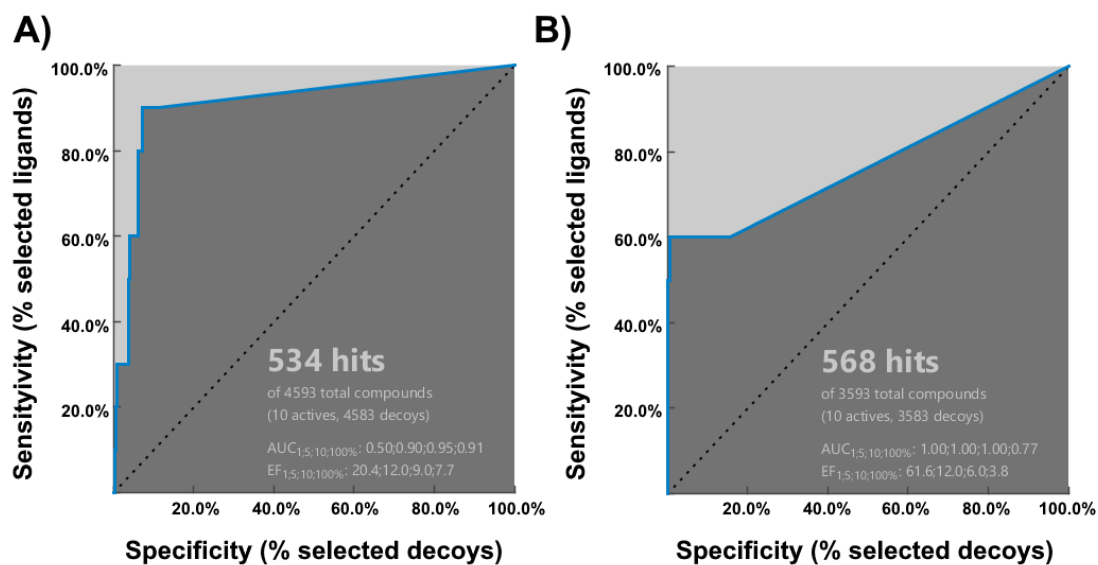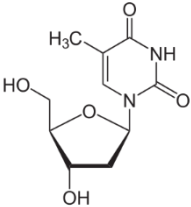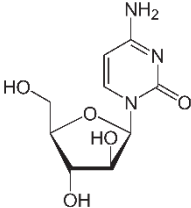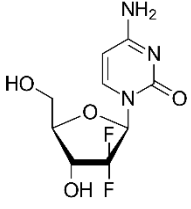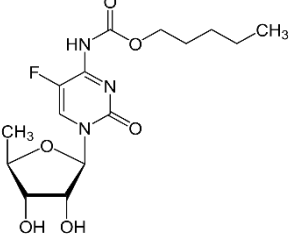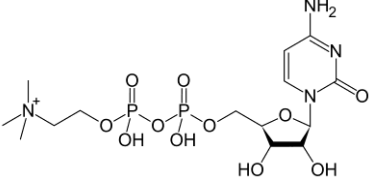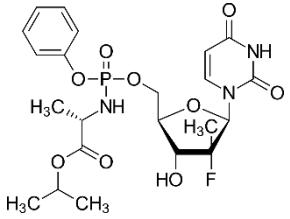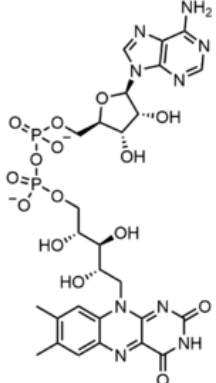
Figure 24 ROC plot of 3D-PBS applied on HBV-RT with **A)** 3TC and **B)** ZDV. The enrichment factor (EF) and the AUC are given for 1, 5, 10, and 100% of database.

Table 2 List of 10 screened compounds obtained from similarity search for 50% of nucleoside-based scaffold and the removal of HIV/HBV drugs.

Name	Target	M.W.	2D structure
Anti-herpesvirus drug			
Trifluridine (TFT)	- Primary keratoconjunctivitis - Recurrent epithelial keratitis	296.20	
Idoxuridine	- Herpes simplex keratitis	354.10	
Brivudine	- Varicella zoster virus which causes herpes simplex virus type 1	333.13	

Anti-cancer drug			
Floxuridine (FdU)	- Colorectal cancer	246.19	
Cytarabine	- Acute myeloid leukemia (AML) - Acute lymphocytic leukemia (ALL) - Chronic myelogenous leukemia (CML)	243.21	
Gemcitabine	- Breast cancer - Ovarian cancer - Non-small cell lung cancer - Pancreatic cancer - Bladder cancer	263.20	
Capecitabine	- Breast cancer - Gastric cancer - Colorectal cancer	359.35	
High molecular weight drug (M.W >400)			
Citicoline	- Alzheimer's disease - Other types of dementia, head trauma and cerebrovascular disease	488.32	
Sofosbuvir (SOF)	- Hepatitis C virus (HCV)	529.46	

FAD	- An ophthalmic treatment for vitamin B2 deficiency	785.55	

The *in vitro* HBV DNA inhibition assay (**Fig. 25A**) revealed that anti-HIV drug d4T, which was obtained from free energy calculations, displayed good anti-HBV activity (~25%) at 1 µM, whereas the activity of this drug was dramatically decreased at higher concentrations (10 and 100 µM). This finding likely suggested that acquired drug resistance might be occurred at high doses. Interestingly, our screened drug FdU, derived from 3D-PBS, showed the highest anti-HBV activity (~10-20%) and suppressed HBV DNA in a concentration-dependent manner. However, the other two screened drugs (TFT and SOF) did not cause a dramatic decrease in the HBV DNA at all three concentrations; therefore, only FdU was selected to examine the cell viability using MTT assay (**Fig. 25B**).

The MTT results indicated that FdU decrease cell viability by ~40% at all concentrations. As compared to 1 µM HBV drug TDF, the FdU at 1 µM displayed the lower cell viability by ~20%, indicating that FdU can cause the toxicity toward hepatocytes. Taken together, FdU could possibly reduce the HBV DNA but still had toxicity; thus, the reduction of HBV DNA might partly come from cell death. Unfortunately, the cell based assay was the indirect inhibition for inhibiting RT, indicating that the inhibition activity of 3TC and d4T might come from other factors such as the ability to transport medicines into cells or the environment of cells. To evaluate the drug activity against HBV-RT, the enzyme assay was the one choice for direct inhibition, which the results would agree with binding free energy calculation.

Hence, the enzyme assay was an interesting method for further testing the ant-HIV drug against HBV-RT.
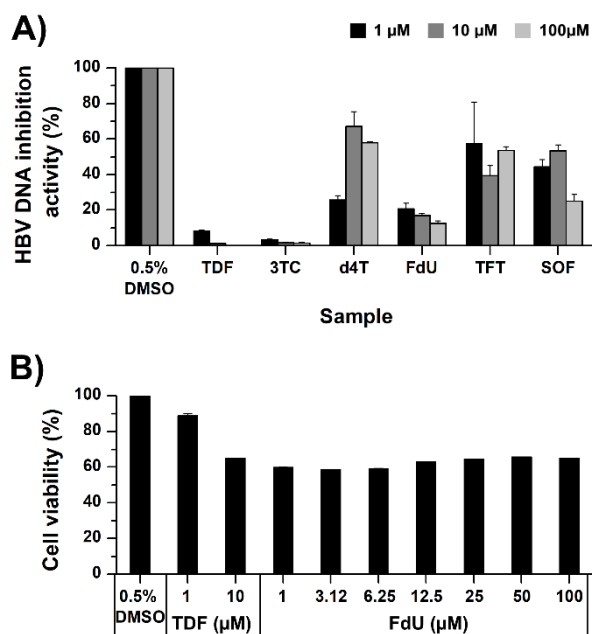
Figure 25 **A)** Inhibition activity of HBV DNA in HepG2.2.15 cells with TDF, 3TC, d4T, FdU, TFT and SOF. Cells were treated with 1, 10 and 100 μM of each drug. **B)** Effect of FdU on HepG2.2.15 cell viability. The number of viable cells was determined using MTT assay. Data are expressed as mean ± SD of two independent experiments.

CHAPTER V

CONCLUSION

5.1 HCV NS3/4A protease

In this work, 100-ns MD simulations were applied on the WT and on the R155K and D168A single point mutations of the NS3/4A protease in the apo form and in complex with ASV, a current drug in phase III clinical trials. According to the PCA, these two mutations dramatically converted the direction of motion of the key binding residues (123, 155 and 168) to move outwards from ASV, resulting in a loss of electrostatic attraction and H-bond formation. The QM/MM-GBSA binding free energy calculations at different levels of QM theory (AM1, RM1, PM3 and PM6) suggested that both mutations caused a significant reduction in the binding affinity, ranked in the order of WT< R155K < D168A, which agreed with the observed experimental EC_{50} values. In addition, it can be postulated that other related mutations might have a similar pattern in drug binding and the loss of interactions like in our cases present here. Consequently, our findings can allow us to predict the effect of mutations on the binding affinity at least without simulations. Taken together, the theoretically obtained information can be used as a rational guide for antiviral drug design and development towards the HCV genotype 1, which is the most prevalent genotype worldwide.

5.2 HBV reverse transcriptase

In this work, 200-ns MD simulations were performed on the HBV-RT in complex with dTTP substrate, anti-HBV drug (3TC) and three anti-HIV drugs (d4T, DDI and ZDV). According to the QM/MM-GBSA and MM/GB(PB)SA binding free energy calculation, the binding affinity of anti-HIV drug d4T was significantly higher than that of anti-HBV drug 3TC, suggesting that d4T could possibly inhibit this HBV-RT. The calculations of FESs revealed that the ligand conformations in bound form were more stable than those

of unbound state. Based on 3D-PBS and *in vitro* anti-HBV activity, the obtained FdU could possibly reduce the intrahepatic HBV DNA. However, this drug still had toxicity toward hepatocytes. Altogether, the obtained information can be used as a rational guide for designing and developing a novel derivative of FdU to diminish its cytotoxicity.

REFERENCES

1.	Organization, W.H., *Global hepatitis report 2017*. 2017: World Health Organization.

2.	Suzuki, T., et al., *Hepatitis C viral life cycle.* Advanced drug delivery reviews, 2007. **59**(12): p. 1200-1212.

3.	Douam, F., D. Lavillette, and F.-L. Cosset, *The mechanism of HCV entry into host cells,* in *Progress in molecular biology and translational science.* 2015, Elsevier. p. 63-107.

4.	Chevaliez, S. and J.-M. Pawlotsky, *HCV genome and life cycle.* Hepatitis C viruses: genomes and molecular biology, 2006: p. 5-47.

5.	Messina, J.P., et al., *Global distribution and prevalence of hepatitis C virus genotypes.* Hepatology, 2015. **61**(1): p. 77-87.

6.	Seifer, M., et al., *Telbivudine, a nucleoside analog inhibitor of HBV polymerase, has a different in vitro cross-resistance profile than the nucleotide analog inhibitors adefovir and tenofovir.* Antiviral research, 2009. **81**(2): p. 147-155.

7.	Franco, S., B. Clotet, and M.A. Martínez, *A wide range of NS3/4A protease catalytic efficiencies in HCV-infected individuals.* Virus research, 2008. **131**(2): p. 260-270.

8.	Sheng, X.C., et al., *Discovery of GS-9451: an acid inhibitor of the hepatitis C virus NS3/4A protease.* Bioorganic & medicinal chemistry letters, 2012. **22**(7): p. 2629-2634.

9.	Fried, M.W., et al., *Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection.* New England Journal of Medicine, 2002. **347**(13): p. 975-982.

10.	Welbourn, S. and A. Pause, *HCV NS2/3 Protease,* in *Hepatitis C Viruses: Genomes and Molecular Biology.* 2006, Horizon Bioscience.

11.	Venkatraman, S., *Discovery of boceprevir, a direct-acting NS3/4A protease inhibitor for treatment of chronic hepatitis C infections.* Trends in pharmacological sciences, 2012. **33**(5): p. 289-294.

12.   Malcolm, B., et al., *SCH 503034, a mechanism-based inhibitor of hepatitis C virus NS3 protease, suppresses polyprotein maturation and enhances the antiviral activity of alpha interferon in replicon cells.* Antimicrobial agents and chemotherapy, 2006. **50**(3): p. 1013-1020.

13.   Howe, A.Y. and S. Venkatraman, *The discovery and development of boceprevir: A novel, first-generation inhibitor of the hepatitis C virus NS3/4A serine protease.* Journal of clinical and translational hepatology, 2013. **1**(1): p. 22.

14.   Perni, R.B., et al., *Preclinical profile of VX-950, a potent, selective, and orally bioavailable inhibitor of hepatitis C virus NS3-4A serine protease.* Antimicrobial agents and chemotherapy, 2006. **50**(3): p. 899-909.

15.   Kwong, A.D., et al., *Discovery and development of telaprevir: an NS3-4A protease inhibitor for treating genotype 1 chronic hepatitis C virus.* Nature biotechnology, 2011. **29**(11): p. 993.

16.   Pilot-Matias, T., et al., *In vitro and in vivo antiviral activity and resistance profile of the hepatitis C virus NS3/4A protease inhibitor ABT-450.* Antimicrobial agents and chemotherapy, 2015. **59**(2): p. 988-997.

17.   Lin, T.-I., et al., *In vitro activity and preclinical profile of TMC435350, a potent hepatitis C virus protease inhibitor.* Antimicrobial agents and chemotherapy, 2009. **53**(4): p. 1377-1385.

18.   Lenz, O., et al., *In vitro resistance profile of the hepatitis C virus NS3/4A protease inhibitor TMC435.* Antimicrobial agents and chemotherapy, 2010. **54**(5): p. 1878-1887.

19.   McCauley, J.A., et al., *Discovery of vaniprevir (MK-7009), a macrocyclic hepatitis C virus NS3/4a protease inhibitor.* Journal of medicinal chemistry, 2010. **53**(6): p. 2443-2463.

20.   Lawitz, E., et al., *Characterization of vaniprevir, a hepatitis C virus NS3/4A protease inhibitor, in patients with HCV genotype 1 infection: safety, antiviral activity, resistance, and pharmacokinetics.* Antiviral research, 2013. **99**(3): p. 214-220.

21.   Jiang, Y., et al., *Discovery of danoprevir (ITMN-191/R7227), a highly selective and potent inhibitor of hepatitis C virus (HCV) NS3/4A protease.* Journal of medicinal chemistry, 2013. **57**(5): p. 1753-1769.

22.   Deutsch, M. and G.V. Papatheodoridis, *Danoprevir, a small-molecule NS3/4A protease inhibitor for the potential oral treatment of HCV infection.* Current opinion in investigational drugs (London, England: 2000), 2010. **11**(8): p. 951-963.

23.   Summa, V., et al., *MK-5172, a selective inhibitor of hepatitis C virus NS3/4a protease with broad activity across genotypes and resistant variants.* Antimicrobial agents and chemotherapy, 2012. **56**(8): p. 4161-4167.

24.   Zeuzem, S., et al., *Faldaprevir and deleobuvir for HCV genotype 1 infection.* New England Journal of Medicine, 2013. **369**(7): p. 630-639.

25.   Scola, P.M., et al., *The discovery of asunaprevir (BMS-650032), an orally efficacious NS3 protease inhibitor for the treatment of hepatitis C virus infection.* Journal of medicinal chemistry, 2014. **57**(5): p. 1730-1752.

26.   McPhee, F., et al., *Preclinical profile and characterization of the hepatitis C virus NS3 protease inhibitor asunaprevir (BMS-650032).* Antimicrobial agents and chemotherapy, 2012. **56**(10): p. 5387-5396.

27.   Yang, H., et al., *Preclinical characterization of the novel hepatitis C virus NS3 protease inhibitor GS-9451.* Antimicrobial agents and chemotherapy, 2014. **58**(2): p. 647-653.

28.   Romano, K.P., et al., *The molecular basis of drug resistance against hepatitis C virus NS3/4A protease inhibitors.* PLoS pathogens, 2012. **8**(7): p. e1002832.

29.   McPhee, F., et al., *Resistance analysis of the hepatitis C virus NS3 protease inhibitor asunaprevir.* Antimicrobial agents and chemotherapy, 2012. **56**(7): p. 3670-3681.

30.   Lagacé, L., et al., *In vitro resistance profile of the hepatitis C virus NS3 protease inhibitor BI 201335.* Antimicrobial agents and chemotherapy, 2012. **56**(1): p. 569-572.

31.   Sorbo, M.C., et al., *Hepatitis C virus drug resistance associated substitutions and their clinical relevance: Update 2018.* Drug Resist Updat, 2018. **37**: p. 17-39.

32.   de Leuw, P. and C. Stephan, *Protease inhibitor therapy for hepatitis C virus-infection.* Expert Opin Pharmacother, 2018. **19**(6): p. 577-587.

33.   Xue, W., et al., *Molecular modeling study on the resistance mechanism of HCV NS3/4A serine protease mutants R155K, A156V and D168A to TMC435.* Antiviral Res, 2012. **93**(1): p. 126-37.

34.   Pan, D., et al., *Understanding the drug resistance mechanism of hepatitis C virus NS3/4A to ITMN-191 due to R155K, A156V, D168A/E mutations: a computational study.* Biochim Biophys Acta, 2012. **1820**(10): p. 1526-34.

35.   Guan, Y., et al., *Exploring resistance mechanisms of HCV NS3/4A protease mutations to MK5172: insight from molecular dynamics simulations and free energy calculations.* Mol Biosyst, 2015. **11**(9): p. 2568-78.

36.   Soumana, D.I., A. Ali, and C.A. Schiffer, *Structural analysis of asunaprevir resistance in HCV NS3/4A protease.* ACS Chem Biol, 2014. **9**(11): p. 2485-90.

37.   Butler, P.J.G., *Hepatitis B Virus (HBV): The Life-cycle and Assembly of a Complex Virus*, in *Supramolecular Structure and Function 9*. 2007, Springer. p. 131-151.

38.   Grimm, D., R. Thimme, and H.E. Blum, *HBV life cycle and novel drug targets.* Hepatology international, 2011. **5**(2): p. 644-653.

39.   Tang, L.S.Y., et al., *Chronic Hepatitis B Infection: A Review.* JAMA, 2018. **319**(17): p. 1802-1813.

40.   Verma, I.M., *The reverse transcriptase.* Biochimica et Biophysica Acta (BBA)-Reviews on Cancer, 1977. **473**(1): p. 1-38.

41.   Verma, I.M., et al., *DNA polymerase activity from two temperature-sensitive mutants of Rous sarcoma virus is thermolabile.* Nature, 1974. **251**(5470): p. 27.

42.   Battula, N. and L. Loeb, *On the fidelity of DNA replication. Lack of exodeoxyribonuclease activity and error-correcting function in avian myeloblastosis virus DNA polymerase.* Journal of Biological Chemistry, 1976. **251**(4): p. 982-986.

43.   Kramvis, A., *Genotypes and genetic variability of hepatitis B virus.* Intervirology, 2014. **57**(3-4): p. 141-150.

44. Baltimore, D. and D.F. Smoler, *Association of an endoribonuclease with the avian myeloblastosis virus deoxyribonucleic acid polymerase.* Journal of Biological Chemistry, 1972. **247**(22): p. 7282-7287.

45. Verma, I.M., *Studies on reverse transcriptase of RNA tumor viruses III. Properties of purified Moloney murine leukemia virus DNA polymerase and associated RNase H.* Journal of virology, 1975. **15**(4): p. 843-854.

46. Keller, W. and R. Crouch, *Degradation of DNA RNA hybrids by ribonuclease H and DNA polymerases of cellular and viral origin.* Proceedings of the National Academy of Sciences, 1972. **69**(11): p. 3360-3364.

47. Berkower, I., J. Leis, and J. Hurwitz, *Isolation and characterization of an endonuclease from Escherichia coli specific for ribonucleic acid in ribonucleic acid· deoxyribonucleic acid hybrid structures.* Journal of Biological Chemistry, 1973. **248**(17): p. 5914-5921.

48. Brewer, L.C. and R.D. Wells, *Mechanistic independence of avian myeloblastosis virus DNA polymerase and ribonuclease H.* Journal of virology, 1974. **14**(6): p. 1494-1502.

49. Lee, W.M., *Hepatitis B virus infection.* New England journal of medicine, 1997. **337**(24): p. 1733-1745.

50. Moreno-Cubero, E., et al., *Is it possible to stop nucleos (t) ide analogue treatment in chronic hepatitis B patients?* World journal of gastroenterology, 2018. **24**(17): p. 1825.

51. Kmet, N.L., M. Poljak, and M. Matičič, *Distribution of hepatitis B virus genotypes in Europe and clinical implications: a review.* Acta dermatovenerologica Alpina, Pannonica, et Adriatica, 2018. **27**(3): p. 141-146.

52. Sunbul, M., *Hepatitis B virus genotypes: global distribution and clinical importance.* World journal of gastroenterology: WJG, 2014. **20**(18): p. 5427.

53. Schaefer, S., *Hepatitis B virus genotypes in Europe.* Hepatology Research, 2007. **37**: p. S20-S26.

54. Ozaras, R., et al., *Epidemiology of HBV subgenotypes D.* Clinics and research in hepatology and gastroenterology, 2015. **39**(1): p. 28-37.

55. Revill, P., et al., *Global strategies are required to cure and eliminate HBV infection.* Nature reviews Gastroenterology & hepatology, 2016. **13**(4): p. 239.

56. Kewn, S., et al., *Lamivudine (3TC) phosphorylation and drug interactions in vitro.* Biochemical pharmacology, 1997. **54**(5): p. 589-595.

57. Qaqish, R.B., K.A. Mattes, and D.J. Ritchie, *Adefovir dipivoxil: a new antiviral agent for the treatment of hepatitis B virus infection.* Clinical therapeutics, 2003. **25**(12): p. 3084-3099.

58. Kearney, B.P., J.F. Flaherty, and J. Shah, *Tenofovir disoproxil fumarate.* Clinical pharmacokinetics, 2004. **43**(9): p. 595-612.

59. Gibson, A.K., et al., *Tenofovir alafenamide: a review of its use in the treatment of HIV-1 infection.* Annals of Pharmacotherapy, 2016. **50**(11): p. 942-952.

60. Santantonio, T. and M. Fasano, *Resistance profile of entecavir in patients with chronic hepatitis B.* Digestive and Liver Disease Supplements, 2008. **2**(2): p. 11-15.

61. Xu, X., et al., *Modeling the functional state of the reverse transcriptase of hepatitis B virus and its application to probing drug-protein interaction.* BMC bioinformatics, 2016. **17**(8): p. 280.

62. van Hemert, F.J., B. Berkhout, and H.L. Zaaijer, *Differential binding of tenofovir and adefovir to reverse transcriptase of hepatitis B virus.* PLoS One, 2014. **9**(9): p. e106324.

63. Lin, P.-F., et al., *Stavudine resistance: an update on susceptibility following prolonged therapy.* Antiviral therapy, 1999. **4**: p. 21-28.

64. Perry, C.M. and J.A. Balfour, *Didanosine.* Drugs, 1996. **52**(6): p. 928-962.

65. Berk, L., S.W. Schalm, and R.A. Heijtink, *Zidovudine inhibits hepatitis B virus replication.* Antiviral research, 1992. **19**(2): p. 111-118.

66. Kayaaslan, B. and R. Guner, *Adverse effects of oral antiviral therapy in chronic hepatitis B.* World journal of hepatology, 2017. **9**(5): p. 227.

67. Palumbo, E., *Lamivudine for chronic hepatitis B: a brief review.* Brazilian Journal of Infectious Diseases, 2008. **12**(5): p. 355-357.

68.   Fontana, R.J., *Side effects of long ‐ term oral antiviral therapy for hepatitis B.* Hepatology, 2009. **49**(S5): p. S185-S195.

69.   Liver, E.A.F.T.S.O.T., *EASL clinical practical guidelines: management of alcoholic liver disease.* Journal of hepatology, 2012. **57**(2): p. 399-420.

70.   Margolis, A.M., et al., *A review of the toxicity of HIV medications.* Journal of Medical Toxicology, 2014. **10**(1): p. 26-39.

71.   Hurst, M. and S. Noble, *Stavudine.* Drugs, 1999. **58**(5): p. 919-949.

72.   Floridia, M., et al., *A randomized trial (ISS 902) of didanosine versus zidovudine in previously untreated patients with mildly symptomatic human immunodeficiency virus infection.* Journal of Infectious Diseases, 1997. **175**(2): p. 255-264.

73.   Kinloch-de Loës, S., et al., *A controlled trial of zidovudine in primary human immunodeficiency virus infection.* New England Journal of Medicine, 1995. **333**(7): p. 408-413.

74.   Chiappini, E., et al., *Preventable zidovudine overdose during postnatal prophylaxis in healthy children born to HIV-1-positive mothers.* Aids, 2008. **22**(2): p. 316-317.

75.   Livshits, Z., et al., *Zidovudine (AZT) overdose in a healthy newborn receiving postnatal prophylaxis.* Clinical Toxicology, 2011. **49**(8): p. 747-749.

76.   Hitchcock, M., *In vitro antiviral activity of didanosine compared with that of other dideoxynucleoside analogs against laboratory strains and clinical isolates of human immunodeficiency virus.* Clinical infectious diseases, 1993. **16**(Supplement_1): p. S16-S21.

77.   Nordenfelt, E., et al., *Inhibition of hepatitis B virus DNA polymerase by 3′ ‐ azido ‐ 3′ ‐ deoxythymidine triphosphate but not by its threo analog.* Journal of medical virology, 1987. **22**(3): p. 231-236.

78.   Saxena, A., R.S. Sangwan, and S. Mishra, *Fundamentals of homology modeling steps and comparison among important bioinformatics tools: An overview.* Sci. Int, 2013. **1**(7): p. 237-252.

79. Meeprasert, A., S. Hannongbua, and T. Rungrotmongkol, *Key binding and susceptibility of NS3/4A serine protease inhibitors against hepatitis C virus.* Journal of Chemical Information and Modeling, 2014. **54**(4): p. 1208-1217.

80. Hospital, A., et al., *Molecular dynamics simulations: advances and applications.* Advances and applications in bioinformatics and chemistry: AABC, 2015. **8**: p. 37.

81. Hollingsworth, S.A. and R.O. Dror, *Molecular dynamics simulation for all.* Neuron, 2018. **99**(6): p. 1129-1143.

82. Duan, Y., et al., *A point ‐ charge force field for molecular mechanics simulations of proteins based on condensed ‐ phase quantum mechanical calculations.* Journal of computational chemistry, 2003. **24**(16): p. 1999-2012.

83. Paquet, E. and H.L. Viktor, *Molecular dynamics, monte carlo simulations, and langevin dynamics: a computational review.* Biomed Res Int, 2015. **2015**: p. 183918.

84. MacKerell Jr, A.D., *Empirical force fields for biological macromolecules: overview and issues.* Journal of computational chemistry, 2004. **25**(13): p. 1584-1604.

85. Hou, T., et al., *Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations.* J Chem Inf Model, 2011. **51**(1): p. 69-82.

86. Hou, T., et al., *Characterization of domain-peptide interaction interface: a generic structure-based model to decipher the binding specificity of SH3 domains.* Mol Cell Proteomics, 2009. **8**(4): p. 639-49.

87. Gohlke, H., C. Kiel, and D.A. Case, *Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes.* J Mol Biol, 2003. **330**(4): p. 891-913.

88. Wang, W. and P.A. Kollman, *Free energy calculations on dimer stability of the HIV protease using molecular dynamics and a continuum solvent model.* Journal of molecular biology, 2000. **303**(4): p. 567-582.

89.   Hou, T., S. Guo, and X. Xu, *Predictions of binding of a diverse set of ligands to gelatinase-A by a combination of molecular dynamics and continuum solvent models.* The Journal of Physical Chemistry B, 2002. **106**(21): p. 5527-5535.

90.   Wang, J., T. Hou, and X. Xu, *Recent advances in free energy calculations with a combination of molecular mechanics and continuum models.* Current Computer-Aided Drug Design, 2006. **2**(3): p. 287-306.

91.   Phanich, J., et al., *Glycan binding and specificity of viral influenza neuraminidases by classical molecular dynamics and replica exchange molecular dynamics simulations.* Journal of Biomolecular Structure and Dynamics, 2018: p. 1-12.

92.   Chen, J., et al., *A comparative study of trypsin specificity based on QM/MM molecular dynamics simulation and QM/MM GBSA calculation.* Journal of Biomolecular Structure and Dynamics, 2015. **33**(12): p. 2606-2618.

93.   Meli, M. and G. Colombo, *A Hamiltonian replica exchange molecular dynamics (MD) method for the study of folding, based on the analysis of the stabilization determinants of proteins.* International journal of molecular sciences, 2013. **14**(6): p. 12157-12169.

94.   Periole, X. and A.E. Mark, *Convergence and sampling efficiency in replica exchange simulations of peptide folding in explicit solvent.* The Journal of chemical physics, 2007. **126**(1): p. 01B601.

95.   Sugita, Y. and Y. Okamoto, *Replica-exchange molecular dynamics method for protein folding.* Chemical physics letters, 1999. **314**(1-2): p. 141-151.

96.   Langer, T. and G. Wolber, *Pharmacophore definition and 3D searches.* Drug Discovery Today: Technologies, 2004. **1**(3): p. 203-207.

97.   Langer, T., *Pharmacophores in drug research.* Molecular informatics, 2010. **29**(6‐7): p. 470-475.

98.   Soumana, D.I., A. Ali, and C.A. Schiffer, *Structural analysis of asunaprevir resistance in HCV NS3/4A protease.* ACS chemical biology, 2014. **9**(11): p. 2485-2490.

99.   Visualizer, D.S., *Release 3.5.* Accelrys Inc, San Diego, CA, USA, 2012.

100. Dolinsky, T.J., et al., *PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations.* Nucleic acids research, 2004. **32**(suppl_2): p. W665-W667.

101. Salomon－Ferrer, R., D.A. Case, and R.C. Walker, *An overview of the Amber biomolecular simulation package.* Wiley Interdisciplinary Reviews: Computational Molecular Science, 2013. **3**(2): p. 198-210.

102. Case, D.A., et al., *Amber 14.* 2014.

103. Phanich, J., et al., *Role of R292K mutation in influenza H7N9 neuraminidase toward oseltamivir susceptibility: MD and MM/PB (GB) SA study.* Journal of computer-aided molecular design, 2016. **30**(10): p. 917-926.

104. Mahalapbutr, P., et al., *Molecular insights into inclusion complexes of mansonone E and H enantiomers with various β-cyclodextrins.* Journal of Molecular Graphics and Modelling, 2018. **79**: p. 72-80.

105. Jorgensen, W.L. and J. Tirado-Rives, *Potential energy functions for atomic-level simulations of water and organic and biomolecular systems.* Proceedings of the National Academy of Sciences, 2005. **102**(19): p. 6665-6670.

106. Miyamoto, S. and P.A. Kollman, *Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models.* Journal of computational chemistry, 1992. **13**(8): p. 952-962.

107. Di Pierro, M., R. Elber, and B. Leimkuhler, *A Stochastic Algorithm for the Isobaric–Isothermal Ensemble with Ewald Summations for All Long Range Forces.* Journal of chemical theory and computation, 2015. **11**(12): p. 5624-5637.

108. Paquet, E. and H.L. Viktor, *Molecular dynamics, monte carlo simulations, and langevin dynamics: a computational review.* BioMed research international, 2015. **2015**.

109. Panman, W., et al., *Computational screening of fatty acid synthase inhibitors against thioesterase domain.* Journal of Biomolecular Structure and Dynamics, 2018. **36**(15): p. 4114-4125.

110. Nutho, B., et al., *Binding mode and free energy prediction of fisetin/ β - cyclodextrin inclusion complexes.* Beilstein journal of organic chemistry, 2014. **10**(1): p. 2789-2799.

111. Waterhouse, A., et al., *SWISS-MODEL: homology modelling of protein structures and complexes.* Nucleic acids research, 2018. **46**(W1): p. W296-W303.

112. Das, K., et al., *Structures of HIV-1 RT-RNA/DNA ternary complexes with dATP and nevirapine reveal conformational flexibility of RNA/DNA: insights into requirements for RNase H cleavage.* Nucleic acids research, 2014. **42**(12): p. 8125-8137.

113. Huang, H., et al., *Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance.* Science, 1998. **282**(5394): p. 1669-1675.

114. Studio, A.D., *1.7, Accelrys Software Inc., San Diego, CA, USA. 2006.*

115. Case, D., et al., *AMBER16 Package.* University of California, San Francisco, 2016.

116. Cheng, X., et al., *Modified replica exchange simulation methods for local structure refinement.* The Journal of Physical Chemistry B, 2005. **109**(16): p. 8220-8230.

117. Wishart, D.S., et al., *DrugBank 5.0: a major update to the DrugBank database for 2018.* Nucleic acids research, 2017. **46**(D1): p. D1074-D1082.

118. Rakers, C., et al., *In silico prediction of human sulfotransferase 1E1 activity guided by pharmacophores from molecular dynamics simulations.* Journal of Biological Chemistry, 2016. **291**(1): p. 58-71.

119. Sehgal, S.A., M. Hassan, and S. Rashid, *Pharmacoinformatics elucidation of potential drug targets against migraine to target ion channel protein KCNK18.* Drug design, development and therapy, 2014. **8**: p. 571.

120. Trinh, Q. and L. Le, *An investigation of antidiabetic activities of bioactive compounds in Euphorbiahirta Linn using molecular docking and pharmacophore.* Medicinal Chemistry Research, 2014. **23**(4): p. 2033-2045.

121. Mysinger, M.M., et al., *Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking.* Journal of medicinal chemistry, 2012. **55**(14): p. 6582-6594.

122. Soumana, D.I., et al., *Molecular and dynamic mechanism underlying drug resistance in genotype 3 hepatitis C NS3/4A protease.* Journal of the American Chemical Society, 2016. **138**(36): p. 11850-11859.

123. Haider, S., G.N. Parkinson, and S. Neidle, *Molecular dynamics and principal components analysis of human telomeric quadruplex multimers.* Biophysical journal, 2008. **95**(1): p. 296-311.

124. Meeprasert, A., et al., *Effect of D168V mutation in NS3/4A HCV protease on susceptibilities of faldaprevir and danoprevir.* Molecular BioSystems, 2016. **12**(12): p. 3666-3673.

125. Xue, W., et al., *Computational study on the drug resistance mechanism against HCV NS3/4A protease inhibitors vaniprevir and MK-5172 by the combination use of molecular dynamics simulation, residue interaction network, and substrate envelope analysis.* Journal of chemical information and modeling, 2013. **54**(2): p. 621-633.

126. Wang, H., et al., *Computational study on the molecular mechanisms of drug resistance of Narlaprevir due to V36M, R155K, V36M+ R155K, T54A, and A156T mutations of HCV NS3/4A protease.* Biochemistry and Cell Biology, 2014. **92**(5): p. 357-369.

127. Daga, P.R., J. Duan, and R.J. Doerksen, *Computational model of hepatitis B virus DNA polymerase: molecular dynamics and docking to understand resistant mutations.* Protein Science, 2010. **19**(4): p. 796-807.

128. Maggs, D.J., *Antiviral therapy for feline herpesvirus infections.* Veterinary Clinics: Small Animal Practice, 2010. **40**(6): p. 1055-1062.

129. Lv, J.W., et al., *Hepatitis B virus screening and reactivation and management of patients with nasopharyngeal carcinoma: A large ‐ scale, big ‐ data intelligence platform ‐based analysis from an endemic area.* Cancer, 2017. **123**(18): p. 3540-3549.

130. Honda, S., et al., *Tegafur-uracil-induced rapid development of advanced hepatic fibrosis.* World journal of gastroenterology, 2017. **23**(31): p. 5823.

131. Fischer, M.A., et al., *Flaviviruses are sensitive to inhibition of thymidine synthesis pathways.* Journal of virology, 2013. **87**(17): p. 9411-9419.

132. Matsuoka, K., et al., *Trifluridine/tipiracil overcomes the resistance of human gastric 5-fluorouracil-refractory cells with high thymidylate synthase expression.* Oncotarget, 2018. **9**(17): p. 13438.

133. Mumtaz, N., et al., *Cell-line dependent antiviral activity of sofosbuvir against Zika virus.* Antiviral research, 2017. **146**: p. 161-163.

APPENDICES

APPENDIX A

Amino acid sequences alignment of reverse transcriptase from HBV (NCBI sequence database reference AY233278) and HIV-1 (PDB code: 4PQU):

Identity: 20.9%

Similarity: 32.8%

```
HBV   32 KNPHNT----------AESRLVVDFSQFSRGSTRVSWP-KFAVPNLQSLT     70
         :||:||          .:.|.:|||.:.:: .|:..|. :..:|...|.
HIV   55 ENPYNTPVFAIKKKDSTKWRKLVDFRELNK-RTQDFWEVQLGIPHPAGLK    103

HBV   71 NLLSSNLSWLSLDVSAAFYHIPLHPAAMPHLLIGSSGLSRYVARLSSNSR    120
         ...|..:     |||..|::.:||           .....:|.|  .:...
HIV   04 KKKSVTV----LDVGDAYFSVPL-----------DEDFRKYTA--FTIPS    136

HBV   21 IHNN------QYGTLQNLHDSCSRQLYVSLMLLYKTYGWKLHLYSHPIIL    164
         |:|.       ||..|..                         |||
HIV   37 INNETPGIRYQYNVLPQ-------------------GWK----------    156

HBV   65 GFRKIPMGVGLSPFLLAQFTSAICSVVRRAFPHCLAFSYMDDVVLGAKSV    214
                  | ||.:......:.|...::.|..:.:.||||:.:|:.
HIV   57 ---------G-SPAIFQSSMTKILEPFKKQNPDIVIYQYMDDLYVGSD--    194

HBV   15 QHLESLYTAVTNFLLSLGIHLNPNKTK---------RWGYS---------    246
                       |.:|.|    :||          |||.:
HIV   95 --------------LEIGQH----RTKIEELRQHLLRWGLTTPDKKHQKE    226

HBV   47 --LNFMGYLIGSWGTLPQDHIVQKIKHCFKKLPVNRPIDWKV---CQRIV    291
           ..:|||.:.     |....||.|     .||  ....|.|   | ::|
HIV   27 PPFLWMGYELH-----PDKWTVQPI-----VLP--EKDSWTVNDIC-KLV    263

HBV   92 GLLGFAAPFTQCGYPAL--MPLYACIQAKQAFT--FSLTYKAFL    331
         |.|.:|:..     ||.:  ..|...:.:.:|.|  ..||.:|.|
HIV   64 GKLNWASQI----YPGIKVRQLSKLLRGTKALTEVIPLTEEAEL    303
```
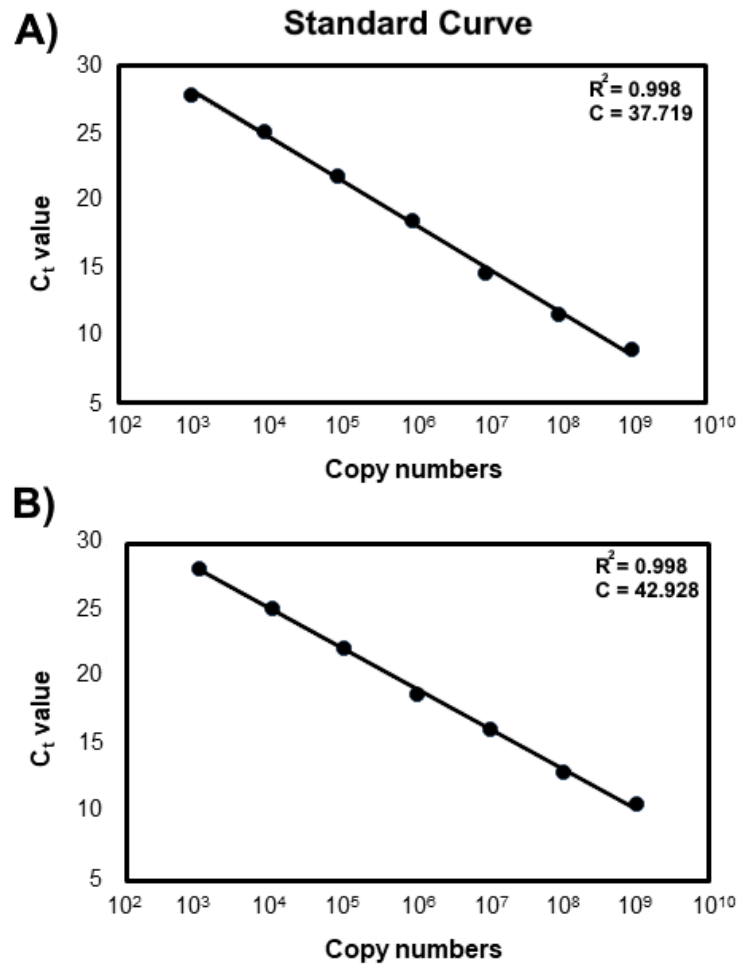
APPENDIX B

H-bond orientation of anti-HIV and anti-HBV drugs in complex with HBV-RT taken from the last MD snapshot.
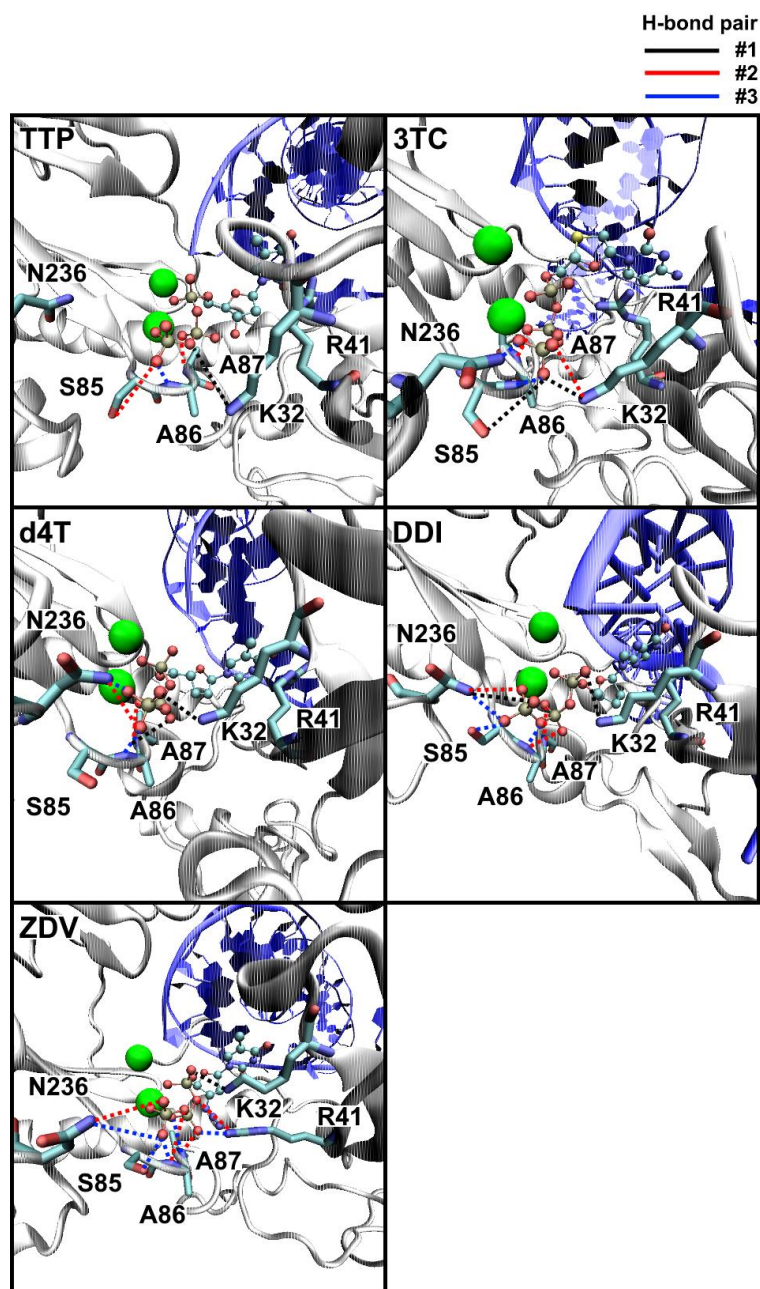
APPENDIX C

A standard curve of A) a HBV plasmid (pHBV^{WT}) with 7-fold serial dilutions (10^3 to 10^9 copies per reaction) and B) a β-globin plasmid with 7-fold serial dilutions (10^3 to 10^9 copies per reaction)

VITA

NAME	Jirayu Kammarabutr
DATE OF BIRTH	18 July 1993
PLACE OF BIRTH	Bangkok, Thailand
INSTITUTIONS ATTENDED	Chulalongkorn University
HOME ADDRESS	1044 / 7-8 Charoen Krung Road, Bang Rak, Bang Rak, Bangkok 10500