TARGET-BASED DESIGN, STRUCTURAL OPTIMIZATION AND CHARACTERIZATION OF NOVEL HEPATITIS B VIRUS CAPSID ASSEMBLY MODULATORS

Elena Detta

UNIVERSITAT ROVIRA i VIRGILI

ICIQ
Institut Català d'Investigació Química

Target-based Design, Structural Optimization and Characterization of Novel Hepatitis B Virus Capsid Assembly Modulators

Elena Detta

DOCTORAL THESIS
2021

UNIVERSITAT ROVIRA I VIRGILI
TARGET-BASED DESIGN, STRUCTURAL OPTIMIZATION AND CHARACTERIZATION OF NOVEL HEPATITIS B VIRUS CAPSID ASSEMBLY MODULATORS
Elena Detta

UNIVERSITAT ROVIRA I VIRGILI
TARGET-BASED DESIGN, STRUCTURAL OPTIMIZATION AND CHARACTERIZATION OF NOVEL HEPATITIS B VIRUS CAPSID ASSEMBLY MODULATORS
Elena Detta

Abridged version of the thesis due to confidentiality agreement.

The thesis is the result of VIRO-FLOW Industrial Doctorate Programme. This project was supported by a research grant from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No. 766058.

UNIVERSITAT ROVIRA I VIRGILI
TARGET-BASED DESIGN, STRUCTURAL OPTIMIZATION AND CHARACTERIZATION OF NOVEL HEPATITIS B VIRUS CAPSID ASSEMBLY MODULATORS
Elena Detta

Target-based Design, Structural Optimization and Characterization of Novel Hepatitis B Virus Capsid Assembly Modulators

Doctoral Thesis by

Elena Detta

Developed under the supervision of

Prof. Miquel A. Pericàs Brondo

Dr. Helmut Buschmann

Departament de Química Analítica i Química Orgànica (URV)

Institute of Chemical Research of Catalonia (ICIQ)

Tarragona

2021

UNIVERSITAT ROVIRA I VIRGILI
TARGET-BASED DESIGN, STRUCTURAL OPTIMIZATION AND CHARACTERIZATION OF NOVEL HEPATITIS B VIRUS CAPSID ASSEMBLY MODULATORS
Elena Detta

UNIVERSITAT ROVIRA I VIRGILI

ICIQ
Institut Català d'Investigació Química

Prof. Miquel A. Pericàs Brondo, Group Leader of the Institute of Chemical Research of Catalonia (ICIQ) and,

Dr. Helmut Buschmann, Lecturer of University of Aachen (RWTH Aachen University) in the field of Medicinal Chemistry and Drug Discovery for Master and PhD Students

STATE, that the present Doctoral Thesis entitled: **"Target-based design, structural optimization and characterization of novel Hepatitis B Virus capsid assembly modulators"**, presented by Elena Detta to receive the degree of Doctor, has been carried out under our supervision at the Institute of Chemical Research of Catalonia (ICIQ) and at AiCuris Anti-infective Cures GmbH.

Tarragona, 30 April 2021

PhD Thesis Supervisor

Prof. Miquel A. Pericàs Brondo

PhD Thesis Co-supervisor

Dr. Helmut Buschmann

UNIVERSITAT ROVIRA I VIRGILI
TARGET-BASED DESIGN, STRUCTURAL OPTIMIZATION AND CHARACTERIZATION OF NOVEL HEPATITIS B VIRUS CAPSID ASSEMBLY MODULATORS
Elena Detta

Acknowledgements

First of all, I would like to express my sincere and warm gratitude to my PhD supervisors, **Prof. Miquel A. Pericàs Brondo** and **Dr. Helmut Buschmann**, who placed their faith in my abilities, supported my ideas and provided me with the opportunity to develop as a scientist.

Next, my deepest appreciation goes to **Dr. Alastair Donald** for his patient guidance, enthusiastic encouragement and constructive support during the planning and development of this research work. His constant dedication and useful insights have been very much appreciated. I would like to offer my special thanks to **Dr. Esther Alza** for her advice and valuable assistance in the process of writing and successfully completing this scientific project.

I would also like to thank **Dr. Thomas Goldner**, **Dr. Andreas Urban, Dr. Susanne Bonsmann and Dr. Florian Engel** for their scientific insights and useful suggestions in our group discussions. I am particularly grateful for the assistance given by **Dr. Anna Maria Banet**, who greatly coordinated the VIRO-FLOW program and ensured it running according to plan.

My profound gratitude goes to my fellow students, **Justine Raymond** and **Tamás Vermes**, who shared with me this great experience in the VIRO-FLOW training network with passion, perseverance and hard work.

I wish to gratefully acknowledge the remarkable computational work performed by **Dr. Mauro Fianchini**. His contribution to the development and writing of the last chapter of this thesis was very much appreciated.

I am also very thankful to **Dr. Angelica M. Corcuera** for her excellent technical support with the experimental work at AiCuris and for her diligent editorial processing of our manuscript. I would like to especially thank **Dr. Laura Amenós Núñez** for her daily support and guidance in the chemistry laboratory. I greatly appreciated our fruitful discussions and her encouragement during my stay at ICIQ.

My special thanks are extended to my colleague and friend **Arancha López Pérez** for our long talks and the amazing adventures we shared in and outside the office. These acknowledgements would not be complete without mentioning my research laboratory colleagues in ICIQ: **Dr. Patricia Llanes**, **Dr. Marco M. Mastandrea**, **Dr. Parijat Borah**, **Dr. Leijie Zhou**, **Dr. Moreshwar Chaudhari**, **Nicola Zanda**, **Stefania Perulli**, **Adrian J. Brenes Rucinski** and **Arianna Brandolese**. It was a great pleasure for me to work in such a joyful and multicultural environment, surrounded by brilliant, dedicated and supportive people, who motivated me every day to give my best contribution.

Finally, I would like to remind myself where I come from. I would not have reached this *destination* without the unconditional love and support of my mother **Lucia**, my father **Francesco** and my brother **Domenico**, who I hope to hug tightly very soon.

UNIVERSITAT ROVIRA I VIRGILI
TARGET-BASED DESIGN, STRUCTURAL OPTIMIZATION AND CHARACTERIZATION OF NOVEL HEPATITIS B VIRUS CAPSID ASSEMBLY MODULATORS
Elena Detta

Summary

Chronic hepatitis B is a severe liver infection caused by the Hepatitis B Virus (HBV). Despite the availability of a prophylactic vaccine since 1982, HBV infection still remains a serious global health issue with more than 250 million carriers worldwide. Currently, HBV infection is treated with nucleos(t)ide analogues (NUCs) and, less commonly, interferon-based therapy (IFN-α). However, the standard of care (SOC) does not provide a functional cure. Thus, there is a significant need for novel therapeutics aiming towards the complete eradication of the virus from infected hepatocytes. Small-molecule capsid assembly modulators (CAMs) have been recently recognized as promising antiviral agents for curing chronic HBV infection.

The VIRO-FLOW project aims at the fast and efficient identification of novel curative agents for HBV, integrating the advantages of continuous flow chemistry with microfluidic technologies. A combination of ligand- and structure-based approaches were explored in the research work described in this thesis for the design and optimization of focused libraries of HBV capsid modulators and to guide and support the synthesis of those molecules either in batch or flow.

In the second chapter of this thesis, an *in silico* workflow is reported, which was established in order to identify novel HBV CAMs, using a known active oxalyl-piperazine compound as starting point. Two virtual libraries of derivatives were created via bioisosteric replacement methods. A complementary structure-based computational approach was adopted to predict water molecules at the HBV core binding site, leading to the identification of a potentially "unhappy" water, which could be displaced to enhance ligand-protein hydrophobic interactions.

A computer-aided scaffold hopping study aiming towards the identification of potential replacements to the indole ring of a known active Class I HBV CAM is described in the third chapter. Two novel series of CAMs were generated. This research work resulted in the discovery of a lead compound with *in vitro* antiviral activity in the nanomolar range.

An extensive target-based screening study is depicted in chapter four. An *in silico* workflow was established with the main aim of identifying novel small

molecules that misdirect HBV capsid formation, with a secondary goal of discovering new non-HAP-like Class II CAMs, based on the hypothesis that the 3D-protein bound structure of HAP-like modulators could explain their unique mode of action. Four weakly active screening hits emerged from the structure-based study and served as a basis for the design and synthesis of two focused libraries of HBV inhibitors. A lead compound was identified with *in vitro* potency in the sub-micromolar range along with good physico-chemical and safety profile *in vitro*. The novel molecule was then further evaluated in molecular docking and mechanism of action studies. In the attempt to optimize its biological activity, PK and toxicological properties, a diversity collection of derivatives was prepared, leading to the identification of sulfonamide-containing analogues with increased antiviral potency compared to the reference lead compound.

As a result of the work developed in the present thesis, two series of highly active HBV CAMs were identified. From the obtained results and further computational studies, it was determined that the [1,2,4]triazolo[1,5-*a*]pyridine core could be efficiently used as a bioisosteric replacement of the *N*-rich bicyclic rings present in the initial hits, leading to the discovery of a novel chemical scaffold with anti-HBV activity. The fifth chapter describes the development of a new method for the synthesis of ethyl [1,2,4]triazolo[1,5-*a*]pyridine-2-carboxylate 3-oxide. The process was implemented in continuous flow in order to minimize the formation of three identified side products obtained when the reaction was carried out in batch. DFT calculations were performed to optimize the overall synthetic procedure and clarify the mechanism of N-oxide formation by reaction of sulfilimine and ethyl 2-chloro-2-(hydroxyimino)acetate (oxime).

List of Abbreviations

aa	amino acid
CAN	acetonitrile
ADV	adefovir
AGL	antigenic loop
ALT	alanine aminotransferase
aSEC	Analytical size exclusion chromatography
AUC	Area Under the Curve
Boc	*tert*-butyloxylcarbonyl
CAM	capsid assembly modulator
Cbz	benzyl chloroformate
cccDNA	covalently closed circular DNA
Cp	core protein
CpAM	capsid assembly allosteric modulator
Cp149	N-terminal domain of Cp (residues 1-149)
Cp183	full-length core protein (183 residues)
Cp185	full-length core protein (185 residues)
CTD	C-terminal domain of Cp
D	distribution coefficient
DCE	1,2-dichloroethane
DCM	dichloromethane
DFT	Density Functional Theory
DMEM	Dulbecco's modified eagle medium
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNAse	deoxyribonuclease
DPBS	Dulbecco's phosphate buffered saline
dsDNA	double stranded DNA
EC_{50}	half-minimal effective concentration in cell culture
EtOH	ethanol
ETV	entecavir
FCS	Fetal calf serum
HAP	heteroaryldihydropyrimidine
HATU	Hexafluorophosphate Azabenzotriazole Tetranethyl Uronium
HBcAg	hepatitis B core antigen
HBeAg	hepatitis B e antigen
HBsAg	hepatitis B surface antigen
HBV	hepatitis B virus
HbxAg	hepatitis B x protein
hERG	Human Ether-à-go-go-Related Gene
HPLC	High performance liquid chromatography
HSPG	heparin sulfate proteoglycans

IFN	interferon
IFM	Immunofluorescence microscopy
kb	kilobases
MSH	O-(mesitylsulfonyl)hydroxylamine
mRNA	messenger RNA
NADPH	Nicotinamide adenine dinucleotide phosphate
NMR	Nuclear Magnetic Resonance
NTCP	sodium taurocholate co-transporting polypeptide
NTD	N-terminal domain of Cp
NZ-4	isothiafludine
ORF	overlapping reading frame
P	partition coefficient
PBS	Phosphate buffer saline
PDB	Protein Data Bank
PFA	Perfluoroalkoxy
PgRNA	pregenomic RNA
PPA	phenylpropenamide
PT	pyrazolyl-thiazole
qPCR	quantitative polymerase chain reaction
rcDNA	relaxed circular DNA
RED	Rapid equilibrium dialysis
RNA	ribonucleic acid
Rt	Retention time
SAR	Structure Activity Relationship
SBA	sulfamoylbenzamide
SOC	standard of care
T½	Half-life
TAF	tenofovir alafenamide fumarate
TDF	tenofovir disoproxil fumarate
TEA	triethylamine
TEMPO	(2,2,6,6-tetramethylpiperidin-1-yl)oxyl
THF	tetrahydrofuran
UPLC-MS	Ultra performance liquid chromatography-mass spectrometer
WHO	World Health Organization
(-)	negative
(+)	positive
5’	5’terminal
3TC	lamivudine

Index

UNIVERSITAT ROVIRA I VIRGILI
TARGET-BASED DESIGN, STRUCTURAL OPTIMIZATION AND CHARACTERIZATION OF NOVEL HEPATITIS B VIRUS CAPSID ASSEMBLY MODULATORS
Elena Detta

Chapter I.

Introduction

I.1 Hepatitis B virus disease as a public health problem

Hepatitis B is a serious liver infection caused by the hepatitis B virus (HBV), which can result in either an acute or chronic disease.[1] Roughly one third of the global population has been infected with HBV during their lives. While the vast majority of adults resolve this acute infection within six months and subsequently develop long-lasting immunity, a consistent number of patients, that have contracted the virus perinatally or during early childhood, become chronically infected. Chronic HBV infection is, in most cases, asymptomatic and generally associated with sustained liver inflammation and activation of fibrotic processes, leading to liver fibrosis, cirrhosis and eventually hepatocellular carcinoma (HCC).[2]

According to the World Health Organization (WHO)'s most recent estimates, approximately 257 million people are chronically infected with HBV worldwide and more than 800,000 patients die annually from HBV-associated hepatic diseases.[3,4] Although a safe and effective prophylactic vaccination has been available since 1982, chronic hepatitis B still remains a severe global health issue. Currently, HBV infection is treated with long-term administration of reverse transcriptase inhibitors, that are nucleotide or nucleoside analogues (NUCs), and, less commonly, interferon-based (IFN) therapy. Profound virological suppression, with associated reduced risk of liver complications, is achieved with continued NUCs or IFN treatments. However, functional cure, defined as serum hepatitis B surface antigen (HBsAg) loss, occurs in only a small number of IFN-treated patients, due to the high incidence of disease reactivation after treatment cessation. Thus, there is a significant need for novel therapeutic approaches to enhance the rate of HBsAg seroclearance and completely eradicate the virus from carriers.

[1] Yuen, M.-F. *et al*. Hepatitis B virus infection. *Nat. Rev. Dis. Prim.* 4, 18035, (2018).
[2] Tang, L. S. Y., Covert, E., Wilson, E. & Kottilil, S. Chronic Hepatitis B Infection: A Review. *JAMA* 319, 1802–1813, (2018).
[3] WHO Hepatitis B Fact Sheet July 2019. https://www.who.int/news-room/fact-sheets/detail/hepatitis-b.
[4] World Health Organization & World Health Organization. Global Hepatitis Programme. Global hepatitis report, 2017.

I.1.1 Epidemiology and transmission

The burden of HBV infection is geographically disparate and its epidemiology is described as the prevalence of serum HBsAg in a population.[5] HBV prevalence is highest in South-East Asia and Sub-Saharan Africa, where chronic HBV infection affects more than 6% of the population, while Northern-Western Europe and America are considered low-prevalence areas with less than 2% of carriers. The virus is mainly transmitted through contact with blood and other body fluids and predominant transmission modes strictly determine the prevalence of the infection. In endemic areas, where roughly 45% of the global population lives, transmission mostly occurs in vertical mode (from mother to child), whereas the infection is generally horizontally acquired through sexual contact during adolescence or adulthood in low-prevalence countries. When infected with HBV, age is a key risk factor for progression to chronicity. It is estimated that 90% of neonates and 30-50% of children infected before the age of 6 years become chronic carriers, while fewer than 5% of adult patients develop chronic infection. It derives that the prevalence of chronic HBV infection in a given area depends on both the predominant mode of transmission in that population and the age at infection, which is by itself inversely proportional to the probability of progression to chronicity.[6]

Infection with different HBV genotypes is also important in determining the risk of chronicity. There are eight major HBV genotypes (A-H) in humans, each with a distinct geographical distribution. Although current FDA-approved vaccines are effective against each genotype, infection with different HBV genotypes is generally associated with different progression to chronicity and response to interferon-based therapy.[7]

[5] MacLachlan JH, Cowie BC. Hepatitis B virus epidemiology. *Cold Spring Harb Perspect Med*.;5(5):a021410 (2015).
[6] Caballero, A., Tabernero, D., Buti, M. & Rodriguez-Frias, F. Hepatitis B virus: The challenge of an ancient virus with multiple faces and a remarkable replication strategy. *Antiviral Research* vol. 158, 34–44, (2018).
[7] Sunbul M. Hepatitis B virus genotypes: global distribution and clinical importance. *World J Gastroenterol*. 20(18):5427-5434, (2014).

I.2 Hepatitis B Virus

Hepatitis B virus is a member of the *Hepadnaviridae* family. Mature and infectious HBV, commonly described as a Dane particle, consists of an icosahedral enveloped capsid containing a partially double stranded, relaxed circular DNA (rcDNA) genome of about 3.2 kilobase (kb) pairs and a viral polymerase, which is covalently attached to the 5' end of the negative-strand DNA.[8]

The HBV genome is organized in a circular form with the positive-strand DNA forming the inner circle and the negative strand forming the outer circle (Fig. I.1). The coding capacity of HBV genome is defined by four partially overlapping open reading frames (ORFs). The S ORF, which is divided into pre-S1, pre-S2 and S regions, encodes the viral surface proteins. The C ORF, including the pre-core and core regions, encodes for both HBV core protein (HBcAg) and serum hepatitis B e antigen (HBeAg). The regulatory X protein (HBxAg) and polymerase are encoded by the X and P ORFs, respectively.[9]

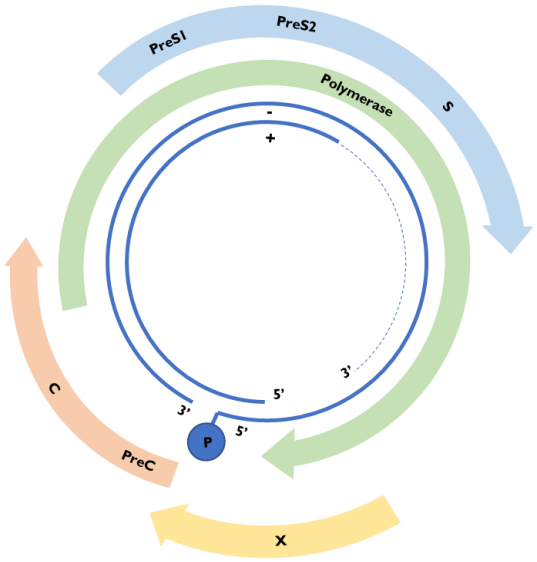


Figure I.1: HBV genome organization. The genome of HBV is a partially double-stranded, relaxed circular DNA (~ 3.2 kb). The two strands are asymmetric: the (+) strand, forming the inner circle, is incomplete, while the (-) strand is completely circular and presents the viral polymerase covalently attached on the 5' end. HBV genome codes for four highly overlapping regions (ORFs): S, C, P and X. The PreS1, PreS2 and S regions (in blue) encode for the viral surface proteins. The PreC and C (in red) genes encode for both HBcAg and HBeAg. The polymerase and the X protein (HBxAg) are encoded by the P (in green) and X (yellow) ORFs, respectively. Adapted from[9].

[8] Lamontagne RJ, Bagga S, Bouchard MJ. Hepatitis B virus molecular biology and pathogenesis. *Hepatoma Res.* 2:163-186, (2016).
[9] Yuen, M.-F. *et al*. Hepatitis B virus infection. *Nat. Rev. Dis. Prim.* 4, 18035, (2018).

I.2.1 HBV replication cycle

HBV is a non-cytopathic virus, which primarily attacks the liver (Fig. I.2).[10] To initiate infection, HBV particles bind to hepatocytes through a reversible interaction between heparin sulfate proteoglycans (HSPG) on the hepatocyte surface and the antigenic loop (AGL) of the HBV capsid proteins. Following adhesion, the HBV receptor, sodium taurocholate co-transporting polypeptide (NTCP), [11], is specifically recognized by the pre-S1 domain of the HBsAg (or large surface protein L-HBsAg). This high-affinity interaction triggers viral internalization via endocytosis, which precedes the release of nucleocapsid into the cytoplasm. The rcDNA genome inside the nucleocapsid is then transported into the host nucleus, where it is converted into the cccDNA using the host cell DNA repair response machinery. cccDNA is packaged into chromatin by histone and non-histone proteins to form a mini-chromosome and its transcriptional activity is finely regulated by epigenetic modifications on histone proteins. This stable mini-chromosome serves as a template for pregenomic RNA (pgRNA) transcription and viral protein production. Its function is regulated by various regulatory factors and viral proteins, such as core proteins, and its long-term persistence in the nucleus of infected liver cells is the major cause of HBV chronicity and disease reactivation after termination of treatment.[12]

cccDNA uses the RNA polymerase II to generate four viral RNA transcripts, including 0.7, 2.4, 2.1 and 3.5 kb mRNAs, which are translated to all seven HBV proteins. The small 0.7 kb RNA encodes the regulatory HBx protein, while the three HBsAg surface capsid proteins, the small surface protein (S- HBsAg), the middle surface protein (M-HBsAg) and the large surface protein (L-HBsAg), are produced from the 2.4 and 2.1 kb RNA transcripts.

[10] Karayiannis, P. Hepatitis B virus: virology, molecular biology, life cycle and intrahepatic spread. *Hepatol. Int*. 11, (2017).
[11] Li, W. The Hepatitis B Virus Receptor. *Annu. Rev. Cell Dev. Biol*. 31, 125–147, (2015).
[12] Schadler S. *et al*., HBV Life Cycle: Entry and Morphogenesis. *Viruses*, 1(2):185-209, (2009).

The 3.5 kb mRNA includes both pregenomic RNA (pgRNA) and precore RNA, that is slightly longer upstream to the pgRNA, encoding for the viral polymerase, the core protein (HBcAg) and the precore protein, which generates the secretory hepatitis B e antigen (HBeAg) through post-translational proteolysis.[13]

Pregenomic RNA (pgRNA), together with all other transcription products, is transported from the nucleus to the cytoplasm. The HBV DNA polymerase binds to the stem-loop structure at the 5' terminus of pgRNA and both are encapsidated within a 120-mer of core protein dimers to form an immature viral nucleocapsid. Inside the nucleocapsid, the pgRNA is converted to rcDNA by reverse transcription catalyzed by the viral polymerase. rcDNA-containing nucleocapsids can either re-enter the nucleus and replenish the cccDNA pool or be enveloped and released from hepatocytes. To egress from cells, mature nucleocapsids are transferred to the endoplasmic reticulum (ER) and/or Golgi apparatus and after eventual post-translational modifications are secreted as 42 nm spherical infectious particles. Interestingly, a large number of non-functional virus-like particles, such as empty capsids or capsids with incomplete RNA or DNA fragments, are also released from hepatic cells. Despite being not infectious, these particles might play key roles in the immune tolerance of HBV infection, since they present a considerable quantity of HBsAg proteins on their surfaces and high levels of those structures are found in the blood of infected patients.[14]

[13] Mitra B. *et al.*, Host functions used by hepatitis B virus to complete its life cycle: Implications for developing host-targeting agents to treat chronic hepatitis B. *Antiviral Research*, 158: 185-198, (2018).
[14] Tong, S. & Revill, P. Overview of hepatitis B viral replication and genetic variability. *Journal of Hepatology* vol. 64, S4–S16, (2016).

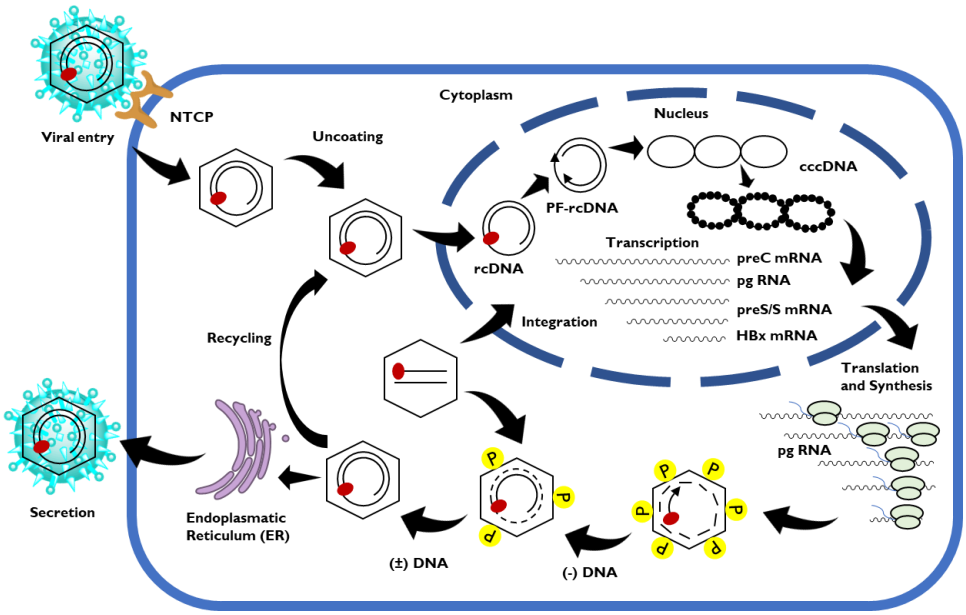


Figure I.2: HBV replication cycle. HBV is a hepatotropic virus. Viral entry occurs via strong binding to the sodium-taurocholate cotransporting polypeptide (NTCP) on the hepatocyte surface. The virion is then internalized, uncoated and the nucleocapsid is transported to the nucleus, where the HBV genome is released. The rcDNA is converted to cccDNA, which serves as a template for transcription of the viral mRNA species. pgRNA is exported to the cytoplasm, where is encapsidated along with the DNA polymerase and reverse transcribed into the rcDNA. The mature nucleocapsid can be then recycled to the nucleus to enrich the pool of cccDNA or it can acquire an envelope and be secreted from the liver cell.

I.3 Acute and chronic HBV infection

In the absence of an adequate host immune response, acute HBV infection becomes a persistent life-long condition, which at present can be treated with suppressive antiviral agents. Current treatments reduce HBV DNA levels, slow the progression of hepatic failure and liver cancer, but are not curative. In most patients, complete suppression of the virus is not achieved, which is mainly attributed to the maintenance of the pool of cccDNA during therapy.[15]

HBV, as it has been previously mentioned, is a non-cytopathic virus, which does not directly kill hepatocytes. However, during infection HBV activates the host immune system, which, in the attempt to eliminate the virus, causes long-term inflammation and necrosis to liver tissue.

[15] Zoulim F, Durantel D. Antiviral therapies and prospects for a cure of chronic hepatitis B. *Cold Spring Harb Prospect Med.*, 5(4):a021501, (2015).

In chronically infected patients, repeated liver injury due to persistent inflammation can result in a wide spectrum of liver diseases ranging from hepatic failure to cirrhosis and ultimately hepatocellular carcinoma (HCC).[16]

I.3.1 Symptoms and diagnosis

Two thirds of patients with acute HBV infection are asymptomatic, while roughly one-third of infected adults develop clinical symptoms, such as fever, fatigue, abdominal pain and jaundice. [17] The detection of specific viral antigens and their respective antibodies in serum of infected patients can be used to diagnose HBV infections and associated liver complications.

The standard serological marker of acute and chronic infection is the detection of HBsAg in serum, which is also used to establish the prevalence of chronic HBV infection in a given population. Persistence of HBsAg (> 6 months), with or without concomitant HBeAg, is a key marker of chronicity and high levels of HBsAg are directly correlated to the risk of hepatocellular carcinoma and viral rebound after NUC treatment cessation. Moreover, variation of HBsAg concentrations can be observed during different stages of chronic HBV disease.[18]

Commercially available serological tests can detect other HBV serological markers. For instance, transition from HBeAg-positive to HBeAg-negative chronic HBV infection, described as HBeAg seroconversion, is generally established after HBeAg serological testing. Viral replication rate is directly assessed by HBV DNA tests that are regularly performed in all chronic carriers. Notably, the efficacy of current HBV therapies can usually be predicted by monitoring HBV DNA concentrations in serum.[19]

[16] Llovet, J.M., Kelley, R.K., Villanueva, A. *et al*. Hepatocellular carcinoma. *Nat Rev Dis Primers* 7, 6 (2021).
[17] Liang, T. J. Hepatitis B: The virus and disease. *Hepatology*, 49, S13-S21, (2009).
[18] Song JE, Kim DY. Diagnosis of hepatitis B. *Ann Transl Med.*, 4(18):338, (2016).
[19] Ghosh M, Nandi S, Dutta S, Saha MK. Detection of hepatitis B virus infection: A systematic review. *World J Hepatol.*, 7(23):2482-2491, (2015).

I.3.2 Natural history of HBV infection

Following HBV infection, there is an incubation period of 4-7 weeks, in which levels of both HBV DNA and HBsAg are non-detectable in serum or liver. While full clinical recovery from acute infection can be achieved in adult patients, it does not generally occur in neonates or children, whose immune system is not yet mature.

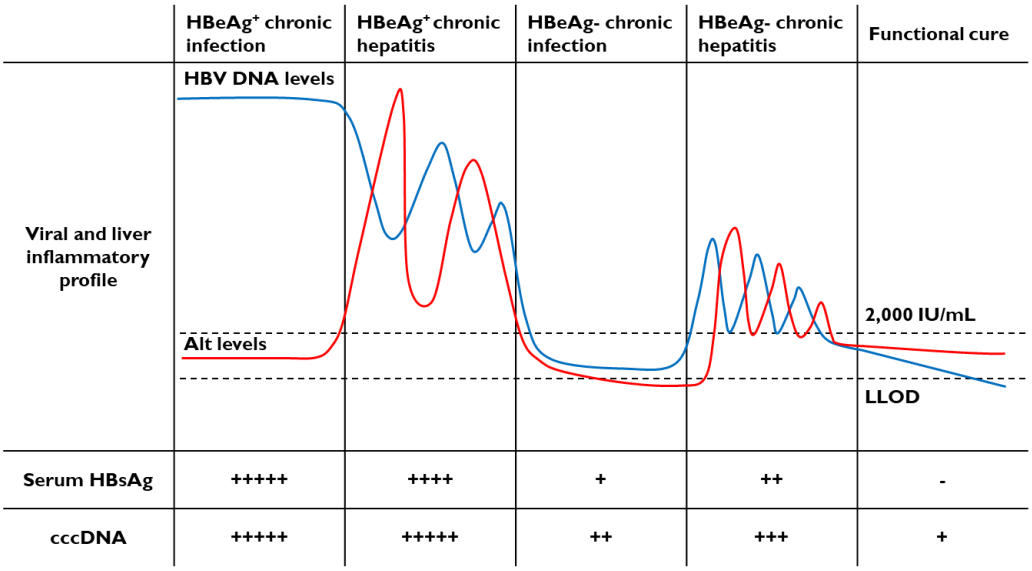


Figure I.3: Natural history of chronic hepatitis B infection. The natural history of chronic HBV infection is described as four (not necessarily sequential) virological phases, ending with a functional cure, characterized by the sustained loss of serum HBsAg with undetectable HBV DNA. High levels of alanine transaminase (ALT, in red line), a sensitive marker of liver damage, are present during both the HBeAg-positive and –negative hepatitis phases, while are normal in the chronic infection phases. cccDNA and serum HBsAg decline over time, but the complete displacement of cccDNA reservoirs is never achieved. (- : not detected; + : detected; ++/+++/++++/+++++ : detected with high frequency; LLOD : lower limit of detection). [20]

The natural history of chronic HBV infection is described as four clinical or virological phases, which are defined by monitoring the levels of HBeAg, HBV DNA and alanine transaminase (ALT). These phases do not necessarily occur in a sequential manner for all patients and strictly depend on the equilibrium between virus replication and host immune response (Fig. I.3). [21]

[20] Fanning, G. C., Zoulim, F., Hou, J. & Bertoletti, A. Therapeutic strategies for hepatitis B virus infection: towards a cure. *Nat. Rev. Drug Discov.* 18, 827–844, (2019).
[21] Idilman, R. The summarized of EASL 2017 Clinical Practice Guidelines on the management of hepatitis B virus infection. *Turkish J. Gastroenterol*. 28, 412–416, (2017).

Phase 1 and 2, also termed HBeAg-positive chronic infection and HBeAg-positive chronic hepatitis B, are both characterized by high levels of HBV DNA and serum positivity for HBeAg. The first phase generally occurs in subjects infected during childhood, while the second is more frequent during adulthood chronic infection. While hepatic inflammation is practically absent during HBeAg-positive chronic infection, high levels of ALT, a sensitive marker of liver damage, are released in phase 2.

Fluctuation of ALT and HBV DNA levels during HBeAg-positive chronic hepatitis B determine the so-called "hepatitis flares", which are responsible for persistent inflammation and long-term liver damage, usually leading to varying degrees of fibrosis and cirrhosis.[22]

Nevertheless, in most cases there is a dramatic reduction of HBV DNA concentrations and patients achieve HBeAg seroconversion to anti-HBe positivity.

High levels of HBeAg antibodies, low or undetectable HBV DNA concentrations and stable or normal ALT are observed during HBeAg-negative chronic HBV infection (phase 3). Patients, who remain in this phase of the disease, have a good prospect of recovery and face low risk of progression to cirrhosis and hepatocellular carcinoma.[23]

However, HBeAg-negative patients frequently develop chronic hepatitis B (phase 4), characterized by serious liver damage and low rate of disease remission. In this phase both ALT and HBV DNA values are elevated, but lower than in HBeAg-positive phases, and there are sustained levels of HBeAg antibodies. Interestingly, chronic HBV infection is characterized by progressive decrease of both cccDNA and HBsAg concentrations during the different stages of the disease.

[22] Ming-Ling Chang, Yun-Fan Liaw, Hepatitis B flares in chronic hepatitis B: Pathogenesis, natural course and management, *Journal of Hepatology*, 61, 6, 1407-1417, (2014).
[23] Liaw YF. HBeAg seroconversion as an important end point in the treatment of chronic hepatitis B. *Hepatol Int*. 3(3):425-433, (2009).

I.3.3 Indications for treatment

According to the most recent clinical practice guidelines on the management of HBV infection, antiviral treatment is indicated in the presence of: (1) HBV DNA concentrations higher than 2,000 IU/mL, (2) ALT values higher than one or two times the upper limit of normal (ULN) and/or (3) cirrhosis.[24]

The ideal therapeutic goal to achieve a definitive cure of chronic hepatitis B disease is the total elimination of the virus from carriers, that is translated into the sustained HBsAg loss combined with undetectable levels of HBV DNA and, more importantly, the complete displacement of cccDNA reservoirs. [25]

I.4 Prevention and current therapies

I.4.1 Vaccination

Over the last 40 years the implementation of systematic infant immunization campaigns has profoundly impacted the global burden of chronic HBV disease, reducing the incidence of liver cancer in children and young adults. The first plasma-based vaccine was approved in 1982 and consisted of inactivated HBsAg extracted from the plasma of HBV chronic infected patients. This vaccine was gradually replaced by safer recombinant vaccines, which do not contain any blood products. At present it is recommended that all infants receive the first dose of vaccine within 24 hours of birth. The efficacy of the HBV vaccination is extremely high, inducing anti-HBsAg antibodies in 95% of individuals and offering lifelong protection. [26]

[24] EASL 2017 Clinical Practice Guidelines on the management of hepatitis B virus infection, European Association for the Study of the Liver, *Clinical Practice Guidelines*, 67, 2, 370-398, (2017).
[25] Kim, W. R. Emerging Therapies Toward a Functional Cure for Hepatitis B Virus Infection. *Gastroenterol. Hepatol.* (N. Y). 14, 439, (2018).
[26] World Health Organization, Hepatitis B vaccines: WHO position paper, July 2017-Recommendations, Vaccine, 37, 2, 223-225, (2019).

I.4.2 Standard of care therapeutics for HBV infection

The standard of care (SOC) against chronic HBV includes six nucleos(t)ide viral DNA polymerase inhibitors (NUCs, Fig. I.4), namely lamivudine, adefovir dipivoxil, entecavir, telbivudine, tenofovir disoproxil fumarate (TDF) and tenofovir alafenamide (TAF), and the immuno-modulator pegylated interferon alpha (IFN-α). Both classes of therapy effectively suppress the replication of the virus and reduce persistent liver inflammation, improving the quality of life and overall survival of chronically infected patients.[27]

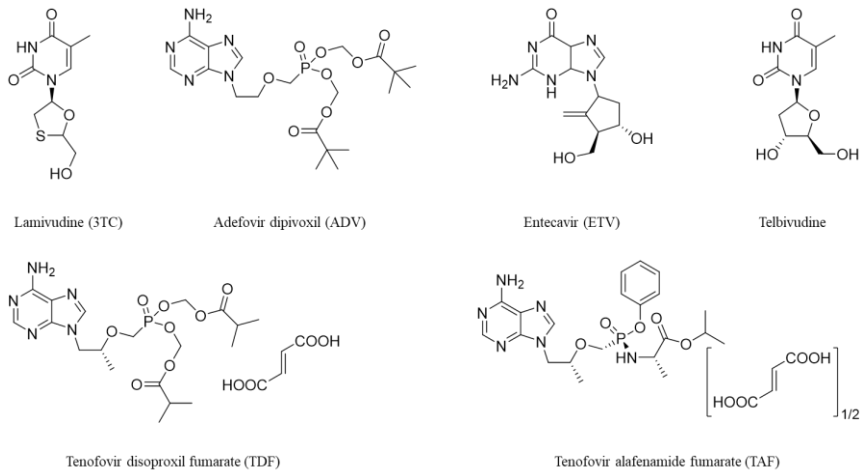


Figure I.4: Approved nucleos(t)ide analogues for the treatment of HBV.

Due to their excellent side effect profile and ease of administration, NUCs are the treatment of choice for most HBV patients. Entecavir, TDF and TAF are considered first-line agents, mainly because of their high efficacy and low resistance rates, whilst the use of lamivudine and adefovir is more limited due the development of resistant variants. In lamivudine-resistant patients, TDF and TAF are the best therapeutic options; however, compared to TDF, TAF treatment results in improved rates of ALT normalization and decreased side effects. In most cases, NUC therapy requires prolonged or lifelong treatments to avoid viral rebound.[28]

[27] Tawada *et al.*, Current and future directions for treating hepatitis B virus infection. *World J Hepatol.*, 7(11):1541-1552, (2015).
[28] Roade L. *et al.*, Long-term efficacy and safety of nucleos(t)ides analogues in patients with chronic hepatitis B. *Therapeutic Advances in Infectious Disease*, (2021).

Two different injectable formulations of pegylated interferon alfa (pegylated interferon alfa-2-a and alfa-2-b) are currently available to treat HBV infection. PEG-IFN-α therapy has both immune modulatory and direct antiviral effects and is indicated only for a small subset of patients with HBV genotypes A or B, which are more responsive to the treatment. Its use is limited by a wide range of severe adverse effects, including flu-like symptoms, cytopenia and depression.[29]

I.4.3 Functional cure: urgent need for novel intervention strategies

Despite their efficacy, current approved drugs do not provide a "functional cure", which is associated with the loss of HBsAg in the blood, and do not affect the transcriptional activity of cccDNA, which represents the main barrier to the complete suppression of the virus. cccDNA is a mini-chromosome that is responsible for viral persistence in the nucleus of infected hepatocytes even when stable suppression of replication is achieved by polymerase inhibitors. [30] Currently, there are no agents that can eradicate the cccDNA viral reservoir; elimination of these reservoirs therefore represents the ultimate goal of current research into novel antivirals.

Several direct antiviral strategies targeting different steps in the HBV replication cycle are being explored, which might directly or indirectly affect the transcriptional activity of cccDNA and therefore lead to a definitive displacement of the virus. [31,32,33] In this context, HBV core protein has been recently recognized as a promising target, since it is not only involved in the processing of nucleocapsid assembly, but it also plays significant roles in the pgRNA encapsidation, reverse transcription, cccDNA maintenance and suppression of immune response.

[29] Santantonio TA, Fasano M. Chronic hepatitis B: Advances in treatment. *World J Hepatol*. 6(5):284-292, (2014).
[30] Allweiss, L. & Dandri, M. The role of cccDNA in HBV maintenance. *Viruses*, 9(6):156, (2017).
[31] Pei, Y., Wang, C., Yan, S. F. & Liu, G. Past, Current, and Future Developments of Therapeutic Agents for Treatment of Chronic Hepatitis B Virus Infection. *Journal of Medicinal Chemistry* vol. 60, 6461–6479, (2017).
[32] Feng, S. et al. Discovery of Small Molecule Therapeutics for Treatment of Chronic HBV Infection. *ACS Infectious Diseases* vol. 4, 257–277, (2018).
[33] Lopatin, U. Drugs in the Pipeline for HBV. *Clinics in Liver Disease* vol. 23, 535–555, (2019).

I.5 HBV core protein as promising target

I.5.1 Capsid structure and function

HBV core protein or core antigen (HBcAg) is a 20 kDa-protein expressed in the cytoplasm of HBV infected hepatocytes. During viral replication HBcAg dimers self-assemble to form the icosahedral virus nucleocapsid, which is the inner virus protein capsid that contains the genomic DNA. 95% of HBV capsids have a T=4 icosahedral symmetry and contain 240 copies of core protein, while 5% have a T=3 symmetry and 180 subunits (Fig. I.5). [34]

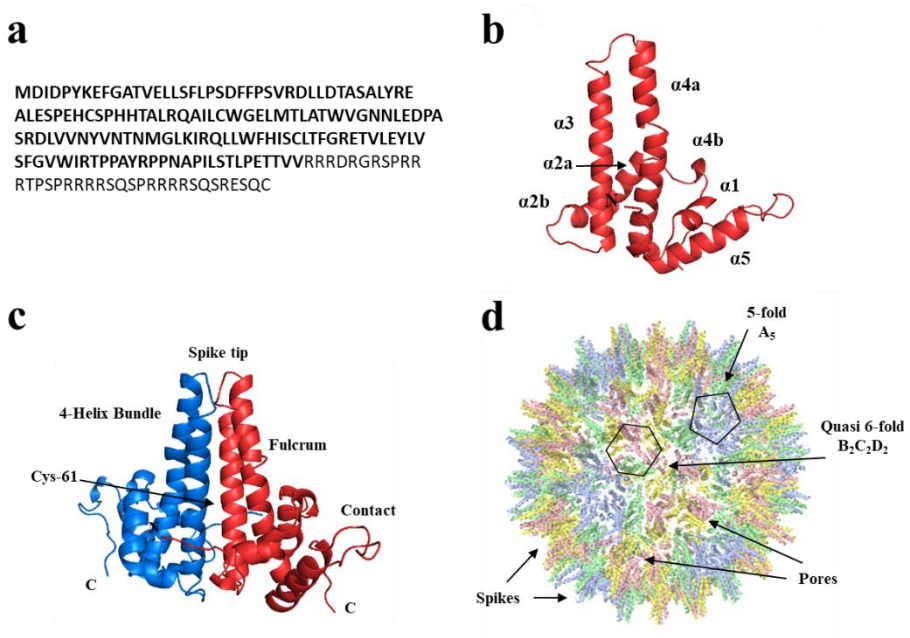


Figure I.5: HBV T=4 capsid structure. a) Full-length human HBV capsid structure. Depending on the genotype, the primary structure of HBcAg monomers presents 183 or 185 amino acids (Cp183 or Cp185). The first 149 residues form the N-terminal assembly domain (Cp149, in bold) and the remaining 34 amino acids constitute the arginine-rich RNA-binding C terminal domain (CTD). **b)** HBV capsid protein monomer fold. **c)** HBV capsid protein dimer. Two adjacent HBcAg monomers (shown in blue and red, respectively) pack together to form a four-helix bundle dimerization interface, characterized by the presence of a disulfide bridge between the Cys-61 residues (indicated by an arrow) of the two monomers. Each dimer presents three subdomains: (1) spike tips, (2) fulcrum and (3) contact. **d)** T=4 icosahedral capsid. In a T=4 virus the capsid consists of a fenestrated protein shell from which protrude spikes. Each capsid contains 240 copies of core protein or 120 dimers. HBcAg subunits A, B, C and D are shown in green, yellow, red and blue. The positions of the icosahedral 5-fold and quasi 6-fold vertices are indicated. (PDB ID 1QGT).

[34] Wynne, S. A., Crowther, R. A. & Leslie, A. G. W. The crystal structure of the human hepatitis B virus capsid. *Mol. Cell*, 3, 6, 771-780, (1999).

In a T=4 virus, the capsid consists of a thin protein shell, with an inner radius of 130 Å and a thickness of 20 Å, from which protrude spikes approximately 25 Å in length and 20 Å in width. Large pores, about 12-15 Å in diameter, fenestrate the capsid surface and are essential to nucleotide trafficking during reverse transcription of pgRNA to dsDNA. Structurally disordered C-terminal tails of each HBcAg monomer are located towards the interior of the capsid. These are believed to interact with the viral nucleic acid and are transiently exposed on the outside through capsid fenestrations.[18]

The predominant form T=4 icosahedral HBV capsid is comprised of 60 asymmetric units, each formed by 4 HBcAg monomers (A-D) or two chemical identical but structurally distinct dimers (AB or CD).[35] According to the principle of quasiequivalence, [36], A monomers form the icosahedral 5-fold vertices, while two sets of B-C, C-D and D-B interfaces form the quasi 6-fold vertices (Fig. I.5d).

I.5.2 HBcAg monomers

Depending on the genotype, the primary structure of HBcAg monomers presents 183 or 185 amino acids (Cp183 or Cp185). In Cp183, the first 149 residues form the N-terminal α-helical assembly domain (Cp149) and the remaining 34 amino acids constitute the arginine-rich RNA-binding C-terminal domain (CTD) (Fig. I.5a).

The N-terminal domain has 5 α-helices connected by loops. Helices 3 (aa 13-17) and 4 (aa 79-110) form a long α-helical hairpin, surrounded on the three sides by helices 1 (aa 13-17), 2 (aa 27-43) and 5 (aa 112-127). The C-terminal domain (aa 150-183) is connected to the N-terminal domain by a proline-rich loop and can be considered as a sequence of arginine-rich clusters (Fig. I.5b).[37] CTD is not strictly indispensable for capsid assembly and is mainly involved in the rcDNA synthesis, pgRNA encapsidation and binding to host protein.

[35] Venkatakrishnan, B. & Zlotnick, A. The Structural Biology of Hepatitis B Virus: Form and Function. *Annu. Rev. Virol.* 3, 429–451, (2016).
[36] Caspar, D. L. and Klug, A. Physical principles in the construction of regular viruses. *Cold Spring Harb. Symp. Quant. Biol.* 27, 1–24, (1962).
[37] Selzer L, Zlotnick A. Assembly and Release of Hepatitis B Virus. *Cold Spring Harb Perspect Med*. 5(12):a021394, (2015).

I.5.3 HBcAg dimer-dimer interface

It has been shown that the structured N-terminal domain spontaneously assembles *in vitro* and *in vivo* to form capsid-like structures identical to nucleocapsids isolated from patient serum.[38] Two HBcAg monomers associate to form a compact dimeric unit, which is the basic soluble element of HBV core protein. Hydrophobic interactions mainly drive the dimerization process and two α-helical hairpins of two adjacent monomers pack together to form a four-helix bundle. Salt bridges and hydrogen bonds further stabilize the highly hydrophobic interface, where an additional disulfide bond can form between Cys61 in helix 3 of two monomers. Structurally, each dimer presents a highly conserved chassis with three distinct subdomains: (1) the spike tips, which form the upper part of the four-helix bundle, (2) the fulcrum, constituted by helices 1 and 2 of each monomer and (3) the C-terminal interdimer contact domain (Fig. I.5c).[38]

During capsid assembly HBcAg dimers are associated by weak hydrophobic interactions between contact domains involving residues Tyr132 (about 10 %), Val124, Arg127, Ile129 and Ile139. Interestingly, mutation of Tyr to Ala at position 132 dramatically reduces buried hydrophobic surface, so that HBcAg Y132A mutant is not able to form complete capsid structures.[39]

A hydrophobic pocket has been identified at the dimer-dimer interface, which can be targeted by small-molecule antiviral agents.[40]

[38] Packianathan C. *et al*. Conformational changes in the hepatitis B virus core protein are consistent with a role for allostery in virus assembly, *Journal of Virology*, 84 (3), 1607-1615, (2010).
[39] Ceres P, Zlotnick A., Weak protein-protein interactions are sufficient to drive assembly of hepatitis B virus capsids, *Biochemistry*, 41(39):11525-31, (2002).
[40] Usha Viswanathan *et al*., Targeting the multifunctional HBV core protein as a potential cure for chronic hepatitis B, *Antiviral Research*, 182, 104917, (2020).

I.6 Capsid Assembly Modulators

Nucleocapsid assembly is a critical step in the HBV replication cycle and its disruption may affect the replication of infectious virions, [41], and regulate the replenishment of the pool of cccDNA. Different factors, such as temperature, ionic strength or small molecules can enhance association energy, strengthen inter-dimer interactions and alter capsid assembly *in vitro*.

Currently, a significant number of compounds, which bind to the HBV core protein at the dimer-dimer interface, misdirect nucleocapsid assembly and/or disregulate the packaging of the pgRNA/DNA polymerase complex, are under clinical development as antiviral agents. [42] These molecules, also known as capsid assembly modulators (CAMs) or HBV core protein allosteric modulators (CpAMs), could represent a valid alternative to the current SOC for chronic HBV infection and maintain their activity against NUC-resistant mutants.

CAMs fall into two distinct categories according to their mechanism of action (MoA). Class II or "HAP-like" CAMs, represented by heteroaryldihydropyrimidines (HAPs), enhance hydrophobic interactions between adjacent core protein dimers and alter the kinetic of capsid assembly, resulting in aberrant dysfunctional HBcAg polymers. On the contrary, Class I or "non-HAP-like" CAMs, including phenylpropenamides (PPAs), sulfamoylbenzamides (SBAs), benzamides (BAs), isothiafludine (NZ-4) and others, [22], induce the formation of morphologically intact capsids that are devoid of viral genome and thus preclude the viral replication.

Interestingly, despite their structural diversity and pharmacological phenotypes, all CAMs share the same binding site at the HBcAg dimer-dimer interface.

[41] Zheng, C. L., Fu, Y. M., Xu, Z. X., Zou, Y. & Deng, K. Hepatitis B virus core protein dimer-dimer interface is critical for viral replication. *Mol. Med. Rep*. 19, 262–270, (2019).
[42] Nijampatnam, B. & Liotta, D. C. Recent advances in the development of HBV capsid assembly modulators. *Current Opinion in Chemical Biology* vol. 50, 73–79, (2019).

I.6.1 HAP-like CAMs

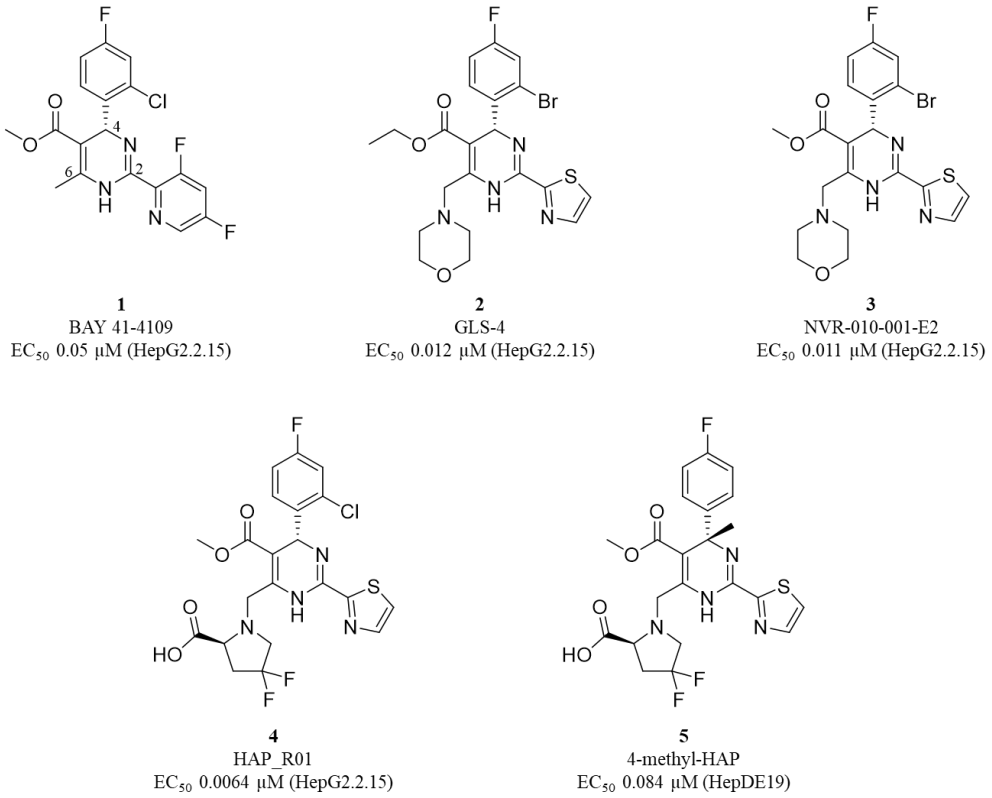


Figure I.6: Representative heteroaryldihydropyrimidine (HAP) series of HBV CAMs. EC_{50} values are for inhibition of HBV DNA replication *in vitro*.

The first example of the HAP series, BAY 41-4109 (**1**, Fig. I.6), was discovered by Bayer scientists in the late '90s. It showed moderate antiviral activity both *in vitro* in stably transfected HepG2.2.15 cells with a EC_{50} of 0.05 µM and *in vivo* in a transgenic mouse model. [43] However, it was reported to induce hepatotoxicity in rats at high doses, [44], and was not further developed in the clinic.

Structural modifications around the central HAP core identified next generation analogues with optimized potency, aqueous solubility, pharmacokinetic (PK) and toxicity profiles. Due to their promising drug-like properties, a number of such HAP compounds were also advanced to clinical trial. [16]

[43] Weber, O. *et al*. Inhibition of human hepatitis B virus (HBV) by a novel non-nucleosidic compound in a transgenic mouse model. *Antiviral Res*. 54, 69–78, (2002).
[44] Shi, C. *et al*. NMR-spectroscopy-based metabonomic approach to the analysis of Bay41-4109, a novel anti-HBV compound, induced hepatotoxicity in rats. *Toxicol. Lett*. 173, 161–167, (2007).

Notably, the introduction of C6 morpholinyl or gem-difluoropyrrolidinyl (with or without an additional carboxyl group) and C2 thiazoyl moieties on the dihydropyrimidine ring led to the discovery of more potent second and third generation HAP analogues (see Fig. I.6), such as GLS-4 (**2**, EC_{50} 0.012 µM in HepG2.2.15), [45,46], NVR-010-001-E2 (**3**, EC_{50} 0.011 µM in HepG2.2.15), [47], and HAP_R01 (**4**, EC_{50} 0.0064 µM in HepG2.2.15). [48]

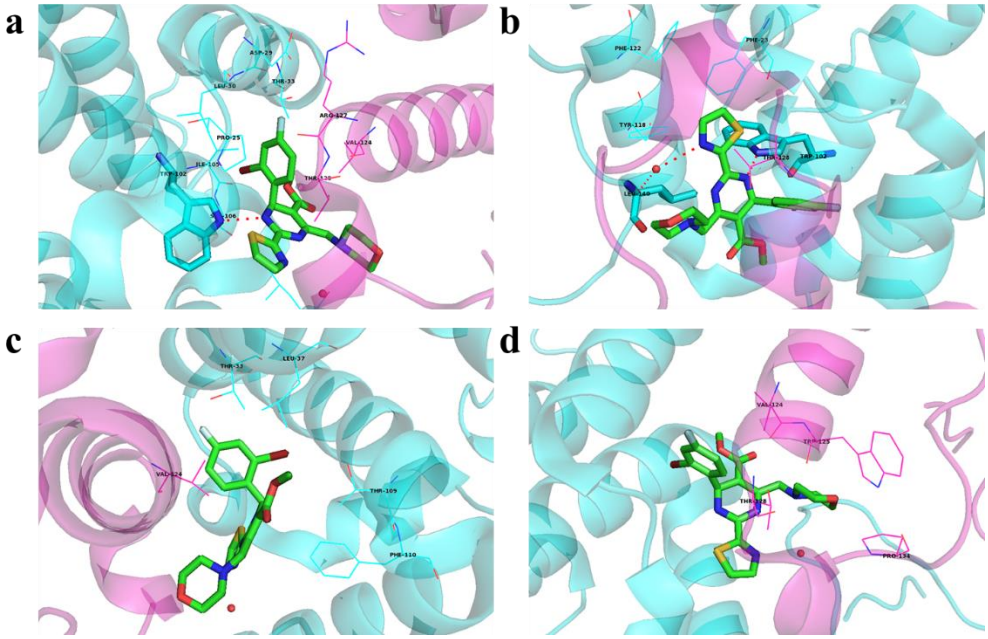


Figure I.7: Crystal structure of HBV core protein Y132A mutant hexamer in complex with NVR-010-001-E2 (3, Fig. I.6). The B-C dimer-dimer interface is shown as an example of the compound binding pocket. Chains B and C are represented as cartoon, colored in cyan and pink, respectively. Key residues on both chains are displayed in lines (or sticks) and labelled. Carbon atoms of compound **3** are highlighted in green sticks. Hydrogen bond interactions are indicated by red dashed lines. **a)** The central HAP core of **3** forms a key H-bond with Trp102 (shown in sticks). The 3-bromo-4-fluorophenyl ring inserts into a deep hydrophobic subpocket formed by Pro23, Asp29, Leu30, Thr33, Trp102, Ile105 and Ser106 from chain B and Val124, Arg127 and Thr128 from chain C. **b)** The 2nd-positioned thiazole moiety interacts with Leu140 through a water-mediated H-bond and sits in a highly hydrophobic environment defined by residues Phe23, Trp102, Tyr118 and Phe122 from chain B and Thr128 from chain C. **c)** The methyl-ester group of **3** is coplanar with both the central HAP core and the thiazole moiety. Main interactions are with Thr33, Leu37, Thr109 and Phe110 from chain B and Val124 from chain C. **d)** The C6 morpholinyl group is partially solvent-exposed, however, minor contacts are with Val124, Trp125, Thr128 and Pro134 from chain C. (PDB ID 5E0I).

[45] Ren, Q. *et al*. Discovery of hepatitis B virus capsid assembly inhibitors leading to a heteroaryldihydropyrimidine based clinical candidate (GLS4). *Bioorganic Med. Chem*. 25, 1042–1056 (2017).
[46] Wu, G. *et al*. Preclinical characterization of GLS4, an inhibitor of Hepatitis B virus core particle assembly. *Antimicrob. Agents Chemother*. 57, 5344–5354, (2013).
[47] Klumpp, K. *et al*. High-resolution crystal structure of a hepatitis B virus replication inhibitor bound to the viral core protein. *Proc. Natl. Acad. Sci. U. S. A*. 112, 15196–15201, (2015).
[48] Qiu, Z. et al. Discovery and Pre-Clinical Characterization of Third-Generation 4-H Heteroaryldihydropyrimidine (HAP) Analogues as Hepatitis B Virus (HBV) Capsid Inhibitors. *J. Med. Chem*. 60, 3352–3371, (2017).

The crystal structures of HBV core protein Y132A mutant hexamer in complex with NVR-010-001-E2 (PDB code 5E0I) and HAP_R01 (PDB code 5WRE) provided understanding on the binding mode of the HAP series (Fig. I.7) and revealed that specific functional groups at the 6-position form additional interactions with the capsid protein (Fig. I.8), contributing to the increased binding affinity and improved anti-HBV activity compared to BAY 41-4109.

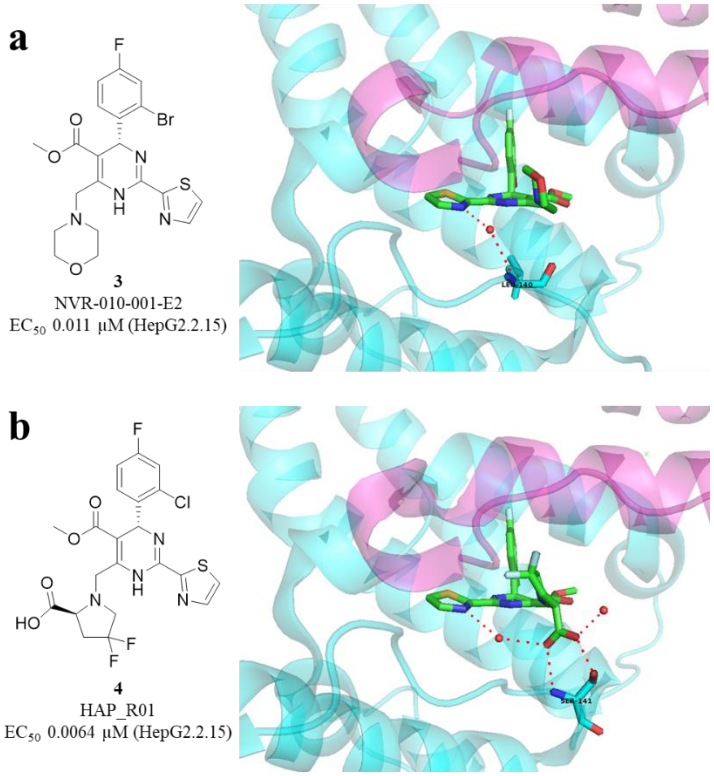


Figure I.8: Comparing the binding mode of NVR-010-001-E2 (3) and HAP_R01 (4) at the solvent-exposed area. The B-C dimer-dimer interface is shown as an example of the compounds binding pocket. Chains B and C are shown in cyan and pink cartoon, respectively. Carbon atoms of compounds **3** and **4** are highlighted in green sticks. Hydrogen bond interactions are indicated by red dash lines. Water molecules are represented as red spheres. **a)** The C6 morpholinyl group is responsible for the increased potency and solubility of the second generation HAPs (compounds **3** and **4**) compared to BAY 41-4109 (**2**); however, it does not form any H-bond interaction at the solvent-exposed area. (PDB ID 5E0I). **b)** The replacement of the C6 morpholinyl with a gem-difluoropyrrolidinyl moiety and the introduction of an additional carboxyl group, typical of the third generation HAPs, increase the binding affinity and therefore the antiviral activity of **4** compared to **3**. A network of H-bond interactions observed between the carboxyl group of **4**, Ser141 (highlighted in sticks) and a pair of water molecules at the solvent-exposed area could contribute to this effect. (PDB ID 5WRE).

In addition, more recent studies in the optimization of the HAP-series molecules have focused on the introduction of a methyl group at the 4-position of the HAP core to block aromatization to the corresponding pyrimidine, resulting in the identification of a novel class of 4-methyl HAPs with improved physicochemical properties. Among them, a (2S,4S)-4,4-difluoroproline substituted analogue **5** exhibited an EC_{50} 0.084 µM in HepDE19 cell line and a good plasma stability, oral bioavailability and liver exposure in mice.

A high-resolution structure of **5** and Cp149Y132A clarified the pattern of ligand-protein interactions of 4-Me-HAPs (Fig. I.9) and, in particular, the structural effect of the inclusion of the C6 gem-difluoropyrrolidinyl group (PDB code 5GMZ). [49]

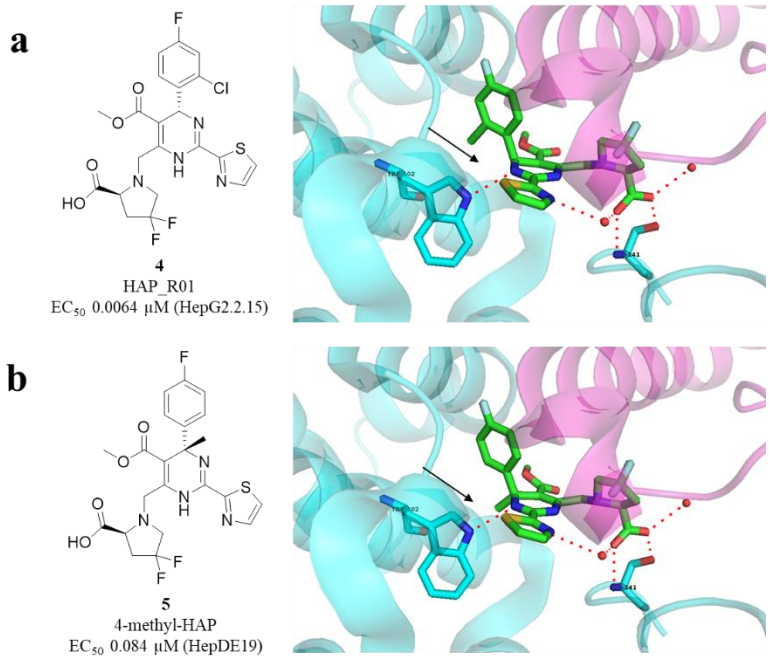
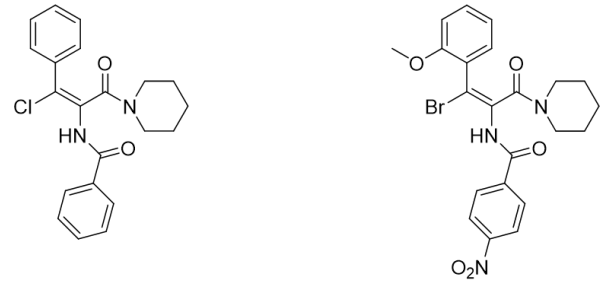


Figure I.9: Comparing the binding mode of HAP_R01 (4) and 4-methyl-HAP (5). Effect of the introduction of a methyl group at the 4 position of the central HAP core. The B-C dimer-dimer interface is shown as an example of the compounds binding pocket. Chains B and C are shown in cyan and pink cartoon, respectively. Carbon atoms of compounds **4** and **5** are highlighted in green sticks. Key residues Trp102 and Ser141 are highlighted in sticks. Hydrogen bond interactions are indicated by red dash lines. Water molecules are represented as red spheres. **a)** The aromatization of the central dihydropyrimidine core of **4** to a pyrimidine, commonly observed for all HAPs, is proved to be a major metabolic pathway. (PDB ID 5WRE). **b)** The introduction of a 4-methyl group on the central HAP core improves the metabolic stability of this series of compounds. Structurally, the presence of the 4-methyl group (indicated with a black arrow) does not alter the conformation of **5** at the binding pocket, which preserves the same pattern of protein-ligand interactions of its HAP analogue (**4**). (PDB ID 5GMZ).

[49] Qiu, Z. *et al*. Design and synthesis of orally bioavailable 4-methyl heteroaryldihydropyrimidine based hepatitis B virus (HBV) capsid inhibitors. *J. Med. Chem.* 59, 7651–7666, (2016).

I.6.2. Non-HAP-like CAMs

Since the discovery of the phenylpropenamide (PPAs) derivatives AT-61 and AT-130 in 1998 (Fig. I.10), [50], several chemical classes, structurally different from HAPs, have been described as forming normal size capsids that lack genetic material. To date, phenylpropenamide (PPAs) and sulfamoylbenzamides (SBAs) are the most studied scaffolds showing a non-HAP-like mode of action.



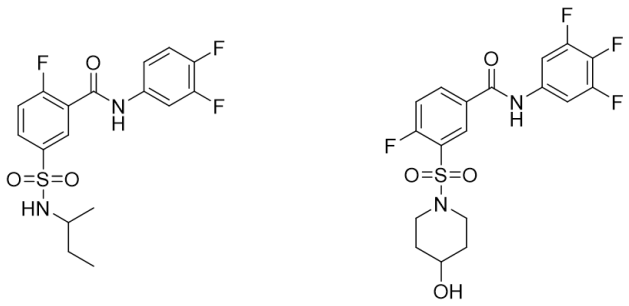
6
AT-61
EC_{50} 1.9 µM (HepAD38)

7
AT-130
EC_{50} 0.13 µM (HepAD38)

Figure I.10: Representative phenylpropenamide (PPA) series of HBV CAMs. EC_{50} values are for inhibition of HBV DNA replication *in vitro*.

AT-61 (**6**) and its optimized analogue AT-130 (**7**) were observed to be highly potent inhibitors of HBV DNA replication in HepAD38 cell line with EC_{50} values in the low micromolar range and to be effective against lamivudine-resistant strains of HBV *in vitro*. [51,52] Unlike HAP modulators, PPAs do not misdirect assembly; rather, they accelerate the rate of capsid assembly and produce morphologically intact but empty particles. [53]

[50] King, R. W. *et al*. Inhibition of human hepatitis B virus replication by AT-61, a phenylpropenamide derivative, alone and in combination with (-) β-L- 2',3',dideoxy-3'-thiacytidine. *Antimicrob. Agents Chemother*. 42, 3179–3186 (1998).
[51] Perni, R. B. *et al*. Phenylpropenamide derivatives as inhibitors of hepatitis B virus replication. *Bioorganic Med. Chem. Lett*. 10, 2687–2690, (2000).
[52] Delaney IV, W. E. *et al*. Phenylpropenamide derivatives AT-61 and AT-130 inhibit replication of wild-type and lamivudine-resistant strains of hepatitis B virus in vitro. *Antimicrob. Agents Chemother*. 46, 3057–3060, (2002).
[53] Katen, S. P., Chirapu, S. R., Finn, M. G. & Zlotnick, A. Trapping of Hepatitis B Virus Capsid Assembly Intermediates by Phenylpropenamide Assembly Accelerators. *ACS Chem. Biol*. 5, 1125–1136, (2010).



8
DVR-23
EC_{50} 0.1 µM (HepDES19)

9
NVR 3-778
EC_{50} 0.26 µM (HepG2.2.15)

Figure I.11: Representative sulfamoylbenzamide (SBA) series of HBV CAMs. EC_{50} values are for inhibition of HBV DNA replication *in vitro*.

Similar to the PPAs, sulfamoylbenzamides (SBAs, Fig. I.11), which contain both an amide and sulfonamide on the core benzene ring, inhibit HBV replication to form genome-free capsids. The first SBA compound, DVR-23 (**8**), was reported in 2013, [54], and since then, various SAR exploration strategies were adopted in the attempt to optimize this promising CAM series. Notably, few SBA molecules, including NVR 3-778 (**9**, EC_{50} 0.26 µM in HepG2.2.15), are currently under clinical development as oral antiviral agents. [22]

Only recently, the X-ray structure of HBV core protein Y132A mutant hexamer in complex with NVR 3-778 was published (PDB code 5T2P, Fig. I.12). [55] Interestingly, when compared to Cp149Y132A in complex with 4-Me-HAP **5** (PDB code 5GMZ), both HAP and non-HAP-like CAMs show similar binding modes at the HBcAg dimer-dimer interface. However, the striking difference lies in the occupancy of a deep hydrophobic subpocket, which is occupied by the C2 thiazoyl group of the HAP but is left unperturbed by NVR 3-778 (Fig. I.12c). [56]

[54] Campagna, M. R. *et al*. Sulfamoylbenzamide Derivatives Inhibit the Assembly of Hepatitis B Virus Nucleocapsids. *J. Virol*. 87, 6931–6942, (2013).
[55] Zhou, Z. *et al*. Heteroaryldihydropyrimidine (HAP) and Sulfamoylbenzamide (SBA) Inhibit Hepatitis B Virus Replication by Different Molecular Mechanisms. *Sci. Rep*. 7, (2017).
[56] Na, H. G. *et al*. Discovery of a New Sulfonamide Hepatitis B Capsid Assembly Modulator. *ACS Med. Chem. Lett.* 11, 166–171, (2020).

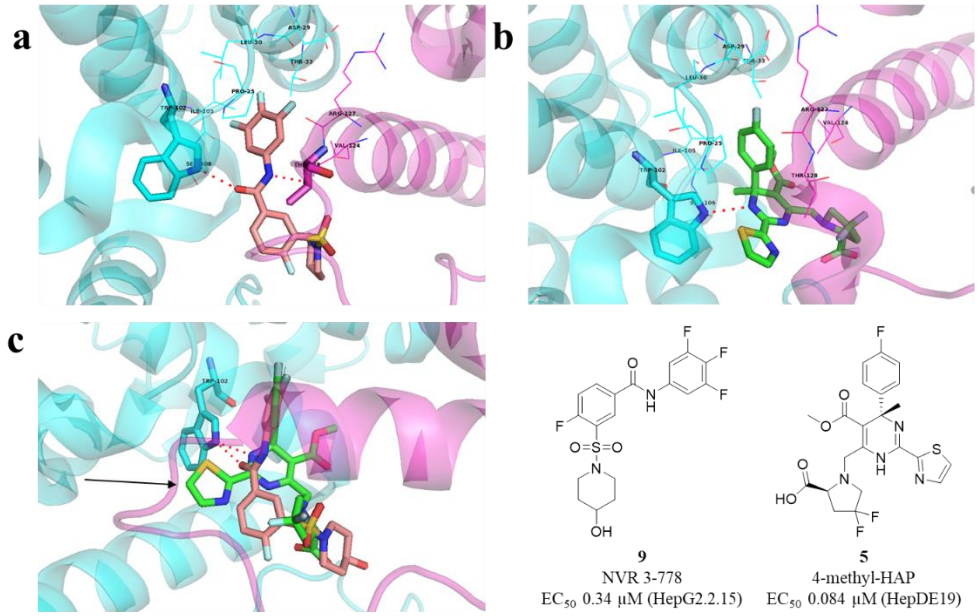


Figure I.12: Comparing the binding mode of HAP and non-HAP-like CAMs: 4-methyl-HAP (5) and NVR 3-778 (9). The B-C dimer-dimer interface is shown as an example of the compounds binding pocket. Chains B and C are shown in cyan and pink cartoon, respectively. Carbon atoms of compounds **9** and **5** are highlighted in pink and green sticks, respectively. Key residues on both chains are displayed in lines and labelled. Trp102 and Thr128 are highlighted in sticks and labelled. Hydrogen bond interactions are indicated by red dash lines. **a)** and **b)** Both the trifluoroaniline of **9** and the 4-fluorophenyl of **5** insert into an halogen-binding subpocket formed by Pro23, Asp29, Leu30, Thr33, Trp102, Ile105 and Ser106 from chain B and Val124, Arg127 and Thr128 from chain C. Both compounds form an H-bond interaction with Trp102. An additional H-bond is observed between **9** and Thr128. **c)** When superimposed, **9** and **5** share a similar binding mode; however, a deep hydrophobic subpocket occupied by the 2nd-positioned thiazole ring of **5** (indicated by a black arrow) is left completely unperturbed by **9**. (PDB IDs 5T2P, 5GMZ).

I.6.3 CAMs binding mode

HAP and non-HAP-like CAMs bind to the same highly hydrophobic pocket at the interface between two adjacent HBcAg dimers. Both classes of compounds generally form a key hydrogen bond interaction to Trp102 and present a halogenated aromatic ring towards a deep hydrophobic sub-pocket formed by Pro25, Asp29, Leu30, Thr33, Trp102, Ile105 and Ser106 from one monomer and Val124, Arg127 and Thr128 from the other monomer. Many compounds also have an H-bond donor paired with the side-chain of Thr128 (Fig. I.12a and I.12b). An H-bonding interaction between ligand and Ser141, either directly or mediated via a water molecule is also frequently observed (Fig. I.8b).

Notably, point mutations of specific amino acid residues at core protein dimer-dimer interface may interfere with the process of capsid assembly and affect the binding and activity of both classes of CAMs. [57,58,59]

I.6.4 CAMs effects on the maintenance of cccDNA

In addition to having an effect on late steps of the viral replication cycle, CAMs can also interfere with *de novo* cccDNA formation. [60] These compounds may induce capsid stabilization and prevent the release of rcDNA to the nucleus and its conversion to cccDNA, thus avoiding the establishment of the cccDNA pool. [61]

CAMs can also distort and disrupt pre-formed capsids. [62] During their entry or after their release into the cytoplasm, CAMs may interact with nucleocapsids in virions and induce premature disassembly. [63] Viral DNA may be released and exposed to cytoplasmatic DNAases, thereby abrogating the subsequent cccDNA synthesis in the nucleus. [64]

As an additional mechanism of action, CAMs may also affect HBcAg nuclear functions: in particular, the ability to bind to the cccDNA mini-chromosome and act as a positive epigenetic regulator of both HBV transcription and replication. [65,66]

[57] Wu, S. *et al*. Discovery and Mechanistic Study of Benzamide Derivatives That Modulate Hepatitis B Virus Capsid Assembly. *J. Virol*. 91, (2017).
[58] Ruan, L., Hadden, J. A. & Zlotnick, A. Assembly Properties of Hepatitis B Virus Core Protein Mutants Correlate with Their Resistance to Assembly-Directed Antivirals. *J. Virol*. 92, (2018).
[59] Böttcher, B. & Nassal, M. Structure of Mutant Hepatitis B Core Protein Capsids with Premature Secretion Phenotype. *J. Mol. Biol*. 430, 4941–4954, (2018).
[60] Lahlali, T. *et al*. Novel Potent Capsid Assembly Modulators Regulate Multiple Steps of the Hepatitis B Virus Life Cycle, *Antimicrobial Agents and Chemotherapy,* 62 (10) e00835-18, (2018).
[61] Berke, J. M. *et al*. Capsid Assembly Modulators Have a Dual Mechanism of Action in Primary Human Hepatocytes Infected with Hepatitis B Virus. *Antimicrobial Agents and Chemotherapy,* 61 (8), e00560-17, (2017).
[62] Kirkegaard, K. *et al*. Hepatitis B virus core protein allosteric modulators can distort and disrupt intact capsids, *Elife*, 7:e31473, (2018).
[63] Qazi, S., Schlicksup, C. J., Rittichier, J., Vannieuwenhze, M. S. & Zlotnick, A. An Assembly-Activating Site in the Hepatitis B Virus Capsid Protein Can Also Trigger Disassembly. *ACS Chem. Biol*. 13, 2114–2120, (2018).
[64] Guo, F. *et al*. HBV core protein allosteric modulators differentially alter cccDNA biosynthesis from de novo infection and intracellular amplification pathways. *PLOS Pathog.* 13, e1006658, (2017).
[65] Guo, Y.-H., Li, Y.-N., Zhao, J.-R., Zhang, J. & Yan, Z. HBc binds to the CpG islands of HBV cccDNA and promotes an epigenetic permissive state. *Epigenetics* 6, 720–726, (2011).
[66] Tu, J., Li, J. J., Shan, Z. J. & Zhai, H. L. Exploring the binding mechanism of Heteroaryldihydropyrimidines and Hepatitis B Virus capsid combined 3D-QSAR and molecular dynamics. *Antiviral Res.* 137, 151–164, (2017).

UNIVERSITAT ROVIRA I VIRGILI
TARGET-BASED DESIGN, STRUCTURAL OPTIMIZATION AND CHARACTERIZATION OF NOVEL HEPATITIS B VIRUS CAPSID ASSEMBLY MODULATORS
Elena Detta

Objectives

The overall goal of the VIRO-FLOW project was the fast and efficient identification of new curative agents targeting HBV core protein, integrating the advantages of continuous flow chemistry with *in vitro* microfluidic bioassay technologies. The end objective of the present thesis was the creation of an integrated system for the generation of structure-activity relationship (SAR) data, where a suite of computational techniques could be utilized to speed up the process of hit/lead identification, to optimize out undesirable pharmacological properties of known actives and to guide and support the synthesis of compound libraries either in batch or flow.

The approach developed in this research work aimed at discovering novel small molecules that misdirect HBV capsid assembly through the application of state-of-the-art computer-aided design methodologies, including virtual screening, scaffold hopping, *in silico* focussed library design and molecular docking studies of potential active compounds.

The objective of the study reported in the second chapter of this thesis was the virtual design of two target-focused libraries of novel HBV CAMs starting from a known active oxalyl-piperazine compound. A key design approach of the virtual libraries was the bioisosteric replacement of the central core of the reference molecule. A water prediction program, WaterFLAP, was used with the aim of identifying specific waters that could be targeted for displacement to optimize the interactions of oxalyl-piperazine ligands with the core protein and to achieve a gain in potency.

A highly active Class I CAM was selected as a template in a computational scaffold hopping study depicted in the third chapter. Recore software was used to replace the indole ring of the given 3D query with the goal of discovering novel molecular structures active against HBV with enhanced potency and generating new ideas for potentially patentable CAMs.

In the fourth chapter, a target-based *in silico* study is reported, aimed towards the identification of new HBV CAMs, with a secondary objective of discovering novel non-HAP-like Class II capsid modulators. Based on recently reported literature, it was hypothesized that the occupancy of the HAP-thiazole subpocket could be the basis for the distinct mode of action of HAP-like molecules. A large database of commercially available compounds was

screened in a computational model based on a HAP-like compound with the goal of retrieving chemical scaffolds that could potentially access this hydrophobic pocket.

The main aim of the study described in the fifth chapter was the development of a more efficient method for the synthesis of [1,2,4]triazolo[1,5-*a*]pyridine-2-carboxylate derivatives in continuous flow with the aid of DFT calculations. Computational studies carried out in this research work aimed at clarifying the mechanism of ethyl [1,2,4]triazolo[1,5-*a*]pyridine-2-carboxylate 3-oxide formation and providing rational pathways to the generation of the by-products experimentally characterized for this reaction.

The interest in the synthesis of the [1,2,4]triazolo[1,5-*a*]pyridine core was based on the rationale that this scaffold could be used as potential bioisosteric replacement of specific N-containing bicyclic rings of highly active HBV CAMs.

Chapter II.

Virtual chemistry: *in silico* design of novel focused libraries active against HBV.

Chapter III.

Computer-aided scaffold hopping to identify novel Hepatitis B Virus Capsid Assembly Modulators.

Chapter IV.

Virtual screening strategies for the identification of novel hepatitis B virus capsid assembly modulators.

Chapter V.

Mechanistic studies and optimization of the synthesis of [1,2,4]triazolo[1,5-*a*]pyridine-2-carboxylate-3-oxide in flow.

General conclusions and future work

The combination of ligand- and target-based design strategies adopted in the research work presented in this thesis led to the identification of novel libraries of HBV CAMs, two of which have recently made subject of patent applications.

As reported in detail in the second chapter, two virtual collections of oxalamide-containing molecules were generated by scaffold replacement methods.

A complementary computational technique was used to predict binding site water molecules, leading to the identification of an energetically unfavourable water in close proximity to the central core of the reference oxalyl-piperazine molecule.

In medicinal chemistry and drug design, computer-aided scaffold hopping refers to the search for molecular structures that share similar activity or properties, but contain different chemical scaffolds. ReCore, a fast and versatile scaffold hopping software, was used in this research work in order to identify potential replacements to the indole ring of a known active Class I HBV CAM. As described in the third chapter of this thesis, the indole ring of the given query was substituted with a bioisosteric core, while all other anchoring points were kept in place. Overall, two novel series of HBV CAMs were discovered with *in vitro* potency in the sub-micromolar range. The outcomes showed to be a good proof of validation of the process of ReCore index creation and the scaffold replacement strategy adopted.

As discussed in the introductory chapter of this thesis, all HBV CAMs share the same binding pocket at the dimer-dimer interface of the core protein and have been traditionally categorized into two classes according to their mode of action. Class I or non-HAP-like CAMs enhance the formation of capsid-like spherical particles devoid of the viral genome, while Class II CAMs induce aberrant capsid structures. When superimposing Class I and II representative molecules at the HBV core protein interface, both share similar binding modes. However, a striking difference lies in the occupancy of a highly hydrophobic subpocket that is filled by the thiazole moiety of HAP-like compounds but is left unperturbed by Class I modulators. It was suggested in this study that the occupancy of the HAP-thiazole subpocket was the basis for the distinct mode

of action of HAP-like molecules. The recent discovery of a new SBA compound with a mechanism of action similar to BAY 41-4109, the first example of the HAP series discovered by Bayer scientists in late '90s, and a specific amino group able to partially access the HAP-thiazole subpocket seemed to support this theory. As depicted in the fourth chapter of this thesis, an extensive target-based *in silico* screening was performed allowing the identification of novel small molecules that misdirect HBV capsid formation. A lead compound exhibited *in vitro* potency in the sub-micromolar range along with good physico-chemical and toxicological profile *in vitro* and was evaluated in molecular docking and mechanism of action studies. Results from these studies suggested that the tested scaffolds can be categorized as non-HAP like or Class I capsid assembly modulators.

In the attempt to further improve its *in vitro* antiviral activity, PK and safety profile, the lead compound was selected as a template to design a diversified collection of derivatives. This SAR exploration study resulted in the identification of two analogues with a 2- fold increase in potency compared to the reference.

Most of the compounds described in the "lead optimization" section of this chapter are still under biological assessment. Future results will surely contribute to the creation of a more detailed SAR map of the series and eventually lead to new interesting conclusions.

In the attempt to discover a novel chemical series of HBV CAMs, in this research work it was envisaged that the [1,2,4]triazolo[1,5-*a*]pyridine core could be used as a potential bioisosteric replacement of *N*-rich bicyclic moieties of known active anti-HBV agents. As described in the fifth chapter of this thesis, an efficient method for the synthesis of the [1,2,4]triazolo[1,5-*a*]pyridine-2-carboxamide scaffold was developed. The second step of the synthesis, yielding ethyl [1,2,4]triazolo[1,5-*a*]pyridine-2-carboxylate 3-oxide was implemented in continuous flow, in order to minimize the formation of three identified side products. While the reaction time was significantly reduced from 3 hours to 3.5 minutes by transferring the process in flow, the optimization process did not result in any improvement of the desired product yield. DFT calculations were

carried out to give a rational explanation to the mechanism of formation of ethyl [1,2,4]triazolo[1,5-*a*]pyridine-2-carboxylate 3-oxide and the other experimentally characterized by-products. Overall, from the combination of experimental and theoretical insights, it was suggested that further optimization of the synthesis of the triazolopyridine precursors could be very challenging, since pathways leading to side products were determined to be kinetically very competitive with the desired product.

UNIVERSITAT ROVIRA I VIRGILI
TARGET-BASED DESIGN, STRUCTURAL OPTIMIZATION AND CHARACTERIZATION OF NOVEL HEPATITIS B VIRUS CAPSID ASSEMBLY MODULATORS
Elena Detta

UNIVERSITAT ROVIRA I VIRGILI
TARGET-BASED DESIGN, STRUCTURAL OPTIMIZATION AND CHARACTERIZATION OF NOVEL HEPATITIS B VIRUS CAPSID ASSEMBLY MODULATORS
Elena Detta

UNIVERSITAT ROVIRA I VIRGILI
TARGET-BASED DESIGN, STRUCTURAL OPTIMIZATION AND CHARACTERIZATION OF NOVEL HEPATITIS B VIRUS CAPSID ASSEMBLY MODULATORS
Elena Detta

UNIVERSITAT ROVIRA I VIRGILI
TARGET-BASED DESIGN, STRUCTURAL OPTIMIZATION AND CHARACTERIZATION OF NOVEL HEPATITIS B VIRUS CAPSID ASSEMBLY MODULATORS
Elena Detta