Interactions of HBV capsid involved in nuclear transport

Lara Gallucci

université de BORDEAUX

THÈSE PRESENTÉE POUR OBTENIR LE GRADE DE:

DOCTEUR DE L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE: des Sciences de la Vie et de la Santé

SPECIALITE: microbiologie et immunologie

Par Lara Gallucci

Etude des interactions de la capside du VHB impliquées dans le transport nucléaire

Sous la direction de Michael Kann

Soutenue le 25 Octobre 2018

Membres du jury:

Pr Thierry Noel	Université de Bordeaux	President
Dr Gualtiero Alvisi	Università di Padova	Rapporteur
Pr Beate Sodeik	Hannover Medical School	Rapporteur
Dr Anne Royou	Université de Bordeaux	Examinateur
Pr Michael Kann	Université de Bordeaux	Superviseur

université de BORDEAUX

DISSERTATION FOR THE AWARD OF THE DEGREE:

DOCTORATE OF THE UNIVERSITY OF BORDEAUX

Doctoral school of health and life sciences

Specialty: microbiology - immunology

Lara Gallucci

Interactions of HBV capsid involved in nuclear transport

Thesis director: Michael Kann

Presented on 25th October 2018

Members of the committee:

Pr Thierry Noel	Université de Bordeaux	President
Dr Gualtiero Alvisi	Università di Padova	Reporter
Pr Beate Sodeik	Hannover Medical School	Reporter
Dr Anne Royou	Université de Bordeaux	Examiner
Pr Michael Kann	Université de Bordeaux	Supervisor

Etude des interactions de la capside du VHB impliquées dans le transport nucléaire

Le virus de l'hépatite B (VHB) est un virus enveloppé composé d'un ADN partiellement double brin (ADNrc) contenu dans une capside icosahédrique. Le VHB est responsable d'infections aiguës et chroniques. VHB est non cytopathique mais l'inflammation chronique entraîne une fibrose hépatique, une cirrhose et un carcinome hépatocellulaire. Le VHB se réplique via un intermédiaire à ARN. La transcription nécessite que l'ADNrc soit convertit en un ADN circulaire clos de manière covalente (ADNccc). Cet ADNccc sert de matrice pour la transcription de l'ARN prégénomique (ARNpg), qui est spécifiquement encapsidé grâce aux interactions entre la polymérase virale, l'ARNpg et la protéine core (Cp) qui forme la capside. La polymérase rétrotranscrit l'ARNpg en ADN monocaténaire puis en ADNrc, conduisant à des matrices de capside matures (MatC). Cp avec 185 aa contient un domaine N-terminal structuré, et un domaine C-terminal (CTD) flexible. Le CTD comprend deux signaux de localisation nucléaire (NLS) et un domaine de liaison avec l'importin β (IBB). Le CTD est orienté vers l'intérieur de la capside de part son interaction avec les acides nucléiques simples brins tandis qu'il est exposé vers l'extérieur dans les capsides vides (EmpC) et les MatC. De plus Cp étant surexprimée, cela conduit à l'assemblage des EmpC. Le VHB doit délivrer son génome dans le noyau des cellules infectées pour sa réplication. Le transport nucléaire est médié par la capside qui interagit avec les récepteurs d'import. L'équipe a démontré préalablement que ce transport a besoin des récepteurs Importin α (Imp.α) et Importin β (Imp.β) en induisant le transport des capsides au panier nucléaire où elle est stoppée par l'interaction avec la nucléoporine 153 (Nup153).

Nous avons démontré que l'Imp.α, mais pas l'Imp.β, se lie aux MatC suggérant que seule la partie du CTD qui contient les NLS est exposée à l'extérieur des MatC. En comparaison, nous avons analysé les EmpC en collaboration avec Adam Zlotnick (Université d'Indiana, États-Unis) et démontré que les EmpC sont capables de lier directement l'Imp.β. Cette interaction qui est plus forte que l'interaction avec l'Imp.α s'effectue via la reconnaissance du domaine IBB du CTD, ce qui implique une exposition totale du CTD à l'extérieur de la capside. Nous avons aussi montré que la liaison avec l'Imp.β à des concentrations très élevées fournit des forces de déstabilisation menant au désassemblage des EmpC.

La libération du génome dans le panier nucléaire implique que l'interaction entre les MatC et Nup153 participe au désassemblage de la capside. Afin de valider cette hypothèse, nous avons exposé des MatC dont le génome est radiomarqué avec un fragment de Nup153 contenant le domaine clé, montré pour interagir avec la capside, en présence de nucléases. Nous avons mis

en évidence qu'en présence de ce fragment, les MatC restent stables. Cela suggère la nécessité de facteurs cellulaires additionnels pour le désassemblage des MatC. Cette conclusion est conforme avec nos résultats sur noyaux isolés, dans lesquels nous avons observé une localisation nucléaire des capsides laissant supposer que les facteurs cellulaires nécessaires au désassemblage des MatC sont nucléaires.

Afin d'étudier plus en détail l'étape de désassemblage et la libération du génome viral, nous avons mis au point un système permettant de suivre en temps réel le devenir du génome viral.

Ce système est basé sur l'interaction coopérative d'une protéine (OR) fusionnée à la DGFP avec une séquence ADN double brin cible (ANCH). Ce modèle a permis de visualiser des génomes viraux isolés après infection avec des pseudo-virions contenant la séquence ANCH de lignées hépatiques humaines exprimant stablement le récepteur spécifique du VHB et la protéine OR-DGFP. Cet outil permettra d'effectuer des criblages avec des inhibiteurs afin de mettre en évidence les facteurs cellulaires impliqués dans la libération du génome et éventuellement la conversion de l'ADNrc en ADNccc.

Mots clès :

VHB, transport nucléaire, libération du génome

UMR 5234 CNRS Université de Bordeaux

Microbiologie Fondamentale et Pathogénicité

Interactions of HBV capsid involved in nuclear transport

The Hepatitis B Virus (HBV) is an enveloped virus containing a partially double stranded DNA genome (relaxed circular; rcDNA). HBV causes acute and chronic infections. HBV is not cytotoxic but chronic inflammation leads to liver fibrosis, cirrhosis and hepatocellular carcinoma. HBV replicates via an RNA intermediate, which is transcribed from a covalently closed circular form of the viral DNA (cccDNA). This pregenomic RNA is specifically encapsidated into the capsid by interaction with the viral polymerase, which also interacts with the core protein (Cp), forming the capsid. The polymerase retrotranscribes the pregenomic RNA into single stranded DNA and subsequently into partially double stranded DNA, which is found in mature capsids (MatC). Cp is a 185 aa long polypeptide comprising a N-terminal assembly domain, and a flexible C-terminal domain (CTD). The CTD includes two overlapping nuclear localization signals (NLS) of eight aa and an Importin β Binding Domain (IBB) of 34 aa. The CTD is fixed in the interior of the capsid by interacting with single stranded nucleic acids but translocates to the exterior in MatC and empty capsids (EmpC). Cp is over expressed leading to assembly of EmpC. HBV has to deliver its genome into the nucleus of infected cells for replication. Nuclear transport is mediated by the capsid that interacts with nuclear import receptors. The group has shown that MatC needs Importin α (Imp.α) and Importin β (Imp.β) for transport of the capsids into the nuclear basket. In this structure where genome liberation likely occurs, the transport of the capsid is arrested by interaction between the capsid and the nucleoporin Nup153 (Nup153).

In the thesis we demonstrate that MatC binds to Imp.α, but not Imp.β, suggesting that only the part of the CTD, which contains the NLSs is exposed on capsids' surface. In collaboration with the Adam Zlotnick's group (Indiana University, U.S.A.) we showed that EmpC, in contrast, bind Imp.β directly without Imp.α acting as an adaptor. This interaction, which is stronger than the one of Imp.α, needs IBB exposure, meaning that the entire CTD becomes externalized. Furthermore, exposure to very high Imp.β concentration led to EmpC destabilization.

The genome release within the nuclear basket implies that Nup153 is involved in genome liberation from MatC. To verify this hypothesis we used MatC with a radioactively labeled genome, which we exposed to the capsid binding-Nup153 fragment. Investigating the accessibility of the genome to nucleases we found that the Nup153 fragment had no impact on capsids stability, suggesting the need of other cellular factors driving disassembly. This conclusion is in agreement with our observation that MatC added to isolated nuclei resulted in nuclear capsid entry, which requires disassembly.

To further study the disassembly step and the consequent release of the viral genome, we developed a system to directly visualize the viral genome allowing investigations of genome uncoating in real time. The system is based on the cooperative binding of a fluorescent fusion protein between the bacterial protein OR with DGFP to a double stranded DNA sequence called Anch. Using this model we showed that infection of OR-DGFP-expressing hepatoma cells with HBV containing a modified Anch genome allowed monitoring genome release into the nucleus. In future, this system may help identifying factors involved in genome release and repair and to decipher their molecular interactions.

Keywords :
HBV, nuclear transport, genome release

UMR 5234 CNRS Université de Bordeaux
Microbiologie Fondamentale et Pathogénicité

Le virus de l'hépatite B (VHB) est un agent pathogène humain majeur avec un ADN partiellement double brin. Le VHB provoque des infections aiguës et chroniques. Le virus est non cytopathique mais la détection immunitaire des hépatocytes infectés par les lymphocytes T CD8 + entraîne une fibrose hépatique, une cirrhose et un carcinome hépatocellulaire suite à une infection chronique.

Comme d'autres virus à ADN et rétrovirus, le VHB doit délivrer son génome dans le noyau des cellules infectées pour sa réplication. L'infection avec le VHB est très efficace, comme le montrent les expériences effectuées chez les chimpanzés dans lesquelles presque chaque particule virale était infectieuse. Cela signifie que le transport intracellulaire du génome, médié par la capside, doit être bien coordonné et que les interactions de la capside virale avec les protéines cellulaires sont cruciales pour cette étape.

Le VHB se réplique par transcription inverse de l'ADN viral nucléaire seulement après conversion du génome partiellement double brin (ADNrc) sous une forme circulaire fermée de manière covalente (ADNccc). Cet ADNccc sert alors de matrice pour la transcription de l'ARN prégénomique (ARNpg), spécifiquement encapsidé grâce aux interactions entre la polymérase virale attachée en 5' de l'ARNpg et la protéine core (Cp) qui forme la capside. La polymérase rétrotranscrit l'ARN prégénomique en ADN monocaténaire (-) encore immature puis en ADN partiellement bicaténaire (ADNrc), conduisant à des matrices de capside matures MatC.

Cp est un polypeptide de 185 aa (183 aa dans certains génotypes) comprenant un domaine N-terminal hautement structuré, responsable de l'assemblage de la capside, et un domaine C-terminal (CTD) flexible. Le CTD comprend un signal de localisation nucléaire (NLS) de 8 aa et un domaine de liaison bêta avec l'Importine β (IBB) de 34 aa. Le domaine CTD est orienté vers l'intérieur de la capside en interagissant avec les acides nucléiques simple brin tandis qu'il est exposé vers l'extérieur dans les capsides vides et matures. Il convient de noter que Cp est surexprimée, conduisant à l'assemblage de capsides vides (EmpC).

Au cours du processus d'infection, après libération de la vésicule endosomale, la capside doit être transportée vers la périphérie nucléaire. Le transport nucléaire est médié par l'interaction avec les récepteurs d'import, Importine α (Imp.α) et Importine β (Imp.β). Cela se produit par la reconnaissance du NLS présent sur les Cp par l'Imp.α qui agit comme un adaptateur pour l'Imp.β.

Contrairement aux protéines cargos caryophiles, le transport nucléaire de la capside du VHB se termine dans le panier nucléaire, où elle est stoppée par l'interaction avec la nucléoporine

153 (Nup153), qui n'implique pas le CTD de Cp. Des résultats indirects suggèrent que la libération du génome se produit dans le panier nucléaire.

En collaboration avec le groupe du Prof. Adam Zlotnick (Université d'Indiana, États-Unis), nous avons montré que l'EmpC est capable de lier directement l'Imp.β sans que l'Imp.α ne serve d'adaptateur. Cette interaction s'effectue par la reconnaissance du domaine IBB présent sur le CTD des Cp des EmpC qui est alors complètement exposé à l'extérieur. Ce résultat diffère de ce qui est observé pour les MatC pour lesquelles la liaison entre les protéines de capsides et les récepteurs d'imports où seul le NLS est exposé.

Nous avons en outre montré que la liaison avec l'Imp.β fournit des forces de déstabilisation agissant sur la capside, conduisant au désassemblage des EmpC. Ces capsides étant dépourvues d'acides nucléiques, elles sont moins stables. De plus, il a été démontré que les dimères Cp ou Cp désassemblés sont dégradés par le protéasome. Nous avons supposé que les fragments protéolytiques entrent dans la voie du CMH de classe I menant à une exposition à la surface cellulaire où ils sont les ligands de la réponse des cellules T CD8 +. Cela peut expliquer que les patients infectés par le VHB et dont les cellules hépatocytaires possèdent majoritairement des capsides dans le cytoplasme, présentent une forte inflammation du foie contrairement aux patients avec des capsides principalement nucléaires. De manière logique, nous avons conclu que les EmpC liées à l'Imp.β sont transportées jusqu'au noyau en évitant une reconnaissance par le système immunitaire.

Nous avons également évalué le rôle du Nup153 dans le désassemblage de la capside par le biais d'une analyse de liaison *in vitro*. Pour cela, nous avons utilisé un fragment de Nup153 étiqueté par Histidine: le fragment C2 contenant une partie du domaine précédemment montré pour interagir avec la capside (HisC2-Nup153). Pour étudier cette étape, nous avons exploité l'activité de la polymérase virale endogène pour marquer le génome viral avec α^{32}P dCTP. Nous avons ensuite exposé ces capsides radiomarquées au fragment HisC2-Nup153, soit en solution, soit fixé sur des billes de Ni / plaques de Ni; ce dernier imitant la fixation de Nup153 dans le panier nucléaire. Afin de déterminer l'exposition au génome viral, nous avons ajouté la nucléase S7 pour suivre le profil de digestion. En effet, la nucléase S7 n'a pas accès au génome viral lorsque la capside est intacte, sa taille trop importante ne lui permet pas de passer à travers le maillage de la capside; ce qui n'est pas le cas lorsque la capside est partiellement ou totalement désassemblée. Nous avons cependant observé que les MatC restaient stables, ce qui suggère que des facteurs nucléaires supplémentaires étaient nécessaires.

Afin d'étudier plus en détail l'étape de désassemblage et la libération du génome viral, nous avons mis au point un système permettant de suivre en temps réel le devenir du génome viral. Ce système est basé sur l'interaction coopérative d'une protéine fusionnée à un fluorochrome (OR-DGFP) avec une séquence ADN double brin, la séquence ANCH. Cela permet de suivre des molécules d'ADN isolées en cellules vivantes. Le nouveau modèle que nous avons mis en place repose sur l'infection de lignées hépatiques humaines exprimant stablement la protéine OR-DGFP avec des pseudo-virions contenant la séquence ANCH. .Lorsque le génome viral est libéré et devient donc accessible, la protéine OR-GFP se lie à la séquence ANCH présente sur les pseudo-génomes, entraînant la formation de points fluorescents dans le noyau des cellules infectées. A terme, cet outil permettra d'effectuer des criblages avec des inhibiteurs afin de mettre en évidence les facteurs cellulaires impliqués dans la libération du génome et éventuellement la conversion de l'ADNrc en ADNccc.

Acknowledgement

I would like to start by thanking the members of my thesis committee, Beate Sodeik, Gualtiero Alvisi and Anne Royou who have kindly accepted to evaluate my manuscript as well as Thierry Noel who kindly agreed to be the president of this committee.

I would like to thank Michael Kann, who gave me the opportunity to pursue my PhD under his supervision and guided and helped me during this time.

I especially thank the members of my team. Christian, your constant presence has been a reference to me; Marie-Lise, your help -as well as the scientific discussion we had- has been important to my personal and professional growth. With you I share this work that is the result of a joint effort.

I want to thank Harald Wodrich. You have been of incredible support and your advice have been really meaningful to me.

I thank all the members of the Adeno team: Fabienne, Benoit and Marie-Edith. You are great colleagues; I could always count on you for anything from bench problems to personal ones.

I would like to thank Jessica, Muriel, Cathy, Sandrine and Valerie. With your constant work, in indifferent ways, you make the life in the lab undoubtedly much easier.

My thanks are for all the members of the unit for their help during my everyday life, especially Paul and Jacqueline for the time spent during long purifications and Nicolas Biteau and Nicolas Landrein for the help and support during lab routine.

Finally, my sincerest thanks to my lab mates from the beginning to the end: Charlotte, Irene, Noemie and Floriane. Your help has been invaluable, scientifically and personally.

To all the friends I met in Bordeaux who made me feel at home and that I cannot list because I have been so lucky that they are too many. I am grateful to you more than I can say.

Alle mie amiche di sempre, Giulia, Monica e Federica che ripetutamente sono venute a trovarmi per condivedere con me questa esperienza incredibile.

Alla mia mamma, al mio papà e alle mie tre sorelle Anna, Elena e Teresa che mi mancano sempre e a cui penso cercando di fare del mio meglio. Mi avete sempre lasciato libera di seguire le mie idee e le mie e le mie aspirazioni, a voi va il grazie piu grande e il risultato di questo traguardo.

Table of content

List of tables

List of figures

Introduction

The Hepatitis B Virus

History of discovery

The hepatitis B virus (HBV) was discovered by Baruch Blumberg in 1963, when he reported the presence of a previously unknown protein in the serum of an Australian aborigine, the Au antigen[1]. Later on, different lines of evidence linked the Au antigen with symptoms of viral hepatitis[2,3]. The combination of separation by cesium chloride gradient centrifugation with immune electro staining, made by Dane end colleagues, proved that the Au antigen was present as spherical and tubular subviral particles named Dane's particles[4](**Figure 1).** The Au antigen was later identified as the surface protein of HBV. By the first middle of the '70s the virus was definitely associated to the hepatitis B and called Hepatitis B virus[5,6]. 13 years later Baruch Blumberg was rewarded with the Noble Prize in Medicine.

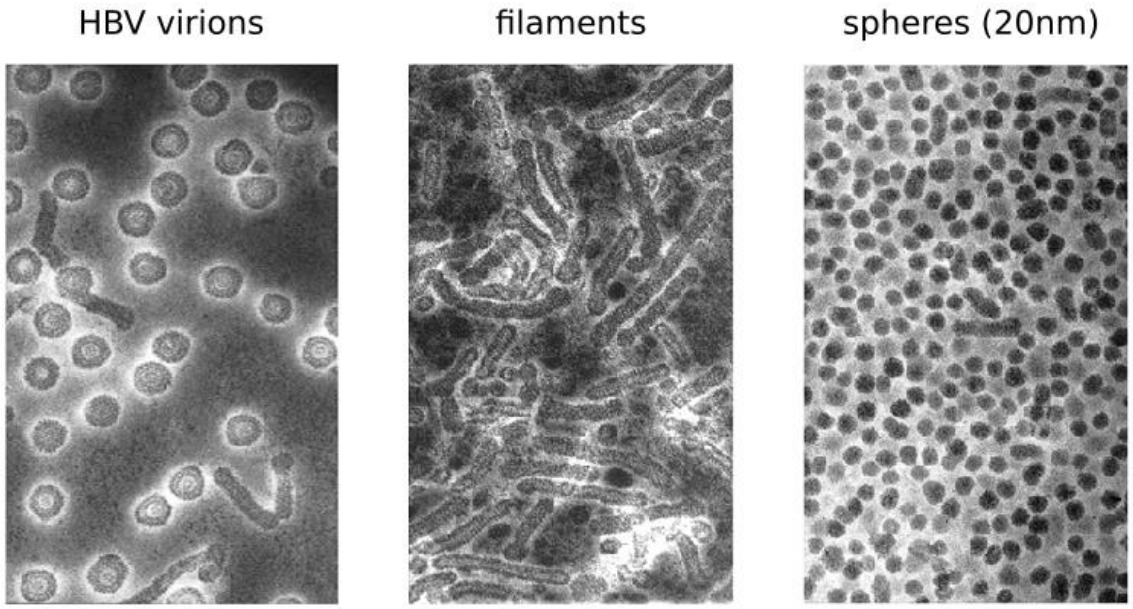


Figure 1: Dane particles
Electron micrographs of HBV particles from Doerr & Gerlich, Medizinische Virologie, 2000, Thieme Verlag. Pictures show the different types of secreted HBV particles: HBV virions on the right panel, and filaments and spherical subviral particles formed by the surface protein in the central and left panels, respectively.

Classification

HBV belongs to the Hepadnaviridae family, a family of viruses that conserve similarity in the organization of the genome and especially in the replication strategy having the unique

characteristic of replicating their DNA genome by an RNA intermediate (reviewed by Saager and Mason[7]).

The Hepadnaviridae family comprises the two genera orthohepadnavirus and avihepadnavirus, infecting mammals and birds respectively[8].

Assignment to Hepadnaviridae family is based on the organization and length of the genome, which is approximately 3 kb, and on the replication strategy by reverse transcription. Within the family the genera are distinguished accordingly to their sequence homology. 40% of divergence is found between orthohepadnavirus and 20% of divergence among avihepadnavirus. Between the two groups there is an almost no sequence homology[8].

Within the two genera different species can be found. The classification of the species is based on the host range. Examples of the orthohepadnavirus genus are the Woolly monkey hepatitis B virus infecting woolly monkeys, the Ground squirrel hepatitis b virus (GSHBV) infecting ground artic and tree squirrels, the Woodchuck hepatitis virus (WHV) infecting woodchucks, and HBV infecting humans, chimpanzees and under certain experimental conditions Tupaia belangeri[9]. An example of avihepadnavirus is Duck hepatitis B virus (DHBV) infecting ducks[10].

HBV is divided into eight genotypes (A to H) with a different geographic distribution and a genomic diversity of 5% to 8%[8]

Clinical features

Overview on HBV infection

It was estimated that 257 million persons are infected with HBV (serum was positive for the presence of the surface antigen) in 2015 resulting in 650,000 death per year[11].

HBV is hepatotropic causing a broad spectrum of liver disease spanning from acute and chronic hepatitis to cirrhosis and hepatocellular carcinoma. The disease is caused by hosts immune response, since HBV infection itself is non-cytopathic[7].

Depending on the patient (age, sex, immunocompetence, etc), HBV infection can be either asymptomatic (two thirds of infected adults) or resulting in acute hepatitis (one third of infected adults)[12,13]. In the majority of the cases the virus is cleared, however in the 5% to 10% of adults the infection becomes chronic[12]. Among the chronic infections, most of them lead to mild liver disease, while in some patients it can progress to cirrhosis and hepatocellular carcinoma (HCC). Reviewed by Seeger and Mason[7].

Immune response to HBV infection

HBV is non-cytopathic and pathogenesis is caused by the host immune response. The infection induces a weak innate immune response, and a strong adaptative immune response, particularly inducing cytotoxic T cells (CD8+ T-cells)[14,15].

In 1991, Bertoletti et al. demonstrated the induction of MHC class I restricted CD8+ T cells specific for a capsid protein epitope during HBV infections in humans. They proposed that this response is characteristic of the acute infection since it is absent in healthy donors[16] and weak in patient affected by chronic hepatitis[17].

In agreement, several other reports showed that acute infection is characterized by a strong T-cell response and high liver injury while chronic infection is associated with a weak T-cell response and only minor liver damage[14,15]. In rare cases (1% of the acute infections), the T-cell response eliminates the entire liver within weeks causing the death of the host (fulminant hepatitis)[13].

Transmission

HBV can be transmitted either horizontally by contact with infected blood or other body fluids such as saliva, menstrual, vaginal, and seminal fluids, or vertically from mother to child during birth. Outside the body HBV can survive at least 7 days[11]. The incubation period depends upon virus intake ranging from 30 to 180 days with an average of 75 days. The detection of the virus is possible within 30 - 60 days after infection[11].

Vaccine and current treatment

A vaccine based on HBV surface proteins is available since 1982[11]. The first vaccine was made from subviral particles purified from the plasma of infected individuals followed by inactivation of the virions. Since 1986 the vaccine contains the major surface protein expressed in *Saccharomyces cerevisiae* becoming the first vaccine made by recombinant DNA technology[18]. The vaccine has been proved to be effective and safe providing immunity in 95 % of infants. The time of protection depends upon the antibody titer, which can last life-long. [11].

So far, there is no an effective cure, but long-life treatments can control the infection. The nucleoside analogue Entecavir and the nucleotide analogue Tenofovir are the mostly used drugs both inhibiting the viral reverse transcriptase[19]. Currently, novel strategies and novel targets are under investigation, targeting viral entry, capsid formation, disassembly, virion formation and the immune system (reviewed by Liang et al[19]).

The virus

HBV is an enveloped virus with a diameter of 42 nm. The envelope encloses the nucleocapsid (NC) which is composed by the proteinaceous capsid and the viral genome. Host cellular proteins such as kinase C (PKC) and the molecular chaperon complex hsp90 are also encapsidated[20,21]. The genome is formed by a partially double stranded DNA. Replication occurs trough reverse transcription mediated by the viral reverse transcriptase (RT). RT is encapsidated inside the nucleocapsid and it is covalently bound to the minus strand DNA (**Figure 2**) (reviewed in[7,22,23]).

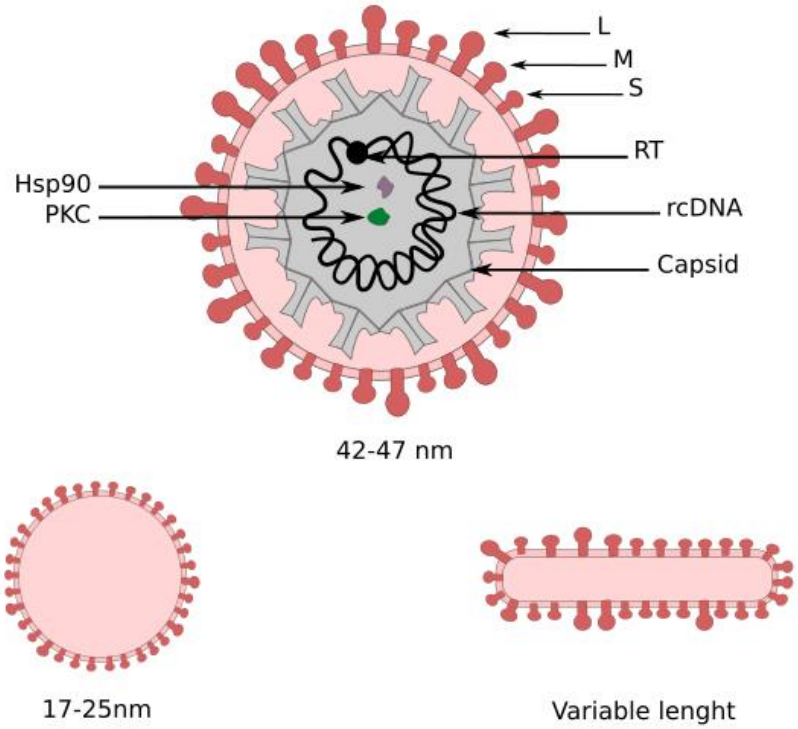


Figure 2: HBV particles
HBV virions with a diameter ranging from 42-47 nm are formed by the envelope, comprising the surface proteins L, M and S, that encloses the nucelocapsid (NC). The NC is composed by the capsid and the viral genome formed by a partially completed relaxed DNA (rcDNA). Reverse transcriptase (RT) is covalently linked to the minus strand DNA Host cellular proteins such as Hsp90 and protein kinase C (PKC) are encapsidated during capsid assembly. Along with completed virions, subviral particles, spheres or filaments, formed by the surface proteins are secreted

The envelope

The viral envelope is composed of tree surface proteins termed Large (L/LHBs), Middle (M/MHBs) and Small (S/SHBs). They are encoded from the same open reading frame (ORF) and they share the same C-terminal domain formed by S (226 aa). The M protein is composed of the S domain and the preS2 domain (55 aa), which is fused to the N terminus of S. The L protein comprises the S and preS2 domain plus the preS1 domain (108 or 119 aa depending on the genotype) at its N terminus[24,25].

The L protein is myristylated at the N-terminus of the PreS1 domain (Gly-2), while M protein is N-glycosylated within the preS2 domain (Asn-4). The preS2 domain of genotypes B-H is O-glycosylated at Thr-37[26].

The preS1 domain is responsible for the envelopment of NC and participates at the binding with the sodium taurocholate cotransporting polypeptide (NTCP) receptor during the entry step. Indeed, L protein exists in two different topologies, that face either the inner or the outside of the virion: in the cytoplasm, during encapsidation, the preS1 domain is directed toward the capsid to direct envelopment, while during the entry step is on the virion surface to mediate the contact with the receptors. [27–29]

Envelope proteins are produced in 100 to 1000 fold excess and form subviral particles: spheres of 22 nm or filaments that differ for their length. Virions and subviral particles contain lipids derived from the post-ER, pre-Golgi compartment[30].

The capsid

HBV capsid (HBcAg) is formed by the capsid or core protein (Cp). Cp can be expressed in *Escherichia coli* (*E. coli*) where it arranges in capsids that are indistinguishable by EM from the ones purified from infected patients[31]. *E. coli* produced capsids, hereafter referred as rHBc, contain *E. coli* RNA with the same length than the physiological HBV pregenomic RNA. Cp arranges either in 180 copies to form capsids with a T=3 symmetry or in 240 copies to form T=4 capsids[32]. The T=4 form is dominant (85%) upon expression in *E. coli* but the T=4/T=3 ratio depends upon the kinetics of assembly[32]. In patients, Cp can also spontaneously assemble to empty capsid (EmpC)[33].

Depending on the genotype, Cp is a 183-185 aa long polypeptide. The first 149 aa form the N-terminal domain (NTD) and the last 34-36 aa form the C-terminal domain (CTD). The NTD in necessary and sufficient for capsid assembly and the mostly basic CTD interacts with the nucleic acids. Expression of the first 149 aa in *E. coli* results in capsids with the proper morphology but which are empty[34].

The CTD comprises four arginine-reach repeats with seven conserved serine and one threonine residues that can be phosphorylated (reviewed in[35]). One the primary sequence, two nuclear localization signal (NLS) sequences have been mapped (aa 158–168 and 165–175)[36].
The structure of the HBV capsid has been solved with a 3.3 Å resolution by crystallization[37] showing that the NTD is highly ordered with a mainly helical structure, folding a major α-helical hairpin. CTD is disordered and was visualized by cryo electron microscopy after nanogold labelling[38]. Capsid assembly is initiated by rapid Cp dimerization of the two major α-helical hairpins of the NTD of two monomers forming a four-helix cluster (**Figure3**). Cp dimers then trimerize in a slow reaction, followed by association of these Cp hexamers to capsids without defined assembly intermediates[37,39]. Assembled capsids appear as spherical shell (260 Å in diameter) being fenestrated by holes of different size: 14 Å in diameter at the three fold axes, 12-15 Å at the two fold axes and 3 Å at the 5 fold axes. From the surface, spikes protrude for a length of 25 Å. These spikes are formed by the four-helix cluster of dimers[37].
After leaving the reducing environment of the cytosol, different disulfide bonds are formed.
At the dimer interface a disulfide bridge is present between the Cys 61 of the two monomers, although C61-C61 disulfide bonds are not essential for capsid assembly. A second intermolecular disulfide bond has been observed by Zheng et al involving Cys 48, however considering the distance between the Cys 48 of two monomers (grater then 20Å), and the icosahedral symmetry of the capsid, this disulfide bridge is not formed in native shell **(Figure3)**[37]. An additional, a disulfide bond was described for the last Cp aa (Cys 185 or 183, dependent upon the HBV genotype). When Cp is expressed in eukaryotic cells, this bridge links the Cys within one dimer, while when it is expressed in *E. coli* frequently leads to aberrant bonds between neighboring[40].
In the assembled capsid the CTD is present in the inside of capsid by nucleic acid interaction, or can be transiently exposed on the capsid exterior, as shown by trypsin digestion studies[41]. The localization of the CTD is associated with genome maturation but also with phosphorylation[41].

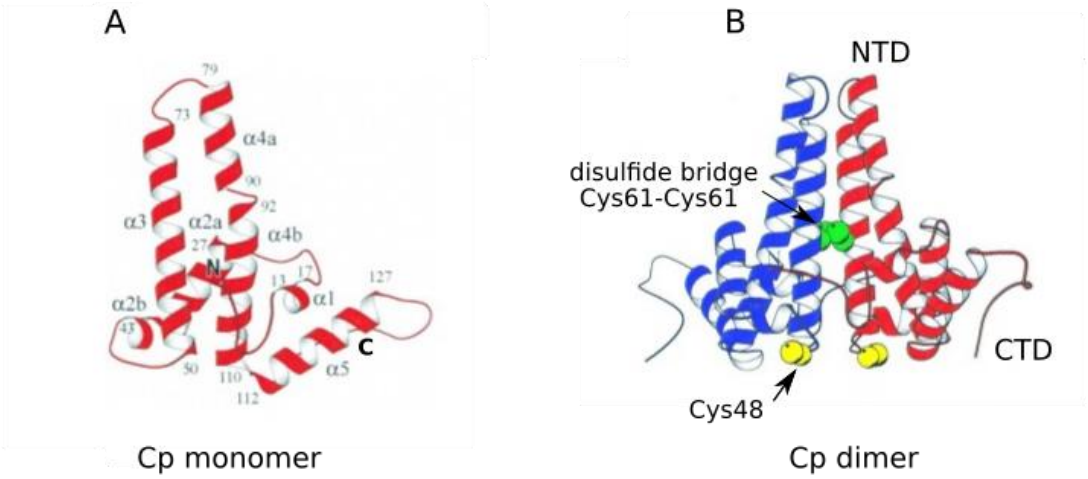


Figure 3: Cp monomer and dimer
A. Cp monomer shown in red. The NTD is highly ordered with a mainly helical structure (5 helix), folding a major α-helical hairpin formed by α3 and α4. The CTD is formed by α5. **B.** Cp dimer is made by the four-helix cluster of two monomers, where one monomer is represented in bleu and the other in red. Cys61 which forms a disulfide bridge is shown in green while Cys48 that does not in native shell is shown in yellow. Picture modified from Wynne et al. [37]

The viral genome

The HBV genome is 3.2 kb long and exists in three forms. In HBV virions, the genome is circular, relaxed and partially double stranded (rcDNA). In this form, the minus strand is complete with an 8 nt terminal redundancy, while the plus strand has a gap being partially synthetized. In the virion, RT is covalently bounded to the 5' end of the minus strand (reviewed in[12]). Upon infection, the rcDNA is converted in an episomal form called covalently closed circular DNA (cccDNA) in a multistep process involving the release of the polymerase, removal of the terminal redundancy, completion of the plus strand DNA by a host polymerase and ligation of 5' and 3' ends (reviewed by Schreiner and Nassal[42]). This process has not been completely elucidated yet, and the order by which these events happen and the host factors involved remain unclear. The cccDNA is the form present in the nucleus of infected cells where serve as a template for viral transcription. The cccDNA is organized as a minichromosome by association to host cellular histones proteins[43]. From studies performed with DHBV and WHV, the number of copies is estimated between 5 and 50[44,45]. For HBV the cccDNA copy number is less clear, but seems to be below that of DHBV and WHV. Cp and the viral protein HBx were described to be associated with it[46,47].

The third DNA form is double stranded linear and is derived from aberrant plus strand DNA synthesis. It is found in the virion and becomes integrated into host chromatin upon infection.

The genome is tightly organized with all nucleotides being protein encoding. Four overlapping ORF are present encoding seven different proteins: RT, L, M, S protein, HBx, Cp, and the preCore (preC) coding for e-antigen (HBeAg). The four promoters are regulated by two enhancers (Enh1, Enh2) and a glucocorticoid-responsive element (GRE)[48]. In addition, two direct repeats (DR1, DR2) exist, which are sites important for minus and plus strand DNA priming. Both, promoters and enhancer are overlapping the ORFs (**Figure4**).

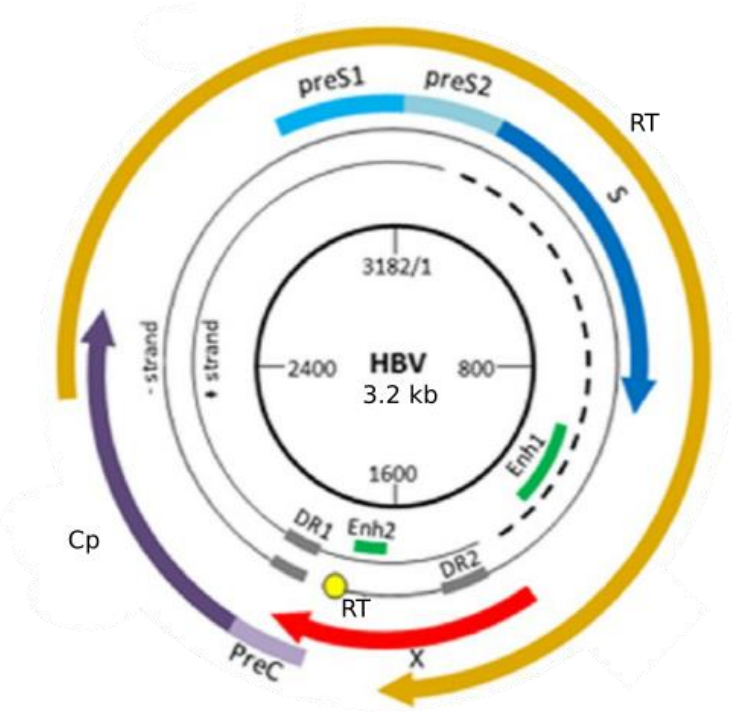


Figure 4: HBV genome organization
HBV contains a 3.2 kb, partially double-stranded (~dsDNA) genome (inner black circles) formed by full-length minus strand and an incomplete (dashed lines) plus strand. The genome is tightly organized with all nucleotides being protein encoding. It contains four promoters, two enhancer regions (Enh1, Enh2), and two direct repeats (DR1, DR2). The four ORFs are shown by the colored arrows. (Picture simplified from Minor et al[49])

Reverse Transcriptase

RT is a multifunctional protein. In addition to the reverse transcriptase activity, it possesses an RNase domain that degrades the RNA upon DNA synthesis. It comprises a so-called priming domain, also called terminal protein (TP), a spacer domain between TP and the domain involved in the reverse transcription (RT domain).

HBV replicates via an RNA intermediate called pregenomic RNA (pgRNA), which is encapsidated together with RT in newly forming capsid by interaction between RT and Cp

and RT and an encapsidation signal called epsilon (ε) on the pgRNA[50]. Within the capsid TP serves as a primer for reverse transcription[51]. During priming the first 5' nt becomes covalently linked to TP.

Hepatitis B virus X protein; HBx

HBx is encoded by the smallest, 154 aa long, ORF and has a molecular weight of 17.5 kDa.

Many different functions have been attributed to HBx: it has been suggested to affect viral replication, as well as host cell functions, by acting on a wide range of cellular processes. It has been described to be implicated in viral transcription, DNA damage repair, and cell cycle progression[52–54].

Most important is its function in replication, which is only needed in differentiated cells and not in hepatoma cell lines. Decorsiere et al. demonstrated that HBx hijacks the cellular DDB1-containing E3 ubiquitin ligase to target the 'structural maintenance of chromosomes' (Smc) complex Smc5/6 for degradation. Smc5/6 complex binds episomal HBV cccDNA inhibiting it transcription. Hence, HBx, by targeting the complex for degradation, relieves cccDNA inhibition allowing productive gene expression[55]. An alternative model was postulated by Alarcon et al. showing that HBx regulates the regulation of the chromatin structure of cccDNA. In the nucleus of infected cells, the cccDNA is present associated to host cellular histones[43] being regulated by the chromatin state. The authors showed that HBx recruits histone lysine-specific demethylase 1 (LSD1) to HBV viral promoters leading to a reduction of methylation on lysine 9 on histone H3 and therefore to an increased transcriptional activation. The authors also demonstrated that HBx recruits Set1A that is responsible for the trimethylation of lysine 4 on histon H3, thus bringing the cccDNA to an active chromatins state[56].

HBeAg

The HBeAg is encoded by the same promoter as Cp but the preC ORF contains 29 residues upstream Cp, leading to a primary translation product of 25 kDa[57]. The additional aa lead to insertion into the endoplasmic reticulum (ER) membrane, where the signal peptide is cleaved off. The protein dimerizes followed by transport to the Golgi compartment where furine proteins remove the C terminus[58]. The secreted HBeAg has a molecular weight of ca. 15 kDa (monomer). It is secreted by all the member of the *Hepadanaviridae* family[35].

HBeAg expression seems not to be required for productive infection[59] but it has been shown to have an immune-modulatory role that may help immune evasion[60]. As HBeAg is

transcribed under control of the same promoter than RT and Cp and its presence is correlated to the viral load except in patients exhibiting a preC stop mutation[12,61].

The viral life cycle

The HBV life cycle takes place exclusively in hepatocytes that provide the specific receptors and specific factors needed for the viral replication as e.g. HNF4 (reviewed in[62,63] and summarized in **figure 5**).

Entry

The entry step is a multistep process that involves a first reversible binding to heparan sulphate proteoglycans (HPSGs) and a second higher affinity interaction with the NTCP[64,65]. HSPG binding is mediated by the antigenic loop on the S domain and NTCP binding is mediated by the preS1 domain of L protein[66]. HPSGs are expressed in the extracellular matrix and on the plasma membrane of many cell types including endothelial and dermal cells. However, in hepatocytes surface, they are higher sulfated[67]. It is generally believed that the first interaction with HPSGs is necessary to concentrate the virus at the right site and to stabilize the interaction with the main receptor (NTCP)[27,68].

It has been recently proposed that HBV undergoes conformational changes of the preS1 domain after secretion, that prevent the virus from not specifically bind to other unspecific tissue. Directly after leaving the cell, the preS1 domain is on the interior due to its interaction with the capsid. Then, a maturation process translocates preS1 to the virion surface allowing binding HPSGs and NTCP. However, the molecular mechanism remains under investigation[69].

The NTCP was identified as main HBV receptor in 2012 by Yan et al. They firstly showed that preS1 binds NTCP with high affinity, and that silencing human NTCP (hNTCP) in HepaRG cells reduces the production of HBeAg and viral mRNA. In addition, the expression of hNTCP in HepG2 cells, that are normally resistant to HBV infection, confers them susceptibility[64]. NTCP is a transmembrane glycoprotein and it is expressed on the basolateral membrane of hepatocyte. The N-terminal domain is located at the extracellular space while the C-terminal domain is in the cytoplasm. Physiologically, it is responsible for the sodium-dependent bile acid uptake[70]. The N-terminal domain contains two N-glycosylation sites at aa 5 and 11 and it has been shown that mutations of both these sites, that make the NTCP glycosylation deficient, impair HBV infection reducing production of HBeAg and cccDNA[71]. Therefore it has been suggested that the N-glycosylation is required for NTCP mediated HBV

infection. Two aa sequence in hNTCP, 157-165 and 84-87 have been shown to be important for binding to preS1[64] and for mediating the entry[72] respectively.

Although the binding properties are well characterized, and although it is known that the binding with the NTCP induces endocytosis, the exact mechanism that mediates the entry, remains unknown. Whether HBV uses a caveolae-mediated endocytosis or a clathrin-mediated endocytosis is unsolved. Macovei et al. proposed the first[73] while two different other studies have been proposed the second to be the main mechanism that mediate the entry of HBV[74,75]. To note, preS1 was shown to interact with clathrin heavy chain and AP2, which is a clathrin-mediated endocytosis adapter[74].

Endosomal trafficking

After the entry step, HBV enters in the endocytic pathway, where the pH drops from 6.2 in the early endosome to 5.5 in the late endosome. For many enveloped viruses, acidification is needed for endosome escape by activating an envelope fusion protein that mediates the fusion with the endosome membrane. However, whether the pH effects HBV infection and how HBV escapes the endosomal pathway remains to be clarified. It has been reported that in HepaRG cells, HBV infection depends on Rab5 and Rab7, which are GTPases involved in endosome maturation, supporting that HBV passes from the early to the late endosome[76].

In experiments using the DHBV as surrogate model, where the pH was raised by ammonium chloride, the infection remained unaffected, suggesting that at least for DHBV the endosomal escape is pH independent[77,78]. However, in another study, where the pH was neutralized by Bafilomycin A1 the infection of DHBV[79] and HBV[80] was inhibited. Thus, the potential role of pH remains controversial. Although a fusion function has been proposed for the preS1[81], how HBV escapes endosome pathway remains to be fully understood.

Cytoplasmic transport

Following the endosomal escape, the capsid is released in the cytoplasm and is translocated by retrograde transport to the nuclear periphery. As most viruses, the HBV capsids use microtubules and cytoplasmic dynein from transport. Our group observed that after microinjection in *Xenopous laevis* oocytes, the capsids accumulate at the nuclear pore complex (NPC), within 15 min requiring capsid interaction with the dynein light chain 1[82]. Interaction and transport required exposure of the CTD[82].

Nuclear transport and capsid disassembly

Once at the nuclear periphery, the capsid has to pass the NPC and disassemble to release the genome inside the nucleus. It was shown that the nuclear localization signal (NLS) present on

the CTD of Cp binds the import receptor Importin α (Imp.α)[83], but how and where the capsids switch from the interaction with the dynein light chain 1 to Imp.α remains to be investigated. Imp.α serves as an adapter for the binding with Importin β (Imp.β) that mediates the interaction with NPC.

Being 36 nm in diameter, NC is just below the upper limit of the nuclear pore complex (NPC), 39 nm[84], and it has been shown that NC cross the NPC intact reaching the nucleoplasmic side of it, termed nuclear basket, where NC interacts with Nucleoporin 153 (Nup153)[85]. The subsequent step of NC disassembly, leading to diffusion of Cp dimers[86], and genome[41] deeper into the nucleus remains unknown.

rcDNA conversion to cccDNA

Once released in the nucleoplasm the rcDNA is converted in cccDNA.

Guo et al. reported the presence of DHBV capsids containing deproteinized rcDNA in the cytoplasm of infected duck hepatitis cells, and they suggested that for DHBV the deproteinization may lead to capsid destabilization with consequent disassembly[87]. However, Cui et al. showed that destabilized HBV capsids contain only rcDNA bounded to RT, arguing against a role for a deproteinization-induced capsid destabilization. They suggested that the deproteinization likely occurs after the genome release step rather than before[88]. Considering that the disassembly is thought to take place at the nuclear side of the nuclear basket, according to Cui et al.[88] the deproteinization likely occurs in the nucleus.

The group of Michael Nassal proposed that the TPN2 protein, which is a cellular enzyme implicated in host cell genome repair, is involved in the releasing of RT from the rcDNA. TPN2 is known to repair DNA adducts induced by the topoisomerase (TOP) by releasing TOP from DNA breaks. This suggests that RT bounded to rcDNA is structurally similar to TOP adducts and therefore is recognized by TPN2[89].

Early studies on DHBV showed that the viral polymerase does not participate in completing the plus strand DNA[90]. Recent investigations showed that several host polymerases as pol κ, pol λ and pol η are involved, although with different efficiency. Their knockout impaired the formation of cccDNA but however it did not completely abolished HBV infection suggesting that other factor need to be implicated in the process[91].

Viral DNA integration

Although being a dead end of infection, HBV DNA but also the DNA of other hepatitis B viruses integrates into host chromosomes. HBV DNA has been found integrated in tissue

originated from HCC patients, and in cells lines derived from HCC[23]. Thus, it has been hypothesized integration may contribute to the development of the carcinogenesis.

In contrast, studies on WHV, often used as surrogate model, showed in woodchuck liver tumor that integrations frequently occurs closed to the N-Myc family of proto-oncogenes causing their activation. This seems to be the leading cause for the developing of WHV-associated HCC[92]. However, studies of human HBV-associated HCC have shown random integration with less clear evident mechanisms[93,94].

Transcription and protein synthesis and assembly

cccDNA serves as a template for viral transcription, which is facilitated by the host polymerase II. Five mRNAs are synthesized: the pregenomic RNA (pgRNA), one further mRNA of supergenomic length coding for HBeAg, and three subgenomic mRNAs. In addition, two defined spliced versions of the pgRNA are synthesized[95]. The three subgenomic mRNAs encode for HBx and for the three surface proteins. pgRNA is the RNA intermediate necessary to replicate the viral genome and which is the template for reverse transcription. In addition, it encodes for Cp, and RT. pgRNA contains a stem loop structure at the 5' end, which is the encapsidation signal ε (reviewed in [23]).

It has been shown that the binding of RT to ε triggers the recruitment of the RT and pgRNA inside newly assembled capsid[96].

To explain how the assembly process occurs, a new model has been recently proposed, where pgRNA plays an active role in Cp enucleation. Using RNA SELEX, Patel et al. identified Cp binding sites (Packaging signals PS) on the pgRNA forming stem loops structures. They also observed that at low concentration, and without RNA, only Cp lacking CTD and not full-length Cp assemble into capsids. This implies that in these conditions, the CTD inhibits the enucleation of Cp and therefore its assembly. However, this inhibition can be suppressed either by phosphorylation of CTD, or by the presence of RNA, thus arguing for electrostatic repulsion of Cp due the positive charged CTD. Therefore, the authors suggested that the recognition of PS by Cp may favor the dimer formation (the first assembly intermediate) by reducing the electrostatic repulsion between the CTD of Cp and consequently start the enucleation process that lead to the formation of assembled NC.[97] In addition, one PS has been mapped adiacent to a pgRNA region, called (φ). φ, being the complementary region of ε can pair with it circularizing the pgRNA.

Thus, the model suggests that the RT-ε/φ complex may guide the folding of PS1 in its stem loop motif which then binds to Cp, leading to assembly initiation; this would also assure that Cp assembly only occurs on a pre-genome that is bounded to RT[97].

Similar mechanisms have been proposed for other viruses RNA such as parechovirus[98].

It has been further shown that packaging also depends on host factors, as e.g. Hsp90, which is thought to stabilize the interaction between RT and ε[99].

Reverse transcription and morphogenesis

Genome maturation starts with RT-ε complex formation, which triggers the start of the reverse transcription that takes place inside the capsid. The required nucleotides get access to the inside of the capsid via the fenestrations in the capsid surface. After priming and synthesis of the first 3 nucleotides, which are copied from the bulge of ε, the complex switches to DR1 closed to the 3' end of the pgRNA from where reverse transcription continues towards the 5' end of the pgRNA. The degradation of the pgRNA by the RNaseH domain occurs in parallel to reverse transcription but 11 nucleotides behind. This delay leaves the capped 5' end of the pgRNA undegraded and this oligonucleotide serves as the primer for plus strand DNA synthesis after translocation of the DR2. Circularization leading to the rcDNA requires three not-well characterized elements called M, 3E and 5E[100].

In 30 - 50% the RNA primer translocation does not succeed and the plus strand DNA synthesis results in a double stranded linear HBV genome. (Reviewed in[23]).

In vitro, the single strand gap on the rcDNA may be repaired by addition of few nucleotides by the endogenous DNA polymerase. The endogenous polymerase reaction carried out in the presence of (α^{32}P) dNTP is generally used to radiolabel HBV genome and was firstly described by Summers et al.[101] to screen serum samples of different host species for the presence of a hepatitis B like virus[102].

The entire genome maturation process seems to occur prior to secretion and at least reverse transcription is a requirement for secretion[103]. Thus, the cytoplasm of infected cells contains capsids at different stages of replication. Capsid containing replication intermediates, e.g. early RNA containing capsid, and ssDNA containing capsid are referred as Immature capsid (ImmC). When the formation of the rcDNA is completed, capsids reach their "maturation" and are therefore called mature capsid (MatC).

Secretion

Once reached their maturation, MatC can either re-enter in the nucleus, leading to the amplification of cccDNA pool, or acquired the lipid envelope together with the surface proteins and be secreted.

It has been proposed that the cccDNA pool in the nucleus is maintained by recycling of newly formed MatC. Although the recycling into the nucleus of MatC is generally accepted for DHBV, experiments on HBV are not fully clear as the half-life of cccDNA remains controversial[104,105].

In the cytoplasm of infected cells, MatC, but not ImmC[103] can be enveloped and eventually secreted[106]. It has been shown that MatC acquire the host-derived lipid envelope, where surface proteins are embedded, by budding into the lumen of an intracellular membrane believed to be MVBs for extracellular secretion[107].

To explain how only MatC are recognized, the maturation signal model has been proposed. According to this hypothesis, rcDNA formation and, therefore, capsid maturation leads to conformational changes in the capsids that lead to its recognition by the envelope proteins[108]. The preS1 domain of L seems to be involved in the interaction with the capsid to guide the envelopment[28]. The structural capsid changes required for MatC envelopment seem to also occur on EmpC leading to empty virions with a concentration of up to 10^{11} particles/mL[109,110]. The secretion and envelopment of EmpC argues against the maturation signal induced by the formation of rcDNA. Recently, Ning et al. proposed the "ssDNA or pgRNA dependent blocking signal" hypothesis, where, an inhibitory signal is present in ImmC[111].

Nevertheless, even if in lower amount, ssDNA and pgRNA containing particles have been recently found in the blood of infected patients (100 to 1000 lower than rcDNA containing virions)[112,113], although their secretion is under debate and their presence needs further characterization.

The secretion of filaments also occurs through the ESCRT machinery and the mutivescicular body secretion pathway[107,108]. While the secretion of spheres occurs through the host constitutive secretory pathway[114]

In infected patients, HBsAg spheres and filaments outnumber rcDNA virions of 100,000-fold with a concentration of 10^{14} particles /mL[115] *versus* 10^{9} particles/mL. The reason for such subviral particles secretion remains unclear. The leading hypothesis regards the modulation of the immune system; and it has been proposed that HBsAg particles would bind circulating antibodies preventing their neutralization of infectious virions[116].

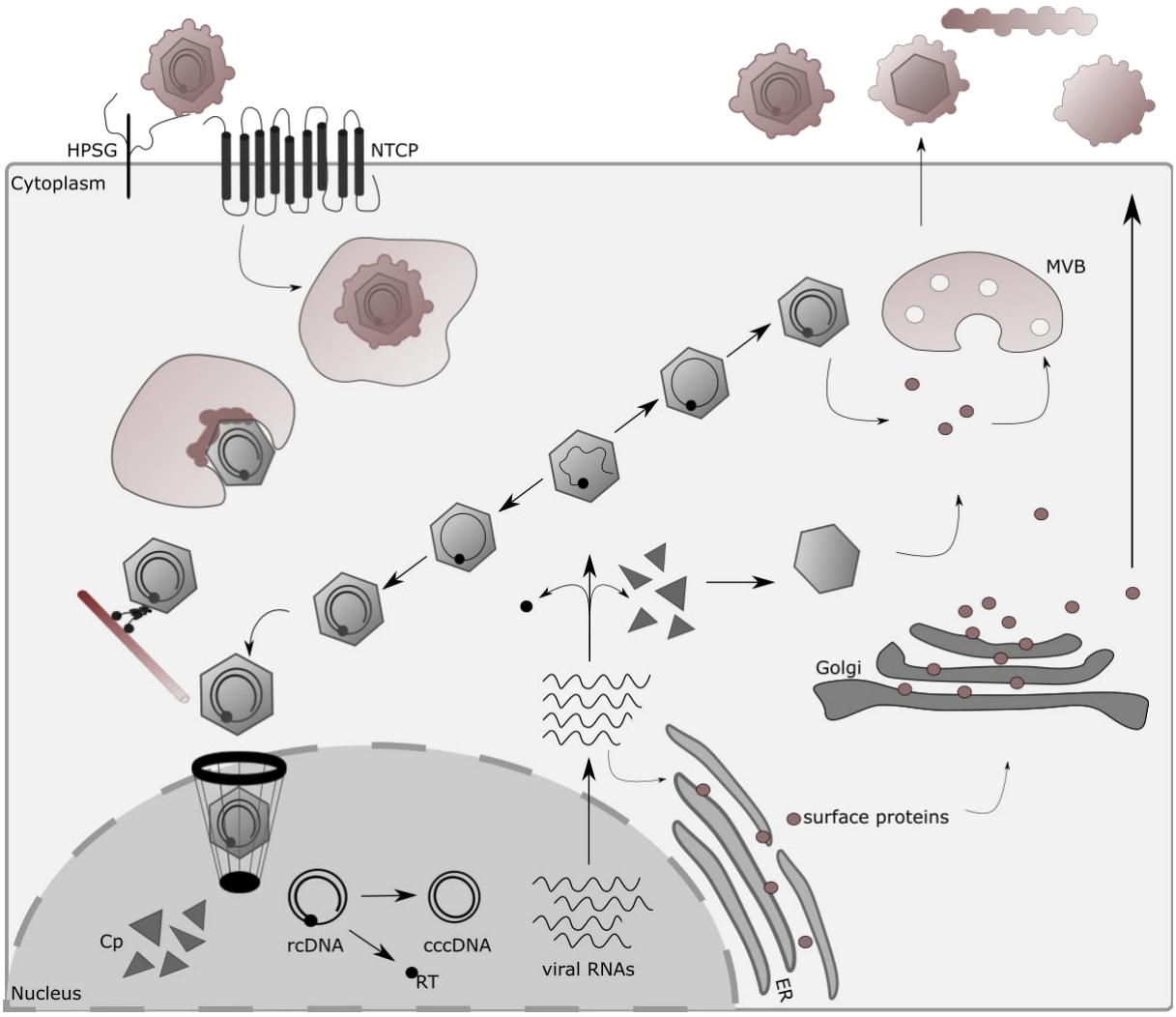


Figure 5: HBV life cycle
HBV replicates via reverse transcription in hepatocytes. Infection starts with the binding to HPSG and NTCP receptors and leads to receptor-mediated endocytosis. NC is released in the cytoplasm and translocated to the nuclear periphery through active transport involving microtubules and association with dynein light chain 1. Nuclear transport ends in the nuclear basket. The genome is released and following the remove of RT repaired forming the cccDNA form. Then, transcription occurs and viral RNAs are brought to the cytoplasm. pgRNA is encapsidated along with RT that start the reverse transcription. Eventually rcDNA containing capsid are formed that can be either recycled back into the nucleus or recruited for envelopment and secreted. In addition to rcDNA containing virions, in the extracellular medium, HBsAg particles are secreted as spheres or filament, as well as EmpC containing particles.

Capsid protein and the maturation process

The maturation process is a tightly organized and involves structural changes of the CTD including its phosphorylation and consequent exposure on the capsid surface.

Localization of the CTD

The CTD of Cp is a highly flexible and has a disordered structure. In 1997, Zlotnick et al. coupled the CTD to gold particles and visualized it by electron microscopy of rHBc

containing *E. coli* RNA. They showed that in this type of capsid, mimicking ImmC containing pgRNA, the CTD is located inside the capsid lumen, probably by interacting with the RNA. The CTD is localized at the quasi-six-fold vertex thus adjacent to a fenestration[38].

Rabe et al, reported that in MatC all the CTDs are sensitive to trypic digestion, implying that the CTDs is exposed on the capsid surface. In ImmC only a small fraction of CTDs (30%) is sensitive to tripic digestion, while in ImmC derived from cells which were treated with a polymerase inhibitor (foscarnet), even a smaller fraction was digested (15%). These data argue for a maturation dependent exposure of the CTD requiring the activity of RT. In agreement, in rHBc, devoid of RT, all CTDs are protected by the activity of trypsin supporting that the initial synthesis of the minus strand DNA is required.

In rHBc, which were *in vitro* phosphorylated, the phosphorylated CTDs were trypsin sensitive. As the introduction of negative phosphate charges, between the positively charged arginine cluster of the CTD, repulse RNA, it was assumed that a reduced affinity between capsids and RNA in the lumen of the capsid allows CTD exposure[41]. This hypothesis is in agreement with a reduced affinity of Cp to dsDNA compared to single stranded nucleic acids Localization of CTD is summarized in **Figure 6.**

Phosphorylation of the CTD

The aforementioned experiments argue for a phosphorylation dependent exposure of the CTD. The CTD contains seven conserved serine and one threonine that are potentially phosphorylated. Point mutations of the serine residues exhibited that Ser162 is crucial for pgRNA packaging[117]. Consistently, mutated Ser164 Cp failed to interact with TP[118]. Precipitations of capsids from patient sera using an antibody against the phosphorylated CTD at Ser170/172 showed that this modification is also found *in vivo* and is maintained in virions[119]. However, it has been observed that phosphorylation decreases with genome maturation, which was also shown for DHBV capsids[120]. As pgRNA encapsidation can be potentially driven by RT interaction with one single phosphorylated Cp these data do not allow drawing a conclusion about the number of phosphorylated sites in assembled capsids so that the data are not necessarily contradictious.

In addition, the protein kinase involved in the phosphorylation has not been unequivocally identified. Protein kinase C (PKC) was found in HBV virions from cell culture and from patients[20,40,121]. Inhibition of protein kinase C in cells expressing HBV results in decrease virion formation and intracellular accumulation, although does not impairs genome maturation.

Additionally, altering the CTD localization, phosphorylation may induce the signal leading to envelopment and seems to influence capsid stability. Selzer et al, reported that the substitution of serine 155, 162 and 170 with glutamate results in increased stability of the capsid[122]. However, it remains unclear if this finding is relevant for capsid disassembly.

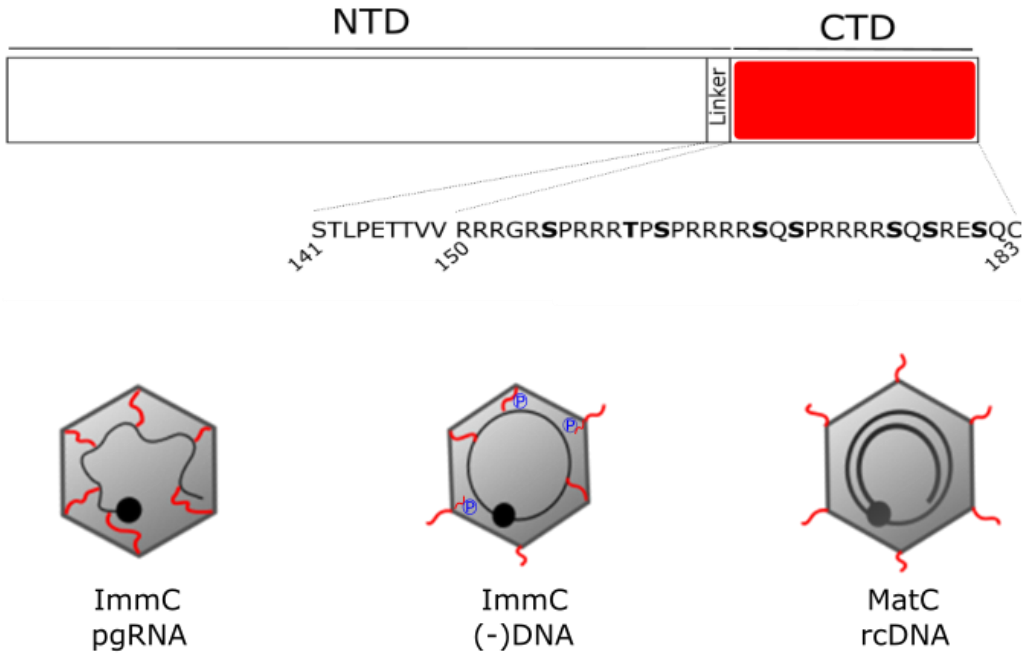


Figure 6: Localization and phosphorylation of CTD according to the maturation degree of the capsid
CTD (in red) has seven conserved serine and one threonine potentially phosphorylated (in bold characters). In pgRNA containing capsid the CTD in situated in the capsid lumen. Synthesis of the minus strand DNA induced Phosphorylation of the CTD (in blue), however, the extent of phosphorylation remains unknown. Phosphorylation causes partial exposure of CTD on the capsid surface. In MatC CTD is present dephosphorylated on the capsid surface

HBV tropism

HBV host tropism

HBV presents a narrow host range infecting only humans, chimpanzee and under experimental conditions *Tupaia belangeri*[9]. Commonly used animal models such as mice and rats are resistant to the infection[123]. The host tropism, together with the liver tropism, is determined on different levels. The first one is the entry step mediated by NTCP. Two aa sequences have been mapped on hNTCP to be crucial for HBV tropism: aa 157-165[64] and aa 84-87[65] mediating binding and entry respectively. It has been shown that the homologue

mouse NTCP conserves the first sequence, while present an alteration in the second one. In agreement, although being able to bind a synthetic L-protein derived peptide, Myrcludex B[64,65], mouse NTCP cannot efficiently mediate HBV entry[65]. However, at least in the case of mice, the host tropism seems to be determined also at later steps. Murine cells expressing hNTCP, support HDV entry, which have the same envelope as HBV[124], but these cells do not support later steps of the HBV life cycle as cccDNA formation and HBeAg secretion[65]. This restriction can be overcome by fusing hNTCP expressing mouse cell with HepG2 human cells that do not express NTCP[125], showing that human cells provide additional factors necessary to the steps after the entry.

On the contrary, it seems not to be the case for other animals normally resistant to the infection as macaque and pig. It has been reported that the only expression of hNTCP in hepatocytes deriving from cynomolgus and rhesus macaque and pig is enough to restore cccDNA and HBV transcripts levels comparable to human hepatocytes[126], suggesting that for these species the only level that determine the host restriction is the entry.

As expected, expression of NTCP from *Tupaia belangeri*, as well as hNTCP in HepG2 cells renders these cells susceptible for HBV infection, confirming the importance of NTCP in defining host specificity and the possibility of *Tupaia belangeri* to be infected[64]. *Tupaia belangeri* hepatocytes however do not support productive HBV synthesis.

Hepatic tropism

The NTCP receptor plays a pivotal role also in determine the hepatic tropism. NTCP expression is regulated by liver-specific transcription factors, hormones and cytokine and is expressed exclusively in the basolateral membrane of hepatocytes. However, as for host tropism, also hepatic tropism is determined at different levels. For example, to promote viral transcription, Enh1 and Enh2 has binding sites for specific liver transcription factors including hepatocyte nuclear factor (HNF) 4 and HNF3 where HNFs are typical example of liver-enriched transcription factor (reviewed in[127])

Research on HBV: challenges

Animals models

Studies on host immune response, as well as studies on the course of the infection have been done mainly on the chimpanzee model. However; the use of high primate in clinical research as chimpanzee raises ethical concerns and it is hampered by the limited animal availability and by their high maintenance costs.

As an alternative, human liver chimera mice have been used. These mice are made by repopulating the partially destroyed liver of nude mice with primary human hepatocytes[128]. These mice are permissive to HBVinfection, but infection rate depends upon the quality of the donor and the outcome of infection is highly variable[128]. Double chimera mice have been developed engrafting human hepatocytes and human immune system and it has been reported that such mice develop human specific liver fibrosis[129].

Cell culture model

Hepatoma-derived cells line, such as HuH-7 and HepG2 cells are not susceptible to HBV infection as they do not express the NTCP. An exception is the hepatoma cell line HepaRG, which is kept differentiated with addition of dimethyl sulfoxide (DMSO)[130].

Nevertheless, hepatoma cell lines in general support later steps of HBV replication. This has been shown the HBV particles derived from transfection of HBV genomes into HepG2 cells (HepAD38 and HepG2.2.15) are infectious in chimpanzees[131].

HuH-7 and HepG2 cell line stably expressing NTCP are currently used as model to study the entry step. To note, HepG2-NTCP cells seems to be more highly susceptible to HBV than Huh7-NTCP cells. These suggest that additional factors, present in HepG2 cells but absent in HuH-7 cells are required to achieve optimal level of viral infection[65].

Human primary hepatocytes (PHH) are the *in cellulo* closest model. However, they present some constrains due to the limited life in the order of weeks in culture and to the limited availability. In addition, their susceptibility is lost three days after taken into culture ref. This seems to depend on the loss of cell polarization caused by the loss of hepatic specific factors during cell culture[132]. One other drawback is the high donor-to-donor variability that renders limited the number of reproducible studies (Reviewed in [133]).

Nuclear transport

The Nuclear Pore Complex

The NPC is macromolecular structure embedded in the nuclear membrane where it forms an aqueous channel that represents the gate that molecules has to pass to get access to or exit the nucleus (**Figure 7**).

The NPC protein complex has a mass of >125 MDa with an octagonal rotational symmetry. It is composed by three major parts: a central ring that fuses the outer nuclear envelope with the inner one, a nuclear face that is made by eight fibers forming a cage-like structure ending with a nuclear ring, called nuclear basket, and a cytoplasmic face with 50–100 nm long flexible filaments that radiate from the central ring toward the cytoplasm.

NPCs are phylogenetically well conserved even between distant species like humans and *Xenopus leavis* (reviewed in[134]).

It is formed by ~30 different proteins called nucleoporins (Nups). All Nups are present as multiples of eight reflecting the highly conserved eight-fold symmetry of NPCs. Aside of their function during import and export they play a role also during cell cycle and intranuclear chromatin distribution[135,136].

The number of NPCs per nucleus varies between organisms, cell types and depends on the metabolic status of the cell. It is higher during G2 phase than G1 phase (e.g., 8.5 vs 5 NPCs/μm^2 nuclear envelope in HeLa cells)[137].

For molecules that shuttle through the nuclear membrane, the barrier to pass is represented by phenylalanine-glycine (FG)-repeats present on one third of the Nups[138–140] and which form a hydrophobic exclusion mesh within the pore.

Analyzed by electron microscopy, the central ring has a diameter of 40 nm and it has been demonstrated that the upper limit size for cargo molecules that can pass through it is 39 nm.[84] Molecules with a size of 5-8 nm and a molecular weight of 20-40 kDa can freely pass the NPC by diffusion[141–144], but diffusion also depends on shape and surface charge. Large molecules need to be actively transported through the NPC.

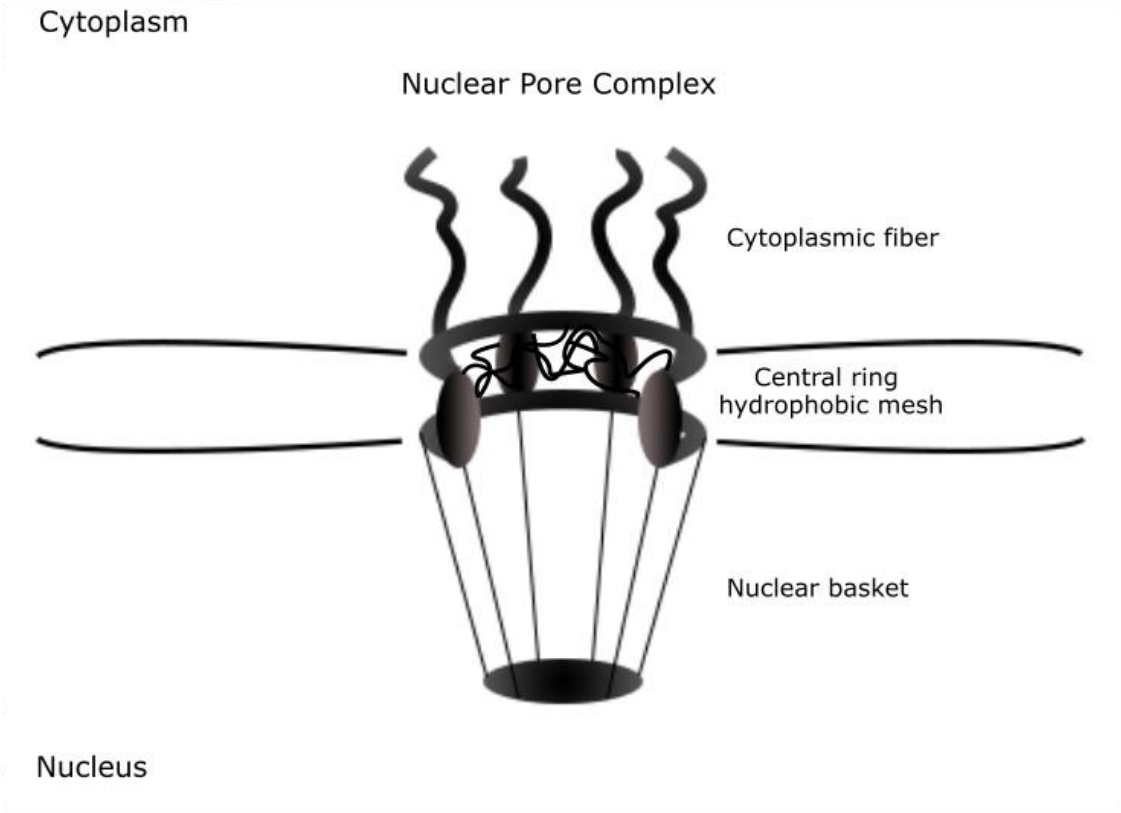


Figure 7: NPC
It is a protein complex of complexes of >125 MDa having an octagonal rotational symmetry. It is composed by three major parts: a central ring, embedded in the nuclear membrane, a nuclear side that is made by eight fibers forming a cage-like structure ending with a nuclear ring, called nuclear basket, and a cytoplasmic face with 50–100 nm long flexible fibers that radiate from the central ring toward the cytoplasm. It is formed by 30 Nups, one third of them having FG repeats forming the hydrophobic barrier of the central ring

Import receptors and nuclear localization signals

Cargos that have to go inside the nucleus harbor a nuclear localization signal (NLS), while cargos that have to go from the nucleus to the cytoplasm harbor a nuclear export signal (NES). NLS and NES are recognized by Karyopherin family of transport receptors that mediate the active transport through the NPC. Although there are a few exception[145], the member of the karyopherin family mediate the transport of cargos unidirectionally with importins facilitating the transport into the nucleus and exportins facilitating the transport into cytoplasm[146].

The NLS is recognized by import receptors. Import receptors either bind their cargo directly through the interaction with NLS, or *via* an adaptor. One prototype is transportin-1 (Kapβ2)

that directly binds an NLS named M9 from its characterization in the C-terminal M9 domain of the Influenza virus nuclear ribonucleoprotein A1 (hnRNP A1)[147]. When transporting cargos with a classical NLS importin β (Imp.β) requires an adapter, Importin α (Imp.α). The prototype of a classical NLS was firstly described on the T-antigen of the polyomavirus Simian virus 40 (SV40Tag)[148,149] where it has the aa sequence PKKKRKV. Classical NLSs are formed by clusters of positively charged aa, usually lysine or arginine. They can be mono- or bipartite. In the monopartite signal there is only one stretch of highly basic aa, while in the bipartite two basic aa sequence are separated between each other by a linker region of usually 10-12 aa. Imp.α serves as adaptor by having an importin beta binding domain (IBB) that allows its binding with Imp.β. Canonical IBB consists of 13 basic amino acids in seven clusters scattered over 39 residues[150,151].

Nuclear import

Classical import pathway

Classical import pathway is mediated by the ternary complex formed by cargo with classical NLS, Imp.α and Imp.β. Imp.α binds to classical NLSs through a binding groove formed by ten armadillo (Arm) repeats, while it binds Imp.β by the IBB on its N-terminal domain[152].

To regulate cargo binding, the IBB on Imp.α has an auto-inhibitory function. Crystals structural studies performed on mammalian Imp.α showed that, when Imp.α is not bounded to a cargo, part of the IBB occupies the NLS binding groove[153]. Therefore, when Imp.α is unbound to a cargo, its IBB is not exposed thus avoiding import of unloaded adaptor. However, it has been shown that aa responsible of Imp.β binding in the IBB remains outside the binding groove free to bind Imp.β[154]. Additionally, it has been shown that the IBB inhibits the binding of NLS-cargos by decreasing the affinity interaction between Imp.α and NLS[155], suggesting that this conformation reduces the affinity of Imp.α for the NLS-bearing cargo when Imp.β is not available. Thus, formation of the trimeric import complex likely occurs when both NLS-cargo and importin β are available in the cytoplasm.

Models for translocations in the nucleus

Binding to a cargo triggers conformational changes of the import receptor that allows the interaction with Nups localizing on the cytoplasmic side of the NPC. In the case of the classical import pathway, the ternary import complex localizes at the nuclear periphery where Imp.β interacts with the FG repeats of Nups present in the cytoplasmic fibers of the NPC[156].

How the translocation across the NPC occurs is still under investigation and several models have been proposed involving FG-repeats of Nups in the central channel (Nups 98, 93, 62, 58, 54 and 45).

Data from atomic force microscopy suggested the so called "collapse model" where, upon binding of import receptor, the hydrophobic mesh in the central channel, collapse opening the passage to the import complex[157]. The "hydrophobic gel model" or "saturated model" proposes that the FG repeats are cross linked with each other creating a dense "gel". According to this model, import receptors would rupture the phenylalanine cross-links dissolving the gel[138,158].

Another described model is the "polymer brush model", where Nups containing FG-repeats form brush-like structures in constant movement[159]. This movement would sweep away macromolecules. In this model would be the shape and the movement of the Nups containing FG repeats to create an entropic barrier for unspecific cargoes rather than FG-FG interactions.

Finally, the "reduction of dimensionality model", proposed by Peters et al., proposes that translocation across NPC occurs via progressive and continuous interaction of import receptors with the FG repeats along the central channel.[160]

The analysis of native nuclear pores of *Xenopus laevis* oocyte cells by high-speed atomic force microscopy favor the polymer brush model[161].

In summary, the majority of Nups, containing FG-repeats, form a selectivity barrier and are localized in the central channel of the NPC. However, some are asymmetrically distributed at the periphery of the NPC. For example, Nup 214 and 358 are major components of the cytoplasmic filaments and that are the first contact of import receptors serving as docking site[162,163].

Nup153 has a N-terminal domain that anchor it to the nuclear basket and a C-terminal domain that is unfolded and highly flexible being ~200 nm long it potentially reaches through to the cytoplasmic side of the NPC. In addition, the C-terminal domain is enriched in FG repeats and therefore it is thought to be involved in the interaction with import receptors[164] (**Figure 9**).

Nuclear import ends in the nuclear basket where the import-complex binds to Nup153 for then being dissociated by interaction between the import receptor and the member of Ras-related nuclear protein, Ran, in its GTP-bound form. Nup153 has also a crucial role in import pathways as its depletion leads to defect in importin α/β-mediated nuclear import and it is thought to participate in the dissociation of importin cargo-complexes[165]. However, it is not essential for M9 mediated import[164].

Ran GTP cycle

The nucleus-cytoplasmic transport mediated by Imp.β requires metabolic energy. Directionality of the transport is caused by the gradient of RanGTP, which is 1000-fold more concentrated in the nucleus then in the cytoplasm[166].

RanGTP is one of the two forms in which the Ras GTPase family member Ran can be found in the cellular environment. Ran can either be present bounded to GTP or to GDP; the latter as a result of GTP hydrolysis. In the nucleus RanGTP binds to import receptors causing the dissociation and the release of the cargo. Imp.α is recycled back to the cytoplasm by the nuclear exporter CAS and RanGTP[167], while Imp.β is recycled back in the cytoplasm in complex with RanGTP. In the cytoplasmic side, GTP hydrolysis is mediated by RanGAP1 that causes the release of Imp.β that can participate in a new import cycle[168].

Recycling of Ran requires nuclear import of RanGDP using the nuclear transport factor 2 (NTF2)[169]. This is followed by the exchange of GDP by GTP catalyzed by the chromatin-bound Ran guanine nucleotide exchange factor (RanGEF; also termed regulator of chromosome condensation, RCC1)[168]. The classical import pathway is schematized in **figure 8.**

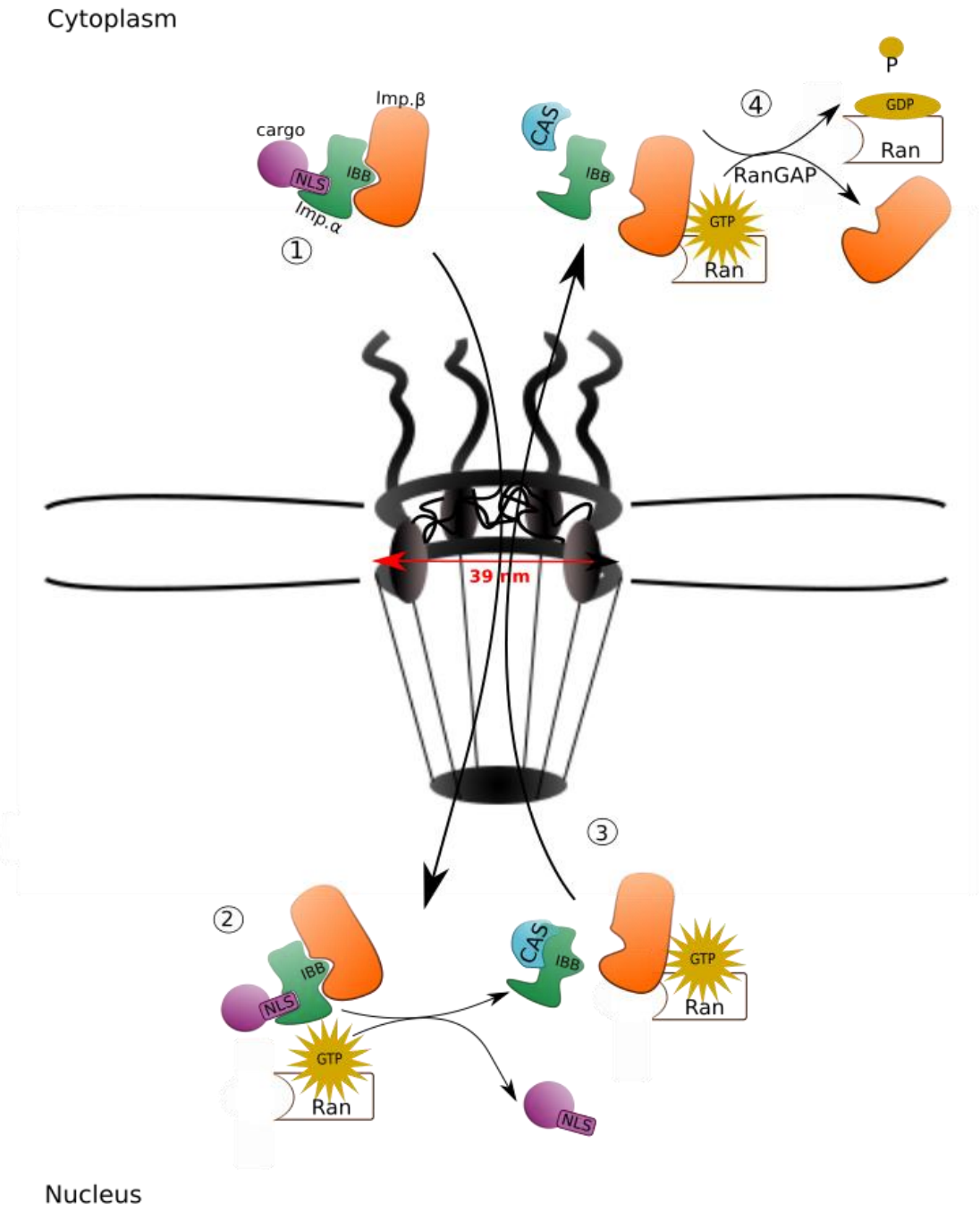


Figure 8: Classical import pathway
1. NLS present on the cargo protein is recognized by Imp.α that in turn binds to Imp.β. The import complex is translocate across the NPC. **2.** In the nucleus RanGTP mediates the release of the cargo from the import receptors. **3.** Imp.α is recycled back in the cytoplasm by CAS, while Imp.β remains associated with RanGTP. **4.** In the cytoplasm, RanGTP is converted in RanGDP by RanGAP1 leading to the dissociation of Imp.β

Nup153

Nup153 is a 1475 aa long polypeptide (**Figure 9**). Using *Xenopus* egg extract in which Nup153 was depleted from reconstituted nuclei, it has been demonstrated that Nup153 is crucial for maintaining the correct architecture of the NPC[164]. Walther et al, showed that

depletion of Nup153 results in unstable association of components of the nuclear basket and of filamentous structures associated to NPC. In addition, the authors showed that lack of Nup153 leads to mobilization of NPCs in the NE[164] and Nup153 was found interacting with lamin A and lamin B[170].

The N-terminus domain of Nup153 is responsible for recruiting another Nup, Tpr, to the nuclear basket and depletion of Nup153 results in the loss of Tpr at the NPC[164,171]. As Nup153, Tpr is involved in maintaining the correct NPC structure[164] and it localizes between the NPC and the underlying chromatin.

The zinc finger domain on Nup 153 interacts with DNA[172] and with RanGDP[173].

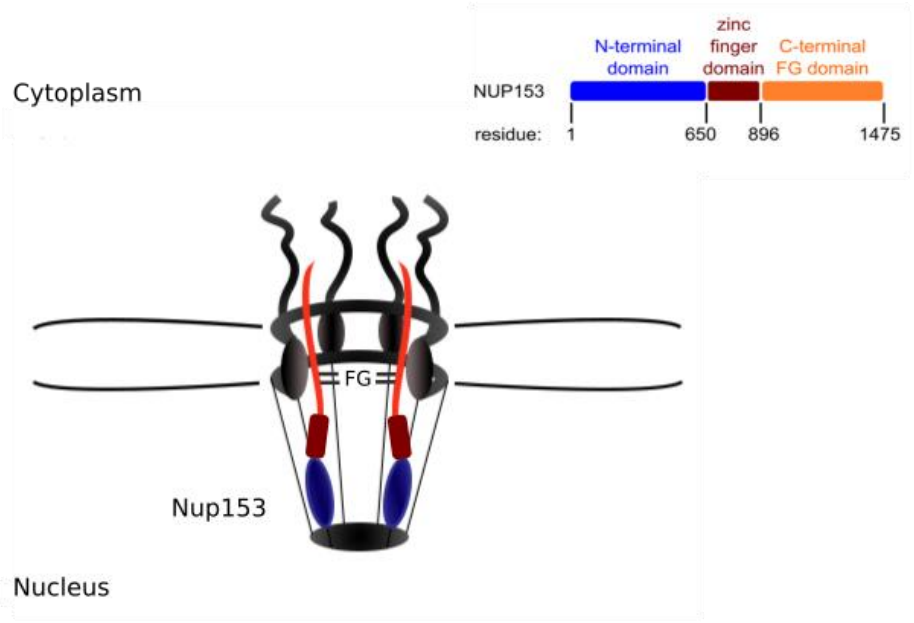


Figure 9: Nup 153
Nup153 has a N-terminal domain that anchor the nucleoporin to the NPC, a zinc finger domain and a CTD domain containing FG-repeats and forming fiber that extrude toward the cytoplasm

Nuclear import of viral genome and capsids

DNA viruses with a nuclear replication step have to deliver their genome in the nucleus. As DNA is not karyophilic it needs transport mostly via attached karyophilic proteins. They can form a shell around the genome having the advantage the genome can is sensed by cytosolic or membrane-bound pattern recognition receptors (PRRs).

Nuclear transport in such a shielded way can be exemplified by the herpes simplex virus 1 (HSV-1). After fusion of the viral envelope with the plasma- or endosomal membrane[174], the

capsid stays attached to several tegument proteins. It has been shown that the nuclear import is mediated by a direct interaction between the inner tegument protein pUL25 of the HSV capsid with the cytoplasmic filaments of Nup214[175]. A direct interaction between pUL36 and the filaments of Nup358 was also described[176]. Ojala et al. showed that the docking of the capsids to NPCs is mediated by Imp.β *via* tegument protein. In addition, they showed that the docking site involves Imp.β association to Nup358 and it is sensitive to RanGTP[176,177].

Having a diameter of 120 nm in diameter herpes viral capsid cannot pass the NPC intact. The capsid opens at the vertex opposing the NPC. This opening leads to ejection of the genome by repulsion of the condensed viral DNA, which resembles to bacteriophages. This mechanism was suggested using *in vitro* assay involving capsids and nuclear envelopes from *Xenopus laevis* oocytes visualized by atomic force microscopy[178]. After relaxation of the DNA inside the capsid, the remaining part also have to enter the nucleus, but by a different mechanism. It was hypothesized that the DNA may be pulled out by the activity of RNA polymerases, a mechanism that has been proposed also for the bacteriophage T7[179]. This assumption is in agreement of observations of Newcomb et al., who observed that the part of the genome entering the nucleus comprises the immediate-early genes[180].

Another example is the *parvoviridae* family. They are small non-enveloped viruses with a genome formed by single strand DNA and a capsid ranging from 18 to 28 nm in diameter. The size of the capsid would allowed parvoviruses to pass the NPC intact[181].

A NLS has been mapped on the large parvovirus capsid protein VP1, which becomes exposed following endosomal acidification[182]. Functionality of the NLS during nuclear import of the genome is however unclear. Parvovirus H1, canine parvovirus, and different adeno-associated viruses (AAV) cause localized transient disintegration of the nuclear envelope (nuclear envelope break down (NEBD) in somatic cells and *Xenopus laevis* oocytes, which could allow nuclear entry without passing the nuclear pore[183]. The function of the NLS could thus also be the nuclear import of VP1-VP2 trimers[184], which is required for virus assembly, which occurs inside the nucleus.

Nuclear import of HBV

Classical import pathway mediates HBV nuclear import

Because of the lack of efficient infection in cell culture, HBV nuclear import has been mostly studied using permeabilized cells and *Xenopous laevis* oocytes. The latter has the advantage of providing a larger cytoplasm allowing a better study of the cytoplasmic and nuclear transport.

Subjecting HBV capsids to digitonin permeabilized cells, it was observed that only *in vitro* phosphorylated rHBc and not unphosphorylated rHBc are able to reach the nuclear periphery and to bind the NPC[83]. Co immune precipitations provided evidence of an Imp.α-mediated Imp.β interaction of MatC, purified from HepG2.2.15 cells, and phosphorylated rHBc[83]. Further, replacing the cytoplasmic extract added to permeabilized cells by Imp.α and Imp.β allowed binding to the NPC (in phosphorylated rHBc[83]) and nuclear import of the viral genome (in MatC). In agreement, a bipartite NLS has been mapped on the CTD of Cp[36]. Moreover, Panté et al. showed that *in vitro* phosphorylated rHBc, but also MatC and ImmatC, microinjected into the cytoplasm of *Xenopous laevis* oocytes localize intact at the nuclear side of the NPC[84].

Together these evidences suggest a model where the phosphorylation of Cp and/or genome maturation induces conformational changes that lead to the CTD exposure on the capsid surface allowing recognition of the bipartite NLS by Imp.α that in turn binds to Imp.β. Regulation of nuclear transport through phosphorylation has been described e.g. for SV40, where phosphorylation of the T antigen can either support nuclear import (aa 111/112[185]) or inhibit nuclear import (aa 124 [186]).

Nup153 involvement in HBV interaction with NPC and disassembly

In the nuclear basket, only MatC disassemble and release the genome inside the nucleus. This conclusion relies on the observation that ImmC and MatC produced in HepG2.2.15 cells and added to digitonin permeabilized cells or which were microinjected into the cytosol of *Xenopus laevis* oocytes reach the nuclear side of the NPC but only MatC result in intranuclear capsid staining and nuclear HBV genomes[41]. Cross-linked MatC entered the nuclear basket but stayed arrested[85]. Formation of intranuclear capsids is thus a surrogate marker for capsid disassembly. Arrest of the capsids in the nuclear basket is caused by their interaction with Nup153 as shown by co immune precipitations[85] and partial RNAi silencing of Nup153.

Binding occurred to the C-terminal domain of Nup153 without requirement of the CTD and was~200 times stronger than the interaction of Imp.β with Nup153[85].

Aim of the work

So far only MatC have been characterized with regard to nuclear transport and transport receptors interactions. Nuclear import of MatC is mediated by the classical import pathway and therefore by Imp.α and Imp.β. Whether EmpC use the same import pathway is an open question and nuclear import of EmpC has been essentially unexplored.

MatC have also been characterized to interact with Nup153. Following the interaction with Nup153 the genome is released inside the nucleus, however how the disassembly occurs to enable the liberation of the genome is unknown. We hypothesized that the interaction with Nup153 may enhance the disassembly of the capsid and we aimed to investigate this possibility.

The lack of suitable detection methods has hampered our investigation of the disassembly step, thus we decided to develop a method that would have allowed us to follow the liberation of the genome directly and in real time, and which would allow identifying factors implicated in genome release

.

Materials and Methods

Buffers used

PBS 1 X (pH 7.4)

0.137 M sodium chloride

0.0027 M Potassium Chloride

0.010 M Disodium phosphate

0.0018 M Potassium phosphate

Endogenous polymerase reaction buffer 1 X

0.05 M tris(hydroxymethyl)aminomethane (Tris)

0.04 M magnesium chloride

0.05 M ammonium chloride

IF buffer 1 X in PBS

10% FCS

0.1% Saponin

Ripa buffer

150 mM sodium chloride

1.0% NP-40

0.5% sodium deoxycholate*

0.1% sodium dodecyl sulfate (SDS)

50 mM Tris, pH 8.0

Supplemented with 0.001 M phenylmethane sulfonyl fluoride (PMSF)

SCC 10X pH 7

1.5 M sodium chloride

0.15 M sodium citrate

TAE 50X

2 M Tris

57 % Acetic Acid

0.5 M pH8 Ethylenediaminetetraacetic acid (EDTA)

TGS 10X

0.5 M Tris

1.92 M Glycerol

1% SDS

TNE 1X

10 mM Tris

100 mM sodium chloride

1 mM EDTA

25 % Saccharose

0.75% NP40

Transport buffer (TB) 1X pH 7.3

2 mM magnesium acetate

110 mM potassium acetate

5 mM sodium acetate

1 mM EGTA

Laemml buffer

200mM Tris pH 6.8

8% SDS

40% glycerol

0.4% Bromophenol Blue

TN 1 X

20 mM Tris pH 7.4

140 mM sodium chloride

Hypotonic buffer

5 mM Tris pH 7.4

5 mM potassium chloride

1.5 mM magnesium chloride

0.1 mM EGTA

1 mM DTT and 1mM PMSF

Cell culture

Cell Culture Maintenance

Cell lines used are listed in **Table 1.**

Table 1: Cells lines

Cell line	Reference	Description
Huh-77	Nakabayashi et al., 1982 [187]	Human liver derived cell line. Kindly provided by Jean Rosenbaum, Université de Bordeaux, Bordeaux, France.
HepG2	ATCC HB-8065	Human liver derived cell line. Kindly provided by Jean Rosenbaum, Université de Bordeaux, Bordeaux, France.
HepG2 NTCP	Konig et al., 2014[188]	HepG2 cell line expressing NTCP under the control of a doxycycline inducible promoter. Provided by Dieter Glebe, Justus-Liebig-Universitat Giessen, Germany
HepG2.2.15	Sells et al.,1987[131]	HepG2 cell line with stably integrated HBV genome. Provided by Wolfram Gerlich Hygiene-Institut, Göttingen, Germany.

During the cell culture routine, unless otherwise indicated, cells were maintained in Dulbecco's modified Eagle Medium (DMEM), supplemented with 10% fetal calf serum (GIBCO), 100 U/ mL of penicillin and 100 µg/ mL of streptomycin (PS) (GIBCO), in 5 % v/v CO_2 at 37°C. For culture of HepG2-NTCP or HepG2-NTCP cells, we used the same medium but certified to be tetracycline free. Cells were regularly splitted twice a week in a 2/10 ratio, by firstly washing them in sterile 1 x PBS (GIBCO), detaching them from the cell flask with 0.05% trypsin/EDTA (GIBCO) and diluting the resuspended cells in fresh medium before adding them to a new dish.

Establishment of stable cell lines

HepG2 NTCP OR-DGFP and HuH-7 OR-DGFP were obtained by transduction with OR-DGFP VSV-G pseudotyped lentivectors MOI 2. After 2 weeks and 4-5 passages, expression of OR-DGFP was verified by microscopy, using an epifluorescent Leica DMI6000 B microscope. Cells were sorted by FACS at the Flow cytometry platform (FR TransBioMed) before populations with similar OR-DGFP expression were expanded.

HuH-7 and HepG2 EAP Δε as well as HuH-7 and HepG2 LAP Δε were established by Dr. Marie Lise Blondot, (Université de Bordeaux). Briefly, cells were obtained by transduction with VSV-G pseudotyped lentivectors harboring the EA, LA or Δε construct MOI 1. Double positive cells were sorted by FACS at the Flow cytometry platform (FR TransBioMed).

Transfection

Transfection of EAP or LAP in HuH-7 OR-DGFP cells

2*10^6 cells were seeded on a 10 cm dish in order to be 60-80% confluence the day after. The next day, 2µg of EAP or LAP were transfected using FugenHD transfection reagent. The reaction mix was prepared by adding 100µL of Optimem medium (Opti-MEM, GIBCO) without antibiotics to the plasmid DNA and incubated for 5 min at RT. 12µL of FugenHD were added and the mix was incubated for 30 min at RT. During the incubation time the cell medium was changed with fresh one. Finally, the transfection mix was added drop wise to the cell culture medium. Formation of green dots indicating internalization of EAP or LAP was verified by microscopy after 24 h.

Double transfection of EAP or LAP constructs with Δε in HuH-7 cells

2*10^6 cells were seeded on a 10 cm (diameter) cell plate. The day after 2µg of EAP or LAP and 2µg of Δε plasmid were transfected using FugenHD transfection reagent. 2µg of a DGFP expressing plasmid were used as transfection control. Transfection was performed as described above. Transfection efficiency was analyzed by verifying the expression of DGFP by microscopy with the epifluorescence microscope Leica DMI6000 B.

Transduction

3.0 x 10^4 HuH-7 LAP Δε were seeded into each channel of an ibidi µ slide IV$^{\alpha}$. The day after, cells were transduced with OR-DGFP VSV-G pseudotyped lentivectors MOI 2. Then cells were verified at the epifluorescence microscope Leica DMI6000 B. 48h after transduction, cells were treated with Demecolcine (Sigma Aldrich) 10 µg/mL to arrest cells in mitosis allowing a clear vision of condensed chromatin. To visualize the chromatin, cells were live

stained with Hoechst diluted 1:500 in the culture medium. After staining, cells were rapidly visualized at the microscope.

Capsids preparation

MatC preparation from HepG2.2.15 cell supernatant

Cells were grown in eight T150 flasks until confluence in DMEM complemented with 5% FCS/ 1% PS at 37°C 5% CO_2. Once at confluence the medium was replaced with DMEM 1% FCS and cells were kept for 2 to 3days. Cell culture supernatant was collected and cleared from cell debris at 4000 rpm for 15 min at 4°C (sigma 4-16k centrifuge Swing-out rotor 11150). 38.5 mL were added on top of a 3 mL of sucrose cushion of 1 x TNE/-0.75 % NP40 buffer and centrifuged at 28000 rpm for 22h at 4°C using Beckman SW28ti rotor. Pellets were resuspended in 1 mL 1 x PBS/0.7% NP40, transferred to 1.5 mL Eppendorf tubes and then incubated for 1 h at 37°C to remove the viral envelope, allowing transferring the preparation to a BSL1 lab. Pellets were dissociated with a dissociater-grinder, and incubated overnight under slow agitation. Samples were centrifuged at 13000 rpm for 15 min at 4°C to remove debris and aggregates. Supernatants were then collected, pooled, and added in 5 mL fractions on top of a 1 mL of sucrose cushion 1 x TNE/0.75 % NP40 for centrifugation at 55000 rpm for 2h at 10°C using a Beckman SW55ti rotor. Pellets containing MatC were resuspended in 100 µL of 1 x PBS, followed by centrifugation at 13000 rpm for 15 min at 4°C with a bench centrifuge. The supernatant was stocked at 4°C.

MatC quantification

To quantify the amount of MatC, different dilutions of the production were loaded a 0.7% agarose/1 x TAE gel together with serial dilutions of a rHBc standard. After the electrophoresis in 1 x TAE, proteins in their native state were transferred to a PDVF membrane by capillary transfer using 10 x SCC buffer overnight. Following preactivation of the membrane by 1 min incubation in absolute EtOH followed by 1 min incubation in distilled water and incubation in 10 x SSC he membrane was blocked with 5% Fat Free Milk/ 1 x PBS for 1h. For capsid detection, polyclonal rabbit anti HBV Core Antigen antibodies (Dako B0586) were added in a1:5000 dilution in blocking solution: 5% Fat Free Milk/ in 1 x PBS for 2 h, followed by three washing steps of the membrane in 0.5% Fat Free Milk/1 X PBS.
As secondary antibody, peroxidase conjugated AffiniPure Donkey Anti Rabbit antibody (Jackson ImmunoResearch Laboratories, 711035152) was used in a 1:10000 dilution in

blocking solution for 1h. Afterward, the membrane was washed three times with 0.5% Fat Free Milk/1 x PBS. Bound antibodies were detected by chemiluminescence using the Millipore Immobilion Western Chemiluminescent HRP substrate (Millipore) and a chemiluminescence imaging system (Las 4000). Exposure times were chosen according to the strength of the signal. All steps were performed at RT.

The acquired image was analyzed with ImageJ to obtain the value of the optical density (OD) of the bands corresponding to rHBc and MatC. A standard curve was generated from the bands of the rHBc serial dilution and the amount of the MatC was obtained by interpolation of the MatC OD values from the standard curve.

EmpC preparation

EmpC were kindly provided by Dr. Andris Dishlers of the Latvian Biomedical Research and Study Centre, Riga, Latvia.

RNA was removed from E. coli-expressed capsids (rHBc), using the same method used by Prof. Zlotnicks. Briefly, particles were biochemically dissociated and reassociated as described by Porterfield et al[189].

Cryo-EM and 3D reconstruction

This part of the work was done in collaboration with the team of Dr. Rémi Fronzes at the Institut Européen de Chimie et Biologie (IECB) in Pessac, France.

To obtain the 3D reconstruction, 6.7µg of EmpC in 1 x PBS were used. The cryo fixation was done in ethane using Vitrobot™ (Thermo Scientific). Electron micrographs were taken at the Talos Arctica 200kV-FEG (FEI) electron microscope. Images were acquired with 26 frame number, after 1.32 second of exposure time with a 120000X magnification for 1.24 Armstrong as pixel size. 3D reconstruction was made by the team of Dr. Rémi Fronzes using the software Relion .

Anch HBV virus purification

EAP, LAP and Δε constructs

EAP and LAP were made by Dr. Marie Lise Blondot, (Université de Bordeaux). Briefly, the Anch sequences, 983 bp for the EAP and 399 bp for LAP, were inserted into HBV genome, genotype Adw2. Δε construct was obtained by deletion of the sequence TTCAAG in the ε region

Virus production

Virus harboring the EA sequence were produced in HuH-7 or HepG2 cells stably transfected with EAP and Δε, while virus harboring the LA sequence were produced in HuH-7 or HepG2 cells stably transfected with LAP +Δε. Cells were grown in eight T150 flask until confluence in DMEM/5% FCS/1% PS and were incubate with 1% DMSO at 37°C 5% CO_2. Once at confluence the medium was replaced with DMEM 1% SCF and cells were cultivated under the above conditions for 2-3days. Cell culture supernatant was collected, centrifuged at 4000 rpm for 15 min at 4°C (sigma 4-16k centrifuge Swing-out rotor 11150) to remove cells debris before the supernatant was passed through a 0.25 µm filter. Fresh medium was added to the cells and collected once again after 2-3 days.

Heparin Affinity Chromatography

Heparin Affinity Chromatography was performed using the Akta chromatography system (GE Healthcare) equipped with 5 mL HiTrap Heparin HP columns (GE Healtcare), according to[69]. Runs were live monitored following the UV absorption at 254 and 280nm (OD_{254} and OD_{280}). Up to 450 mL were applied per 5 mL Column. Then the column was washed with 5 column volumes of 1 x TN. Elution was made using a linear gradient of NaCl over 10 column volumes ranging from 140 to 2140 mM in 20 mM Tris-Cl (pH 7.4) NaCl.

Virus quantification

The amount of produced virus was quantified by quantitative PCR (qPCR) measuring the genome equivalents.

DNA was purified from 50µL of virus preparation using the kit High Pure Viral Nucleic acid (Roche 11858874001) following the manufacturer instructions.

For the qPCR reaction Perfecta Sybr Green super mix Quantabio was used. Primers were used accordingly to Jursch et al.[190]. They were designed against the X region: HBV X2s [nucleotide (nt.) 1413-1436]: GAC GTC CTT TGT YTA CGT CCC GTC and HBV X2as (nt. 1601-1578): TGC AGA GGT GAA GCG AAG TGC ACA.

The program of the light cycler CFX96 was the follow: denaturation and activation of the Polymerase at 95°C for 3min, and 39 cycles of amplifications made by 95°C for 10sec and 60°C for 30sec.

As standard, six dilution of a plasmid containing the full length HBV genome was used.

Proteins Purification

Imp.α purification

Purification of Imp.α was made by Dr Mildred Delaleau. The protein was expressed as His tag protein using E. coli XL1-blue transformed with pQE70-Imp.α (Kindely provided by Dirk Gorlich). The His fragment was eluted from Ni++ agarose according to manufacturers' instructions with 500 mM NaCl and 500 mM imidazole

Imp.β purification

Purification of Imp.β was made by Dr Mildred Delaleau. Imp.β was expressed as His tag protein using E. coli M15 e transformed with pQE60-Imp.β (kindely provided by Dirk Gorlich)[191]. The His fragment was eluted from Ni++ agarose according to manufacturers' instructions and dialyzed against: 50 mM Tris-HCl, 50 mM NaCl 5% glycerol, 250 mM sucrose and 2 mM DTT

His-C_2Nup153 fragment

The C-terminal domain of Nup153 was purified as described by Schmitz et al[85]. It was expressed as His tag protein using *E. coli Rosettte* transformed with pET28-153C2 (kindely provided by K. Ullamn). The His fragment was eluted from Ni^{++} agarose according to manufacturers' instructions and dialyzed against 0.2 M NaCl, 0.1 M NaH_2PO_4 pH7.6

Gel electrophoresis, transfers and antibody staining

Sodium Dodecyl Sulfate - Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Proteins were diluted in 1 x Laemmli buffer/100 mM DTT and boiled for 5 min at 100°C. Polyacrylamide gels consisted of staking gel (4% acrylamide/bis-acrylamide solution (Biorad), 125 mM Tris pH 6.8 and 0.1% SDS), the separating gel of 11% acrylamide/bis-acrylamide, 375 mM Tris pH 8.8 and 0.1% SDS. Gel electrophoresis was run at constant voltage of 90 V in 1 x TGS. A protein ladder was loaded onto the gel to determine the MW of the proteins (PageRuler Plus Prestained Protein Ladder; Thermo Scientific).

Western Blot (WB) after SDS PAGE

Separated proteins were transferred to a nitrocellulose membrane by wet electrotransfer using a Trans-Blot Cell apparatus (Biorad) and filled with transfer buffer (composition) at 600 mA

for 90 min at RT. Following the transfer, the membrane was blocked with 5% Fat Free Milk in 1 X PBS for 1h RT. Antibody staining and detection were performed as described before. Antibodies used and their dilutions are summarized in **Table 2**.

Table 2: Antibodies used for Western blot (WB) and Immunofluorescence (IF)

Primary Antibody	Immunogen	Organism	Dilution	Source
Anti HBc Ag	HBV capsid	Rabbit	WB 1:2000	Dako B0586
Fab 3105	HBV capsid	Mouse	IF 1:200	Institute of Immunology Co LTD 2AHC21
Anti Karyopherin alpha	Imp.α	Mouse	WB 1:1000	BD bioscience 610985
	Imp.β	Rabbit	WB 1:4000	
Anti GST	GST	Mouse	WB 1:500	BD bioscience 554805
Mab 414		Mouse	IF 1:200 WB 1:2000	Ab 24609
Anti His	His Tag	Mouse	WB 1:1000	Novagen 707964
Anti NTCP	NTCP	Rabbit	WB 1:250 IF 1:700	Sigma HPA042727
Anti Actin	Actin	Mouse	WB 1:2000	MerkMillipore Mab1501
Secondary Antibody	**Immunogen**	**Organism**	**Dilution**	**Source**
Alexa fluor 647 donkey anti mouse IgG	IgG of mouse	Donkey	1:500	Invitrogen A31573
Alexa fluor 555 donkey anti mouse IgG	IgG of mouse	Donkey	1:500	Invitrogen
Peroxidase conjugated anti mouse	IgG of mouse	Donkey	1:10000	Jackson ImmunoResearch 715035151
Peroxidase conjugated anti rabbit	IgG of rabbit	Donkey	1:10000	Jackson ImmunoResearch 715035151

In vitro interaction studies

Native Agarose Gel Electrophoresis (NAGE) of Multiprotein complexes

Interaction between MatC and EmpC with Importins

2 ng of MatC or EmpC were incubated alone or with either 8 µg of Imp.β, either 4 µg of Imp.α, or both, in the presence of 1 x TB supplemented with 2 mM DTT unless otherwise indicated, for 2 h at 37°C to allow the formation of complexes.

The mixes were separated by NAGE, followed by capillary transfer and detection of the proteins of interest.

Titration assay

A constant amount of the His-C_2Nup153 fragment was incubated with different amounts of EmpC for 2 h at RT in 1 x TB. The ratio used ranged from one His-C_2Nup153 fragment for each Cp to one His-C2Nup153 fragment for 0.03 Cp The reaction was carried on in 20 µL of finale volume. After the incubation, 5 µL of each condition were loaded onto a 0.7% agarose gel for a NAGE followed by capillary transfer and anti His staining.

Imp.β Pull down

10 µL of Glutathione Sepharose beads (GE Health Care Life Science, 17075601) were taken per sample and washed three times with 1 x PBS. 2.5µg of GST-IBB was added to each sample and incubated with the beads for 1 h at RT in a rotating wheel. Next, beads were washed three times in 1 x PBS and 1 µg or 0.5 µg of Imp.β were added. Incubation was carried on overnight at 4°C in the rotating wheel. Beads were then spin down and washed three times in 1 x PBS. Beads were resuspended in 20 µL of 1 x Laemmly buffer/100 mM DTT and boiled for 5 min at 100°C. Separation of the bounded fraction was done by SDS page as described above. Imp.β and GST-IBB were detected by Western blot as described before.

Radiolabeling

MatC radiolabeling

The genome of MatC was radiolabeled using the endogenous polymerase reaction. The viral polymerase can incorporate into the viral genome radiolabeled nucleotides that enter inside the capsid trough the fenestration present on the capsid surface. Previous experiments of the lab showed that four nucleotides were incorporated in average.

10 µL of MatC, corresponding to 200 ng, were incubated together with 1 mM dATP, dGTP, dTTP 1 x endogenous polymerase buffer and 10µCi [α-^{32}P] dCTP for 2 h at 37°CAfter the incubation time non radio labeled dCTP was added at 1 mM final concentration. MatC were purified from the excess of nucleotides by column purification using the kit Illustra Probe Quant G-50 Micro Columns and following manufacturer instructions.

To verify the successful labelling, an aliquot was separated by NAGE. The gel was fixed in 5% Trichloroacetic acid (TCA) for 30 min and dried with Biorad 583 electrophoresis gel dryer prior to exposure to a Fuji screen imaging plate. Signals on the screen were visualized using a Biorad PharosFX phosphorimager and Quantity One.

Viral genome purification by phenol/chloroform

Phenol was added to the samples in the same volume as the sample, samples were vortexed vigorously and centrifuged one minute at the maximum speed. After the centrifuge samples were separated in an aqueous and in an organic phase. The aqueous phase was recovered, while the organic one was discarded

Chloroform was added to the sample was added to the samples in the same volume as the sample, then samples were vortexed vigorously and centrifuged one minute at 17000g. As above, the aqueous phase was recovered.

One out of ten sample volumes of 3 M sodium acetate, ph5.2, and EtOH 100% ice cold were to the samples together with 1 µL of Glycogen as DNA carrier. Samples were incubated overnight at -20°C. Following the incubation, samples were centrifuged at maximum speed for 1h at 4°C. The pellet was rinsed with 70 % EtOH and centrifuged again for 30 min. Finally, the supernatant was discarded and the pellet was left air dry before being suspended in the appropriate buffer of water

Southern Blot

Capsid production and quantification was done as described above. Viral DNA was purified by phenol/chloroform extraction as described in the above paragraph. 10 µL of purified DNA were loaded onto 0.7% agarose gel. After the electrophoretic run, DNA was transferred into a nitrocellulose membrane under vacuum for 1h and 30 min. Cross-linking was carried on for 2h at 80°C. For the hybridization, the membrane was preincubated for 30 min at 65°C in Rapid Hyb buffer (GE-Healtcare). Incubation with radiolabeled DNA probes, prepared according to Wittkop et al.[192], was maintained for 2h at 65°C. The membrance was washed with SCC buffer with decreasing salt concentration. The first washed was carried on at RT for

10min with 2 x SSC / 0.1% SDS, the second at 65°C for 15min with 1x SSC /0.1% SDS while the the third one at 65°C for 15 min with 0.1 % SSC 0/0.1 % SDS. The final rinse was made with 0.1 x SSC at RT for 10 min. The membrane was let dry at 37°C and finally imaged with Biorad PharosFX phosphorimager and Quantity One.

Viral genome degradation by Nuclease S7

15 ng of MatC containing α-^{32}P were incubated in the presence or absence of Nuclease S7/ TB at pH 7 if not otherwise stated, for 2 h at 37°C if not otherwise stated. 15 U/µL Nuclease S7 (Roche 10107921001) corresponding to 20 µg, was added and $CaCl_2$ was adjusted to 3 mM. After the incubation, an aliquot of the sample was passed through a Wallac 1400 liquid scintillation system for separating non-incorporated nucleotides, followed by determination of the radioactivity. Then, samples were separated by NAGE, fixed in 5% trichloroacetic acid (TCA), dried, and exposed to a Fuji screen imaging plate. Fuji screen was then imaged with Biorad PharosFX phosphorimager and Quantity One.

Immobilization of the His-C_2Nup153 fragment on Nickel beads

10 µL of beads (Admatech beads Ni magnetic 04500) were aliquoted for each sample in a lower binding Eppendorf cup, afterwards 20 µL corresponding to 8 µg of His-C_2Nup153 were added to the beads and incubated for 4 h at RT. At the end of the incubation time, beads were sedimented, the supernatant was removed and the beads were washed in 1 x PBS three times. Then, beads were saturated with bovine serum albumin (BSA) 0.1% over night at 4°C.

Immobilization of the His-C_2Nup153 fragment on Nickel plates

For this assay, Pierce Nickel Coated Plates (Thermo Scientific) were used. 2 µg of His-C_2Nup153 were added to the plate and incubated for 1 h at RT. At the end of the incubation, the plates were washed three times with 1 x PBS.

Transport assay

Transport assay in Digitonin permeabilized cells

12 mm coverslips were washed with 2-propanol for 10 min and subsequently dried under the hood. Then they were coated with 100 µL of 0.01% (v/v) Poly-L-Lysine (Sigma) and incubated for 30 min at RT. Finally, they were washed with sterile water, and EtOH, and dried before use.

For each sample, 8.0*10^4 HuH-7 cells were seeded. The day after, cells were washed three times before being permeabilized with Digitonin. For this, 50 µL of 0.005% (v/v) Digitonin (Calbiochem 300410) in transport buffer were added per coverslip for 10 min at 37°C. After cell permeabilization, coverslips were washed three times with 1x TB at RT. Then, coverslips were incubated for 30 min in 1 x TB for 30 min at 30°C in humidified environment. After, the reaction mix was added to each sample. Components of the reaction mix are summarized in **Table 3**.

The transport reaction was carried on for 30 min at 37°C. Finally, cells were washed three times in 1 x TB.

Table 3: Reaction Mix for transport assay in Digitonin permeabilized cells

Sample	Substrate µg	Substrate µL	Energy Mix µL	RRL µL	TB 1X µL
MatC pH 5,5	0.2	7	4	13,25	25,75
MatC pH 7	0.2	7	4	13,25	25,75
BSA-NLS		1	4	13,25	31,75

Energy Mix	Stock	µL
ATP 50mM	50mM	1
Creatine Phosphate (CP)	100mM	2.5
Creatin Phosphate Kinase (CPK)	2000U/mL	0.5

Transport assay in isolated nuclei

Nuclei were isolated from HepG2 cells following the protocol described by Chang et al., 2012[193]. Briefly, cells were incubated with hypotonic buffer on ice for 1 h, harvested by scraping and homogenized by several passages through a 27 gauge needle. Then cells were centrifuged at 500 g for 5 min at 4C°(sigma 4-16k centrifuge Swing-out rotor 11150) and the pellet was washed once in hypotonic buffer. After isolation nuclei were verified for their integrity with a light transmission microscope and counted. ~2*10^4 nuclei were aliquoted per sample, the transport mix was added and the transport reaction was carried on for 30 min at 37°C. Sample mixes are summarized in **Table 4**. At the end of the incubation, nuclei were separate by centrifugation at 800g for 10 min. Then, both, the supernatant and the nuclei pellet were treated with 15 U of Nuclease S7, corresponding to 20 µg, with 0.1% Triton as

final concentration to allow the nuclease to enter in the nucleus, for 30 min at 37°C. Afterwards, samples were treated with 5µg of proteinase K, 6 mM EGTA and 2% SDS for 1 h at 65°C to inactivate the Nuclease S7 and remove capsid and polymerase from the viral DNA. Viral nucleic acids were extracted by phenol/chloroform and precipitated by EtOH. The radioactivity was determined before samples were loaded onto 0.7% agarose/1 x TAE gel. Following electrophoresis, the gel was fixed dried as described above, exposed to a Fuji screen imaging plate. After incubation time in the cassette, which depended on the amount of radioactivity in the samples, the Fuji screen was imaged with Biorad PharosFX phosphorimager and Quantity One software.

Table 4: Reaction Mix for transport assay in Digitonin permeabilized cells

Sample	Substrate µL	Substrate µg	Energy µL	RRL	WGA (2mg/mL)	TB 1X
MatC-WGA	12	60ng	3,6	12		22,4
MatC-WGA +S7	12	60ng	3,6	12		22,4
MatC+ WGA	12	60ng	3,6	12	20	2,4
MatC+WGA+S7	12	60ng	3,6	12	20	2,4

Immunofluorescence (IF)

Cells were fixed using 4% (v/v) paraformaldehyde/1 x PBS for 15 - 20 min at RT and then washed with 1 x PBS. Cells were then incubated with 100 µL IF-buffer per coverslip for 15 - 20 min at RT. For antibodies staining, primary antibodies were diluted in IF-buffer (see Table 2) and centrifuged at 17000g for 2 min to remove aggregates. Next, each coverslip was incubated with 100 µL of the primary antibody solution at 37 ºC for 1 - 2 h in a humid chamber, then washed three times for at least 15 min in 1 x PBS. Secondary antibodies were diluted in IF-buffer and centrifuged at 20000g for 2 min as well and each coverslip was incubated with 100 µL of the secondary antibody solution 37 ºC for 30min -1 h in a humid chamber. Coverslips were washed as before. Antibodies used and their dilutions are summarized in **table 2**. Finally, coverslips were rinsed in distilled water and absolute EtOH and dried. Moviol mounting medium containing DABCO as anti bleaching agent and 1 µg/mL DAPI (Sigma) was used for mounting.

Doxycycline (dox) inducible NTCP expression

Western Blot to verify NTCP expression followed by dox Treatment

2*10^6 HepG2 NTCP cells were seeded into a poly L Lysin-pretreated 10 cm dish. The day after, 5µg/mL doxycycline (doxycycline hydrocloride, Fisher scientific BP2653-1) were added. Doxycycline induction was maintained for 24h, 48h or 72h. At each time point cells were lysed by washing with ice cold 1 x PBS, and then incubated on ice with 1 mL of Ripa Buffer for 2 min. Without removing the Ripa Buffer, cells were detached using a cell scraper, transferred to a 1.5 mL Eppendorf Tube, and incubated for 30 min on ice with constant agitation. Finally, the debris was removed by centrifugation for 20 min at 12000 rpm in a bench centrifuge. The supernatant was kept and stored at -20°C

Proteins in the cell lysate were quantified using the BSA Pierce kit according to the manufactures instructions. 20 µg of total cell lysate added per lane of an SDS PAGE and a Western blot using anti NTCP antibodies were performed as described before.

Immunofluorescence to follow NTCP expression followed by dox Treatment

8*10^4 HepG2 NTCP cells were seeded into a 12mm coverslip pretreated with Poly L Lysin for each sample. The day after, 5µg/mL of Doxycycline (Doxycycline Hydrocloride Fisher scientific BP2653-1) were added to the cell medium. Doxycycline induction was maintained for 24h, 48h 72h or 96h. At each time point cells were fixed with 4% (v/v) paraformaldehyde in 1 x PBS for 15 - 20 min at RT and then washed with PBS 1X. Immunofluorescence using anti NTCP was done as described above.

Anchor HBV capsid lipofection and infection

Capsid lipofection

Lipofection was performed using the pulsin reagent (Polyplus 501-01).

10 µL of pseudo HBV capsid harboring the EAP construct, corresponding to 5 ng of capsids were incubated with 100 µL of Hepes (20 mM pH 7.4) for 5 min at room temperature. Following the incubation 2 µL of pulsin were added to the mix. The mix was incubated 20 min at RT before being added dropwise to 2 x 10^5 of HuH-7 OR-DGF seeded the day before in a µ- dish 50 nm (ibidi). Pictures were taken each 12 min over 4h.

Infection

8 x 10^4 HepG2 NTCP OR-DGF cells were seeded in each well of a 12 well plate falsk. The day after NTCP expression was induced by doxycycline (5 µg/mL) and cells were infected with HBV Anch (EA contruct) at MOI 10. Pictures were regularly taken (each our during the first 6 h post infection and each day for three days post infection) at the epifluorescent microscope using 20 x objective to monitor the kinetic of the formation of green dots.

Microscopy images acquisition and processing

Images acquisition

Images were acquired either using a laser scanner confocal microscope Leica DMI6000 TCS SP5 or an epifluorescence wide field microscope Leica DMI6000 B. The first was provided by the Bordeaux Imaging Center (BIC) platform.

Image processing

Images were processed using Image J software (National Institutes of Health).

Results

EmpC interaction with Imp.β

The phosphorylation of the CTD of Cp as well as genome maturation have been associated with its exposure on the capsid surface [36,122,194]. It is hypothesized that this structural change is caused by a reduced affinity between CTD and the encapsidated nucleic acids[40], allowing that CTD passes adjacent holes in the center of the capsids' quasi 6-fold symmetry. Further, it has been shown that *in vitro* phosphorylated rHBc and purified MatC are targeted to the NPC via the canonical nuclear import pathway, by an eight aa-long NLS present on the CTD of Cp[83]. This however requires exposure of the NLS on the capsid surface, allowing binding by Imp.α that in turn binds to Imp.β by an Imp.β binding domain (IBB), which is classically ~38 aa long. It remains however unclear how much of the CTD is exposed. Noteworthy all import studies done until now contained Imp.α and Imp.β so that a full exposure of the CTD with an intact IBB could not be excluded.

During HBV infection, EmpC are produced as a result of the spontaneous association of Cp, which is produced in excess with respect of the amount needed [111]. Little is known about the role of EmpC during HBV infection, and the interaction between EmpC and import receptors is an open question that is essentially unexplored. We investigated this question in collaboration with the group of Prof. Adam Zlotnick (Indiana University, USA).

Through a series of size exclusion chromatography experiments, the group of Prof. Zlotnick showed that the full length Cp and EmpC are able to bind Imp.β alone without Imp.α [110]. This was complemented by results of our group showing that the entire CTD binds Imp.β preferentially to Imp.α or Imp.α/Imp.β.

Through size exclusion chromatography our collaborators further showed that more than more than 50 molecules of Imp.β could bind to each EmpC. Considering that the expected binding sites for Imp.β are only 30, they proposed that during the capsid breathing, Imp.β molecules could also be swallowed inside the capsid shell. This was investigated by cryo electron microscopy (Cryo-EM).

Figure 10A shows a Cryo-EM micrograph of EmpC incubated in NaCl buffer with a ratio of 205 molecules of Imp.β for each capsid. Four different types of capsids were found: normal empty capsid (white arrow), defective particles (stars) a minor population of T3 particles (black arrowed) and darker particles (black arrows). The latter showed a darker interior that might be due to an internal content. The presence of an internal content was confirmed by image reconstruction of the darker particles (**Figure 10B**). As this contrast was not visible in

absence of Imp.β, the inner density represents internal Imp.β allowing the calculation that the internal volume could accommodate 48 closely packed Imp.β copies. Given that 30 external Imp.β binding sites were found, a theoretical number of 78 capsid-bound Imp.β molecules can be concluded. This is consistent with the more than 50 molecules founded experimentally.

To visualize the β-binding, which is linked to nuclear import, the structure of Imp.β was fitted into the external Imp.β-mediated densities. The 3D structure in **Figure 10C** shows the external presence of Imp.β (red) arranged at the quasi-six-fold vertex and thus directly opposing the 12-15 Å holes in the capsids, through which the CTDs might be traslocated.

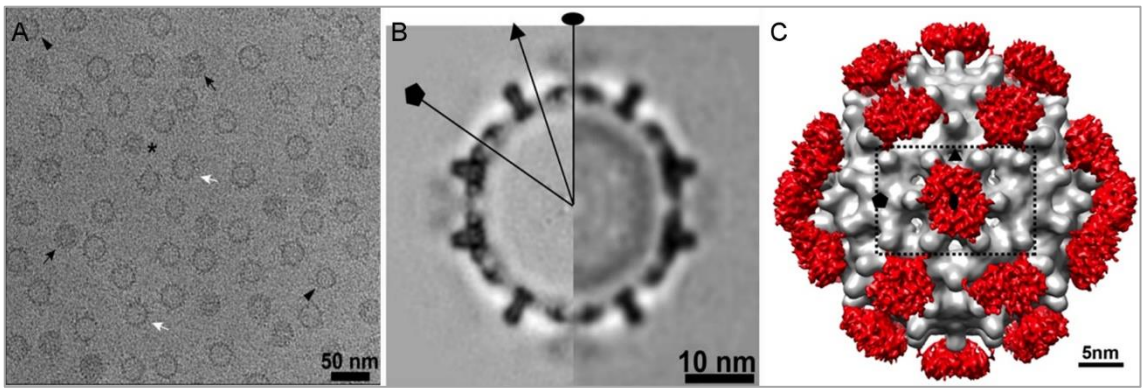


Figure 10: Cryo-EM micrograph and 3D structure of EmpC bounded to Imp.β in NaCl
A. Representative Cryo-EM micrograph of EmpC in complex with Imp.β. with a ratio of 205 molecules of Imp.β for each capsid. White arrows indicate empty capsids, stars indicate defective particles, black arrowheads indicate T3 capsids and black arrows indicate darker particles. **B**. 3D reconstruction of darker particles on the right and all particles from the left viewed from the central section. The reconstructions were determined to 8.9 Å resolution for all particles and 15.9 Å resolution for the dark particles. Oval shows the quasi-six-fold symmetry axis, triangle the 3-fold symmetry axis and pentagons the 5-fold symmetry axis **C**. Imp.β density in red superimposed to the 3D structure of EmpC in grey. Oval triangles and pentagon represent the quasi-six, three and five symmetry respectively. The additional surface-attached Imp.β-derived density is found on top of the holes in the quasi-six-fold symmetry [110].

EmpC Characterization

3D structure of EmpC

Complementary to these findings, we investigated the interaction between EmpC and Imp.β by biochemical assays. During these analyses we also investigated the MatC and Imp.α and Imp.β binding as control.

First, we characterized the EmpC preparation we would use during all the following experiments. We analyzed the *E. coli*-expressed capsids, which were biochemically dissociated and reassociated for removal of the *E. coli* RNA encapsidated upon expression. These particles were kindly provided by Dr. Andris Dishlers of the Latvian Biomedical Research and Study Centre, Riga, Latvia.

As shown in **Figure 11A**, the capsid preparation was devoid of significant amounts of contaminating proteins showing the one dominant Cp band of 21.5 kDa in Coomassie stain after SDS PAGE. **Figures 11B**, **C** show a migration of these capsids on an agarose gel under native conditions (native agarose gel electrophoresis, NAGE) followed by Coomassie stain (**B**) and ethidium bromide stain (**C**). In addition to EmpC, *E. coli*-expressed capsids without RNA removal were added as control. The panel shows that both capsids migrated as a defined band indicating no major differences in their diameter or surface charge. The loss of ethidium bromide stain in panel **C** indicates the absence of at least ds RNA into which the dye intercalates.

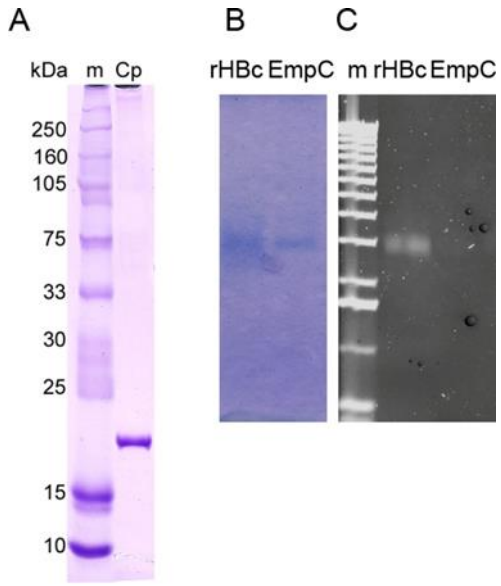


Figure 11: Separation of 2 µg of core particles..
Separation of 2 µg of core particles by **A**. SDS PAGE, **B**, and **C**, NAGE. **A**, **B**: Coomassie brilliant blue stain, **C**: ethidium bromide stain. EmpC: empty capsids, rHBc: control capsids purified from *E. coli*. m: protein molecular weight markers (**A**) dsl DNA marker (**C**). The figure shows the purity of the capsid preparation and confirms the absence of RNA in EmpC.

Next, we analyzed the EmpC by Cryo-EM and by reconstructing their 3D structure. Achieving a resolution of 6.5Å (**Figures 12A, B, C**). **Figures 12D** and **E** show a similar reconstruction of RNA filled capsid from Wang et al., 2012[195].

In both reconstructions, the characteristic spikes, formed by the two α helixes of the assembly domain of two monomers, can be seen on the capsid surface, as well as the fenestrations at the 3, 5 and quasi-6-fold axes. As reported by others, the hole at the 5-fold symmetry is smaller than the others[37]. EmpC showed in contrast to the non-reassembled capsids no additional densities inside the capsid shell, confirming the absence of nucleic acids, leading to the conclusion that the density is derived from packaged RNA. The RNA (**Figure 12E**, in blue)

was condensed at the inner walls of the capsid, which agrees with the RNA-binding of the CTD.

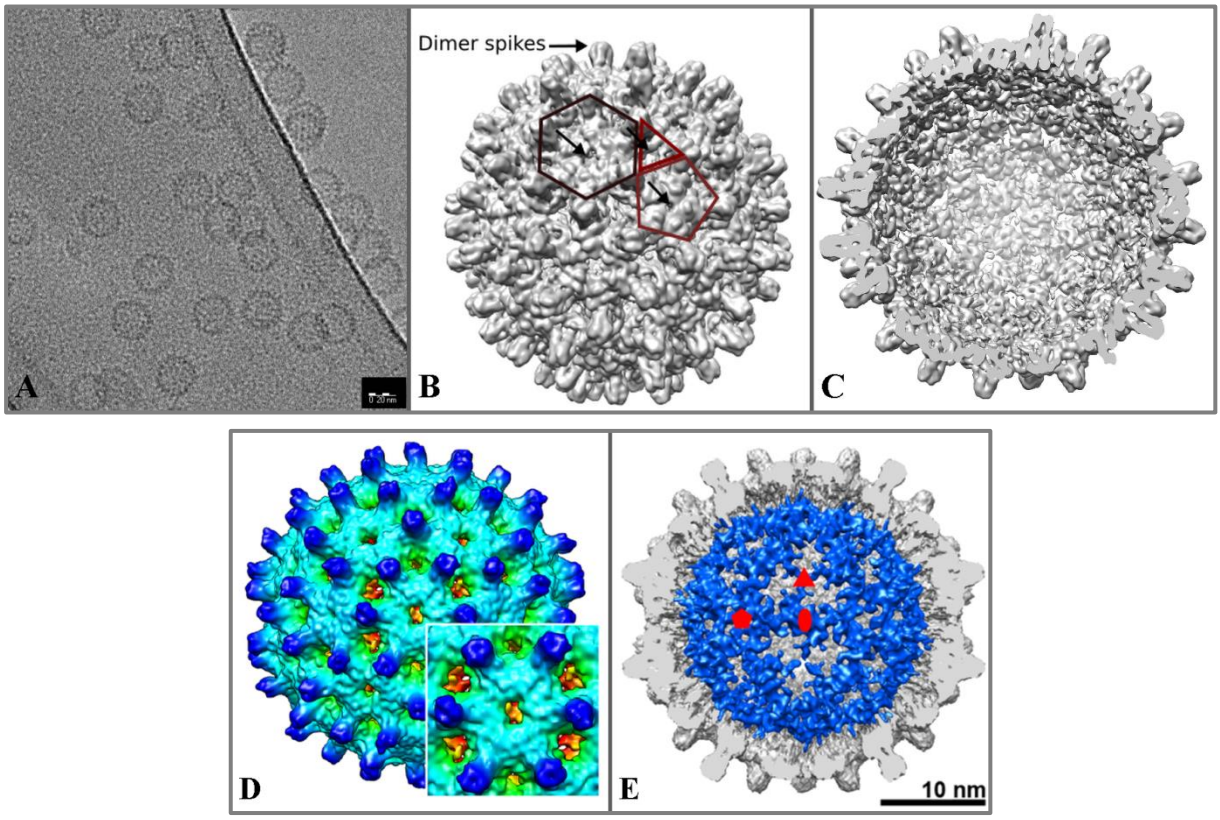


Figure 12: EmpC characterization by Cryo-EM and 3D reconstruction.
EmpC were incubated 2h in 1 x PBS at RT before the samples were Cryo fixed in liquid ethane. **A**. Images were acquired with 26 frames of 1.32 second of exposure time with a 120000 x magnification for 1.24 Angstrom as pixel size. Darker grey represents the carbon phase while the lighter grey represents the vitreous phase. **B**. 3D structure of EmpC showing 5-fold axes (pentagon), 3-fold axes (triangle), and quasi-6-fold axes (hexagon). Arrows inside pentagon, triangle and hexagon indicate the fenestrations of the capsid. **C**. EmpC seen from the inside. **D** and **E**. RNA filled capsid from Wang et al.,2012[195]. In **E** central section in which in red the symmetry axes and in blue the RNA

SDS-page of EmpC in the presence of 2 mM Dithiothreitol (DTT)

To investigate the disulfide bonds in our preparation, we subjected reduced and non-reduced capsids to an SDS-PAGE followed by Coomassie brilliant blue staining. **Figure 13** shows that for both samples, which were exposed to SDS without heat, a major band is present at 17.5 kDa regardless to the presence of 2 mM DTT. This indicates that the Cp is still folded to a certain extent but that the majority of Cp is not covalently linked. This hypothesis is in agreement with the additional bands at 30 kDa in the presence of DTT and 27 kDa in the absence, likely representing Cp dimers as observed in other capsid preparations. Without DTT, this band migrated faster than in the presence of DTT, suggesting a more compact form, likely by keeping the two CTDs of a dimer together. This is in agreement with the observation

that the C terminal disulfide bonds are better accessible. Higher assembled forms could not be detected indicating that the Cp were not linked by aberrant disulfide bonds between Cys 185 as observed previously[40]. The experiment further indicates a high stability of the Cp dimers as SDS alone did not result in complete unfolding resulting in disulfide bond access for DTT. Quantification of surface plots (**Figure. 13B**) revealed that 10 and 2% of Cp are dimers. The signal strengths did not allow a precise quantification so that a significant difference between the samples cannot be supposed. However, the data allow the conclusion that the majority of Cp are not linked by disulfide bonds and that at least the majority of capsids are thus equivalents to intracellular capsids.

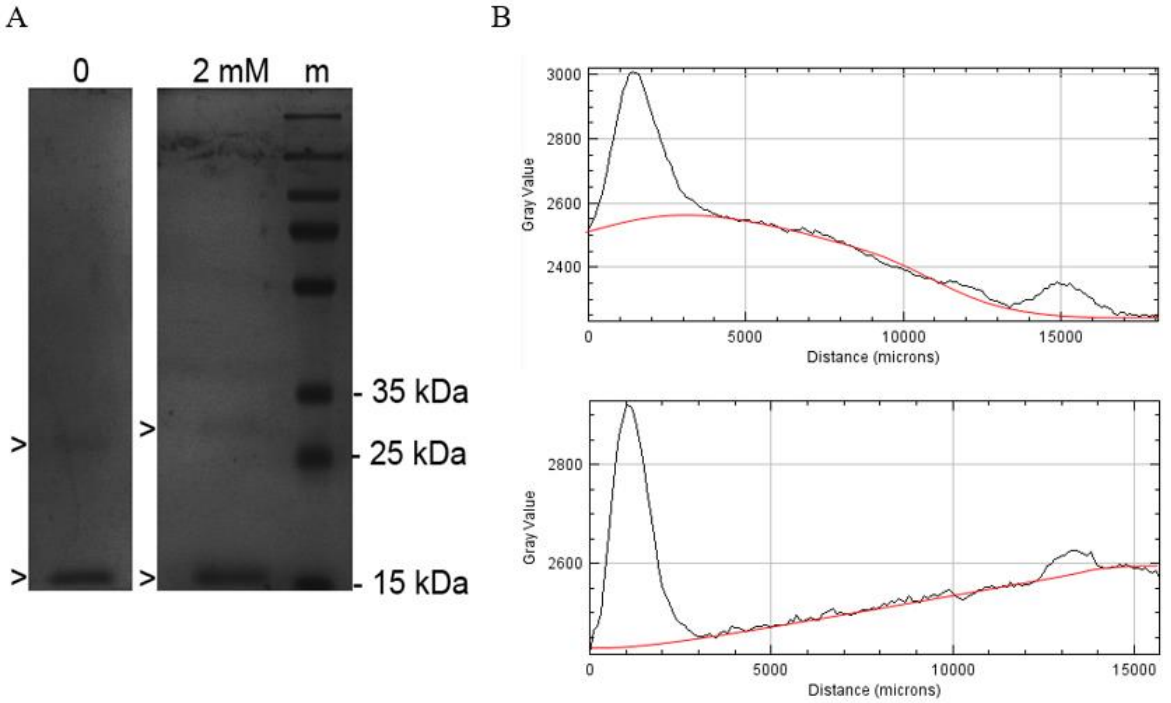


Figure 13: SDS PAGE of EmpC
A. 1 µg of EmpC were diluted in 1 x PBS buffer containing 2 mM or no DTT. Afterwards, samples were resuspended in 1 x Laemmli buffer and separated on an 11% acrylamide gel. Proteins were stained with Coomassie brilliant blue. M: molecular weight marker; the MW of the bands is given on the right. 0: no DTT, 2 mM: 2 mM DTT. The arrows indicate the migration of putative Cp monomers and Cp dimers with a migration corresponding to 17.5 kDa and 30 kDa (2 mM DTT) and 27 kDa (no DTT). **B**. Surface plot of A. The background baseline is depicted as a red line. Upper panel: no DTT, lower panel: 2 mM DTT. The plots allow the estimations that 10% (no DTT) and 2% (2 mM DTT) of the Cp are dimers.

NAGE of EmpC

NAGE separates macromolecules by their quotient of surface charge and diffusion diameter. This, for instance, does not allow the separation of HBV capsids with a T=3 and a T=4 symmetry as the larger diameter of the T=4 capsids (36 nm versus 32 nm) is compensated by

less charge-exposing Cp (240 versus 180). Attachment of other proteins change, however, the surface charge and – if a significant binding occurs – also the diameter.

We first analyzed the migration of EmpC, rHBc, and MatC (**Figure 14**). **Figure 14A** shows that MatC migrated slightly faster than rHBc and EmpC (**Figure 14B**).

As all three capsid types exhibit identical diameters we conclude that the surface charge of MatC differs, which can be explained by the fact that MatC and EmpC sequences were from different HBV genotypes (MatC: D; EmpC: A).

Of note, a second, slower migrating band was observed for rHBc, which is well reported in the literature. It corresponds to two core particles, which are linked by RNA.

The attachment of macromolecules to MatC, which were co-purified, was considered to be unlikely as they would cause an increase in the diameter and thus a retarded migration. Confirming the same capsid structure, we observed that all capsid types showed the same reactivity using a polyclonal antibody (DAKO), which mainly reacts against the tips of the spikes formed by the two α helices of the Cp dimers.

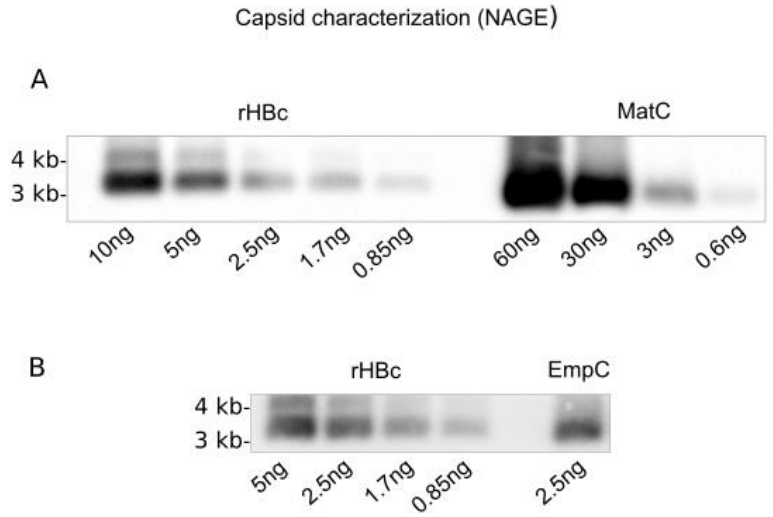


Figure 14: NAGE of MatC and EmpC
Different amounts of MatC as indicated below the panel **(A)** and 2.5 ng of EmpC **(B)** were loaded onto a 0.7% Agarose gel in 1 x TAE together with different dilutions of rHBc (indicated below the panel) as migration control. Proteins were transferred to a PDVF membrane and polyclonal anti capsid antibodies (Dako B0586) were used for detection. The figure shows that MatC migrated slightly faster and that all the capsids type have the same antibody reactivity using a polyclonal antibody (DAKO)

Imp.α and Imp.β purification and functionality

Imp.α and Imp.β gel migration properties

To verify the correct purification, different dilutions of both Imp.α and Imp.β were analyzed by Western blot using anti Imp.α and Imp.β antibodies. **Figures 15A** and **C** show that both proteins migrated at their expected size. We then investigated the migration property under non-denaturing conditions by NAGE followed by antibody staining to Imp.α and Imp.β, showing that Imp.β migrated as capsids while Imp.α migrated much slower (**Figures 15B** and **D**).

Imp.β functionality

To confirm functionality of the purified Imp.β, we performed a pulldown assay using Glutathione S-transferase fused to IBB (GST-IBB). Different amounts of Imp.β were incubated with constant amounts of GST-IBB, which was immobilized on Sepharose Beads. Input fractions as well as beads bounded fractions were analyzed by Western blot with anti-Imp.β and anti-GST antibodies.

In **Figure 15E**, the upper panel represents the different amounts of Imp.β added to the pull down, the bottom panel shows the beads-bounded fractions after the pull down. Anti-GST antibody staining was used as loading control.

The result shows that ~400 ng of Imp.β were precipitated by 2.5 µg GST-IBB when 1 µg of purified Imp.β was added. Given that GST-IBB was in excess in the reaction mixture, we conclude that~ 40% of Imp.β was functional.

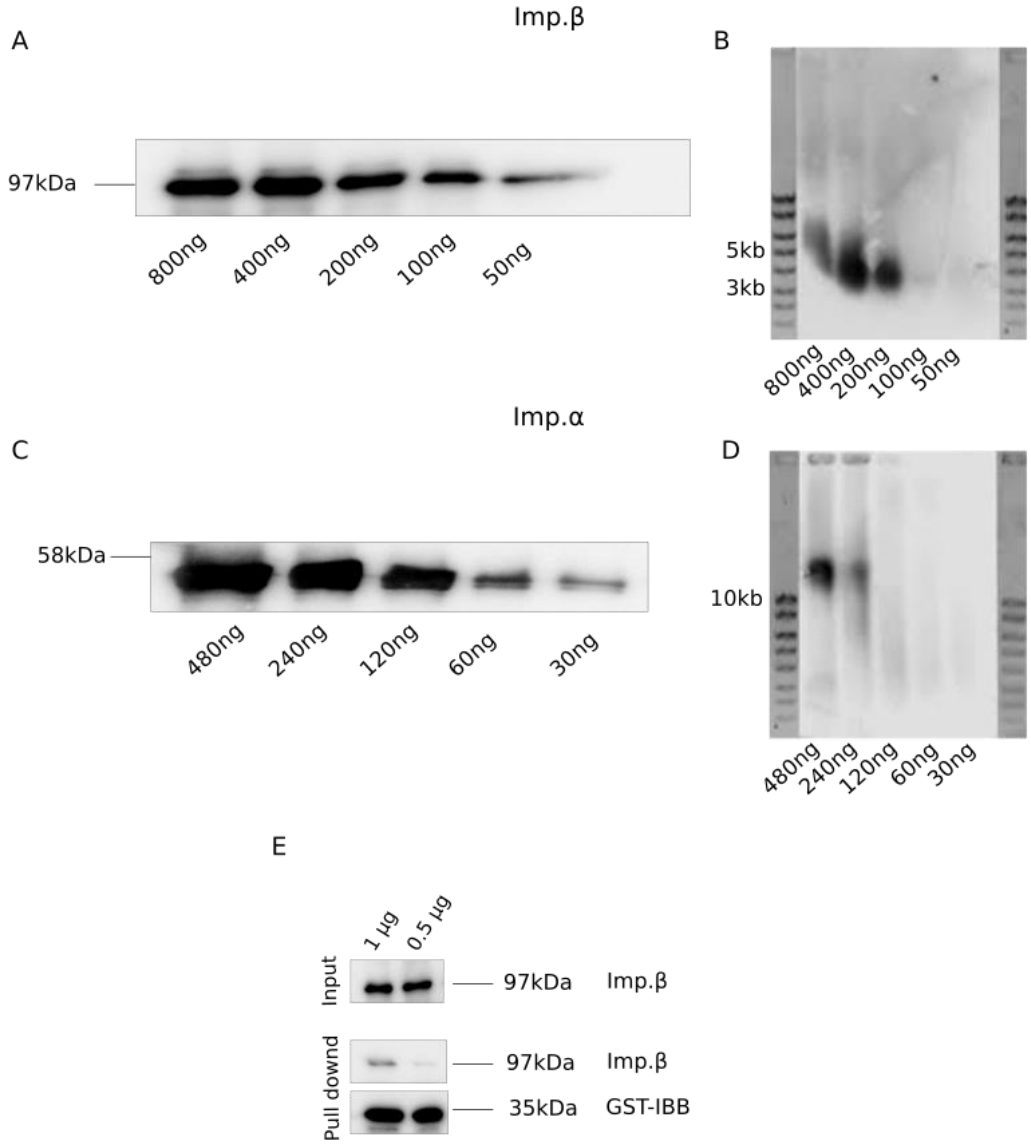


Figure 15: Migration property of Imp.α and Imp.β and Imp.β functionality
Western Blot after **A**. SDS PAGE and B**.** NAGE of Imp.β. The amount of Imp.β is indicated below each lane. Imp.β was detected using rabbit anti Imp.β antibodies. Western Blot after **C**. SDS PAGE and **D.** NAGE of Imp.α. The amount of Imp.α is indicated below each lane. Imp.α was detected using rabbit anti Imp.α antibodies **E.** Pull down of Imp.β. 1 µg or 0.5 µg of Imp.β were added.to 10µL of Glutathione Sepharose beads preincubated with 2.5 µg GST-IBB. The upper panel shows the input represented by the different amount of Imp.β added to the beads. The lower panel shows the beads fraction: Imp.β was detected using rabbit anti Imp.β antibody while GST was detected using anti GST mouse antibody.

Interaction between EmpC and Imp.β by NAGE

Two ng of MatC and EmpC were incubated for 2h at 37°C with either 8 µg of Imp.β alone, or with 4 µg of Imp.α or with both importins. These corresponded to a stoichiometry of 887

Imp.β or Imp.α molecules per Cp. After NAGE, proteins were transferred to a PDVF membrane by capillary transfer and the change in the migration pattern was analyzed by antibody staining against the capsid protein using anti capsid antibodies.

Figure 16 shows that the presence of Imp.α alone or in combination with Imp.β induced a retardation of MatC migration, suggesting that Imp.α binding to the capsid NLS does not require stabilization by Imp.β. The absence of retardation when Imp.β alone was added showed that only an NLS became exposed but not a functional IBB. The retardation was however limited despite of the slow migration of Imp.α alone, suggesting that just a few Imp.α molecules bind per capsid.

In contrast, migration of EmpC was retarded by Imp.β alone (**Figure 16A**) resulting in a smear (**Figure 16B**). Retardation and inhomogeneity indicate that either the number of Imp.β molecules bound to capsids differs, increasing their size, or that a part of EmpC disassembled, which also causes a slower inhomogeneous migration[36]. We further observed that the presence of Imp.β resulted in increased signals by the polyclonal antibody but only when using EmpC.

Collectively, these results confirm that, in contrast to MatC, EmpC are able to bind Imp.β directly thus expose the entire CTD. We further concluded that Imp.β-binding induced conformational changes that lead to increased antibody binding. This conclusion is consistent with a capsid destabilization and the entry of Imp.β into capsids' lumen as observed by cryo-EM.

Figure 16C shows that in the absence of 2 mM DTT, EmpC incubated with Imp.β at the same concentration as before migrate as EmpC alone. This implies that the reducing environment provided by the DTT helps the formation of the EmpC-Imp.β complex.

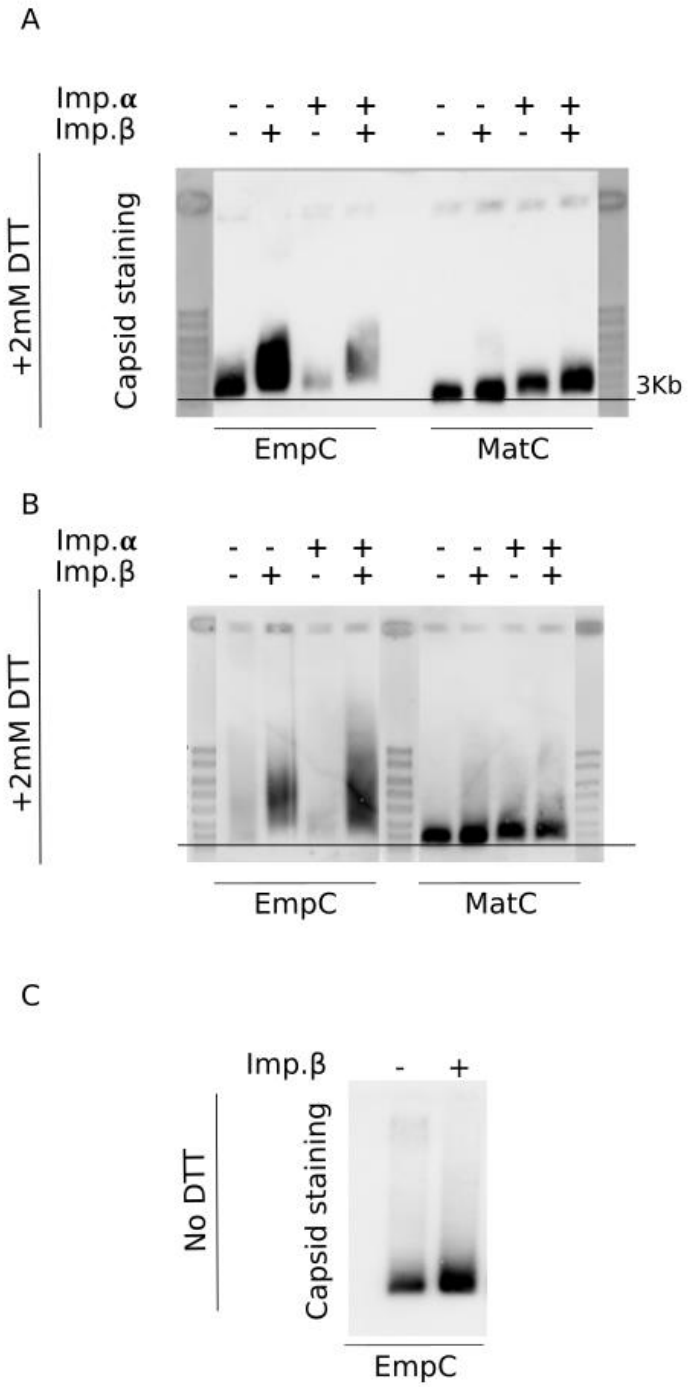


Figure 16: NAGE of MatC and EmpC complexes with Importins
A. EmpC and MatC were incubated for 2h at 37°C either alone either with Imp.α or Imp.β or both (combination showed above each panel). Change in the migration pattern was analyzed by NAGE followed by capillary transfer and antibody staining using anti HBcAg. **B.** The same experiment was repeated leaving the gel migrate for a longer time. **C.** The same experiment was carried on incubating capsid with importins in transport buffer without DTT.

Analysis of HBV capsid disassembly

As for other DNA viruses, HBV needs to release its genome inside the nucleus. Previous studies have shown that both, ImmC and MatC reach the nuclear basket of the NPC, but only MatC disassemble[41,82]. Both the type of capsid interact with the C-terminal portion of Nup153[85] in an CTD-independent manner. The disassembly of MatC occurs via Cp hexamers, which disassemble to Cp dimers. The capsid subunits diffuse into the nucleus, where they reassemble to genome-devoid capsids. However, the exact mechanism by which the viral capsid disassembles to release the viral genome remains unclear.

Viral genome degradation by Nuclease S7

To investigate the disassembly, we used an *in vitro* model based on the radiolabeling of the viral genome and the Nuclease S7. Labelling by ^{32}P was chosen as capsid concentrations allowing EtBr staining of the DNA can be hardly achieved by eukaryotic expression. The experiment relies on the idea that the Nuclease S7 has no access to the viral genome when the capsid is intact, being too large to pass by the fenestrations of the capsid surface (**Figure 17**). Nuclease S7 is a globular protein having a calculated Stokes diameter of 3.6 nm (http://www.calctool.org/CALC/prof/bio/protein_size). It must be thus concluded that the removal of one Cp hexamers leaving a hole of ~8 nm would be sufficient for allowing entry S7 nuclease entry into capsids' lumen.

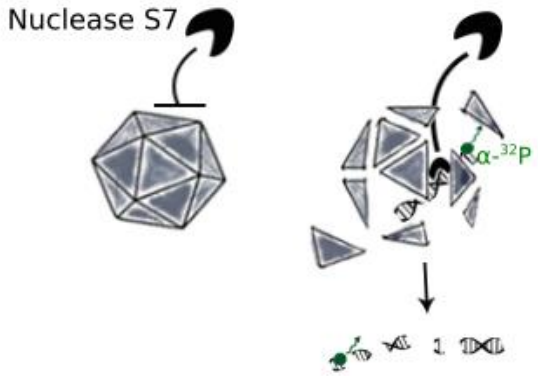


Figure 17: Nuclease S7 access to viral genome
The Nuclease S7 has no access to the viral genome when the capsid is intact, being too large to pass by the fenestrations present on the capsid surface, on the contrary when the capsid is disassembled or partially disassembled, the Nuclease S7 gain access to the viral genome depredating it and, when the viral genome is radiolabeled, its degradation can be used as read out for the disassembly step.

Proof of concept: stability of the capsid at 37°Cand 56°C

To prove our method, we purified MatC capsids from the supernatant of HepG2.2.15 cells and we radiolabeled the viral genome inside using α-^{32}P dCTP which is incorporated into the viral genome by the activity of the viral polymerase. We incubated the radiolabeled MatC with Nuclease S7 in transport buffer for 2h at 37°C and 56°C, and we analyzed the degradation of the α-^{32}P labeled genome by phosphoimaging after NAGE.

Figure 18 shows that nuclease S7 almost completely degraded the viral genome at 56°C, while no degradation was observed at 37°C.

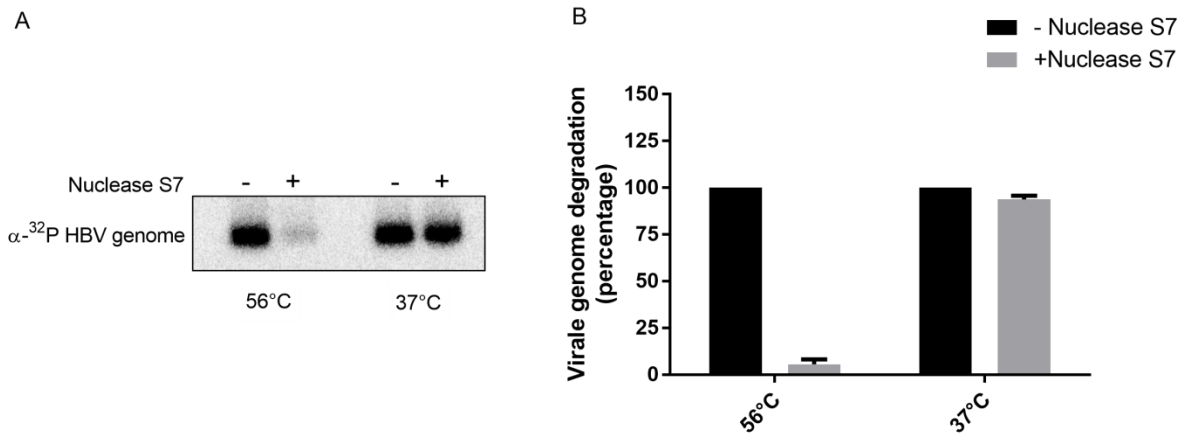


Figure 18: Viral genome degradation at 56°C by Nuclease S7
A. MatC capsids were purified from HepG2.2.15 and their genome was radio labeled with α-^{32}P dCTP exploiting the polymerase endogenous reaction. MatC with α-^{32}P labeled genome were incubated in the presence or absence of Nuclease S7, in transport buffer at pH7, for 2h at 37°C and 56°C. Samples were loaded on a NAGE and viral genome degradation was determined by phosphoimaging. **B**. Quantification of viral genome degradation was done using ImageJ after background subtraction. Normalization was achieved by using the negative control without Nuclease S7 as 100% value. Error bars represent the quantification made from two independent experiments.

Stability of the capsid at pH 7 pH6.5 and pH5.5

The so-called internal recycling pathway leading to cccDNA accumulation in infected cells via newly synthesized capsids from the cytosol is – despite of shown in textbooks – not established for HBV. Further, the entry pathway of HBV in terms of virus exposure to mildly acidic pH (~ 6.5) is unclear. To analyze if the pH has an impact on capsid stability favoring genome liberation we investigated the genome accessibility for nuclease S7. We subjected the capsids with α^{32}P labeled genome in transport buffer to pH 7, pH 6.5 and pH 5.5 for 10 min at 37°C before the incubation with the Nuclease S7 for 2h at 37°C at pH 7.

Figures 19A and **B** show that the capsids remained stable at pH 7, 40% of viral genomes were degraded when capsids were incubated at pH 6.5 and 60% of the genomes were

degraded at pH 5.5. This shows a pH-dependent capsid stability and allows drawing the conclusion that a moderate capsid acidification upon virus entry could increase infectivity. For investigating if this had an effect on genome transport into the nucleus we performed a transport assay using digitonin permeabilized HuH-7 cells, subjecting capsids, which were preincubated with the different pHs. As read-out, we used the arrival of intranuclear capsids/Cp, which is a surrogate marker for capsid liberation. As control of the transport reaction we used BSA coupled to NLS, which was efficiently imported. Nuclear import occurred actively through the nuclear pores as wheat germ agglutinin (WGA) blocked the reaction[196,197] (**Figure 19C**, **lower panels**). **Figures 19C** (**upper panels**) and **D** shows that the same amount of capsids were imported, regardless of the pH treatment (~30%). This suggests that low pH-induced destabilization does not affect the capsid arrival in the nucleus and consequently not enhance genome liberation from the capsids. Considering the stability of the genomes at 37°C, these findings further indicate that capsid disassembly must be strongly enhanced by a cellular factor.

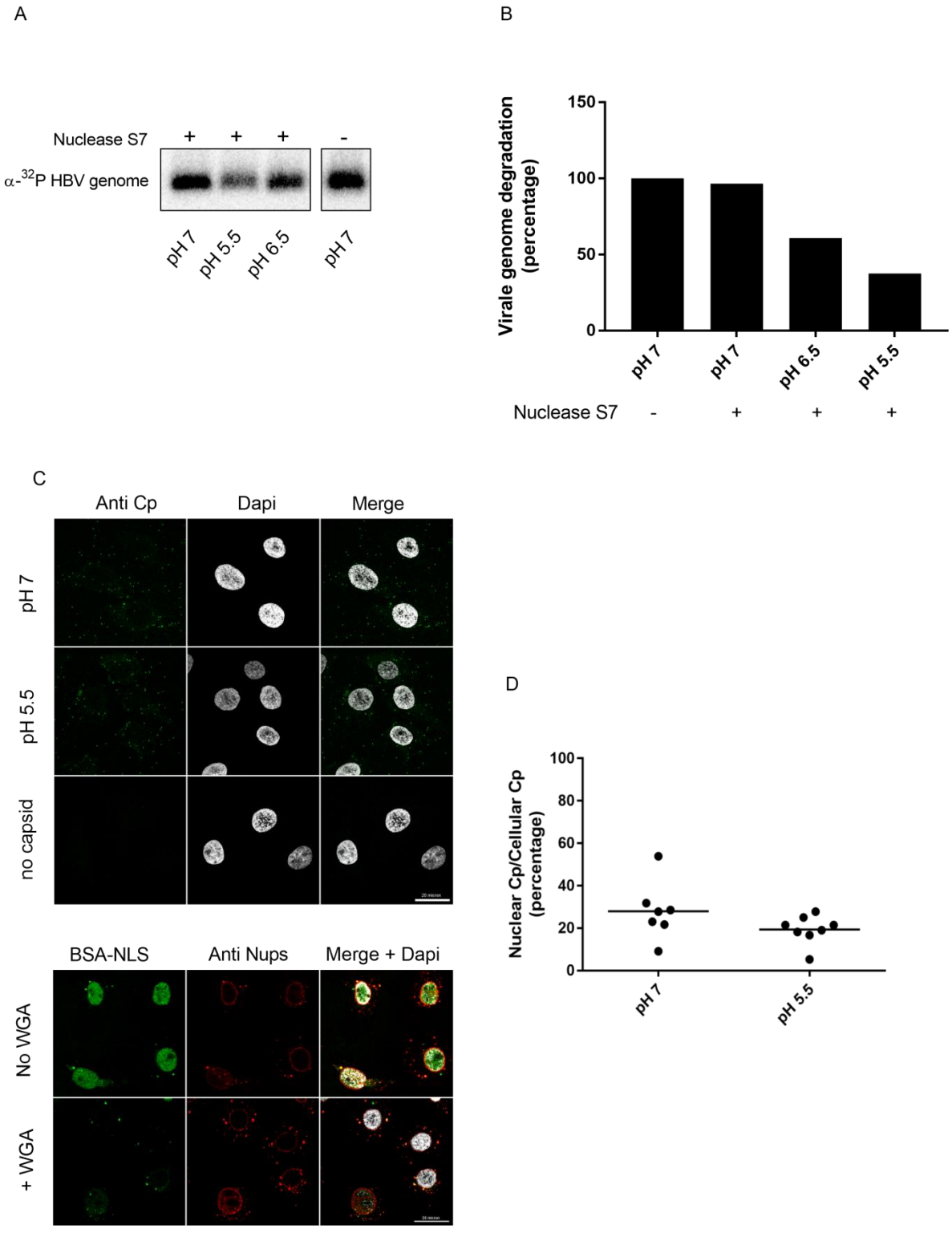


Figure 19: Capsid stability and viral genome degradation by Nuclease S7 at different pHs
A. MatC capsids were purified from HepG2.2.15 and their genome was radio labeled with α-^{32}P dCTP .MatC with α-^{32}P labeled genome were incubated in transport buffer at different pH for 10 min before being incubated in the presence or absence of the Nuclease S7. After the incubation, samples were loaded on a NAGE and viral genome degradation was determined by phosphoimaging. **B**. Quantification of viral genome degradation was done using ImageJ after background subtraction. Normalization was achieved by using the negative control without Nuclease S7 as 100% value. **C**. Representative confocal images of Digitonin permeabilized cells after transport assay. MatC were incubated in transport buffer at different pHs for 10 min before the transport reaction

(specific pH treatment is shown on the left of each row). MatC were stained with Fab 3105 acting against the HBV capsid, Nucleoporins were stained with Ab414 antibod. BSA-NLS was coupled to the A488 fluorophore (green) and DAPI staining was used to visualize the chromatin. Images were acquired using a 63X object. Maximal projection images are shown for the capsid part, while focal plane is shown for the BSA-NLS. **D**. Quantification of HBV capsid inside the nucleus was done using Image software, by counting nuclear dots corresponding to the capsid inside the nucleus and normalizing for the total number of cellular dots. Length bar 20 µm.

Capsid disassembly during transport assay with isolated nuclei

To reduce the number of potential cellular interaction partners involved in capsid disassembly, we replaced the permeabilized cells by isolated nuclei. We incubated MatC containing ^{32}P labeled genome with a nuclei preparation from HepG2 cells in the presence or absence of WGA and in the presence or absence of the Nuclease S7 for 30 min at 37°C. After incubation, we separated the nuclei fraction from the supernatant by sedimenting the nuclei. The pellet thus contained capsids inside the nucleus but also those attached to the inner and outer part of the NPC, or other parts of nuclear envelope. Both the fractions were treated by proteinase K, followed by phenol-chloroform extraction for purifying the viral genome. **Figure 20A** shows the nuclei preparation by light microscopy. As shown in **Figure 20B**, the majority of the non-degraded genomes remained in the supernatant. In this fraction the Nuclease S7 had a limited effect on the viral genome degradation (~30% of viral genome degradation). In the nuclear fraction, the Nuclease S7 had a stronger impact of more than 50% of genome degradation (**Figure 20C**). We interpret this finding in that capsid opening does not exclusively occur in the nuclear basket but also happens at an earlier step of nuclear transport, possibly at the outer NPC surface. Considering the need of Nup153 for genome arrest and its potential function in genome release[85], we conclude that the externally exposed Nup153 fraction could be involved. Moreover, in the viral genome degradation was decreased in both fractions by WGA suggesting that glycosylated Nups enhance the disassembly process.

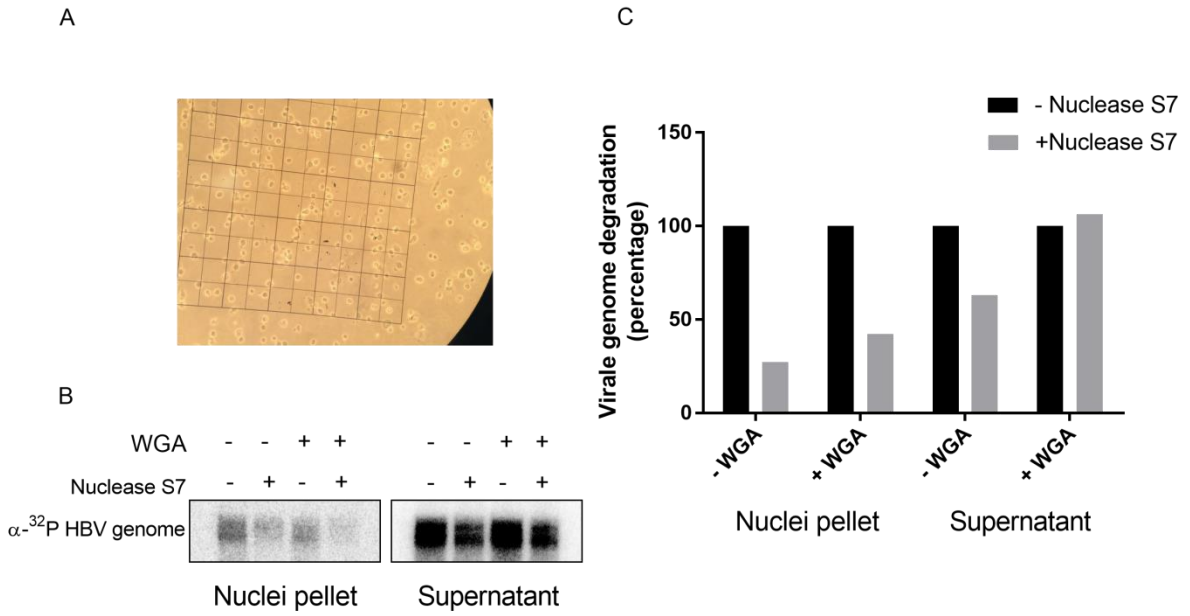


Figure 20: MatC disassembly during transport assay with isolated nuclei
A. Representative picture of Nuclei purified from HepG2 cells. Nuclei were imaged trough a light transmission microscope. **B**. MatC capsids were purified from HepG2.2.15 and their genome was radio labeled with α-^{32}P dCTP MatC with α-^{32}P labeled genome were incubated with the isolated nuclei in the presence or absence of wheat germ agglutinin (WGA) and in the presence or absence of the Nuclease S7 for 30 min at 37°C. At the end of the incubation the nuclei fraction was separated from the supernatant by centrifugation at 8000g for 10 min. Both the fractions were subjected to proteinase K digestion and to phenol-chloroform to purify the viral genome α-32P labeled. Finally, samples were loaded on a 0.7% agarose. Viral genome degradation was determined by phosphoimaging. **C**. Quantification of viral genome degradation was done using ImageJ after background subtraction. Normalization was achieved by using the negative control without Nuclease S7 as 100% value

Role of Nup153 C-term during capsid disassembly

Characterization of the binding between EmpC and Nup153

To explore the possibility that Nup153 impact on capsid opening, we characterized the binding between EmpC and a soluble fragment of Nup153 by NAGE and cryo-EM. We used EmpC because of their structural similarity with MatC and the need for the high amount of capsid required for cryo-EM. As soluble fragment we used the His-tagged C_2 fragment, which corresponds to the C-term portion of Nup153 implicated in capsids binding (His-C_2Nup153)[85]. This fragment contains most of the ~ 30 FXFG repeats and corresponds to the aa 992-1219[197].

To characterize the binding between EmpC and His-C_2Nup153, we analyzed the binding ratio between the two partners. In a titration assay, we incubated a constant amount of His-C_2Nup153 with a decreasing amount of EmpC, for 2h at 37°C, and we analyzed the complexes by NAGE. As shown in **figure 21A**, His-C_2Nup153 migrates towards the anode, in the opposite direction than the capsids. **Figure 21B** shows the dilution of Cp and the constant amounts of Nup153-C2 fragments by Coomassie staining after SDS PAGE.

Figure 21C shows the NAGE of the capsid His-C_2Nup153 fragment complexes. At ratio 4 Cp: 1 His-C_2Nup153 one band was present migrating to the cathode, as the capsids do. No band was visible at the migration of His-C_2Nup153 suggesting that all C_2 fragments were bound to the capsid. At a ratio of 1:2 two bands were present: one migrating in the direction of the capsid and the other migrating like the C_2 fragment. At lower ratios, the C_2-like migration occurred only. We interpret this finding in that from a ratio of 1:1 Cp per Nup153-C_2 fragment on, all capsids were covered with the Nup153 fragment. Consistently, approximately half of the Nup153-fragments were associated with capsids at the 1:2 ratio, leading to the conclusion that one Cp binds to one Nup153 fragment.

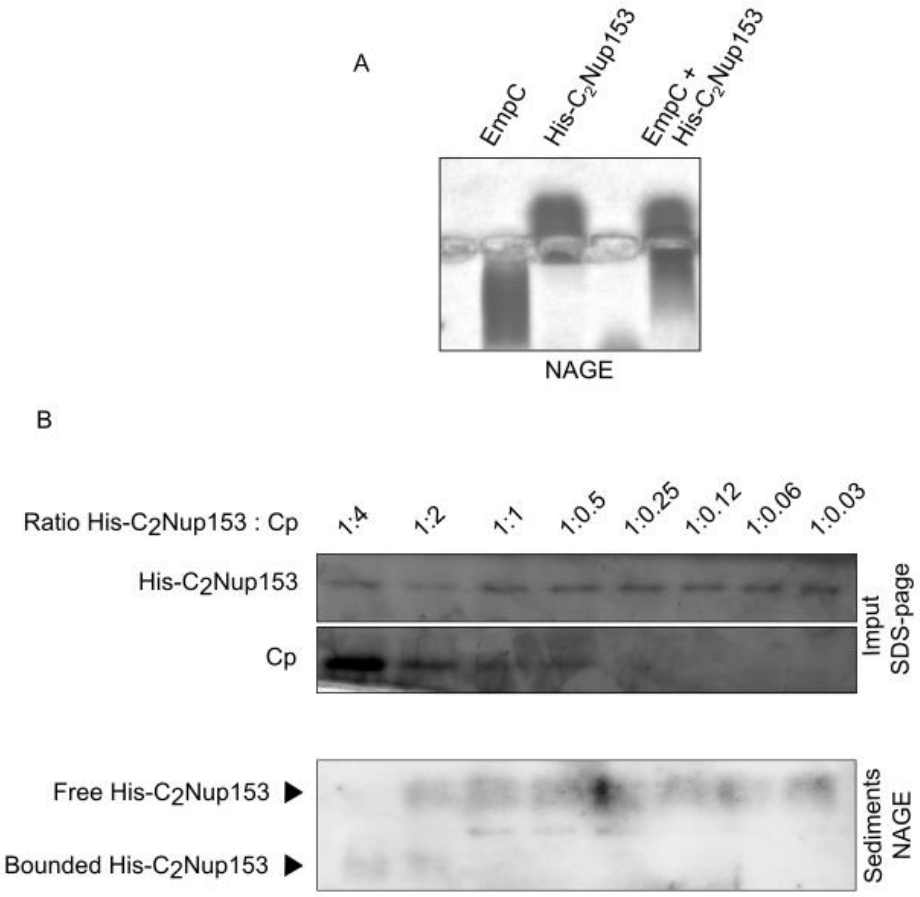


Figure 21: NAGE of EmpC and His-C_2Nup153 and titration assay
A. 10µg of His-C_2Nup153 and 3 µg were loaded onto 0.7%agarose gel in TAE1X, either alone either together after 1h incubation at 37°C to allow the formation of the complexes. NAGE was performed as described before. After the electrophoresis, the gel was stained with Comassie blue and sequentially destained in destaining solution until clear bands over a dark background were observed. **B.** Titration was made keeping constant the amount of His-C_2Nup153 and decreasing the amount of EmpC. Therefore, His-C_2Nup153 fragment and EmpC were incubated together at different ratio (indicated above the upper panel) for 2h at RT. At the end of the incubation, an aliquot was taken from each sample as control and immediately loaded on acrylamide gel for SDS page followed by SyproRed staining (Upper panel). NAGE was performed as described before and was followed by antibody staining with Anti His Novagen 707964 directed against the His tag of the His-C_2Nup153 fragment. The experiment was made twice with the same outcome.

To characterize the binding sites and to possibly obtain 3D structures we analyzed cryo-EM images of EmpC bounded to His-C_2Nup153 at 1:1 ratio.

The EM micrographs in **figure 22** show a different behavior of EmpC when in complex with the His-C_2Nup153fragment (**Figure 22A**): capsids were no longer attracted to the carbon phase, as shown for EmpC alone (**Figure 22B**), but diffused in the vitreous ice phase. This indicates a difference in the capsid surface charge, confirming the difference in the migration behavior during NAGE. However, we were not able to obtain regular additional Nup153-caused densities suggesting that binding is heterogeneous.

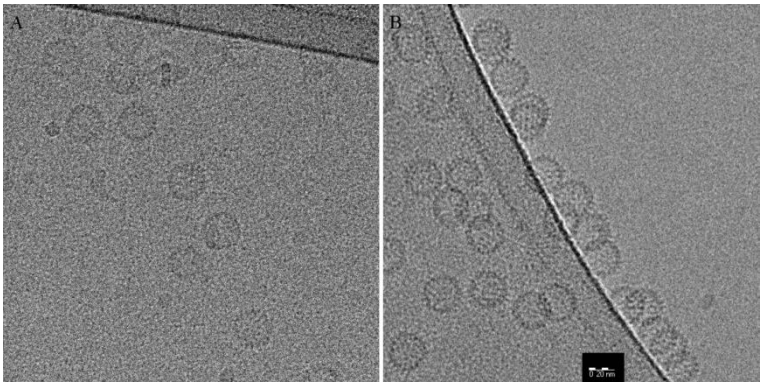


Figure 22: Cryo-Em images of EmpC and EmpC bounded to the His-C_2Nup153 fragment
EmpC were incubated in 1x PBS 2h at RT, either alone either with His-C_2Nup153 fragment. Afterwards samples were cryo fixed in liquid ethane. Images were acquired with 26 frames of 1.32 second of exposure time with a 120000 x magnification for 1.24 Angstrom as pixel size. Darker grey represents the carbon phase while the lighter grey represents the vitreous phase. In **A**. EmpC with the His-C_2Nup153 fragment are shown while in **B**. EmpC alone are shown to compare.

Role of the C-term portion of Nup153 in solution during capsid disassembly

For further evaluating a potential function of the C-terminal portion of Nup153 in enhancing capsid disassembly, we incubated MatC with ^{32}P labeled genome with Nuclease S7, in the presence or absence of His-C_2Nup153 for 2h at 37°C in transport buffer at pH 7. We observed slight degradation in the presence of Nuclease S7 that was not increased in the presence of the C-terminal portion of Nup153 (**Figure 23**). We thus concluded that at least under these conditions the C-terminal portion of Nup153 did not impact the accessibility of the viral genome.

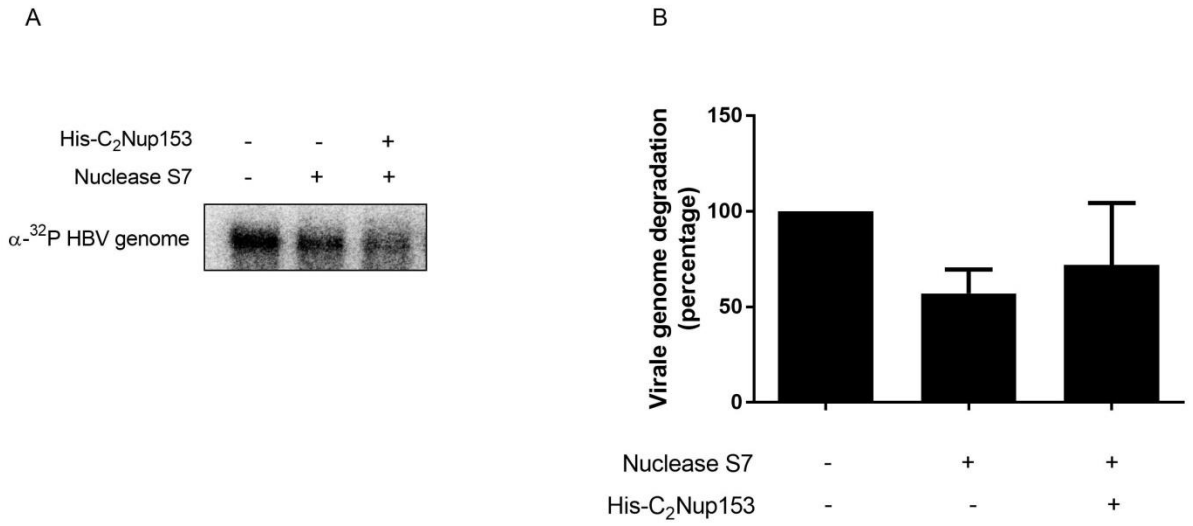


Figure 23: Viral genome degradation by the Nuclease S7 in the presence of soluble His-C_2Nup153 fragment

A. MatC capsids were purified from HepG2.2.15 and their genome was radio labeled with α-^{32}P dCTP exploiting the polymerase endogenous reaction. MatC with α-^{32}P labeled genome were incubated in the presence or absence of Nuclease S7, and in the presence or absence of the soluble His-C_2Nup153 fragment in transport buffer pH7, for 2h at 37°C. After the incubation, samples were loaded on a NAGE. Viral genome degradation was determined by phosphoimaging. **B**. Quantification of viral genome degradation was done using ImageJ after background subtraction. Normalization was achieved by using the negative control without Nuclease S7 as 100% value Error bars represent the quantification made from two independent experiments.

Role of the C-term portion of Nup153 fixed on Nickel beads during capsid disassembly

According to the current model, the capsids should release the genome within the nuclear basket but not already inside the cytoplasm where free Nup153 molecules are present. Thus, we wondered whether having the C_2 fragment immobilized on a solid phase might alter the geometry of the binding and enhance the disassembly of the capsid. With this aim we immobilized the His-C_2Nup153 on Nickel beads.

Firstly, we verified the successful binding of His-C_2Nup153 to Nickel beads by a pull down followed by Western blot and anti Nup staining. As input control, we used different amount of the C_2 fragment. **Figure 24A** shows that, for each dilution most of the fragment was precipitated on the beads fraction indicating the correct immobilization of the fragment to the beads.

To test the accessibility of the viral genome, we repeated the same experiment as before by incubating MatC containing ^{32}P labeled genome with Nuclease S7, in the presence or absence of immobilized His-C_2Nup153 for 2h at 37°C in transport buffer at pH 7. We observed only a slight degradation by Nuclease S7 in the presence of the C-terminal portion of Nup153 (10%), very similar to the degradation induced by the soluble fragment (**Figure 24B** and **C**). Hence,

we concluded that in these conditions the C-terminal portion of Nup153 did not significantly affect the accessibility of the viral genome.

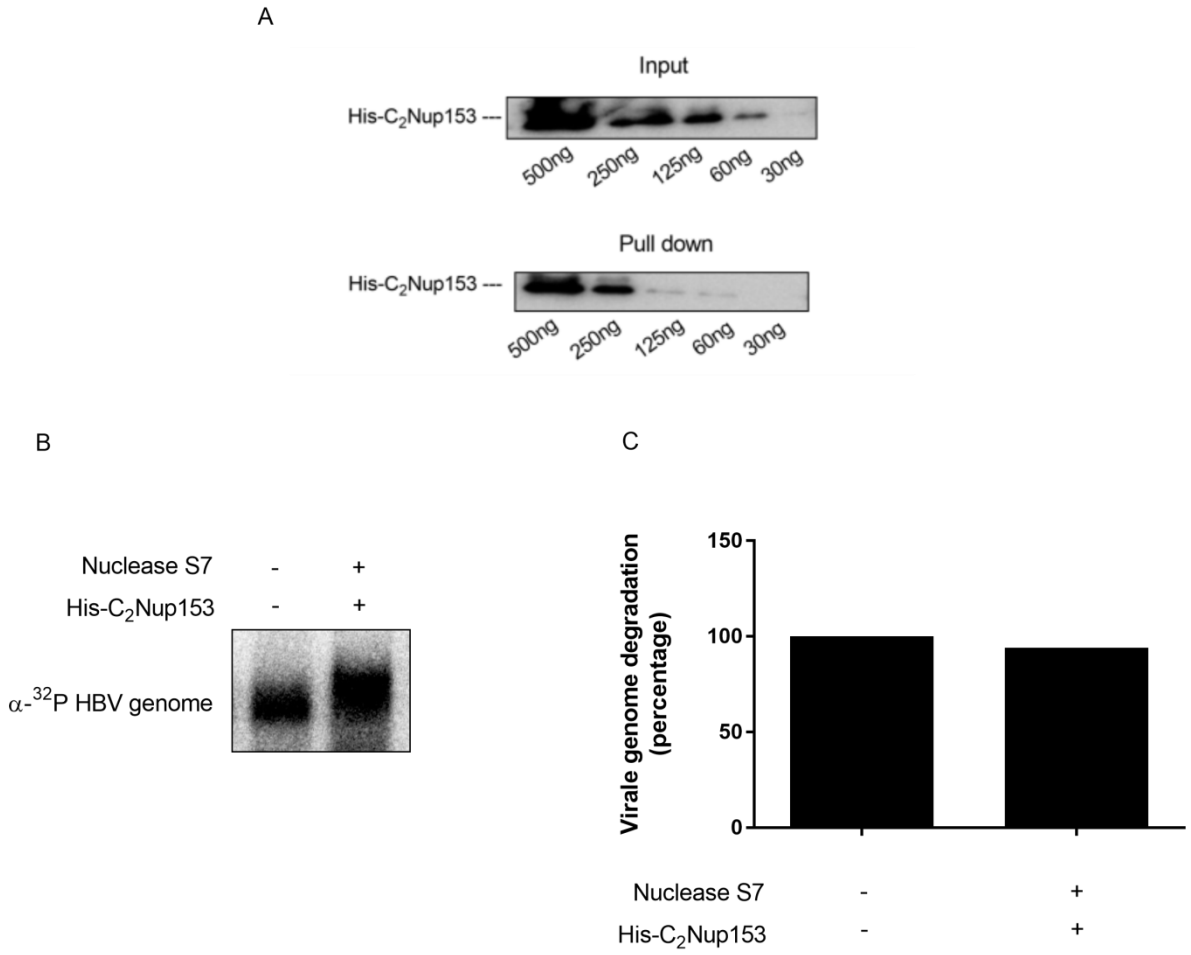


Figure 24: Viral genome degradation by the Nuclease S7 in the presence of the His-C2Nup153 fragment immobilized on Nickel beads.
A. Several amounts of the His-C_2Nup153 indicated below the panel were added to 10 µL of nickel beads and incubate for 4h RT. Imput and beads fractions were loaded on a acrylamide gel for a western blot. His-C_2Nup153 was detected using anti Nup antibody (mAb414). **B.** MatC capsids were purified from HepG2.2.15 and their genome was radio labeled with α-32P dCTP. MatC with α-32P labeled genome were incubated in the presence or absence of Nuclease S7 and in the presence or absence of the His-C2Nup153 fixed on nickel beads in transport buffer pH7, for 2h at 37°C. After the incubation, samples were loaded on NAGE. Viral genome degradation was determined by phosphoimaging. **C**. Quantification of viral genome degradation was done using ImageJ after background subtraction. Normalization was achieved by using the negative control without Nuclease S7 as 100% value.

Role of the C-term portion of Nup153 fixed on Nickel beads during capsid disassembly of MatC pretreated at pH5.5

pH 5.5 represents the environment of the late endosome, and MatC that were incubated for 10 min at pH 5.5 showed partial degradation of the viral genome, indicating less stability of the

capsid. However, these capsids were still capable to reach the NPC. Thus, we asked whether the combination of a previous incubation at pH 5.5, mimicking the late endosome, and the C-terminal portion of Nup153, mimicking the interaction at the NPC, would have provided the right cues for the disassembly of the capsid. To asses this possibility, we incubated MatC with ^{32}P labeled genome in transport buffer at pH 7, pH 6.5 and pH 5.5 for 10 min at 37°C, and sequentially with the Nuclease S7, in the presence of the His-C_2Nup153 immobilized on nickel beads, for 2h at 37°C in transport buffer at pH 7.

Figure 25 shows that the addition of the C-terminal portion of Nup153 does not increase the instability of the capsid at pH 5.5 rather, when it is added the degradation of the viral genome induced by the Nuclease S-7 decreases.

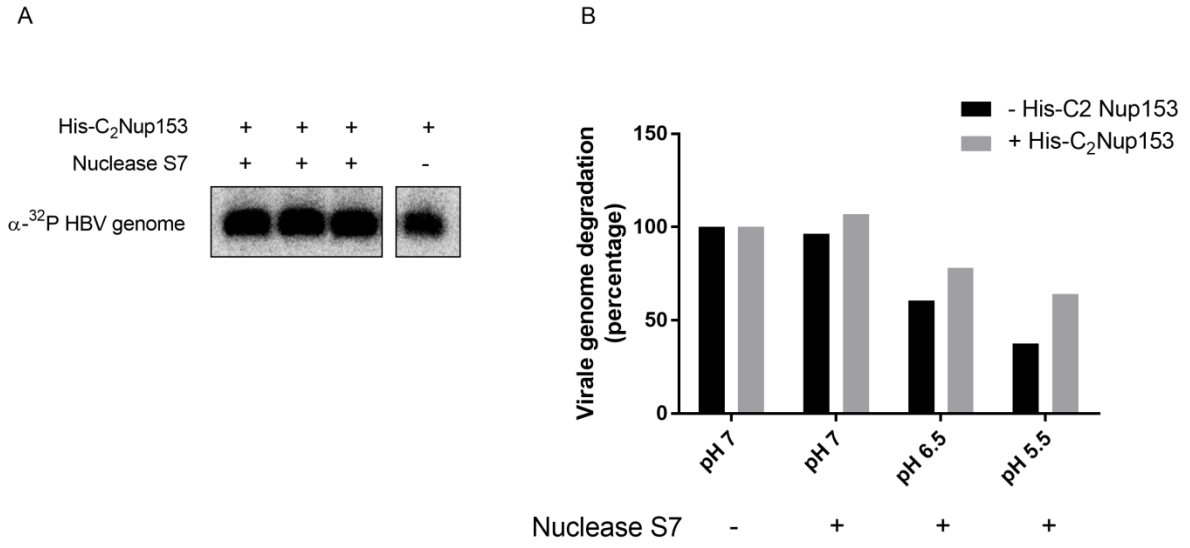


Figure 25: Viral genome degradation by the Nuclease S7 of MatC pretreated at different pH in the presence of the His-C_2Nup153 fragment fixed on Nickel beads.
A. MatC capsids were purified from HepG2.2.15 and their genome was radio labeled with α-^{32}P dCTP. MatC with α-^{32}P labeled genome were incubated in transport buffer at different pH for 10 min before being incubated with the Nuclease S7 in the presence of the His-C_2Nup153 fragment fixed on Nickel beads. After the incubation, samples were loaded on NAGE. Viral genome degradation was determined by phosphoimaging. **B**. Quantification of viral genome degradation was done using ImageJ after background subtraction. Normalization was achieved by using the negative control without Nuclease S7 as 100% value. In the graph this results is put in comparison with the previus experiment with different pHs without the Nup fragment.

Role of the C-term portion of Nup153 fixed on Nickel plates during capsid disassembly

To verify that the presence of the His-C_2Nup153 does not inhibit the Nuclease S7, we incubated 1 µg of control plasmid with Nuclease S7 in the presence or absence of the fragment. As shown in **figure 26,** the plasmid became completely degradated in the presence of the Nuclease S7 and in the presence of the His-C_2Nup153.

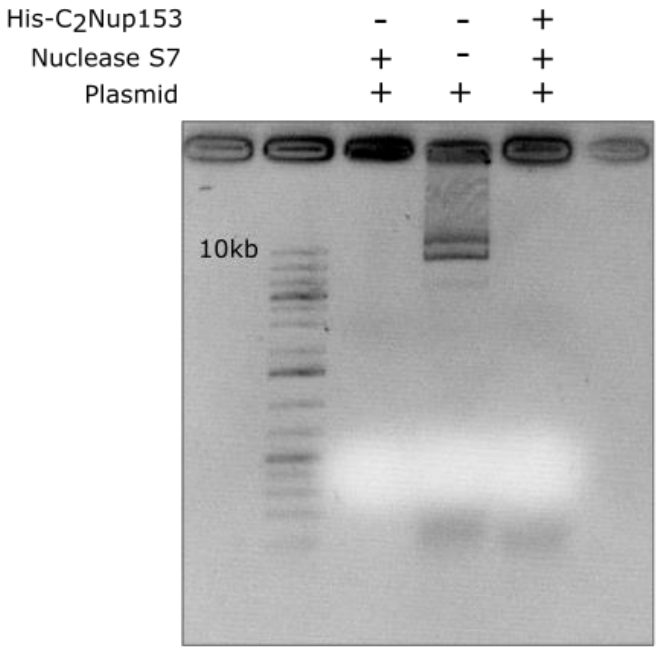


Figure 26: His-C_2Nup153 does not inhibits the Nuclease S7:
1µg of plasmid DNA was incubated with the Nuclease S7 for 15 min at 37°C in the presence or absence of the HisC_2-Nup153 fragment. After the incubation samples were loaded on a 0.7% agarose gel and DNA was stained with EtBr. Figure shows that the His-C_2Nup153 fragment does not inhibit the degradation activity of the nuclease S7.

We then asked whether the presence of the His-C_2Nup153 fragment, either in solution either fixed on Nickel beads, was hindering the complete disassembly of the capsid. Probably it could bind to it from all the directions without respecting the physiological environment and geometry of the NPC. To verify this hypothesis, we fixed the His-C_2Nup153 fragment on the bottom of a Nickel plate so that the binding would have come from only one direction and to closer mimic the NPC environment. However, as shown in **figure 27**, not even under these conditions the His-C_2Nup153 fragment was unable to support complete disassembly of the capsid.

Together these results suggest that this portion alone is not sufficient to promote the disassembly of the capsid and that other factors probably coming from the host cell are involved in the HBV disassembly step.

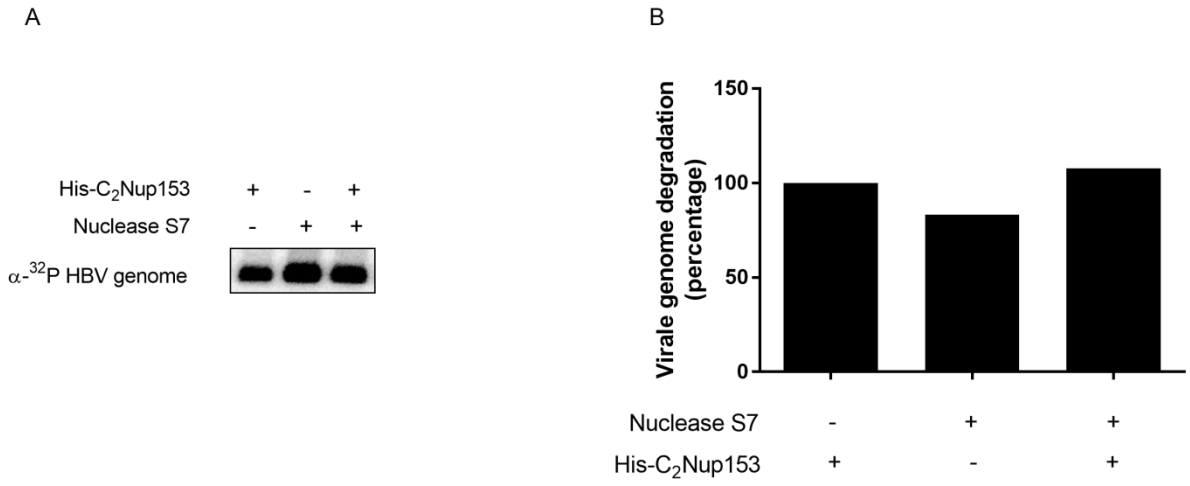


Figure 27: Viral genome degradation by the Nuclease S7 in the presence of the His-C_2Nup153 fragment immobilized on Nickel plate.

A. MatC capsids were purified from HepG2.2.15 and their genome was radio labeled with α-^{32}P dCTP exploiting the polymerase endogenous reaction. MatC with α-^{32}P labeled genome were incubated in the presence or absence of Nuclease S7 and in the presence or absence of the His-C_2Nup153 previously added to Nickel plate in transport buffer pH7, for 2h at 37°C. After the incubation, Supernatant was collected and the His-C_2Nup153 was eluted from the plate with elution buffer. Both elution samples and supernatants were loaded on a NAGE. **B**. Viral genome degradation was quantified by the ratio between the addition of the optical density of the bands corresponding to the radio labeled genome in the elution and supernatants, calculated with the ImageJ software, and the addition of CPMA of the radioactivity in the elution and supernatant in the same sample, collected before loading the samples in the gel. Normalization was made on the negative sample, in the absence of the Nuclease S7 and in the absence of the His-C2Nup153 fragment.

HBV Anchor™

Our results indicate that the presence of Nup153 does not enhance the capsids disassembly rather it seemed to stabilize them. at least with an *in vivo*-like kinetics and efficiency. However, our in vitro model did not provide sufficient information to explain how the disassembly of HBV capsid occurs in vivo. For following genome liberation in living cells allowing to investigate individual genomes we introduced the Anchor technology developed at the Neovirtech by Dr. Frank Gallardo (http://neovirtech.com/) into HBV. The Anchor technology exploits the specific binding of OR fused to a fluorescent protein, to a DNA sequence called Anch, which has a minimal length of 200 bp. (**figure 28A).** OR is a bacterial partition protein (ParB) and its interactions with the Anch leads to an accumulation of up to 500 molecules for each sequence, that nonspecifically associate to the adjacent DNA.

In infection with a pseudo-HBV in which the Anch has replaced part of the genome, single green dots corresponding to the viral Anch sequence recognized by OR become visible. Recognition is only possible when the genome is accessible and therefore released by the viral capsid (**figure 28B)**. As binding of OR requires dsDNA, the Anch must be either inserted into the ds part of the genome or the single stranded part of the genome must be filled.

Interaction between OR and Anch is highly dynamic with a turnover of 57s allowing the expression of adjacent genes (**figure 28C)**.

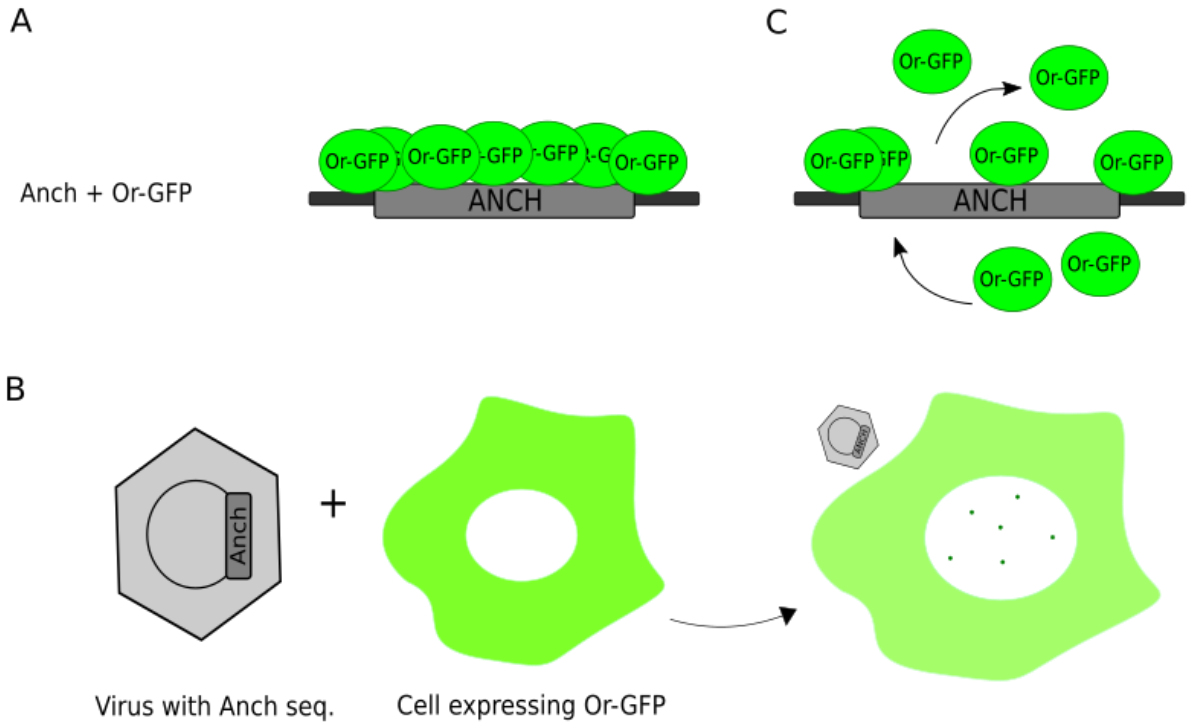


Figure 28: Schematic presentation of the AnchorTM Technology
A. The Anchor™ Technology is based on two partners: a DNA sequence called Anch and a bacterial protein called OR fused to DGFP. OR-DGFP specifically binds to the Anch sequence and the binding serves as enucleation site for the recruitment of up to 500 molecules of OR-DGFP for each Anch sequence. **B**. Example of a model system made by a virus that harbors in its genome the Anch sequence and cells expressing OR-DGFP. In this model, the formation of green dots is used as read out of genome released since they represent the recognition of the Anch sequence, present in the viral genome by OR-DGFP. This recognition is only possible when the viral genome is accessible and therefore released from the capsid. **C.** The binding of OR-DGFP to the Anch sequence is highly dynamic with a t1/2 of 57s as determined by FRAP (personnel information Neovirtech) allowing protein expression from OR-DGFP-covered ORFs.

HBV Anch plasmid constructs

For adapting the Anch-OR system to HBV we used hepatic cells line expressing OR fused to Dasher Green Fluorescent Protein (OR-DGFP). We first inserted the Anch sequence into the HBV genome (HBV Anch). In the specific case of HBV this is particularly challenging, because of the limited size of the genome, 3.2kb and because of its tight organization. We solved these problems by making three plasmid constructs.

Design of the constructs

In the first construct we placed the Anch sequence in the double strand portion of the HBV genome allowing genome detection directly after release (Early Anchor Plasmid, EAP).

In the second construct, we placed the Anch sequence in the single stranded region of the viral genome. In this case, the Anch sequence will be recognized only once that the genome is released and repaired (Late Anch Plasmid, LAP).

As the Anch insertion led to disruption of at least two viral ORFs, we constructed a trans complementation vector, termed Δε. It misses a functional ε encapsidation signal ensuring that only the pseudoviral genome harboring the Anch becomes encapsidated (**Figure 29)**

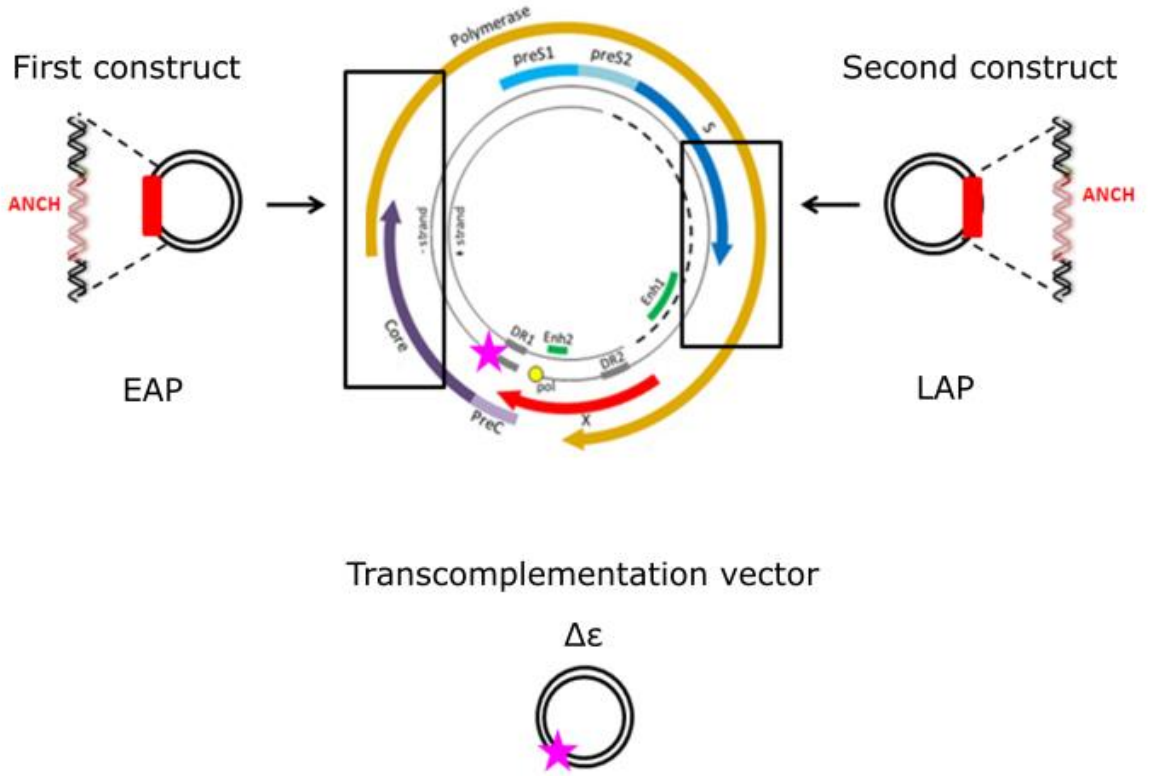


Figure 29: Insertion sites of theAnch in the HBV genome
The Anch sequence was inserted into two different regions of the viral genome. In the first construct it is inserted were the genome is double stranded, allowing HBV pseudogenome detection directly after genome release from the capsid (Early Anchor Plasmid, EAP). In the second construct the Anch sequence is placed in the single stranded part of the genome between the defined 5' end and the undefined 3' end of the plus strand DNA. In this position the HBV pseudogenomes will not be recognized by OR-DGFP until the plus strand DNA will be (nearly) completed (Late Anch Plasmid, LAP). For production of HBV pseudoviruses, we constructed a trans complementation vector. This construct codes for all viral proteins but not for the encapsidation signal ε. Pregenomic RNA coded from vector Δε is thus not encapsidated. The transexpression of the proteins, including protein over expression derived from co-transfection of EAP (or LAP) and Δε is not interfering with virus production as shown in the literature.

Validation of Anch constructs recognition by OR-DGFP

We validated the constructs first by plasmid transfection of EAP LAP in HuH-7 cells expressing OR-DGFP. **Figure 30** shows formation of green dots in the transfected cells. The

difference in intensity and size of the dots is consistent with the lipofection used for plasmid transfection since it induces the aggregate formation.

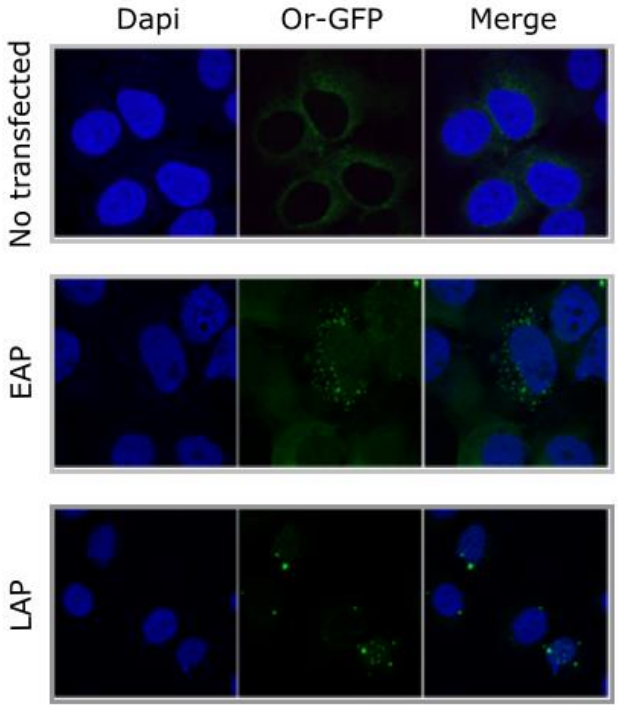


Figure 30: EAP and LAP recognition by OR-DGFP.
250 copies of EAP and LAP per cell were transfected in HuH-7 OR-DGFP. Images were taken at the confocal microscope with 63 x objective. Figure shows formation of OR-DGFP dots upon transfection.

To validate the trans complementation strategy, we double transfected either EAP or LAP with Δε in HuH-7 cells. We purified the pseudovirus from HuH-7 supernatant and characterized the viral genome from purified pseudo HBV by Southern blot using a radio labeled probe against the HBV X region present on EAP and LAP. As shown in **Figure 31A** for both LAP and EAP, two bands were visible corresponding to the rcDNA and to the single stranded DNA. Further, an excess of replication intermediates was observed, which is consistent to Southern blots using HBV genomes extracted from virions purified from patients or from HepG2.2.15 cell supernatants.

For further validation, we analyzed the capsids extracted from the purified pseudovirions by non-ionic detergent treatment by NAGE, followed by anti capsid staining. **Figure 31B** shows that capsid containing the Anch sequence exhibited the same migration behavior and antibody reactivity as MatC purified under the same conditions.

Following the quantification, we estimated that with the double transfection strategy, 0.1 ng of capsid were produced per mL of cell culture medium.

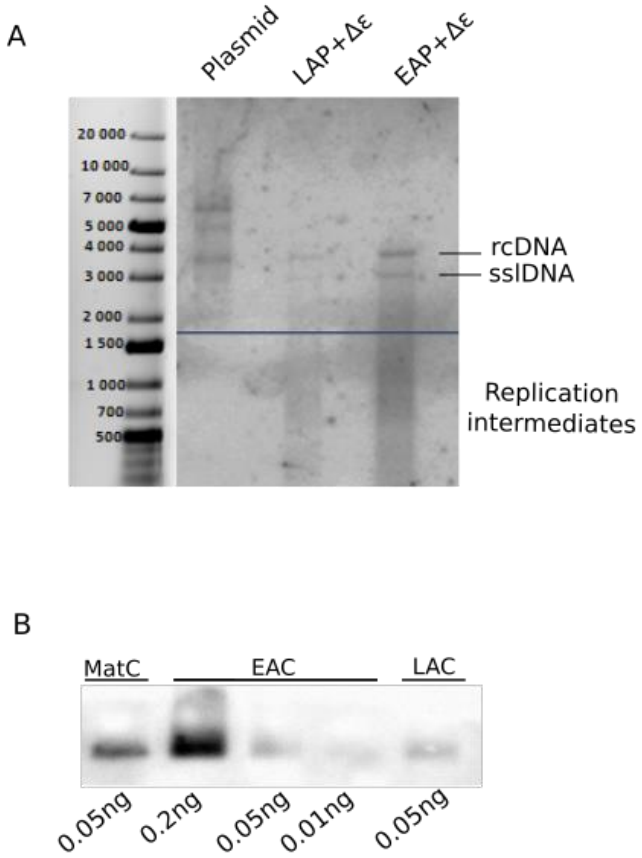


Figure 31: Validation of the trans complementation strategy
A: Southern blot of nucleic acid extracted from purified particles extracted from HuH-7 double transfected with LAP and Δε or EAP and Δε. The figure shows the presence of two bands one corresponding to rcDNA and the other corresponding to sslDNA. **B**. NAGE of different amount (indicated below each lane) of purified particles extracted from HuH-7 double transfected with LAP and Δε or EAP and Δε in comparison to 0.05 ng of MatC. The figure shows that all the particles present the same migration properties and the same reactivity to the Dako polyclonal antibody.

Production of Anch modified virus

Once validated the system, we established cell lines for permanent EAP or the LAP pseudovirus production.

Establishment of stable cell lines constitutively producing HBV Anch

To establish stable cell lines constitutively producing HBV Anch, we made three lentivectors using the three plasmid constructs. Then, from these, VSV-G pseudotyped lentivectors particles were produced by the Vectorology platform in Bordeaux. As transduction marker for the EAP and LAP constructs, we used DGFP, and for the Δε we used tdTomato.

Once obtained the lentiviral particles, we double transduced HepG2 and HuH-7 with either the EAP and the Δε or the LAP and the Δε.

Two weeks after transduction, we verify transduction efficiency by the fluorescence of the cells and sorted the positive cell by FACS next.

Figure 32 shows a representative FACS sorting of HuH-7 double transduced with either EAP and Δε (left side), or LAP and Δε (right side).

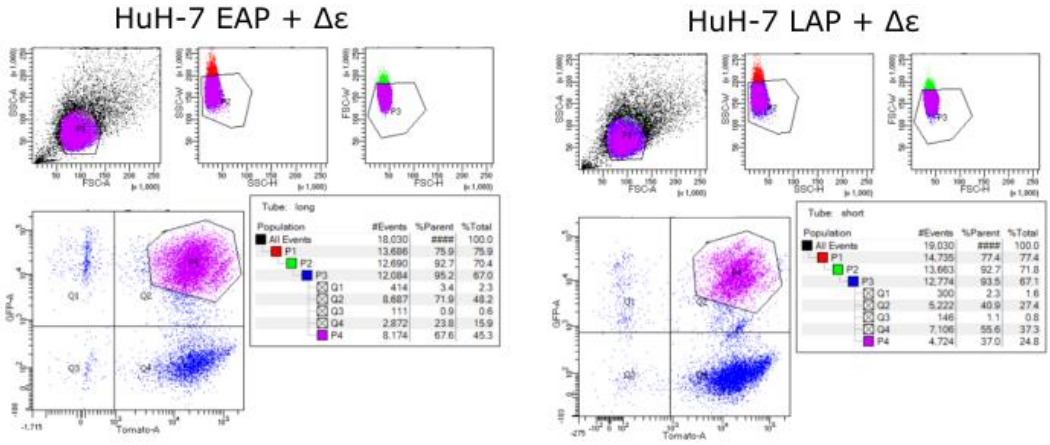


Figure 32: Establishment of stable cell lines: HuH-7 LAP/EAP +Δε
Representative FACS sorting of HUH-7 double transduced with either EAP and Δε (left side), or LAP and Δε (right side).

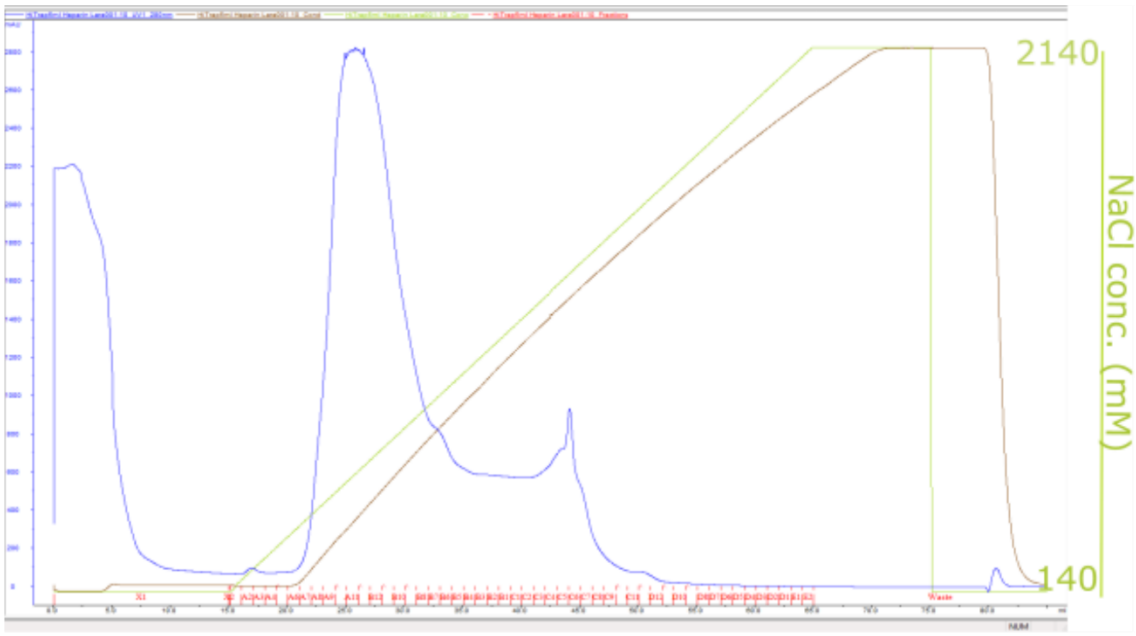
Virus purification

In a recent study it has been shown that HBV that bind to heparin are highly infectious[69]. This is caused by the cell attachment step needing attachment to HSPGs serving as a low-affinity HBV receptor. Purification via heparin columns thus selects infectious HBV and avoids co-purification of defective particles. It further avoids co purification of other impurities from cell culture supernatant we observed using centrifugation-based protocols. We collected twice the supernatant from eight T150 flask of HuH7 or HepG2 cells line stable double transduced previously treated or not with 1% DMSO. DMSO has been reported to enhance the viral production by bringing the cell into a differentiate state and providing the specific factors needed for the viral replication cycle[198]. Purification was performed by binding the pseudovirus to 5 mL heparin columns, followed by elution using a linear gradient of NaCl. **Figure 33A** depicts a chromatogram showing the elution peak at ~300 mM NaCl, in agreement with the study above[69]. Afterwards, the pseudoviruses were concentrated by

sedimentation. After suspension of the (invisible) pellet, the pseudoviruses were quantified by quantitative PCR amplifying the HBV X region, which is present in EAP and LAP.

Figure 33B shows that in these conditions we were able to purify up to 3 x10^3 copies per µL and that as expected, the amount produced was slight higher in HepG2 cells then in HuH-7 cells and in the presence of DMSO. Assuming that one copy corresponds to one viral particles, we estimated that we obtained in our best preparation 3 x10^3 particles of HBV Anch per mL of cell culture medium. This is far below reports of Glebe et al. observing the production of 10^6 HBV genome equivalents per mL from different cell lines stable transfected with HBV DNA construct[198], which can be explained by transcomplementation.

A



B

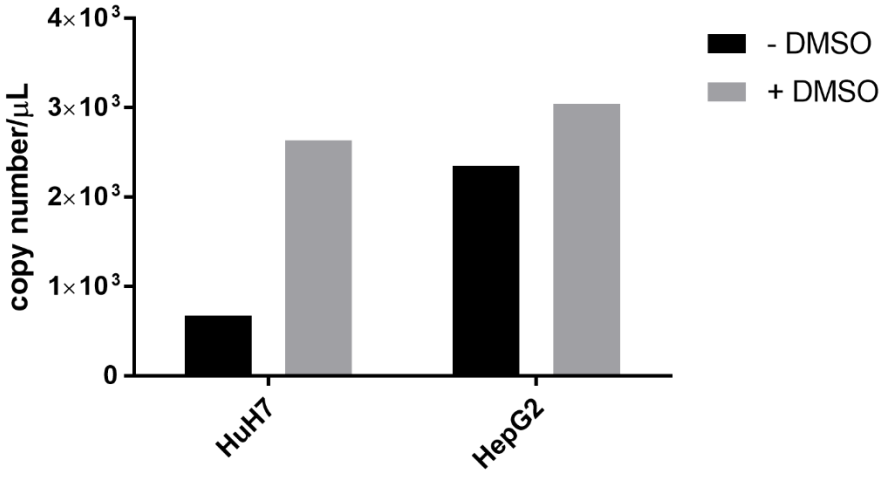


Figure 33: HBV Anch purification
A. Chromatogram obtained during virions purification. Virus purification was performed using Heparin Columns and the Akta purifier that allows the live time monitoring of the chromatography. 240 mL of virus containing HepG2 EAP + Δε supernatant were applied to a 5 mL Hi trap Heparin column. Bounded virions were eluted with a linear gradient of NaCl. NaCl concentrations are shown on the right of the chromatogram. OD at 280 nm is represented by the blue line. The first peak represents unbound material appearing in the flow-through. The majority of the bounded material is eluted at ~340 mM NaCl (second peak). A second small peak is presents at~1000mM. **B.** Quantification of copy number/µL of HBV EAP obtained as described above.

HBV Anch infection of HepG2 inducible NTCP cell line expressing OR-DGFP

HepG2 cell line expressing NTCP

For testing infectivity of the pseudoviruses, we used HepG2 cell expressing NTCP under the control of a TET on system, kindly provided by Dr. Dieter Glebe (Giessen, Germany). In this system, the expression of the NTCP is induced by the presence of doxycycline that enhances the recognition of the TET regulatory elements by the TET transcription factor.

We first tested the correct expression of NTCP by Western blot and IF upon induction of NTCP by adding 5 µg/ mL of doxycycline to the culture medium.

Figure 34A shows a gradual expression of NTCP starting from 24h post induction reaching a maximum at 72h post induction. The NTCP showed an apparent molecular weight of 72 kDa, which is in agreement of others[199]. As the MW of the NTCP is 37 kDa, the retarded migration indicates its glycosylation, which is consistent with reports showing that the NTCP is N-Glycosylated at its N-terminal domain[71]. Next, we confirmed the expression of the NTCP by IF also allowing verifying exposure on the plasma membrane. **Figure 34B** shows NTCP in induced cells only with a dominant localization at the plasma membrane. The kinetic showed that NTCP expression increased until 96h of induction.

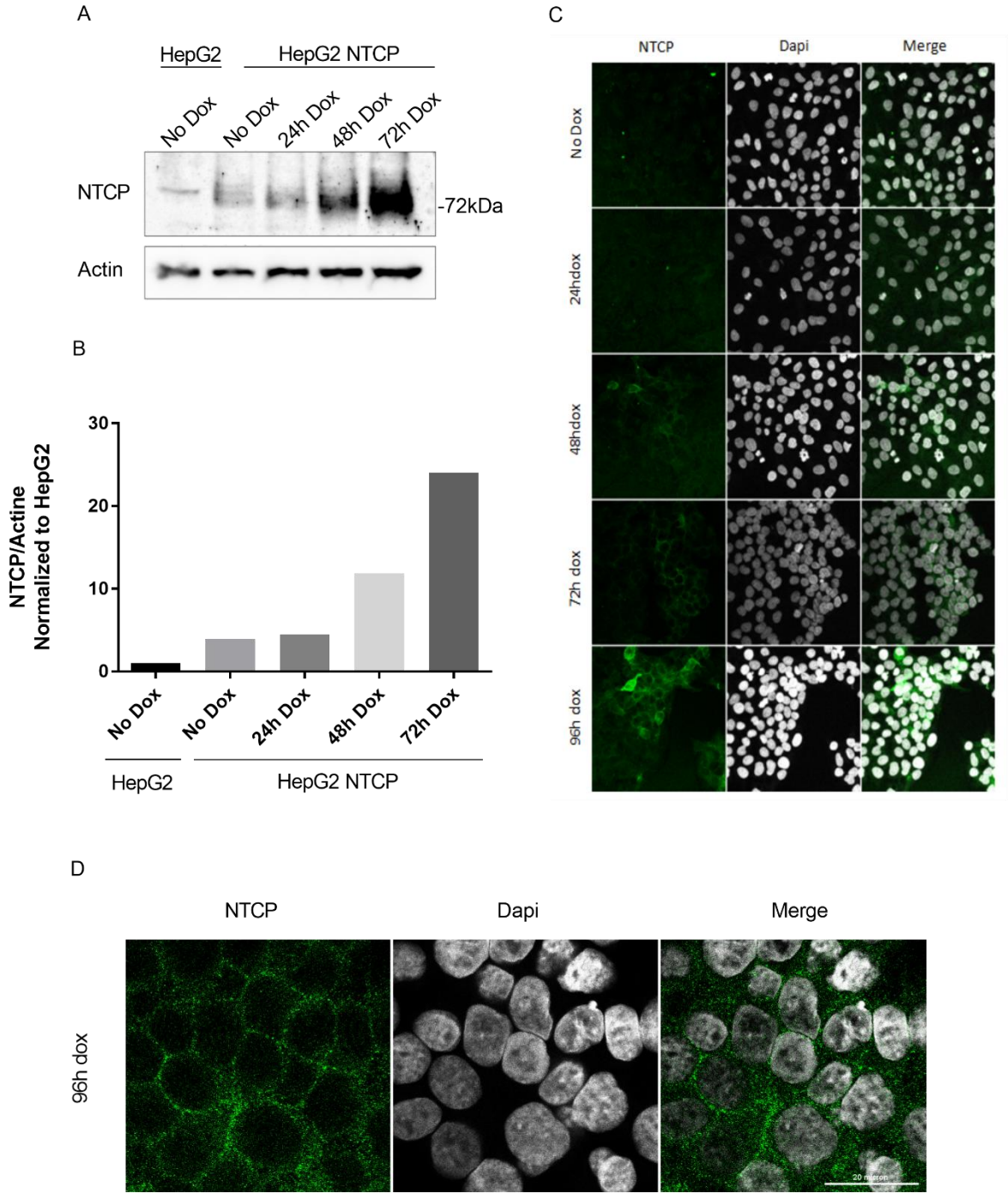


Figure 34: Doxycycline induction of NTCP in HepG2 NTCP inducible cell line
NTCP expression was induced by 5µg/ mL of Doxycycline over different time point. **A**. At each time point indicated in the upper panel, cells were lysate and proteins were separated by SDS-page followed by western blot. NTCP was detected using an anti SLC10A1 antibody, while Actin was detected by an anti Actin antibody lower panel. **B**. quantification of the NTCP expression was made with the software ImageJ calculating the ratio between NTCP and Actin for each sample and normalizing it to the negative control, HepG2 cells. **C**. At each time point indicated left of each row, cells were fixed and the NTCP expression was visualized by IF using the anti SLC10A1 antibody. Images were taken at the confocal microscope Leica DMI6000 TCS SP5 with a 40X objective. **D.** 63 x magnification of HepG2 NTCP after 96h of dox induction showing NTCP expression on the plasma membrane. Length bar 20 µm.

Infection of HepG2 NTCP OR-DGFP cells with HBV Anch (EA)

In order to test genome localization after infection, we next transduced the HepG2 cells with OR-DGFP VSV-G pseudotyped lentivectors particles at MOI 2. Three weeks later, we sorted positive cells by FACS to have an homogeneous population of Or-DGFP expressing cells (**Figure 35**).

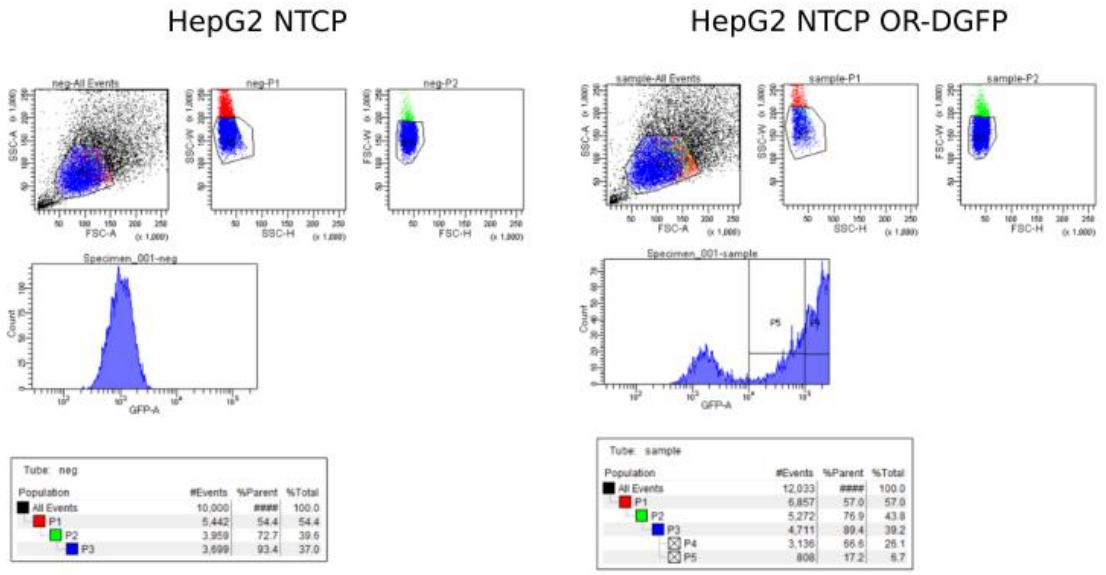


Figure 35: FACS sorting of HepG2 NTCP OR-DGFP
The figure shows the expression of OR-DGF in HepG NTCP transduced with OR-DGFP (in the left panel), in comparison with the negative control represented by HepG2 NTCP (right panle), where only a basal GFP fluorescence is present

Then we induce the expression of NTCP by doxycycline and we infected them with HBV Anch (EA construct) at MOI 10. 72h later we verified we the formation of green dots by microscopy using the epifluorescent microscope. We observed the presence of one positive cell after 2h p.i. The number of positive cells increased after 3 days p.i to 9% (**Figure 36**).

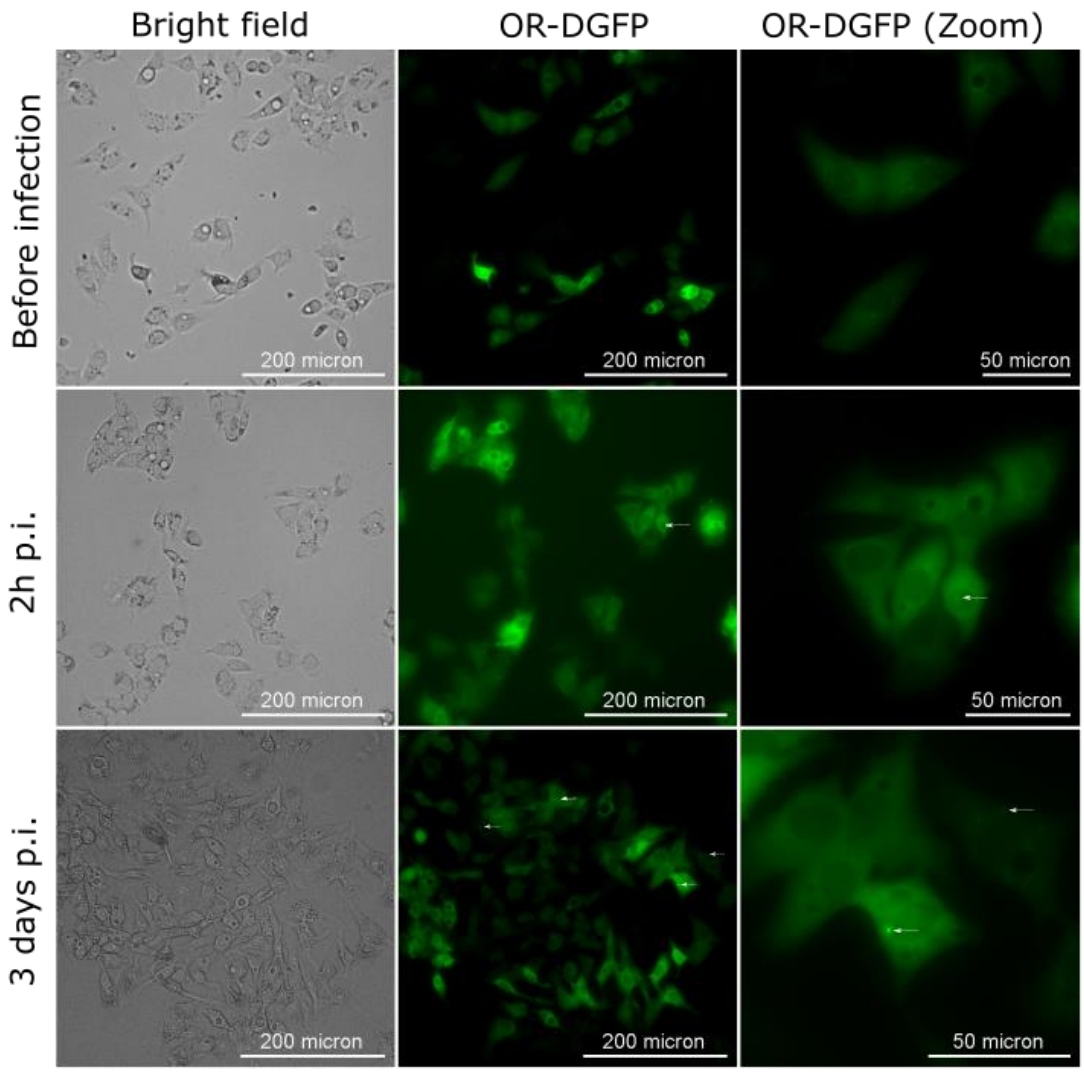


Figure 36: Infection of HepG2 NTCP OR-DGFP cells
NTCP expression was induced with 5µg/µL of Doxycycline and infection was made using HBV Anch (EA contruct) Images were taken 72h pi with epifluorescent microscope Leica DMI6000 B using a 40X Object.

Capsid Lipofection

In cell culture, HBV infection is slow and inefficient. An MOI of 3000 was reported to yield in 80% infection (Prof. Ulrike Protzer, Munich, personal communication). For circumventing this obstacle, also with regard to the relatively low quantities of pseudo HBV we obtained, we used capsid lipofection, which bypasses the initial rate-limiting steps of HBV infection[200]. During capsid lipofection, capsids are lipo-coated also allowing entry of non-susceptible cells. Using capsid containing EAP, we observed the formation of green dots after 2h post lipofection (**figure 37**) confirming that a slow entry kinetic is not the limiting factor in infection.

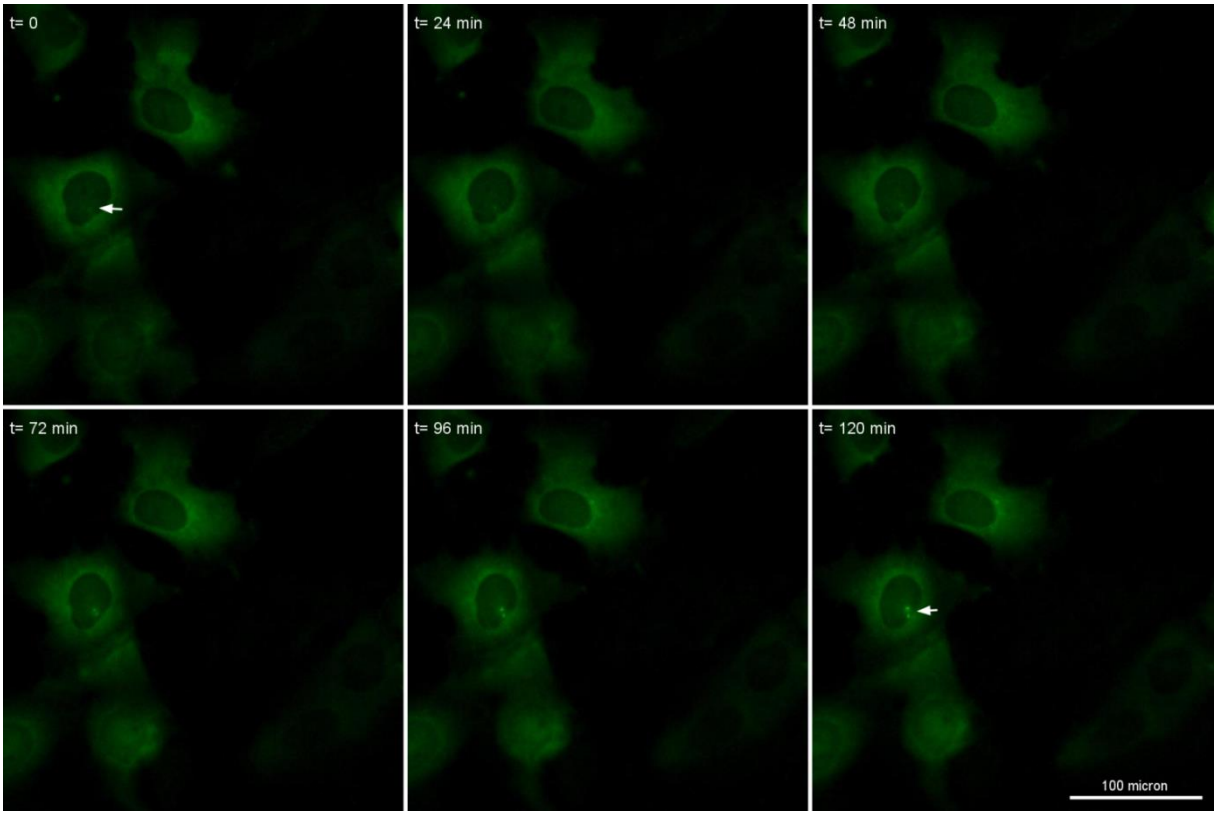


Figure 37: HBV-EAP capsid lipofection
HBV EAP lipofection in HuH-7 OR-DGFP cells. Pictures were taken each 12 min over 4 h. Here six representative time points are shown. Time depicted in the figure refers to the time passed from the start of the acquisition. The figure shows the appearance of two green dots within two hours from the start of the acquisition (white harrows).

HBV pseudogenome visualization after integration

The chromatinization of HBV cccDNA could lead to displacement of OR-DGFP from the genome. This would not only cause disappearance of the signals after infection but would also prevent HBV genome detection after integration. As the loss of genomes after infection or lipofection could also be caused by other reasons as genome degradation, we investigated the detection of pseudo HBV genomes after integration. We investigated the presence of the Anch sequence integrated into the cellular genome by transduction of HuH-7 LAP +Δε cells with OR-DGFP VSV-G pseudotyped lentivectors particles. **Figure 38A** shows that after 10 passages HuH-7 LAP +Δε were still positive for the both transduction marker: DGFP for LAP constructs and tdTomato for the Δε, although with different intensity. **Figure 38B** shows the presence of green dots corresponding to LAP in the 11% of the cells. Considering that the cells underwent 10 divisions prior to transduction, we hypothesize that the Anch sequence remained in the cells in an integrated form and accessible to the OR-DGFP protein after three

weeks. We observed however that the dots were present in the non-condensed open chromatin only. This was particularly evident when cells undergo mitosis (**Figure 38B** lower panel)

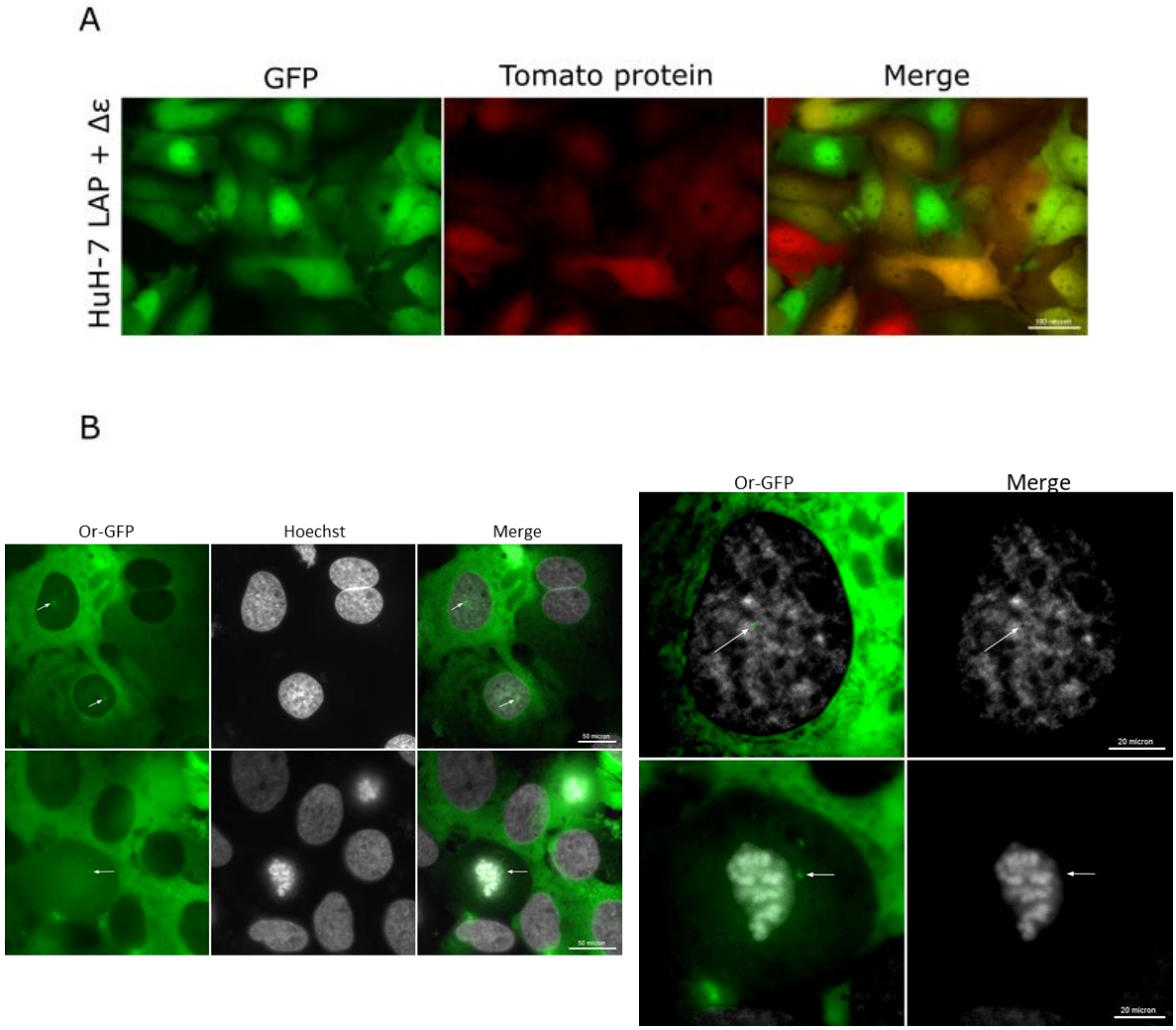


Figure 38: OR-DGFP transduction of HuH-7 LAP +Δε.
A. Huh7 double transduce with LAP and Δε lenti vector the figure shows that all cells were positive for both transduction marker.B. HuH-7 LAP Δε transduced with OR-DGFP lenti vector. The figure shows the presence of the integrate LAP construct. Picture were taken at the epifluorescent microscope length bar in A. 100 µn and in **B**. 20µn

Discussion

EmpC binds directly to Imp.β

HBV has to deliver its genome inside the nucleus of infected cells. To do this it has to hijack the cellular pathways that mediate nuclear import.

Studies of HBV nuclear import have been done mostly subjecting different preparation of capsids to digitonin permeabilized cells or microinjecting them to *Xenopus laevis* oocytes. It has been shown that *in vitro* phosphorylated rHBc but not unphosphorylated rHBc are able to reach the nuclear periphery and to bind to the NPC in digitonin permeabilized HuH-7 cells[36]. Microinjecting these capsids into the cytoplasm of *Xenopus laevis* oocytes confirmed this finding and showed that the capsids pass the nuclear pore and accumulate in the nuclear basket. Further, phosphorylated rHBc and MatC purified from HepG2.2.15 cells are able to coimmuneprecipitate Imp.β. from RRL implying a binding with Imp.β. However, the precipitation of Imp.β is lost when Lamin B2 NLS is added to RRL suggesting that the binding requires Imp.α[36]. Consistently, phosphorylated rHBc interacts with the nucleus, when Imp.α and Imp.β are added and two overlapping NLS have been mapped on the primary sequence of the CTD of Cp (aa 158–168 and 165–175)[36]. Phosphorylation has been linked with the maturation process, even though MatC are thought to be dephosphorylated. Therefore, it has been suggested that the phosphorylation might help the exposure of the CTD from the interior of the capsid starting the maturation process and leading to the recognition of the NLS by Imp.α that then acts as an adaptor for Imp.β. Thus, regarding MatC a canonical import pathway has been proposed.

So far, import studies have not involved EmpC and their interaction with import receptor is still an open question. EmpC derive from spontaneous assembly of Cp[111] and being devoid of nucleic acid they are less stable[88] and prone to disassembly. The intracellular localization of Cp reflects the clinical state of the infection: cytosolic Cp or capsids in patients' liver are linked to hepatocellular injury[201,202] and to a low viral load[203] while nuclear localization correlates with high viral load and minor hepatitis activity[204], showing the importance of understanding the biological role and the intracellular trafficking of EmpC and Cp.

In this work, we investigated the interaction between EmpC and import receptors. Our principal observation was that EmpC bind to Imp.β directly. EM micrographs and 3D structure showed the presence of Imp.β at the quasi-six fold vertex of EmpC and in the internal lumen. This finding is supported by the presence of a putative IBB on Cp, overlapping the NLS but spanning aa 141 to aa 180 thus comprising almost the entire CTD

and the linker region. Such overlapping nuclear import signals were already described earlier[205].

Together these data of HBV capsids imply a difference in the exposure of the CTD between MatC, phosphorylated rHBc and EmpC. Being 39 aa long the IBB domain is larger than the NLS (8 aa) thus in EmpC a larger portion of Cp needs to be switched on the capsid surface.

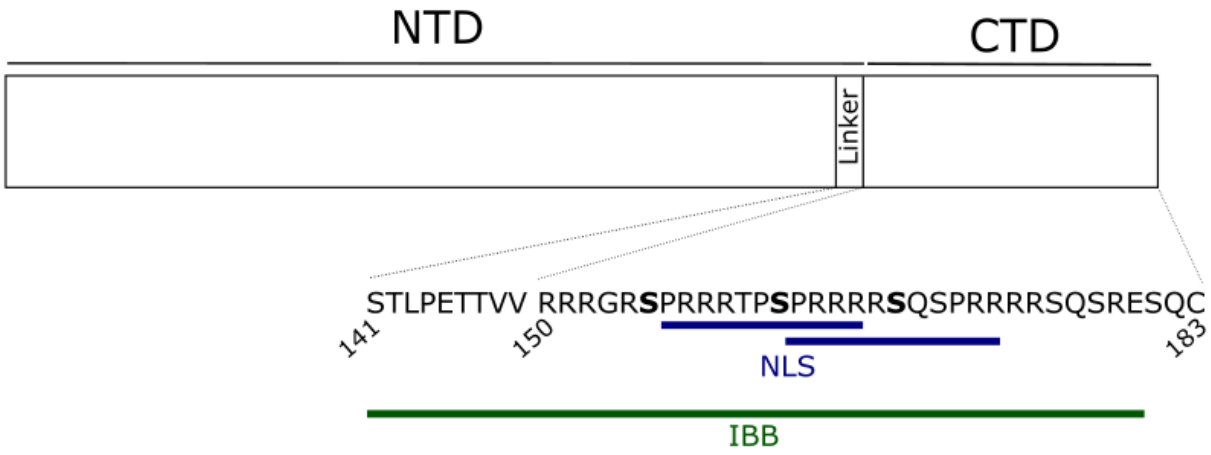


Figure 39: IBB and NLSs on the CTD of Cp (183 aa)
Cp comprises an NTD from aa 1 to aa 149 and a CTD from aa150 to aa183. Two overlapping NLS (blue) have been mapped on the CTD of Cp. IBB (green) has been proposed encompassing aa 140 to aa180, thus almost all the CTD plus the linker region (from aa 141 to aa 149). In the sequence serines 155, 162 and 170 are depicted in bold.

The structural change can be explained by two models. Considering that the CTD is unstructured the polypeptide chain has a diameter of 6 – 7 Å, while the triangular fenestration in the quasi six fold symmetry has diameter of 12 x 15 Å, the CTD could pass the pore. Consistently, the CTD is connected to the N terminal part of Cp directly adjacent to the opening[38]. In the second model, the extrusion can be explained by "capsid breathing"[206] that has been observed for other virus such as poliovirus[207] and rhinovirus [208]. Capsid breathing involves the temporary dissociation of Cp hexamers followed by their re-association, which would open the capsid shell by at least one Cp hexamer, corresponding to a size of ~8 nm. In difference to MatC, ImmC, and rHBc, which contain nucleic acids interacting with the CTD to different extent, EmpC do not exhibit this interaction retracting the CTD inside the lumen of the capsid. This hypothesis is in agreement to the observation that MatC, ImmC and rHBc

are more stable leading to the consequence that capsid breathing would be more seldom resulting in less CTD exposure[88].

The second model also explains the presence of Imp.β in the capsid lumen as its size of 14 x 4.6 nm[209] does not allow diffusion through the fenestrations of the capsid.

After binding to an IBB, Imp.β wraps around the IBB leading to an essentially globular shape with a diameter of 8.5 nm[210]. Capsids lumen has a diameter of 26 nm allowing internalization of <50 Imp.β molecules preventing that additional Cp- Imp.β complexes integrate into the capsid. Association of released Cp hexamers by capsid breathing, which became bound to Imp.β would thus prevent re-association with the capsid, which would result in broken capsids as we observed in EM micrographs. However, only broken particles were seen and not complete dissociated ones, which leads to the hypothesis that Imp.β binding may favor the disassembly by decreasing energetic barriers rather than directly causing it.

The conclusion of an Imp.β-mediated destabilization is also consistent with the finding that EmpC in complex with Imp.β showed an increased antibody reactivity to the polyclonal anti capsid antibody that mainly binds to the tips of the capsid spikes. Further, EmpC did not migrate as a single band but as a smear supporting disintegration. Such an inhomogeneous migration was also observed after acid-denaturation of phosphorylated rHBc[36]. Neither an increase in signal strength nor an inhomogeneous migration was observed in the presence of Imp.α leading to the hypothesis that the increased antibody reactivity is caused by exposure of additional epitope(s). Importantly, a similar increase of signal strength or an inhomogeneous migration was not observed using MatC, suggesting that MatC was not destabilized, neither by Imp.β nor by Imp.α nor by Imp.α/Imp.β.

While a missing destabilization by Imp.α is not surprising as the Imp.α NLS-interaction has the same affinity of 3 kcal/mol[211] as the Cp subunits (3-5 kcal/mol;[212]), a destabilization in the presence of both Imp.α and Imp.β is surprising as Imp.β increases the affinity to an NLS by nine-fold [213]. We interpret this finding in that the import receptors gained access to the externally exposed CTD part comprising the NLSs but not to CTDs of isolated Cp hexamers derived from capsid breathing. This means that capsid breathing occurs less – if at all - in MatC compared to EmpC. This in turn is consistent with the higher stability of MatC and with our observation that the genome of MatC was not degraded upon addition of Nuclease S7. Furthermore, this conclusion is in agreement with the observation that UV crosslink of MatC leads to DNA Cp bonds, indicating DNA Cp-interaction, while EmpC cannot be linked by this method, as shown by Rabe et al[41] and by unpublished observations of our group.

We further observed that the capsid Imp.β-binding was DTT-dependent at 2 mM concentration, which mimics the reducing intracytoplasmic environment. This finding may either indicate a low activity of Imp.β, which is sensitive to oxidization, or may suggest that the capsids expose more Imp.β binding sites upon reduction. Our estimation that 40% of Imp.β was active in the absence of DTT however favors that DTT acted rather on the capsids than on Imp.β.

This assumption is supported by the observation that DTT treatment changed the migration pattern of Cp dimers during SDS PAGE. The finding that monomeric Cp migrated with and without DTT in a 15 kDa band confirms that the folding of only the dimers changed, while the monomers exhibited a stable SDS-resistant folding. The fact that the dimers were not dissociated to Cp monomers suggests that the better accessible disulfide bond between Cys185 were reduced and that the inner bonds between Cys61 were not affected. However, the observations that only a minority of Cp exhibited disulfide bonds but that reduction is important for EmpC Imp.β-interaction implies that the bonds have a dominant negative effect on capsid breathing. Nonetheless, the Imp. binding-properties and the low fraction of disulfide bonds-linked Cp indicated that EmpC in the presence of DTT were in a *in vivo*-like state similar to intracellular capsids.

In summary, our data indicate for the first time that EmpC differs from MatC, ImmatC and phosphorylated rHBc by its nuclear import receptor binding, which is caused by different exposure of the CTD. Our findings indicate that this different exposure is based on two different mechanisms, either full CTD exposure due to capsid breathing or partial exposure by topology change of the CTD through the fenestrations in quasi-six-fold symmetry. Imp.β binding upon capsid breathing can further cause EmpC disassembly at least under in vitro conditions. However, it must be considered that microinjection of EmpC into the cytoplasm of *Xenopus laevis* oocytes led to the accumulation of assembled EmpC at the nuclear pore[82]. Correlating the data of this work (95% NPCs decorated with ~7 capsids after microinjection of 5 x 108 EmpC) and the estimated number of 5 x 107 NPC per oocyte[214] , in fact indicate that – if at all - only a minority of EmpC become disintegrated prior to their arrival at the nuclear pore.

Cytosolic EmpC disassembly, however, may have an important function in hepatitis B. Cp, which have a much shorter half-life than capsids are degraded by the proteasome[215] and the degraded fragments enter the TAP pathway for being transported in an MHC class I-dependent pathway to the plasma membrane. In particular in acute HBV infection these Cp-derived peptides are recognized by the CD8+ T cells[16] activating the inflammatory response.

Consistently, cytoplasmic localization of Cp has been linked to chronic infection and strong hepatocellular injury.

In this scenario the transport of EmpC into the nucleus would be a mechanism by which HBV limits the cellular immune response.

Study of HBV capsid disassembly

For HBV the mechanism by which capsid disassembly is unclear. Rabe et al., in 2003 showed that ImmC and MatC produced in HepG2.2.15 cells reach the nuclear side of the NPC in digitonin permeabilized cells but only MatC generate intranuclear capsid staining and released viral genomes[41], indicating genome maturation-dependent differences. UV cross-linked MatC in contrast were transported into the nuclear basket but failed to enter the nucleus showing that capsid disassembly is required for entering the nucleus[85]. Thus, the presence of intranuclear capsids is a surrogate marker for the disassembly and genome release.

It has further been observed that Cp reassembles inside the nucleus, indicating that both the nuclear and the cytoplasmic environment support the assembly reaction of the capsid[216].

Together, these evidences suggest that the environment where the disassembly takes place is likely after the transport towards the NPC, but before the entry in the karyoplasm, likely the NPC. At the level of the NPC, it has been shown that ImmC, MatC and phosphorylated rHBc interacts with the C-terminal domain of Nup153, that is part of the nuclear basket but which was also described to be cytosolically exposed[85].

We thus investigated the disassembly of MatC, and especially the role of the Nup153. We found that the C-terminal domain of Nup153 interacts with the Cp in ratio 1:1, a ratio implying that Nup153 binds to a different site then Imp.β. In agreement, it has been shown that capsids lacking the CTD were able to bind Nup153 demonstrating that the binding of Nup153 to the capsid is independent from the CTD, region at which Imp.β. binds[85]. Further, we observed that the binding of the His-C_2Nup153 to EmpC was heterogeneous but resulted in a different surface charge of capsids and the capsid-Nup153 complex.

We further observed that the presence of the His-C_2Nup153 fragment did not enhance the capsid disassembly; rather it seemed to stabilize it. When the fragment was added in solution or solid-phase fixed to mimic the conditions of the nuclear basket, the amount of degraded genome by Nuclease S7 was slightly decreased, strongly arguing against an impact on capsid disassembly. As Nup153 attaches to the outer surface of the capsids, this finding is in agreement to our conclusion that Imp.α/β also binding to capsid's exterior did not disassemble MatC.

The stabilization could be explained by the need of HBV capsids to complete their maturation so that only MatC would be able to release their genome. In this scenario, Nup153 would be necessary to arrest the capsid until genome maturation is complete. In agreement with this assumption are conclusions of Dhason et al. who proposed a model in which the formation of

dsDNA facilitates capsid disassembly by acting as a coiled spring[217]. Contradictious to this model is the observation that capsid release is efficient in permeabilized cells in which no deoxynucleotides are added[41,216]

We have to consider, however, that in our experiments all Cp were covered with Nup153, while *in vivo* just 8 – 16 copies of Nup153 are present in the nuclear basket. This would still allow interaction of other cellular factors potentially involved in genome release.

In the control reaction we used MatC, which were preincubated in different acidic pHs. The observation that the genome became accessible after pH 5.5 could be interpreted that acidification is relevant in infection as shown for numerous other viruses as adeno[218,219]- and hepatitis C viruses[220,221]. Such a destabilization upon cell entry would then result in release of the polymerase-DNA genome complex, which is karyophilic at least when extracted capsids of the WHV[222]. This model would be in agreement with the finding of nuclease sensitive HBV genomes after infection[88]. Experiments of others however showed that the treatment with by Bafilomycin A1 blocks HBV infection[80]. Addition of Nup153 to the pre-acidified capsids also inhibited MatC genome degradation although an unspecific inhibition of the nuclease by the Nup153 fragment was excluded. The experiments thus confirmed that Nup153 is rather a stabilizing factor and that capsid disassembly requires additional factors.

The conclusion of a cellular factor was confirmed by using digitonin-permeabilized cells, using nuclear capsids as read-out. pH 5.5 pretreatment did not increase nuclear capsids indicating that a pre-opening of the capsid is not required for nuclear entry of the capsids. WGA which blocks the NPC by interacting with the O-linked β-N-acetyl glucosamine (O-GlcNAc) of glycosylated Nups[196] blocked nuclear entry of the control protein, which is agreement to previous results[41].

Replacing permeabilized cells with isolated nuclei, we observed that the nuclear fraction contained mostly nuclease sensitive genomes, while the genomes in the supernatant were dominantly nuclease resistant. The fact that not all nuclear genomes could be degraded can be explained by a capsid interaction with the nuclear envelope, which was also part of this fraction or by an incomplete separation due to technical obstacles arising from the small volumes. The latter interpretation is supported by the observation that WGA was not fully blocking generation of nuclease sensitivity.

Most of the capsid in this experiment rested however in the supernatant and we conclude that the transport competence of isolated nuclei is much weaker than that of digitonin-permeabilized cells. In fact, quantification of transport and genome liberation by Rabe et al. showed that all MatC released their genome using similar capsid amounts[216].

The experiment further showed that a minor fraction of the genomes in the supernatant became nuclease sensitive. Aside of technical reasons linked to separation of the fractions as mentioned before, these results could also mean that MatC just need contact with the cytoplasmic face of the NPC for capsid opening. This interpretation is not contradictious to the presence of capsids in the nuclear basket as EM in thin sections hardly allow the visualization of capsids integrity. It is however in agreement with findings showing free cytoplasmic HBV DNA after infection although it must be considered that the authors used the same technique for separation of nuclei and cytoplasm[88].

In summary, we conclude that the used *in vitro* approaches were unable to demonstrate the need of Nup153 for genome release or capsid opening. However, our data can not exclude that a larger portion of Nup153 is required to trigger the disassembly of the capsid as we restricted our analysis to the C_2 fragment which only represents one part of the capsid interacting domains identified on this nucleoporin.

Establishment of the Anchor technology to HBV

In the light of the partially conflicting interpretations – genome release outside the nucleus or exclusively inside – and the limited significance of the test system,s we established a system of single genome visualization for HBV.

We first inserted the Anch sequence into the HBV genome. The disruption of at least one ORF required trans-complementation as it was successfully done by others using HBV[223,224]. Aside of controlling the formation of capsids, we also verified the formation of rcDNA, which depends on multiple DNA sequences, which are poorly characterized as 5R, 3R and M important for circularization. Southern blotting revealed a ratio of 90:10 rcDNA molecules: dsl DNA molecules, which is also found in wt HBV from cell culture and patients[225]. Transfection assays confirmed the functionality of the methods by showing the formation of green dots when the Anch constructs plasmids were added to OR-DGFP expressing cells. The dots were of different size and intensity as expected from DNA lipofection, which causes DNA aggregates.

Production of properly folded capsids was proven by migration in NAGE migration and polyclonal antibody reactivity, which were identical as wild type HBV virions produced in HepG2.2.15.

In addition, we established a purification method to purify high infectious particles. It was shown by Seitz et al. that using a heparin-based purification system two population of HBV virion were obtained: not binding to heparin and binding to heparin. The latter was proved to be highly infectious, when added to HepaRG cells[69].

Performing protein lipofection using capsids with the pseudo-typed genome, we observed that incubation of OR-DGFP-expressing HuH-7 cells with HBV EAP pseudo particles resulted in green dots exclusively into the nucleoplasm. Dots appeared at the nuclear membrane and remained immobile throughout the 4 h observation period.

To perform infection experiment we used NTCP expressing cells where we verify the expression of NTCP by IF and WB. The first confirmed the correct localization of the NTCP on the plasma membrane of HepG2 cells, while the second confirmed the glycosylation status of the receptor. In infection we observed only few positive cells after 2h p.i. that however increased after 3 days post infection. This, together with the lipofection experiment confirms that the limiting step during HBV infection is the entry into the host cells rather than the transport and the genome liberation.

HBV genomes undergo genome repair followed by chromatinization. The assays before gave proof that the genome of the Anch-bearing rcDNA is released, but did not give any evidence about detection after chromatinization and in particular integration; the latter important for applying the system to cancer research. We thus investigated if integrated HBV Anch sequences were still detectable. We observed that after lentiviral transduction of the HBV Anch sequence and 10 passages of the cells that, the pseudoviral genome was detectable in 11% of transduced cells. This suggests that the Anch did not allow efficient association of OR-DGFP despite of transcriptional activity of the inserted sequence. Detailed analysis showed that the dots were however still visible in not condensed chromatin. This indicates that not histone-association but the condensation of the chromatin suppressed DGFP-OR binding. We thus conclude that cccDNA could be detected after histone association but that the integration of integrates in particular in mitotic cells is limited.

In summary, our HBV Anch experiments showed that this system is a powerful tool for investigating HBV genome liberation and suggest that unsolved questions like the fate of the genome upon cell division can be addressed. Further, it could serve a tool in the validation of drugs targeting cccDNA but higher yield of infections must be obtained. It is nonetheless the first system allowing HBV genome detection not only on the single molecule level but also in real time. Its use in integration studies remains however more complicated.

List of abbreviations

Au: Australian Antigene

BSA: bovine serum albumin

BSL: biosafety level

cccDNA: covalently closed circular DNA

CD8 + T Cell: cytotoxic T cells

Cp: Core/Capsid Protein

cryo-EM: cryo Electron Microscopy.

CTD: carboxyl-terminal domain of the Core Protein

DAPI: 4',6-diamidino-2-phenylindole

DHBV: duck hepatitis B virus

DHBV: duck hepatitis B virus

DMSO: Dimethyl sulfoxide

DNA: Deoxyribonucleic acid

dNTP: Deoxynucleotide

DR: direct repeat

E. coli: *Escherichia coli*

EAP: Early Anch plasmid

EmpC: empty capsids

Enh: enhancer

FG-repeats: phenylalanine-glycine repeats

GRE: glucocorticoid-responsive element

GSHBV: grand squirrel hepatitis B virus

HBeAg: HBV preCore antigene

HBV: Hepatitis B virus

HBx: HBV protein x

HCC: hepatocellular carcinoma

HIV-1: Human immunodeficiency virus 1

hNTCP: human NTCP

HPSG: heparan sulfate proteoglycans

Hsp90: molecular chaperon complex 90

IBB: Importin β binding domain

IF: Immunofluorescence

ImmC: immature capsids

Imp.α: Importin α

Imp.β: Importin β

L/LHBs: HBV Large surface protein

LAP: Late Anch plasmid

LSD1: lysine-specific demethylase 1

M/MHBs: HBV Middle surface protein

MatC: mature capsids

MOI: Multiplicity of infection

NAGE: Native Agarose Gel Electrophoresis

NC: nucleocapsid of HBV

NEBD: Nuclear envelope breakdown

NES: nuclear export signal

NLS: Nuclear Localization Signal

NPC: Nuclear pore complex

NTCP: sodium taurocholate cotransporting polypeptide

NTD: amino-terminal domain of Cp

Nup153: Nucleoporine 153

ORF: open reading frame

PCR: Polymerase chain reaction

PKC: protein kinase C

RanBP: Ran binding protein

RanGAP: Ran GTP-ase activating protein

rcDNA: relaxed circular DNA

rHBc: recombinant *E. Coli* expressed capsids

RNA: Ribonucleic acid

RRL: rabbit reticulocyte lysate

RT: Reverse Transcriptase

S/SHBs: HBV Small surface protein

Smc: structural maintenance of chromosomes

SV40: simian virus 40

TCA: Trichloroacetic acid

TOP: topoisomerase

TP: terminal protein (domain of RT)

WGA: wheat germ agglutinin

WHV: woodchuck hepatitis virus

Bibliography

1. Blumberg, B. S., Gerstley, B. J., Hungerford, D. A., London, W. T. & Sutnick, A. I. A serum antigen (Australia antigen) in Down's syndrome, leukemia, and hepatitis. *Ann. Intern. Med.* **66,** 924–931 (1967).
2. Sutnick, A. I., London, W. T., Gerstley, B. J., Cronlund, M. M. & Blumberg, B. S. Anicteric hepatitis associated with Australia antigen. Occurrence in patients with Down's syndrome. *JAMA* **205,** 670–674 (1968).
3. London, W. T., Sutnick, A. I. & Blumberg, B. S. Australia antigen and acute viral hepatitis. *Ann. Intern. Med.* **70,** 55–59 (1969).
4. Dane, D. S., Cameron, C. H. & Briggs, M. Virus-like particles in serum of patients with Australia-antigen-associated hepatitis. *Lancet* **1,** 695–698 (1970).
5. Barker, L. F. *et al.* Transmission of type B viral hepatitis to chimpanzees. *J. Infect. Dis.* **127,** 648–662 (1973).
6. Block, T. M., Alter, H. J., London, W. T. & Bray, M. A historical perspective on the discovery and elucidation of the hepatitis B virus. *Antiviral Res.* **131,** 109–123 (2016).
7. Seeger, C. & Mason, W. S. Molecular Biology of Hepatitis B Virus Infection. *Virology* **0,** 672–686 (2015).
8. Schaefer, S. Hepatitis B virus taxonomy and hepatitis B virus genotypes. *World J Gastroenterol* **13,** 14–21 (2007).
9. Mason, W. S. Animal Models and the Molecular Biology of Hepadnavirus Infection. *Cold Spring Harb Perspect Med* **5,** a021352 (2015).
10. Winer, B. Y. & Ploss, A. Determinants of hepatitis B and delta virus host tropism. *Curr Opin Virol* **13,** 109–116 (2015).

11. WHO | Progress toward access to hepatitis B treatment worldwide. *WHO* Available at: http://www.who.int/hepatitis/news-events/cdc-hepatitis-b-article/en/. (Accessed: 29th August 2018)

12. Liang, T. J. Hepatitis B: The Virus and Disease. *Hepatology* **49,** S13–S21 (2009).

13. Trépo, C., Chan, H. L. Y. & Lok, A. Hepatitis B virus infection. *The Lancet* **384,** 2053–2063 (2014).

14. Yang, P. L. *et al.* Immune effectors required for hepatitis B virus clearance. *Proc Natl Acad Sci U S A* **107,** 798–802 (2010).

15. Boni, C. *et al.* Characterization of Hepatitis B Virus (HBV)-Specific T-Cell Dysfunction in Chronic HBV Infection. *J Virol* **81,** 4215–4225 (2007).

16. Bertoletti, A. *et al.* HLA class I-restricted human cytotoxic T cells recognize endogenously synthesized hepatitis B virus nucleocapsid antigen. *Proceedings of the National Academy of Sciences of the United States of America* **88,** 10445 (1991).

17. Penna, A. *et al.* Cytotoxic T lymphocytes recognize an HLA-A2-restricted epitope within the hepatitis B virus nucleocapsid antigen. *Journal of Experimental Medicine* **174,** 1565–1570 (1991).

18. McAleer, W. J. *et al.* Human hepatitis B vaccine from recombinant yeast. *Nature* **307,** 178–180 (1984).

19. Liang, T. J. *et al.* Present and Future Therapies of Hepatitis B: From Discovery to Cure. *Hepatology* **62,** 1893–1908 (2015).

20. Wittkop, L. *et al.* Inhibition of protein kinase C phosphorylation of hepatitis B virus capsids inhibits virion formation and causes intracellular capsid accumulation. *Cellular microbiology* (2010).

21. Hu, J., Toft, D. O. & Seeger, C. Hepadnavirus assembly and reverse transcription require a multi-component chaperone complex which is incorporated into nucleocapsids. *EMBO J* **16,** 59–68 (1997).

22. Nguyen, D. H., Ludgate, L. & Hu, J. Hepatitis B Virus-Cell Interactions and Pathogenesis. *J Cell Physiol* **216,** 289–294 (2008).

23. Seeger, C. & Mason, W. S. Hepatitis B Virus Biology. *Microbiol. Mol. Biol. Rev*. **64,** 51–68 (2000).

24. Heermann, K. H. *et al.* Large surface proteins of hepatitis B virus containing the pre-s sequence. *J Virol* **52,** 396–402 (1984).

25. Glebe, D. & Urban, S. Viral and cellular determinants involved in hepadnaviral entry. *World J. Gastroenterol.* **13,** 22–38 (2007).

26. Schmitt, S. *et al.* Structure of pre-S2 N- and O-linked glycans in surface proteins from different genotypes of hepatitis B virus. *J. Gen. Virol.* **85,** 2045–2053 (2004).

27. Watashi, K. HBV Slow Maturation Process Leads to Infection. *Trends in Microbiology* **24,** 597–599 (2016).

28. Bruss, V. A short linear sequence in the pre-S domain of the large hepatitis B virus envelope protein required for virion formation. *J Virol* **71,** 9350–9357 (1997).

29. Ostapchuk, P., Hearing, P. & Ganem, D. A dramatic shift in the transmembrane topology of a viral envelope glycoprotein accompanies hepatitis B viral morphogenesis. *EMBO J* **13,** 1048–1057 (1994).

30. Huovila, A. P., Eder, A. M. & Fuller, S. D. Hepatitis B surface antigen assembles in a post-ER, pre-Golgi compartment. *J. Cell Biol.* **118,** 1305–1320 (1992).

31. Cohen, B. J. & Richmond, J. E. Electron microscopy of hepatitis B core antigen synthesized in E. coli. *Nature* **296,** 677–679 (1982).

32. Crowther, R. A. *et al.* Three-dimensional structure of hepatitis B virus core particles determined by electron cryomicroscopy. *Cell* **77,** 943–950 (1994).

33. Petit, M. A. & Pillot, J. HBc and HBe antigenicity and DNA-binding activity of major core protein P22 in hepatitis B virus core particles isolated from the cytoplasm of human liver cells. *J. Virol.* **53,** 543–551 (1985).

34. Zlotnick, A. *et al.* Dimorphism of hepatitis B virus capsids is strongly influenced by the C-terminus of the capsid protein. *Biochemistry* **35,** 7412–7421 (1996).

35. Venkatakrishnan, B. & Zlotnick, A. The Structural Biology of Hepatitis B Virus: Form and Function. *Annu Rev Virol* **3,** 429–451 (2016).

36. Kann, M., Sodeik, B., Vlachou, A., Gerlich, W. H. & Helenius, A. Phosphorylation-dependent Binding of Hepatitis B Virus Core Particles to the Nuclear Pore Complex. *The Journal of Cell Biology* **145,** 45–55 (1999).

37. Wynne, S. A., Crowther, R. A. & Leslie, A. G. W. The Crystal Structure of the Human Hepatitis B Virus Capsid. *Molecular Cell* **3,** 771–780 (1999).

38. Zlotnick, A. *et al.* Localization of the C terminus of the assembly domain of hepatitis B virus capsid protein: Implications for morphogenesis and organization of encapsidated RNA. *PNAS* **94,** 9556–9561 (1997).

39. Zhou, S. & Standring, D. N. Hepatitis B virus capsid particles are assembled from core-protein dimer precursors. *Proc Natl Acad Sci U S A* **89,** 10046–10050 (1992).

40. Kann, M. & Gerlich, W. H. Effect of core protein phosphorylation by protein kinase C on encapsidation of RNA within core particles of hepatitis B virus. *J Virol* **68,** 7993–8000 (1994).

41. Rabe, B., Vlachou, A., Panté, N., Helenius, A. & Kann, M. Nuclear import of hepatitis B virus capsids and release of the viral genome. *PNAS* **100,** 9849–9854 (2003).

42. Schreiner, S. & Nassal, M. A Role for the Host DNA Damage Response in Hepatitis B Virus cccDNA Formation—and Beyond? *Viruses* **9,** 125 (2017).

43. Bock, C.-T., Schranz, P., Schröder, C. H. & Zentgraf, H. Hepatitis B virus genome is organized into nucleosomes in the nucleus of the infected cell. *Virus Genes* **8,** 215–229 (1994).

44. Zhang, Y.-Y. *et al.* Single-cell analysis of covalently closed circular DNA copy numbers in a hepadnavirus-infected liver. *Proc Natl Acad Sci U S A* **100,** 12372–12377 (2003).

45. Moraleda, G. *et al.* Lack of effect of antiviral therapy in nondividing hepatocyte cultures on the closed circular DNA of woodchuck hepatitis virus. *J. Virol.* **71,** 9392–9399 (1997).

46. 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).

47. Levrero, M. *et al.* Control of cccDNA function in hepatitis B virus infection. *J. Hepatol.* **51,** 581–592 (2009).

48. Tur-Kaspa, R. *et al.* The glucocorticoid receptor recognizes a specific nucleotide sequence in hepatitis B virus DNA causing increased activity of the HBV enhancer. *Virology* **167,** 630–633 (1988).

49. Minor, M. M. & Slagle, B. L. Hepatitis B Virus HBx Protein Interactions with the Ubiquitin Proteasome System. *Viruses* **6,** 4683–4702 (2014).

50. Bartenschlager, R. & Schaller, H. Hepadnaviral assembly is initiated by polymerase binding to the encapsidation signal in the viral RNA genome. *EMBO J.* **11,** 3413–3420 (1992).

51. Wang, G. H. & Seeger, C. The reverse transcriptase of hepatitis B virus acts as a protein primer for viral DNA synthesis. *Cell* **71,** 663–670 (1992).

52. Gearhart, T. L. & Bouchard, M. J. The hepatitis B virus HBx protein modulates cell cycle regulatory proteins in cultured primary human hepatocytes. *Virus Res.* **155,** 363–367 (2011).

53. Chirillo, P. *et al.* The hepatitis B virus X gene induces p53-mediated programmed cell death. *Proc Natl Acad Sci U S A* **94,** 8162–8167 (1997).

54. Elmore, L. W. *et al.* Hepatitis B virus X protein and p53 tumor suppressor interactions in the modulation of apoptosis. *Proc Natl Acad Sci U S A* **94,** 14707–14712 (1997).

55. Decorsière, A. *et al.* Hepatitis B virus X protein identifies the Smc5/6 complex as a host restriction factor. *Nature* **531,** 386–389 (2016).

56. Alarcon, V. *et al.* The enzymes LSD1 and Set1A cooperate with the viral protein HBx to establish an active hepatitis B viral chromatin state. *Scientific Reports* **6,** 25901 (2016).

57. Standring, D. N., Ou, J. H., Masiarz, F. R. & Rutter, W. J. A signal peptide encoded within the precore region of hepatitis B virus directs the secretion of a heterogeneous population of e antigens in Xenopus oocytes. *Proc Natl Acad Sci U S A* **85,** 8405–8409 (1988).

58. Messageot, F., Salhi, S., Eon, P. & Rossignol, J.-M. Proteolytic Processing of the Hepatitis B Virus e Antigen Precursor CLEAVAGE AT TWO FURIN CONSENSUS SEQUENCES. *J. Biol. Chem.* **278,** 891–895 (2003).

59. Tong, S. P. *et al.* In vitro replication competence of a cloned hepatitis B virus variant with a nonsense mutation in the distal pre-C region. *Virology* **181,** 733–737 (1991).

60. Chen, M. T. *et al.* A function of the hepatitis B virus precore protein is to regulate the immune response to the core antigen. *Proc Natl Acad Sci U S A* **101,** 14913–14918 (2004).

61. Liang, T. J. & Ghany, M. Hepatitis B e Antigen--the dangerous endgame of hepatitis B. *N. Engl. J. Med.* **347,** 208–210 (2002).

62. Schädler, S. & Hildt, E. HBV Life Cycle: Entry and Morphogenesis. *Viruses* **1,** 185–209 (2009).

63. Ganem, D. & Prince, A. M. Hepatitis B virus infection--natural history and clinical consequences. *N. Engl. J. Med.* **350,** 1118–1129 (2004).

64. Yan, H. *et al.* Sodium taurocholate cotransporting polypeptide is a functional receptor for human hepatitis B and D virus. *eLife* **1,** (2012).

65. Ni, Y. *et al.* Hepatitis B and D Viruses Exploit Sodium Taurocholate Co-transporting Polypeptide for Species-Specific Entry into Hepatocytes. *Gastroenterology* **146,** 1070-1083.e6 (2014).

66. Sureau, C. & Salisse, J. A conformational heparan sulfate binding site essential to infectivity overlaps with the conserved hepatitis B virus a-determinant. *Hepatology* **57,** 985–994 (2013).

67. Lindblom, A. & Fransson, L. A. Endothelial heparan sulphate: compositional analysis and comparison of chains from different proteoglycan populations. *Glycoconj. J.* **7,** 545–562 (1990).

68. Hayes, C. N. *et al.* Early events in hepatitis B virus infection: From the cell surface to the nucleus. *Journal of Gastroenterology and Hepatology* **31,** 302–309 (2016).

69. Seitz, S. *et al.* A Slow Maturation Process Renders Hepatitis B Virus Infectious. *Cell Host & Microbe* **20,** 25–35 (2016).

70. Anwer, M. S. & Stieger, B. Sodium-dependent bile salt transporters of the SLC10A transporter family: more than solute transporters. *Pflugers Arch.* **466,** 77–89 (2014).

71. Appelman, M. D., Chakraborty, A., Protzer, U., McKeating, J. A. & Graaf, S. F. J. van de. N-Glycosylation of the Na+-Taurocholate Cotransporting Polypeptide (NTCP)

Determines Its Trafficking and Stability and Is Required for Hepatitis B Virus Infection. *PLOS ONE* **12,** e0170419 (2017).

72. Yan, H. *et al.* Molecular Determinants of Hepatitis B and D Virus Entry Restriction in Mouse Sodium Taurocholate Cotransporting Polypeptide. *Journal of Virology* **87,** 7977–7991 (2013).

73. Macovei, A. *et al.* Hepatitis B Virus Requires Intact Caveolin-1 Function for Productive Infection in HepaRG Cells. *Journal of Virology* **84,** 243–253 (2010).

74. Huang, H.-C., Chen, C.-C., Chang, W.-C., Tao, M.-H. & Huang, C. Entry of Hepatitis B Virus into Immortalized Human Primary Hepatocytes by Clathrin-Dependent Endocytosis. *Journal of Virology* **86,** 9443–9453 (2012).

75. Cooper, A. & Shaul, Y. Clathrin-mediated Endocytosis and Lysosomal Cleavage of Hepatitis B Virus Capsid-like Core Particles. *J. Biol. Chem*. **281,** 16563–16569 (2006).

76. Macovei, A., Petrareanu, C., Lazar, C., Florian, P. & Branza-Nichita, N. Regulation of Hepatitis B Virus Infection by Rab5, Rab7, and the Endolysosomal Compartment. *J Virol* **87,** 6415–6427 (2013).

77. Rigg, R. J. & Schaller, H. Duck hepatitis B virus infection of hepatocytes is not dependent on low pH. *Journal of Virology* **66,** 2829–2836 (1992).

78. Funk, A., Mhamdi, M., Hohenberg, H., Will, H. & Sirma, H. pH-independent entry and sequential endosomal sorting are major determinants of hepadnaviral infection in primary hepatocytes. *Hepatology* **44,** 685–693 (2006).

79. Chojnacki, J., Anderson, D. A. & Grgacic, E. V. L. A Hydrophobic Domain in the Large Envelope Protein Is Essential for Fusion of Duck Hepatitis B Virus at the Late Endosome. *J Virol* **79,** 14945–14955 (2005).

80. Watashi, K. *et al.* Cyclosporin A and its analogs inhibit hepatitis B virus entry into cultured hepatocytes through targeting a membrane transporter, sodium taurocholate cotransporting polypeptide (NTCP). *Hepatology* **59,** 1726–1737 (2014).

81. Núñez, E. *et al.* Interaction of preS domains of hepatitis B virus with phospholipid vesicles. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1788,** 417–424 (2009).

82. Osseman, Q. *et al.* The chaperone dynein LL1 mediates cytoplasmic transport of empty and mature hepatitis B virus capsids. *Journal of Hepatology* **68,** 441–448 (2018).

83. Kann, M., Sodeik, B., Vlachou, A., Gerlich, W. H. & Helenius, A. Phosphorylation-dependent binding of hepatitis B virus core particles to the nuclear pore complex. *J. Cell Biol.* **145,** 45–55 (1999).

84. Panté, N. & Kann, M. Nuclear Pore Complex Is Able to Transport Macromolecules with Diameters of ~39 nm. *Mol. Biol. Cell* **13,** 425–434 (2002).

85. Schmitz, A. *et al.* Nucleoporin 153 Arrests the Nuclear Import of Hepatitis B Virus Capsids in the Nuclear Basket. *PLOS Pathog* **6,** e1000741 (2010).

86. Rabe, B. *et al.* Nuclear Entry of Hepatitis B Virus Capsids Involves Disintegration to Protein Dimers followed by Nuclear Reassociation to Capsids. *PLOS Pathog* **5,** e1000563 (2009).

87. Guo, H. *et al.* Characterization of the Intracellular Deproteinized Relaxed Circular DNA of Hepatitis B Virus: an Intermediate of Covalently Closed Circular DNA Formation. *J. Virol.* **81,** 12472–12484 (2007).

88. Cui, X., Ludgate, L., Ning, X. & Hu, J. Maturation-Associated Destabilization of Hepatitis B Virus Nucleocapsid. *J. Virol.* **87,** 11494–11503 (2013).

89. Königer, C. *et al.* Involvement of the host DNA-repair enzyme TDP2 in formation of the covalently closed circular DNA persistence reservoir of hepatitis B viruses. *PNAS* **111,** E4244–E4253 (2014).

90. Köck, J. & Schlicht, H. J. Analysis of the earliest steps of hepadnavirus replication: genome repair after infectious entry into hepatocytes does not depend on viral polymerase activity. *J. Virol.* **67,** 4867–4874 (1993).

91. Qi, Y. *et al.* DNA Polymerase κ Is a Key Cellular Factor for the Formation of Covalently Closed Circular DNA of Hepatitis B Virus. *PLOS Pathogens* **12,** e1005893 (2016).

92. Fourel, G. *et al.* Frequent activation of N-myc genes by hepadnavirus insertion in woodchuck liver tumours. *Nature* **347,** 294–298 (1990).

93. Wang, J., Chenivesse, X., Henglein, B. & Bréchot, C. Hepatitis B virus integration in a cyclin A gene in a hepatocellular carcinoma. *Nature* **343,** 555–557 (1990).

94. Dejean, A., Bougueleret, L., Grzeschik, K. H. & Tiollais, P. Hepatitis B virus DNA integration in a sequence homologous to v-erb-A and steroid receptor genes in a hepatocellular carcinoma. *Nature* **322,** 70–72 (1986).

95. Soussan, P. *et al.* In vivo expression of a new hepatitis B virus protein encoded by a spliced RNA. *J. Clin. Invest*. **105,** 55–60 (2000).

96. Pollack, J. R. & Ganem, D. An RNA stem-loop structure directs hepatitis B virus genomic RNA encapsidation. *J Virol* **67,** 3254–3263 (1993).

97. Patel, N. *et al.* HBV RNA pre-genome encodes specific motifs that mediate interactions with the viral core protein that promote nucleocapsid assembly. *Nat Microbiol* **2,** 17098 (2017).

98. Shakeel, S. *et al.* Genomic RNA folding mediates assembly of human parechovirus. *Nature Communications* **8,** 5 (2017).

99. Hu, J. & Seeger, C. Hsp90 is required for the activity of a hepatitis B virus reverse transcriptase. *Proc Natl Acad Sci U S A* **93,** 1060–1064 (1996).

100. Loeb, D. D., Hirsch, R. C. & Ganem, D. Sequence-independent RNA cleavages generate the primers for plus strand DNA synthesis in hepatitis B viruses: implications for other reverse transcribing elements. *EMBO J.* **10,** 3533–3540 (1991).

101. Summers, J., O'Connell, A. & Millman, I. Genome of hepatitis B virus: restriction enzyme cleavage and structure of DNA extracted from Dane particles. *Proc Natl Acad Sci U S A* **72,** 4597–4601 (1975).

102. Summers, J., Smolec, J. M. & Snyder, R. A virus similar to human hepatitis B virus associated with hepatitis and hepatoma in woodchucks. *Proc Natl Acad Sci U S A* **75,** 4533–4537 (1978).

103. Gerelsaikhan, T., Tavis, J. E. & Bruss, V. Hepatitis B virus nucleocapsid envelopment does not occur without genomic DNA synthesis. *J. Virol.* **70,** 4269–4274 (1996).

104. Wu, T. T., Coates, L., Aldrich, C. E., Summers, J. & Mason, W. S. In hepatocytes infected with duck hepatitis B virus, the template for viral RNA synthesis is amplified by an intracellular pathway. *Virology* **175,** 255–261 (1990).

105. Tuttleman, J. S., Pourcel, C. & Summers, J. Formation of the pool of covalently closed circular viral DNA in hepadnavirus-infected cells. *Cell* **47,** 451–460 (1986).

106. Wei, Y., Tavis, J. E. & Ganem, D. Relationship between viral DNA synthesis and virion envelopment in hepatitis B viruses. *J. Virol.* **70,** 6455–6458 (1996).

107. Watanabe, T. *et al.* Involvement of host cellular multivesicular body functions in hepatitis B virus budding. *Proc Natl Acad Sci U S A* **104,** 10205–10210 (2007).

108. Hu, J., Liu, K., Hu, J. & Liu, K. Complete and Incomplete Hepatitis B Virus Particles: Formation, Function, and Application. *Viruses* **9,** 56 (2017).

109. Luckenbaugh, L., Kitrinos, K. M., Delaney, W. E. & Hu, J. Genome-free hepatitis B virion levels in patient sera as a potential marker to monitor response to antiviral therapy. *J Viral Hepat* **22,** 561–570 (2015).

110. Chen, C. *et al.* Importin β Can Bind Hepatitis B Virus Core Protein and Empty Core-Like Particles and Induce Structural Changes. *PLOS Pathog* **12,** e1005802 (2016).

111. Ning, X. *et al.* Secretion of Genome-Free Hepatitis B Virus – Single Strand Blocking Model for Virion Morphogenesis of Para-retrovirus. *PLOS Pathog* **7,** e1002255 (2011).

112. Wang, J. *et al.* Serum hepatitis B virus RNA is encapsidated pregenome RNA that may be associated with persistence of viral infection and rebound. *J. Hepatol.* **65,** 700–710 (2016).

113. van Bömmel, F. *et al.* Serum hepatitis B virus RNA levels as an early predictor of hepatitis B envelope antigen seroconversion during treatment with polymerase inhibitors. *Hepatology* **61,** 66–76 (2015).

114. Jiang, B., Himmelsbach, K., Ren, H., Boller, K. & Hildt, E. Subviral Hepatitis B Virus Filaments, like Infectious Viral Particles, Are Released via Multivesicular Bodies. *J. Virol.* **90,** 3330–3341 (2015).

115. Blumberg, B. S. Australia antigen and the biology of hepatitis B. *Science* **197,** 17–25 (1977).

116. Maini, M. K. & Gehring, A. J. The role of innate immunity in the immunopathology and treatment of HBV infection. *J. Hepatol.* **64,** S60–S70 (2016).

117. Gazina, E. V., Fielding, J. E., Lin, B. & Anderson, D. A. Core protein phosphorylation modulates pregenomic RNA encapsidation to different extents in human and duck hepatitis B viruses. *J. Virol.* **74,** 4721–4728 (2000).

118. Deroubaix, A. *et al.* Expression of viral polymerase and phosphorylation of core protein determine core and capsid localization of the human hepatitis B virus. *Journal of General Virology* **96,** 183–195 (2015).

119. Machida, A. *et al.* Phosphorylation in the carboxyl-terminal domain of the capsid protein of hepatitis B virus: evaluation with a monoclonal antibody. *J Virol* **65,** 6024–6030 (1991).

120. Basagoudanavar, S. H., Perlman, D. H. & Hu, J. Regulation of Hepadnavirus Reverse Transcription by Dynamic Nucleocapsid Phosphorylation. *J. Virol.* **81,** 1641–1649 (2007).

121. Duclos-Vallée, J. C., Capel, F., Mabit, H. & Petit, M. A. Phosphorylation of the hepatitis B virus core protein by glyceraldehyde-3-phosphate dehydrogenase protein kinase activity. *J. Gen. Virol.* **79 (Pt 7),** 1665–1670 (1998).

122. Selzer, L., Kant, R., Wang, J. C.-Y., Bothner, B. & Zlotnick, A. Hepatitis B Virus Core Protein Phosphorylation Sites Affect Capsid Stability and Transient Exposure of the C-terminal Domain. *J. Biol. Chem*. **290,** 28584–28593 (2015).

123. Winer, B. Y., Ding, Q., Gaska, J. M. & Ploss, A. In vivo models of hepatitis B and C virus infection. *FEBS Lett*. **590,** 1987–1999 (2016).

124. Taylor, J. M. Hepatitis delta virus. *Virology* **344,** 71–76 (2006).

125. Lempp, F. A. *et al.* Evidence that hepatitis B virus replication in mouse cells is limited by the lack of a host cell dependency factor. *Journal of Hepatology* **64,** 556–564 (2016).

126. Lempp, F. A. *et al.* Sodium taurocholate cotransporting polypeptide is the limiting host factor of hepatitis B virus infection in macaque and pig hepatocytes. *Hepatology* **66,** 703–716 (2017).

127. Kim, D. H., Kang, H. S. & Kim, K.-H. Roles of hepatocyte nuclear factors in hepatitis B virus infection. *World J Gastroenterol* **22,** 7017–7029 (2016).

128. Bissig, K.-D. *et al.* Human liver chimeric mice provide a model for hepatitis B and C virus infection and treatment. *J. Clin. Invest*. **120,** 924–930 (2010).

129. Bility, M. T. *et al.* Hepatitis B Virus Infection and Immunopathogenesis in a Humanized Mouse Model: Induction of Human-Specific Liver Fibrosis and M2-Like Macrophages. *PLOS Pathogens* **10,** e1004032 (2014).

130. Gripon, P. *et al.* Infection of a human hepatoma cell line by hepatitis B virus. *PNAS* **99,** 15655–15660 (2002).

131. Sells, M. A., Chen, M. L. & Acs, G. Production of hepatitis B virus particles in Hep G2 cells transfected with cloned hepatitis B virus DNA. *PNAS* **84,** 1005–1009 (1987).

132. Gripon, P. *et al.* Hepatitis B virus infection of adult human hepatocytes cultured in the presence of dimethyl sulfoxide. *J. Virol.* **62,** 4136–4143 (1988).

133. Verrier, E. R., Colpitts, C. C., Schuster, C., Zeisel, M. B. & Baumert, T. F. Cell Culture Models for the Investigation of Hepatitis B and D Virus Infection. *Viruses* **8,** 261 (2016).

134. von Appen, A. & Beck, M. Structure Determination of the Nuclear Pore Complex with Three-Dimensional Cryo electron Microscopy. *Journal of Molecular Biology* **428,** 2001–2010 (2016).

135. D'Angelo, M. A., Anderson, D. J., Richard, E. & Hetzer, M. W. Nuclear Pores Form de Novo from Both Sides of the Nuclear Envelope. *Science* **312,** 440–443 (2006).

136. Raices, M. & D'Angelo, M. A. Nuclear pore complex composition: a new regulator of tissue-specific and developmental functions. *Nature Reviews Molecular Cell Biology* **13,** 687–699 (2012).

137. Maeshima, K. *et al.* Cell-cycle-dependent dynamics of nuclear pores: pore-free islands and lamins. *Journal of Cell Science* **119,** 4442–4451 (2006).

138. Frey, S. & Görlich, D. A Saturated FG-Repeat Hydrogel Can Reproduce the Permeability Properties of Nuclear Pore Complexes. *Cell* **130,** 512–523 (2007).

139. Patel, S. S., Belmont, B. J., Sante, J. M. & Rexach, M. F. Natively Unfolded Nucleoporins Gate Protein Diffusion across the Nuclear Pore Complex. *Cell* **129,** 83–96 (2007).

140. Strawn, L. A., Shen, T., Shulga, N., Goldfarb, D. S. & Wente, S. R. Minimal nuclear pore complexes define FG repeat domains essential for transport. *Nature Cell Biology* **6,** 197–206 (2004).

141. Paine, P. L., Moore, L. C. & Horowitz, S. B. Nuclear envelope permeability. *Nature* **254,** 109–114 (1975).

142. Ghavami, A., Giessen, E. van der & Onck, P. R. Energetics of Transport through the Nuclear Pore Complex. *PLOS ONE* **11,** e0148876 (2016).

143. Mohr, D., Frey, S., Fischer, T., Güttler, T. & Görlich, D. Characterisation of the passive permeability barrier of nuclear pore complexes. *EMBO J.* **28,** 2541–2553 (2009).

144. Ma, J., Goryaynov, A., Sarma, A. & Yang, W. Self-regulated viscous channel in the nuclear pore complex. *Proc Natl Acad Sci U S A* **109,** 7326–7331 (2012).

145. Yoshida, K. & Blobel, G. The Karyopherin Kap142p/Msn5p Mediates Nuclear Import and Nuclear Export of Different Cargo Proteins. *J Cell Biol* **152,** 729–740 (2001).

146. Weis, K. Importins and exportins: how to get in and out of the nucleus. *Trends in Biochemical Sciences* **23,** 185–189 (1998).

147. Siomi, H. & Dreyfuss, G. A nuclear localization domain in the hnRNP A1 protein. *The Journal of Cell Biology* **129,** 551–560 (1995).

148. Kalderon, D., Roberts, B. L., Richardson, W. D. & Smith, A. E. A short amino acid sequence able to specify nuclear location. *Cell* **39,** 499–509 (1984).

149. Lanford, R. E. & Butel, J. S. Construction and characterization of an SV40 mutant defective in nuclear transport of T antigen. *Cell* **37,** 801–813 (1984).

150. Lott, K. & Cingolani, G. The importin β binding domain as a master regulator of nucleocytoplasmic transport. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1813,** 1578–1592 (2011).

151. Palmeri, D. & Malim, M. H. Importin β Can Mediate the Nuclear Import of an Arginine-Rich Nuclear Localization Signal in the Absence of Importin α. *Mol. Cell. Biol.* **19,** 1218–1225 (1999).

152. Goldfarb, D. S., Corbett, A. H., Mason, D. A., Harreman, M. T. & Adam, S. A. Importin α: a multipurpose nuclear-transport receptor. *Trends in Cell Biology* **14,** 505–514 (2004).

153. Kobe, B. Autoinhibition by an internal nuclear localization signal revealed by the crystal structure of mammalian importin alpha. *Nat. Struct. Biol.* **6,** 388–397 (1999).

154. Harreman, M. T., Hodel, M. R., Fanara, P., Hodel, A. E. & Corbett, A. H. The Auto-inhibitory Function of Importin α Is Essentialin Vivo. *J. Biol. Chem.* **278,** 5854–5863 (2003).

155. Fanara, P., Hodel, M. R., Corbett, A. H. & Hodel, A. E. Quantitative analysis of nuclear localization signal (NLS)-importin alpha interaction through fluorescence depolarization. Evidence for auto-inhibitory regulation of NLS binding. *J. Biol. Chem.* **275,** 21218–21223 (2000).

156. Otsuka, S., Iwasaka, S., Yoneda, Y., Takeyasu, K. & Yoshimura, S. H. Individual binding pockets of importin-β for FG-nucleoporins have different binding properties and different sensitivities to RanGTP. *PNAS* **105,** 16101–16106 (2008).

157. Lim, R. Y. H. *et al.* Nanomechanical Basis of Selective Gating by the Nuclear Pore Complex. *Science* **318,** 640–643 (2007).

158. Frey, S., Richter, R. P. & Görlich, D. FG-Rich Repeats of Nuclear Pore Proteins Form a Three-Dimensional Meshwork with Hydrogel-Like Properties. *Science* **314,** 815–817 (2006).

159. Lim, R. Y. H. *et al.* Flexible phenylalanine-glycine nucleoporins as entropic barriers to nucleocytoplasmic transport. *PNAS* **103,** 9512–9517 (2006).

160. Peters, R. Translocation through the nuclear pore: Kaps pave the way. *Bioessays* **31,** 466–477 (2009).

161. Sakiyama, Y., Mazur, A., Kapinos, L. E. & Lim, R. Y. H. Spatiotemporal dynamics of the nuclear pore complex transport barrier resolved by high-speed atomic force microscopy. *Nature Nanotechnology* **11,** 719–723 (2016).

162. Yokoyama, N. *et al.* A giant nucleopore protein that binds Ran/TC4. *Nature* **376,** 184–188 (1995).

163. Moroianu, J., Hijikata, M., Blobel, G. & Radu, A. Mammalian karyopherin alpha 1 beta and alpha 2 beta heterodimers: alpha 1 or alpha 2 subunit binds nuclear localization signal and beta subunit interacts with peptide repeat-containing nucleoporins. *Proc. Natl. Acad. Sci. U.S.A.* **92,** 6532–6536 (1995).

164. Walther, T. C. *et al.* The nucleoporin Nup153 is required for nuclear pore basket formation, nuclear pore complex anchoring and import of a subset of nuclear proteins. *The EMBO Journal* **20,** 5703–5714 (2001).

165. Ogawa, Y., Miyamoto, Y., Oka, M. & Yoneda, Y. The Interaction Between Importin-α and Nup153 Promotes Importin-α/β-Mediated Nuclear Import. *Traffic* **13,** 934–946 (2012).

166. Görlich, D., Seewald, M. J. & Ribbeck, K. Characterization of Ran-driven cargo transport and the RanGTPase system by kinetic measurements and computer simulation. *The EMBO Journal* **22,** 1088–1100 (2003).

167. Kutay, U., Bischoff, F. R., Kostka, S., Kraft, R. & Görlich, D. Export of importin alpha from the nucleus is mediated by a specific nuclear transport factor. *Cell* **90,** 1061–1071 (1997).

168. Bischoff, F. R. & Ponstingl, H. Catalysis of guanine nucleotide exchange of Ran by RCC1 and stimulation of hydrolysis of Ran-bound GTP by Ran-GAP1. *Meth. Enzymol.* **257,** 135–144 (1995).

169. Ribbeck, K., Lipowsky, G., Kent, H. M., Stewart, M. & Görlich, D. NTF2 mediates nuclear import of Ran. *EMBO J.* **17,** 6587–6598 (1998).

170. Al-Haboubi, T., Shumaker, D. K., Köser, J., Wehnert, M. & Fahrenkrog, B. Distinct association of the nuclear pore protein Nup153 with A- and B-type lamins. *Nucleus* **2,** 500–509 (2011).

171. Lelek, M. *et al.* Chromatin organization at the nuclear pore favours HIV replication. *Nat Commun* **6,** 6483 (2015).

172. Sukegawa, J. & Blobel, G. A nuclear pore complex protein that contains zinc finger motifs, binds DNA, and faces the nucleoplasm. *Cell* **72,** 29–38 (1993).

173. Nakielny, S., Shaikh, S., Burke, B. & Dreyfuss, G. Nup153 is an M9-containing mobile nucleoporin with a novel Ran-binding domain. *EMBO J.* **18,** 1982–1995 (1999).

174. Campadelli-Fiume, G., Menotti, L., Avitabile, E. & Gianni, T. Viral and cellular contributions to herpes simplex virus entry into the cell. *Current Opinion in Virology* **2,** 28–36 (2012).

175. Pasdeloup, D., Blondel, D., Isidro, A. L. & Rixon, F. J. Herpesvirus Capsid Association with the Nuclear Pore Complex and Viral DNA Release Involve the Nucleoporin CAN/Nup214 and the Capsid Protein pUL25. *J. Virol.* **83,** 6610–6623 (2009).

176. Copeland, A. M., Newcomb, W. W. & Brown, J. C. Herpes Simplex Virus Replication: Roles of Viral Proteins and Nucleoporins in Capsid-Nucleus Attachment. *J. Virol.* **83,** 1660–1668 (2009).

177. Ojala, P. M., Sodeik, B., Ebersold, M. W., Kutay, U. & Helenius, A. Herpes Simplex Virus Type 1 Entry into Host Cells: Reconstitution of Capsid Binding and Uncoating at the Nuclear Pore Complex In Vitro. *Mol. Cell. Biol.* **20,** 4922–4931 (2000).

178. Shahin, V. *et al.* The genome of HSV-1 translocates through the nuclear pore as a condensed rod-like structure. *Journal of Cell Science* **119,** 23–30 (2006).

179. Molineux, I. J. No syringes please, ejection of phage T7 DNA from the virion is enzyme driven. *Molecular Microbiology* **40,** 1–8 (2001).

180. Newcomb, W. W., Cockrell, S. K., Homa, F. L. & Brown, J. C. Polarized DNA Ejection from the Herpesvirus Capsid. *Journal of Molecular Biology* **392,** 885–894 (2009).

181. Snoussi, K. & Kann, M. Interaction of parvoviruses with the nuclear envelope. *Advances in Biological Regulation* **54,** 39–49 (2014).

182. Mani, B. *et al.* Low pH-dependent endosomal processing of the incoming parvovirus minute virus of mice virion leads to externalization of the VP1 N-terminal sequence (N-VP1), N-VP2 cleavage, and uncoating of the full-length genome. *J. Virol.* **80,** 1015–1024 (2006).

183. Porwal, M. *et al.* Parvoviruses Cause Nuclear Envelope Breakdown by Activating Key Enzymes of Mitosis. *PLOS Pathog* **9,** e1003671 (2013).

184. Riolobos, L. *et al.* Viral Oncolysis That Targets Raf-1 Signaling Control of Nuclear Transport. *J. Virol.* **84,** 2090–2099 (2010).

185. Hübner, S., Xiao, C.-Y. & Jans, D. A. The Protein Kinase CK2 Site (Ser111/112) Enhances Recognition of the Simian Virus 40 Large T-antigen Nuclear Localization Sequence by Importin. *J. Biol. Chem*. **272,** 17191–17195 (1997).

186. Jans, D. A., Ackermann, M. J., Bischoff, J. R., Beach, D. H. & Peters, R. p34cdc2-mediated phosphorylation at T124 inhibits nuclear import of SV-40 T antigen proteins. *The Journal of Cell Biology* **115,** 1203–1212 (1991).

187. Nakabayashi, H., Taketa, K., Miyano, K., Yamane, T. & Sato, J. Growth of Human Hepatoma Cell Lines with Differentiated Functions in Chemically Defined Medium. *Cancer Res* **42,** 3858–3863 (1982).

188. König, A. *et al.* Kinetics of the bile acid transporter and hepatitis B virus receptor Na+/taurocholate cotransporting polypeptide (NTCP) in hepatocytes. *Journal of Hepatology* **61,** 867–875 (2014).

189. Porterfield, J. Z. *et al.* Full-Length Hepatitis B Virus Core Protein Packages Viral and Heterologous RNA with Similarly High Levels of Cooperativity. *J Virol* **84,** 7174–7184 (2010).

190. Jursch, C. A. *et al.* Molecular approaches to validate disinfectants against human hepatitis B virus. *Med. Microbiol. Immunol.* **190,** 189–197 (2002).

191. Kutay, U., Izaurralde, E., Bischoff, F. R., Mattaj, I. W. & Görlich, D. Dominant-negative mutants of importin-beta block multiple pathways of import and export through the nuclear pore complex. *EMBO J* **16,** 1153–1163 (1997).

192. Wittkop, L. *et al.* Inhibition of protein kinase C phosphorylation of hepatitis B virus capsids inhibits virion formation and causes intracellular capsid accumulation: HBV capsid phosphorylation by PKC. *Cellular Microbiology* **12,** 962–975 (2010).

193. Chang, C.-W. *et al.* Epstein-Barr Virus Protein Kinase BGLF4 Targets the Nucleus through Interaction with Nucleoporins. *J. Virol.* **86,** 8072–8085 (2012).

194. Kann, M., Thomssen, R., Köchel, H. G. & Gerlich, W. H. Characterization of the endogenous protein kinase activity of the hepatitis B virus. *Arch. Virol. Suppl.* **8,** 53–62 (1993).

195. Wang, J. C.-Y., Dhason, M. S. & Zlotnick, A. Structural Organization of Pregenomic RNA and the Carboxy-Terminal Domain of the Capsid Protein of Hepatitis B Virus. *PLOS Pathogens* **8,** e1002919 (2012).

196. Inhibition of in vitro nuclear transport by a lectin that binds to nuclear pores. *J Cell Biol* **104,** 189–200 (1987).

197. Li, B. & Kohler, J. J. Glycosylation of the nuclear pore. *Traffic* **15,** 347–361 (2014).

198. Glebe, D. *et al.* Optimised Conditions for the Production of Hepatitis B Virus from Cell Culture. *INT* **44,** 370–378 (2001).

199. Miyakawa, K. *et al.* Development of a cell-based assay to identify hepatitis B virus entry inhibitors targeting the sodium taurocholate cotransporting polypeptide. *Oncotarget* **9,** 23681–23694 (2018).

200. Rabe, B., Glebe, D. & Kann, M. Lipid-Mediated Introduction of Hepatitis B Virus Capsids into Nonsusceptible Cells Allows Highly Efficient Replication and Facilitates the Study of Early Infection Events. *J. Virol.* **80,** 5465–5473 (2006).

201. Kim, T. H. *et al.* The Degrees of Hepatocyte Cytoplasmic Expression of Hepatitis B Core Antigen correlate with Histologic Activity of Liver Disease in the Young Patients with Chronic Hepatitis B Infection. *Journal of Korean Medical Science* **21,** 279–283 (2006).

202. Naoumov, N. V. *et al.* Detection of hepatitis B virus antigens in liver tissue. A relation to viral replication and histology in chronic hepatitis B infection. *Gastroenterology* **99,** 1248–1253 (1990).

203. Liu, C.-J. *et al.* Hepatitis B Virus Basal Core Promoter Mutation and DNA Load Correlate with Expression of Hepatitis B Core Antigen in Patients with Chronic Hepatitis B. *J Infect Dis.* **199,** 742–749 (2009).

204. Chu, C. M., Yeh, C. T., Chien, R. N., Sheen, I. S. & Liaw, Y. F. The degrees of hepatocyte nuclear but not cytoplasmic expression of hepatitis B core antigen reflect the level of viral replication in chronic hepatitis B virus infection. *J. Clin. Microbiol.* **35,** 102–105 (1997).

205. Moroianu, J., Blobel, G. & Radu, A. Nuclear protein import: Ran-GTP dissociates the karyopherin alphabeta heterodimer by displacing alpha from an overlapping binding site on beta. *PNAS* **93,** 7059–7062 (1996).

206. Ceres, P. & Zlotnick, A. Weak protein-protein interactions are sufficient to drive assembly of hepatitis B virus capsids. *Biochemistry* **41,** 11525–11531 (2002).

207. Li, Q., Yafal, A. G., Lee, Y. M., Hogle, J. & Chow, M. Poliovirus neutralization by antibodies to internal epitopes of VP4 and VP1 results from reversible exposure of these sequences at physiological temperature. *Journal of Virology* **68,** 3965–3970 (1994).

208. Lewis, J. K., Bothner, B., Smith, T. J. & Siuzdak, G. Antiviral agent blocks breathing of the common cold virus. *Proc Natl Acad Sci U S A* **95,** 6774–6778 (1998).

209. Zachariae, U. & Grubmüller, H. Importin-beta: structural and dynamic determinants of a molecular spring. *Structure* **16,** 906–915 (2008).

210. Cingolani, G., Petosa, C., Weis, K. & Müller, C. W. Structure of importin-beta bound to the IBB domain of importin-alpha. *Nature* **399,** 221–229 (1999).

211. Hodel, M. R., Corbett, A. H. & Hodel, A. E. Dissection of a nuclear localization signal. *J. Biol. Chem.* **276,** 1317–1325 (2001).

212. Selzer, L., Katen, S. P. & Zlotnick, A. The Hepatitis B Virus Core Protein Intradimer Interface Modulates Capsid Assembly and Stability. *Biochemistry* **53,** 5496–5504 (2014).

213. Falces, J. *et al.* Recognition of nucleoplasmin by its nuclear transport receptor importin α/β: insights into a complete import complex. *Biochemistry* **49,** 9756–9769 (2010).

214. Maimon, T. & Medalia, O. Perspective on the metazoan nuclear pore complex. *Nucleus* **1,** 383–386 (2010).

215. Deres, K. *et al.* Inhibition of Hepatitis B Virus Replication by Drug-Induced Depletion of Nucleocapsids. *Science* **299,** 893–896 (2003).

216. Rabe, B. *et al.* Nuclear Entry of Hepatitis B Virus Capsids Involves Disintegration to Protein Dimers followed by Nuclear Reassociation to Capsids. *PLOS Pathog* **5,** e1000563 (2009).

217. Dhason, M. S., Wang, J. C.-Y., Hagan, M. F. & Zlotnick, A. Differential assembly of Hepatitis B Virus core protein on single- and double-stranded nucleic acid suggest the dsDNA-filled core is spring-loaded. *Virology* **430,** 20–29 (2012).

218. Seth, P., Fitzgerald, D. J., Willingham, M. C. & Pastan, I. Role of a low-pH environment in adenovirus enhancement of the toxicity of a Pseudomonas exotoxin-epidermal growth factor conjugate. *J Virol* **51,** 650–655 (1984).

219. Wiethoff, C. M., Wodrich, H., Gerace, L. & Nemerow, G. R. Adenovirus Protein VI Mediates Membrane Disruption following Capsid Disassembly. *J Virol* **79,** 1992–2000 (2005).

220. Hsu, M. *et al.* Hepatitis C virus glycoproteins mediate pH-dependent cell entry of pseudotyped retroviral particles. *Proc. Natl. Acad. Sci. U.S.A.* **100,** 7271–7276 (2003).

221. Tscherne, D. M. *et al.* Time- and Temperature-Dependent Activation of Hepatitis C Virus for Low-pH-Triggered Entry. *J Virol* **80,** 1734–1741 (2006).

222. Kann, M., Bischof, A. & Gerlich, W. H. In vitro model for the nuclear transport of the hepadnavirus genome. *Journal of Virology* **71,** 1310 (1997).

223. Chiang, P. W. *et al.* Encapsidation of truncated human hepatitis B virus genomes through trans-complementation of the core protein and polymerase. *Virology* **176,** 355–361 (1990).

224. Yamada, M. *et al.* Three distinct Southern blot hybridization patterns of HBV-DNA in the sera of HBV carriers. *Tohoku J. Exp. Med.* **170,** 219–228 (1993).

225. Hantz, O. *et al.* Persistence of the hepatitis B virus covalently closed circular DNA in HepaRG human hepatocyte-like cells. *Journal of General Virology* **90,** 127–135 (2009).