II INTRODUCTION

II.1 NON VIRAL GENE TRANSFER

Gene therapy describes a variety of approaches that are based on the treatment of diseases by the efficient and safe transfer of genetic material into target cells and tissue. The ideal gene delivery system should be specifically targeting, biodegradable, non-toxic, non-inflammatory, non-immunogenic and stable for storage. It should also have a large capacity for genetic material, allow an efficient transfection and the capacity to be produced in high concentrations at low cost.

The transfer of genetic information into the target cells can be achieved by viral or non- viral transfer systems. Viral transfer systems are characterized by a very high efficiency [3] but on the other hand gene transfer by recombinant viruses can harbor certain risks. Viral vectors have been implicated in the death of at least one patient, leading to the suspension of clinical trials [4,5]. Indeed, these vectors suffer from several drawbacks, including a need for packaging cell lines, and problems with safety, toxicity, the elicitation of an immune response, and the lack of cell-specific targeting. Furthermore, viral vector systems are rapidly cleared from the circulation, limiting transfection to 'first-pass' organs, such as the lungs, liver and spleen [6]. Recombination events that lead to the production of a replicating virus are also a remote danger.

Hence, non-viral carriers have several advantages over their viral counterparts: (1) they are easy to prepare and to scale-up, (2) they are more flexible with regard to the size of the DNA being transferred, (3) they are generally safer *in vivo*, (4) they elicit only a reduced immune response and can therefore easier be administered repeatedly. The low cost, consistent standard of production, greater safety and high flexibility also make these compounds attractive [6]. Furthermore, non-viral agents can be used in different combinations, providing flexibility in achieving particular therapeutic aims.

Few non-viral vectors have reached clinical trials, although many have been described, including naked DNA, DNA–cationic-liposome complexes (lipoplexes), DNA–polymer complexes and combinations of these [7-10]. Non-viral vector systems are usually either composed of a plasmid-based expression cassette alone ('naked' DNA), or are prepared with a synthetic amphipathic DNA-complexing agent [9,10]. However, currently available non-viral vectors are generally less efficient in delivering DNA and in initiating gene expression than their viral counterparts, particularly when used *in vivo* [11]. For reasons of safety and practicality, non-viral vectors

have several advantages compared with viral vectors. Unfortunately, broad clinical application of non-viral vectors has been hampered by low transfection efficiency and stability.

To solve this problem, non-viral vectors were modified with peptides to increase intracellular gene delivery. Many anionic pH-sensitive peptides [12] and cationic fusogenic peptides [13] show an enhancing effect on gene expression mediated by cationic liposomes. These peptides show membrane disrupting activities in weakly acidic conditions, which are similar to that in the endosome compartment, and could enhance the translocation of the DNA to cytosol. Rittner et al [14] reported that a bifunctional peptide with both DNA-binding and membrane-disrupting activities showed significant gene expression in the lung after tail-vein injection. Inefficient entry of DNA into the nucleus is a major limiting step in the development of non-viral gene delivery system [15]. The problem is particularly serious in nondividing cells, where entry into the nucleus is thought to occur only through the nuclear pore complex [16].

To achieve active transport to the nucleus, nucleus localizing signal (NLS) peptides have been widely used. Synthetic NLS, notably those homologous to that of SV40 T antigen, have been shown to direct nuclear import of a wide range of non-karyophilic proteins [17]. Several attempts to target DNA to the nucleus with a similar approach have been published. These include the use of electrostatic binding of DNA to cationic NLS-containing proteins or peptides, as well as covalent attachment of NLS peptides to DNA [18]. Despite the fact that peptide-based modifications can improve the active transport of DNA into the nucleus, the main disadvantages of synthetic vectors are formulation instability and inactivation in blood, a relatively low transfection efficiency and poor targeting [19].

II.1.1 Learning from virus

Because many viruses replicate in the nucleus of their host cells, they have developed several efficient methods for transporting their genome into this compartment [20]. To gain access to the nucleus, viruses make use of two quite different overall strategies. They can stay in the cytosol until the cell undergoes mitosis. Used by many retroviruses, this strategy leads to a major restriction in terms of cell tropism, since they can only replicate in dividing cells [21]. More commonly, the genome of incoming viruses is delivered directly through the nuclear envelope into the nucleoplasm. As described above the mechanisms and strategies by which incoming viruses enter the nucleus differ considerably among virus families. The size of the capsids and nucleoproteins is one of the main factors that determine the

mechanism. Very few viruses are small enough to pass through the pores intact without apparent deformation (parvovirus, see **Figure 13**). Larger capsids, for example from HIV or influenza virus, undergo partial or full disassembly in the cytosol or at the nuclear envelope. The available evidence indicates that viruses enter the nucleus through the nuclear pore complexes (NPCs). Targeting and transport of viral genomes to the nucleus depends on nuclear localization sequences exposed on the surface of the capsid particle. There is evidence that import of viral nucleic acids into nuclei of host cells utilizes pathways similar to those of nuclear protein import [18,22].

The use of unprotected DNA for *in vivo* gene transfer approaches results in a very low efficiency. This is due to the fact that naked DNA *in vivo* is rapidly degraded and furthermore is not able to reach target cells in sufficient amounts [23]. Viral strategies and protein can be used to address these problems. Learning from the viruses, it is clear that an effective transport of plasmid DNA into the nucleus would overcome the limitation of cytoplasmic degradation and improve the efficiency of gene transfer.

Therefore virus like particles (VLPs) could be an alternative. Utilizing evolutionary developed features of viral proteins VLPs combine advantages of both systems. Like in the viral system the genetic material is packaged and by this protected against degradation. Application of VLPs for gene transfer is not unprecedented, i.e. recombinant rabbit hemorrhagic disease virus particles [24], pseudo virus of human papillomavirus type 16 [25] or VLPs build out of HBV surface proteins [26]. Here the gene transfer depends on the presence of a receptor protein specific for surface proteins of these VLPs resulting in receptor mediated endocytosis and the internalized genetic material has to pass the endosomal compartment that represents the bottleneck of this approach.

II.1.2 Cell permeable peptides

The plasma membrane consists of a lipid bilayer into which proteins and glycoproteins are inserted. The hydrophobic nature of these lipids makes it impossible for hydrophilic compounds to cross the membrane unless specific mechanisms have evolved for this purpose. Some hydrophilic cargoes carried through the membrane by active transport. This transport is mediated by transmembrane shuttle proteins and gains direct access to the cytoplasm. For larger proteins, internalization occurs by endocytosis, resulting in uptake of cargo into endocytic vesicles. Subsequent steps are then needed for release into the cytoplasm [27].

Cell-permeable peptides (CPPs) are defined by their ability to translocate the cell membrane and gain access to the cellular interior (cytoplasmic and/or nuclear compartments in live cells after internalization). The term CPP includes synthetic cell-permeable or -penetrating peptides, protein-transduction domains (PTD) and membrane-translocating sequences (MTS). In generally they consist of less than 30 amino acids, have a net positive charge and the ability to transport several different cargoes into the cytoplasm and nucleus in a seemingly energy-independent manner [28]. The initial discovery of CPPs originated from the unexpected observation that certain full-length proteins or protein domains can translocate across the plasma membrane. This was first shown for the HIV Tat transactivator [29] and for the homeodomain of the *Drosophila melanogaster* transcription factor Antennapedia [30], and has since expanded to include 'non-natural' peptides that share this property [27]. Examples of sequences of known CPPs are listed in **Table 1**.

Table 1: Selection of known CPPs
IP = Isoelectric point calculated by WinPep, MW = Molecular weight in Da, Rf = References.

Name	Sequence	aa	IP	MW	Rf
Penetratin, pAntp	RQIKIWFQNRRMKWKK	16	12,4	2247	[31]
HIV TAT	GRKKRRQRRRPPQ	13	12,7	1719	[32]
MAP	KLALKLALKALKAALKLA-amide	18	11,4	1878	[33]
Transportan	GWTLNSAGYLLGKINLKALAALAKKIL-amide	27	10,9	2842	[34]
Transportan 10	AGYLLGKINLKALAALAKKIL-amide	21	10,9	2183	[35]
R7 peptide	RRRRRRR	7	12,8	1111	[36]
pVEC	LLIILRRRIRKQAHAHSK-amide	18	12,5	2210	[37]
MPG peptide	GALFLGWLGAAGSTMGAPKKKRKV-amide	24	11,8	2445	[38]
KALA peptide	WEAKLAKALAKALAKHLAKALAKALKACEA	30	10,7	3132	[39]
Buforin 2	TRSSRAGLQFPVGRVHRLLRK	21	12,2	2435	[40]

Reports show that translocation occurs via a sofar unknown mechanism that is not affected by various endocytosis inhibitors or low temperatures (i.e. +4°C) [41].

Despite the lack of understanding what the mechanism of endogenous process is, these proteins are proving to be an extremely useful tool for targeting heterologous cargoes into cells. Apart from being a mild and effective tool for access to different cellular organelles *in vitro*, CPPs have been used as vectors for biologically active molecules and as biologically active molecules in gene regulation *in vivo*, with promising results (**Figure 1**).

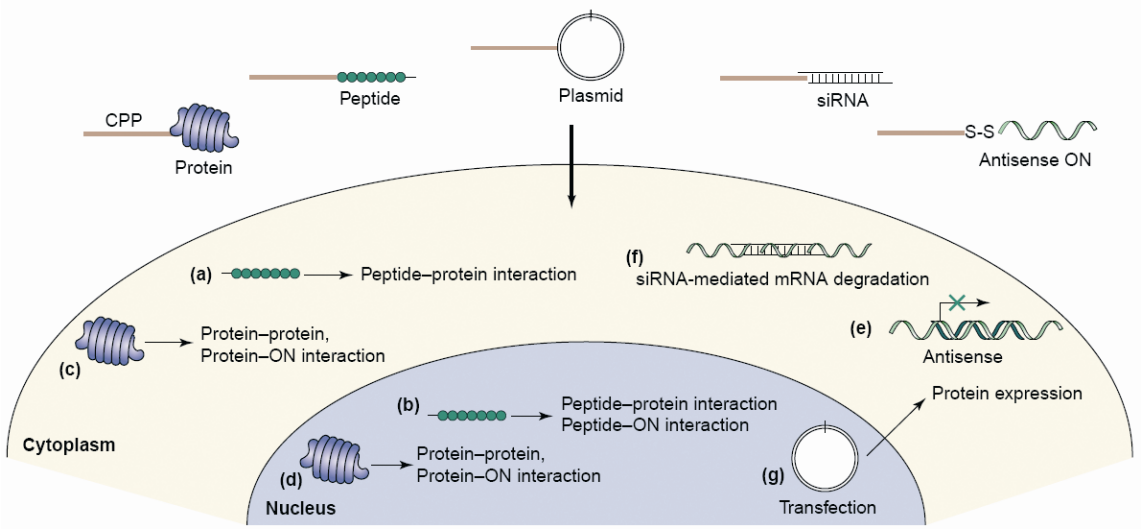

Figure 1: Selection of applications for peptide-mediated translocation.
a) Peptide–protein interaction in the cytoplasm. **b)** Peptide–protein and peptide–ON interaction in the nucleus. **c)** Protein–protein interaction in the cytoplasm. **d)** Protein–protein and protein–ON interaction in the nucleus. **e)** Antisense ON–mRNA hybridization. **f)** siRNA-mediated mRNA degradation. **g)** Transfected plasmid and protein expression. Abbreviations: CPP, cell-penetrating peptide; ON, oligonucleotide. [28]

A critical point is to distinguish between cell-associated peptides and truly intracellular peptides. It is also possible that some peptides might be, in part, taken up by cells after fixation: indeed, this is claimed for VP22 (another described CPP) [42] and may be caused by its high basic residue content. To overcome this potential problem, direct observation by microscopy of live cells pre-incubated with fluorescent peptides is now preferred and can be used to measure a wide range of incubation times (from seconds to hours) or concentrations accurately. Furthermore, recent findings in reevaluation of the mechanism of cellular uptake suggest that the some of CPP translocation occurs via an energy-dependent pathway and that CPP translocation is reduced by known endocytosis inhibitors [41,43-45].

Despite the fact that the actual pathway of CPP uptake has yet to be elucidated, peptides are still of great significance for improving cellular delivery locally. Several successful CPP-application with distinct physiological effects have been described that highlight the diversity of these applications, and their ability to answer physiological questions that are difficult to address by other means [28,46-51]. The efforts made so far indicate that CPPs are a promising tool for future pharmaceutical applications.

II.1.3 Translocation motif - TLM

A novel cell permeability mediating peptide (TLM, translocation motif) was identified in the PreS-domain of HBV [52] (see II.2.7). The TLM mediates the energy-independent and receptor-independent transfer of peptides, nucleic acids and proteins when fused to them, without affecting the integrity of the membrane [52-54]. The functionality of the TLM is highly conserved between all *Hepadnaviridae* suggesting a crucial role for the viral life cycle [52]. The TLM is a 12 amino acid encompassing domain that forms an amphipathic alpha helix (**Figure 2**).

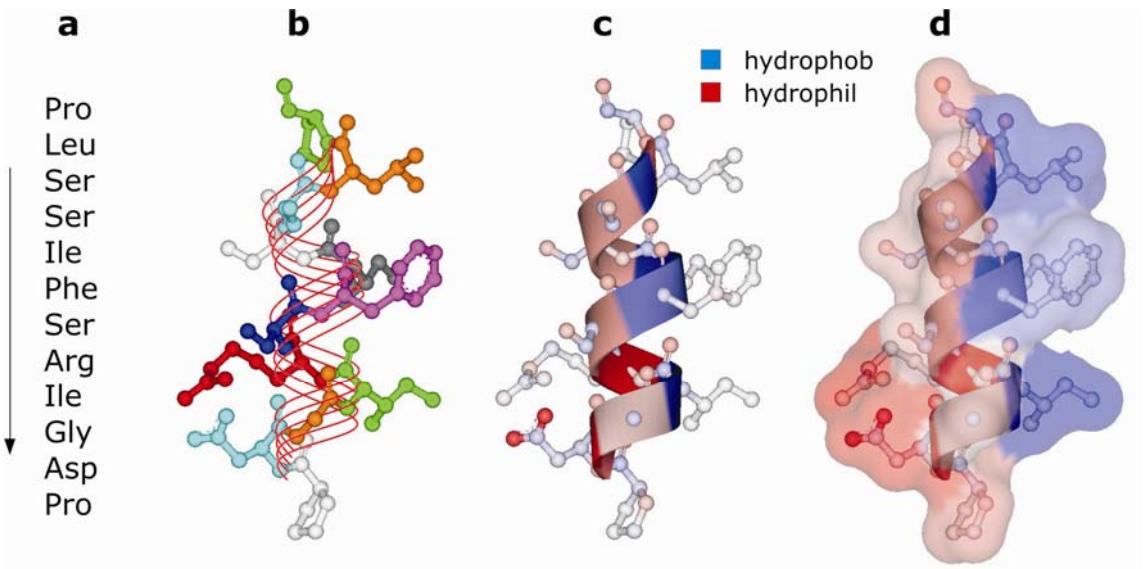

Figure 2: Structure and characteristics of TLM
a) The Translocation motif (TLM) consists of 12 amino acids. **b)** TLM builds an α-helical secondary structure visualized by central line ribbon. The ball and stick model is colored by different amino acids. **c)** Atoms are colored by charge (blue = negative, red = positive and white/gray = neutral) and the central solid ribbon by hydrophobicity. **d)** The solvent and transparent surface of the molecule is colored by hydrophobicity. Images are generated by accelrys ViewerLite 5.

The amphipathic character of cell permeable peptide motifs has repeatedly been described as a requirement for their ability to translocate via biological membranes [55]. In the case of TLM, previous mutagenesis studies and CD spectroscopy revealed the existence of an amphipathic α-helix between amino acids 41 and 52 of the PreS2 domain (PLSSIFSRIGDP; subtype ayw) [56]. Moreover, this amphipatic α-helix has been shown to mediate dimerization of PreS2. Replacement of three amino acid residues within the amphipathic α-helix by serine (PreS2–3S-mutant; PSSSSSSRIGDP) resulted in a loss of the α-helical structure and lead to β-sheet conformation of the sequence of interest, thus considerably reducing the oligomerization

tendency of PreS2 [56]. Further on, destruction of the TLM by these mutations abolishes the cell permeability.
The small size of the TLM of 12 aa, its high transfer efficiency (about 100% of the cells are affected), its low immunogenicity, the lack of any activator function of the TLM, its spreading capacity, and its defined structure function relation make it a helpful tool for the transfer of biomolecules into cells or tissues [52,54].

II.1.4 Relevance of TLM for the life cycle of HBV

The best example for the function of TLM is the hepatitis B virus itself. It was demonstrated that hepadnaviral infections require an endocytic process [57-60]. Little is known about the mechanism the virus, the viral nucleocapsid or at least the viral genome, uses to escape from the endocytic pathway. The translocation motif (TLM) identified within the surface protein of hepatitis B virus is a general feature of all hepadnaviruses and is assumed to play a role in viral life cycle. As it was shown in this study the TLM peptide when fused to the HBV capsid allows the efficient transfer of these capsids across cellular membranes. Detailed analysis of the translocation motif revealed that the functionality does not strictly depend on the primary amino acid sequence, but two parameters are crucial. The amino acids (i) must form an alpha-helix and (ii) the alpha helix must display an amphipatic character, requiring a sequence of hydrophilic and hydrophobic amino acids [52]. In TLM deficient duck hepatitis B virus neither viral morphogenesis nor particle secretion is impaired, but these mutants are non-infectious [61]. Inactivation of the TLMs neither affects virus attachment nor entry in the endosomal compartment, but prevents the escape from the endosome into the cytoplasm. The processing of the viral particle in the endosome results in a conformational change exposing the TLMs to the surface of the viral particle. The exposed TLMs enable the translocation across the endosomal membrane into the cytoplasm. This post entry activation of the TLM ensures the tissue specificity of the viral particle. Permanently surface-exposed TLMs would result in the translocation of viral particles into non hepatic cells since the TLM mediated cell permeability shows no tissue specificity [52].
The unmasking or the generation of entry mediating sequences during a viral infection process is not unprecedented: Influenza virus surface proteins display no fusogenic activity. Endosomal cleavage by a furin like protease results in conformational changes. The new conformation allows a membrane fusion of the virus [62,63]. In case of the poliovirus the interaction with the receptor results in externalization of the

N-myristoylated VP4 peptide and exposure of amphiphatic sequences of VP1 that can insert into membranes mediating the entry in the cytoplasm [64,65]. Hildt and coworkers reported that the processing of surface protein by endosomal proteases unmasks the TLM exposing it on the virus surface. Functional surface-exposed TLMs enable translocation across the endosomal membrane into cytoplasm. Non permissive cell lines can be efficiently infected by patient derived HBV after processing of the viral particle by endosomal lysate. Based on these data, Hildt et al. propose a novel infectious entry mechanism unique for hepadnaviruses which involves virus internalization by receptor-mediated endocytosis, followed by processing of surface protein in endosomes. This induces surface exposure of TLMs that subsequently triggers particle escape across the membrane into the cytoplasm (**Figure 3**).

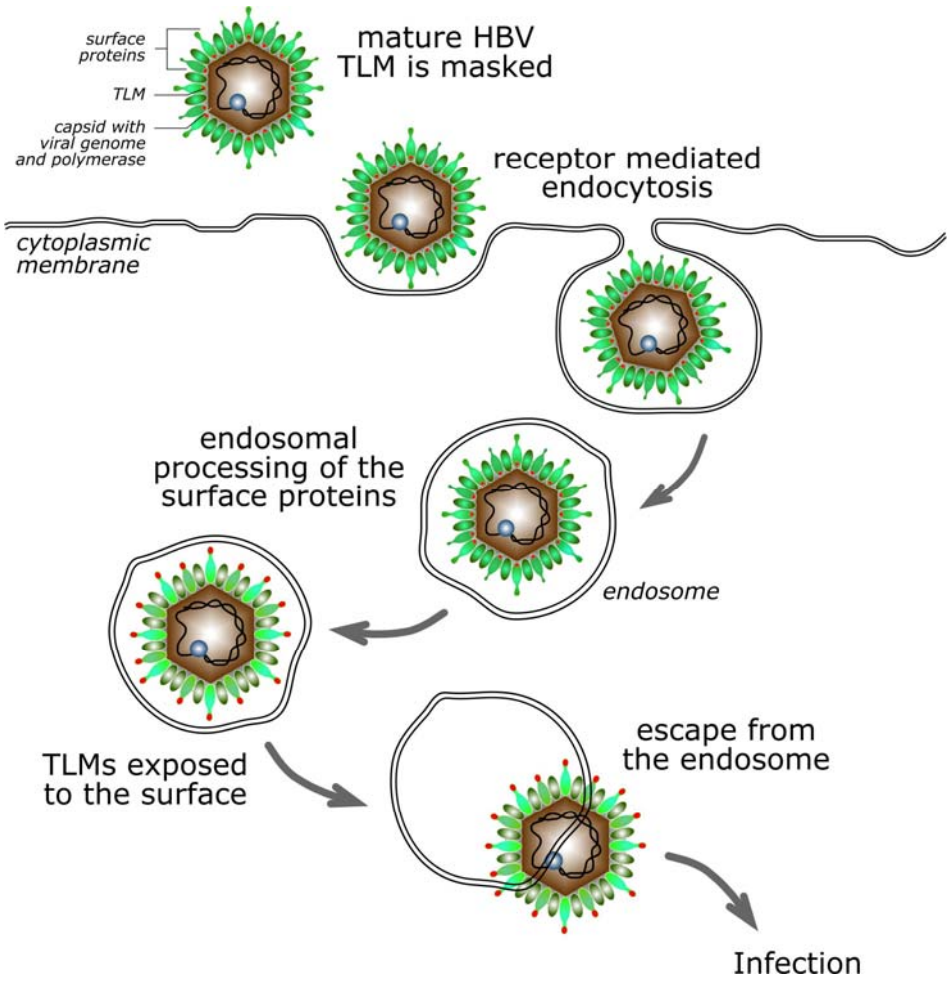

Figure 3: Model of the endosomal processing of hepadnaviral particles
Complete Hepadnaviruses are internalized by receptor mediated endocytosis. In the endosomal compartment proteolytic cleavage of the surface protein occurs. This results in a conformational change that exposes the TLMs on the surface of the viral particle. The endosomal cleavage generates a modified viral particle that is decorated on its surface with TLMs. The high density of TLMs exposed on the surface of the particle allows the escape from the endosome into the cytoplasm to initiate the infection [61].

II.2 HEPATITIS B VIRUS

The capsid protein of hepatitis B virus (HBV) was frequently selected as a target for protein engineering manipulations. Recombinant hepatitis B virus cores (HBc) gave onset to icosahedral virus-like particles (VLPs) as a basic class of non-infectious carriers of foreign immunological epitopes. The recombinant HBc particles were used to display immunodominant epitopes of hepatitis B, C, and E virus, human rhinovirus, papillomavirus, hantavirus, and influenza virus, human and simian immunodeficiency virus, bovine and feline leukaemia virus, foot-and-mouth disease virus, murine cytomegalo-virus and poliovirus, and other virus proteins, as well as of some bacterial and protozoan protein epitopes (for review see [66]). Practical applicability of the HBc particles as carriers was enabled by their ability to high level synthesis and correct self-assembly in heterologous expression systems. The interest in the HBc VLPs was reinforced by the resolution of their fine structure by electron cryomicroscopy and X-ray crystallography, which revealed the existence of long protruding spikes (as shown in **Figure 7**). The tips of the latter seem to be another target for the display of foreign sequences up to 238 amino acid residues in length (the hole GFP, [67]). Combination of numerous experimental data on epitope display with the precise structural information enables a knowledge-based design of diagnostic, and vaccine and gene therapy tools on the basis of the HBc particles.

Hepatitis B virus is a ubiquitous virus with a global distribution. It causes acute, transient and chronic infections of the liver. HBV is transmitted through blood and other body fluids, including semen and saliva. The virus is 100 times more infectious than human immuno-deficiency virus (HIV) and, unlike HIV, it can live outside the body in dried blood for longer than a week [68]. According to WHO, one third of the world's population (2 billion people) has been infected with the hepatitis B virus [69]. About 5 % of the population are chronic carriers of HBV, and nearly 25 % of all carriers develop serious liver diseases such as chronic hepatitis, cirrhosis, and primary hepatocellular carcinoma. HBV infection causes more than one million deaths every year [70]. Furthermore, HBV, hepatitis C virus (HCV) and cirrhosis between them contribute to the genesis of almost all global hepatocellular carcinoma (HCC).
Fortunately, there are very effective vaccines against the virus, which are about 95 % effective. The vaccine is very safe and there is no convincing evidence of any long-term undesirable sequelae. Eradication of HBV infection worldwide is a distinct possibility.

Human HBV is the prototype for a family of viruses, referred to as *Hepadnaviridae*. Those most closely related to HBV have been found in woodchucks [71] and ground squirrels [72]. These viruses have about 70 % sequence homology to HBV but are not known to infect humans or other primates; in contrast, HBV is infectious for the great apes. Because of similarities in DNA sequence and genome organization, the viruses infecting mammals are grouped in the genus *Orthohepadnavirus*. More distantly related viruses, with somewhat similar genome organization but almost no sequence homology, are found in ducks and geese [73,74]. These are grouped in the genus *Avihepadnavirus*.

II.2.1 Systems used for the study of HBV infection

The major limitation to study HBV is the lack of an effective *in vitro* infection system, which can support the entire life cycle of HBV. Hepadnavirus infection is strictly species and tissue specific only chimpanzees and human liver cells (primary hepatocytes) are available for human HBV infection studies. Human primary hepatocyte cultures derived from liver explants have been shown to be susceptible to HBV infection, but only for a limited amount of time following explantation [75]. Besides of the human primary hepatocytes, there are several systems derived from hepatoma cells retaining some of the hepatocyte markers and support HBV replication which were used for the study of HBV infection [76]. However, even these permissive cell lines are not susceptible to *in vitro* HBV infection without manipulation of the cultured cells with dimethyl sulfoxide (DMSO) or polyethyleneglycol (PEG) [77,78]. Therefore, although cultured cells can be shown to support HBV gene expression and replication, so far there does not exist any readily reproducible tissue culture system for HBV infectivity.
In the absence of a suitable and productive human HBV infection system, the study of hepatitis B infection and life cycle benefited from the use of animal models, in particular, woodchuck and duck hepatitis B virus (WHBV/DHBV). As all hepadnaviruses are highly related at the level of amino acid sequence and share similar genome organization and virion structure [79], the data on DHBV are relevant to HBV. In addition, transgenic mice have been engineered, which support HBV replication, but only rarely express disease [80]. All these models have been used to successfully investigate the major steps hepatitis B virus infection and also to develop and study treatments for chronic HBV infections. However, each of these *in vivo* and *in vitro* models has drawbacks, and in particular, specific aspects of human HBV that are difficult to study include: early steps

of virus-cell attachment and entry, specific kinetic parameters of infection, immune responses, and development of liver disease [81].

II.2.2 Structure of HBV

II.2.2.1 Overview

With a partially double-stranded circular genome that is only 3.2 kb in length, hepadnaviruses express a very limited repertoire of proteins. As shown in **Figure 4** the capsid, or core antigen (HBcAg), enclosed the viral DNA and is surrounded by a spherical envelope (host cell-derived lipids and viral surface antigens, HBsAg). The entire virion is known as the Dane particle [82] with a size of 42 nm. In addition to the core and surface proteins, the HBV genome encodes a viral DNA polymerase (P) that also acts as a reverse transcriptase and two gene products are of function which has yet to be resolved: hepatitis X protein (HBx) and hepatitis e-antigen (HBeAg).

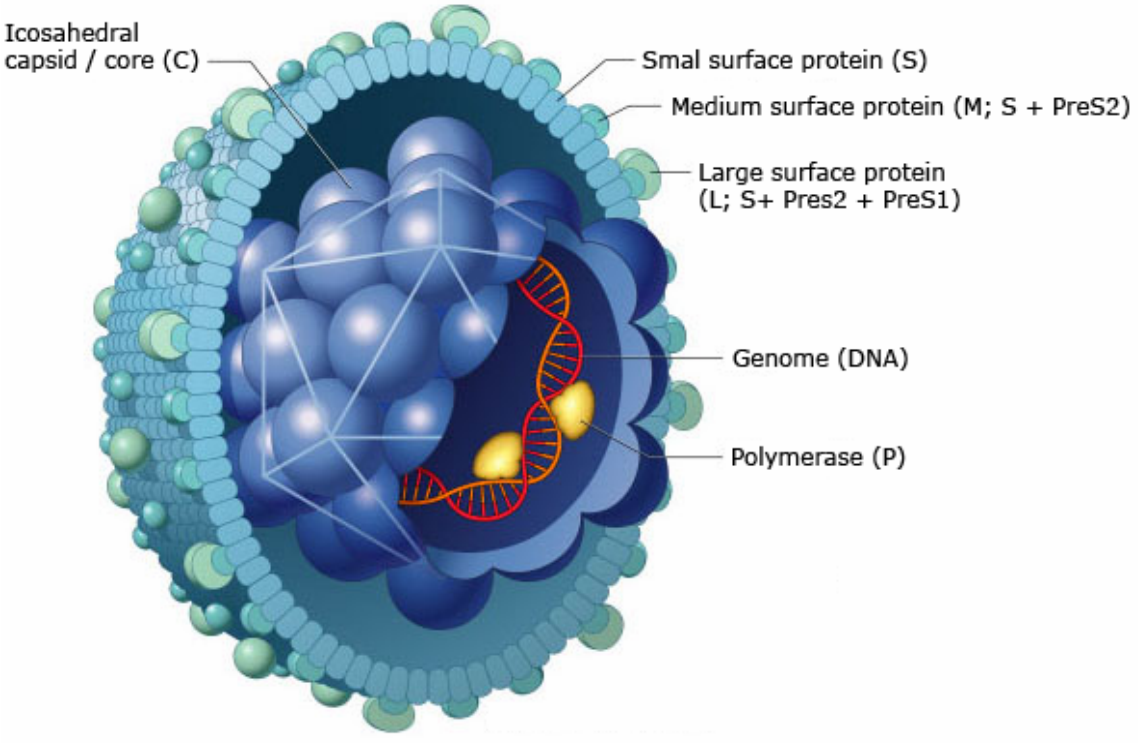

Figure 4: Structure of the hepatitis B virus [83]

II.2.3 Life cycle of HBV

As for HBV [84-86], it is known that DHBV infection is initiated by attachment of the virus particle to the hepatocyte surface via the PreS-domain of the viral surface protein [87]. A reported candidate receptor for DHBV that is able to bind to DHBV PreS is gp180 (carboxypeptidase D, p170) [88,89]. However, gp180 is not selectively found on the surface of liver cells. Moreover, expression of gp180 in chicken derived hepatoma cell line does not render these cells permissive for HBV infection. Therefore the existence of an additional co-receptor is postulated to enable hepatocyte-specific viral entry [58,90].

Previous work suggested that DHBV enters the cell by receptor-mediated endocytosis [60]. The resulting question is about the mechanism that allows the internalized viral particle to escape from the endocytotic pathway.

Similar to other viruses [91], a mechanism exists for HBV by which the viral envelope protein is being released from a preactivated, metastable state into a fusion-active form, given the appropriate conditions. Noticeably, low pH was shown to trigger this mechanism, in accordance with its possible role in virus entry.

Very little information exists regarding the post-endosomal-escape events in the hepadnaviral entry. The viral genome must be transported to the nucleus, where it is transcribed (**Figure 5**). The core particle encapsulating the viral genome mediates the genome nuclear transport. The core protein contains nuclear localization signals (NLSs) at its highly basic C terminus (see **Figure 7**) [92] and has been shown to bind the hepatocyte nuclear membrane [93]. *Escherichia coli*-derived core particles were shown to directly bind the nuclear pore complex (NPC) in a core phosphorylation and importin-dependent manner [94]. Core phosphorylation was presumed necessary to expose C-terminally disposed, luminally sequestered NLSs onto the capsid surface. Interestingly, core phosphorylation was also shown to interfere with the nucleic acid binding activity of core [95]. Therefore, core phosphorylation may play a dual role during the initial stages of infection, promoting both the nuclear targeting and release of the viral genome. Such a model seems particularly attractive in light of the fact that the 30 nm viral capsid may be too large to pass the nuclear pore itself.

Once inside the nucleus, the virus converts its partially double-stranded genome into covalently closed circular DNA (cccDNA). The presence of cccDNA in hepatocytes indicates a successful initiation of infection [96]. CccDNA is template for transcription of genomic and subgenomic viral mRNAs. The greater-than-genome-length preC/pgRNA which is transcribed by cellular RNA polymerase II, represents the mRNAs for the translation of

P and C proteins as well as being the template for reverse transcription of minus-strand DNA [97]. The preC/pgRNA interacts, possibly simultaneously, with both its gene products, C and P. Translation of P and pgRNA packaging are tightly coupled events. Polymerase binds to the 5' end ε, and triggers the addition of core complex dimers, which allows for packaging into capsids (see **Figure 12** and **Figure 5**) [98]. Reverse transcription takes place during capsid formation [99]. HBV surface proteins are initially synthesized and polymerized in the rough endoplasmic reticulum. These proteins are transported to the post ER and pre-Golgi compartments where budding of the nucleocapsid follows [100]. The assembled HBV virion and sub-viral particles are transported to the Golgi for further modification of its glycans in the surface proteins, and then are secreted out of the host cell to finish the life cycle.

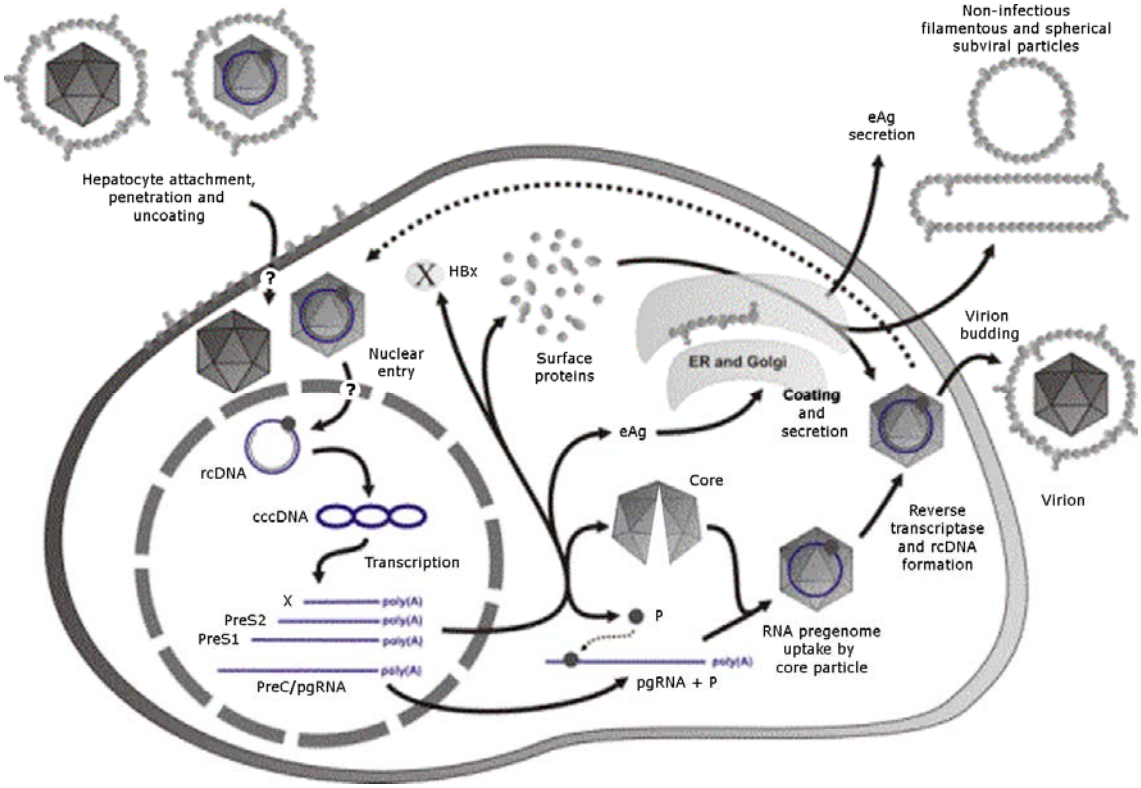

Figure 5: A schematic illustration of the HBV infection and replicative cycle within the hepatocyte.

Infectious virions attach to cellular receptor(s) and uncoat, releasing nucleocapsids that migrate to the cell nucleus. The partially double-stranded DNA genome is converted to cccDNA, which is the template for the transcription of four viral transcripts. Translation occurs following transcript export to the cytoplasm. The pgRNA interacts with two gene products, P and C, to form immature, RNA packaged, nucleocapsids. The pgRNA is reverse transcribed into DNA by P. The DNA genome can be either redelivered to the nucleus, or nucleocapsids can be coated by surface glycoproteins (in the Golgi and endoplasmic reticulum) before being exported as enveloped virions. [101] modified.

HBV replication and viral protein synthesis in the infected cells are fairly well elucidated. However, the early steps of HBV infection including of the penetration of virus, the nucleocapsid trafficking and release of its genome into host cells is uncertain [102].

II.2.4 HBV genome organization

The genome of HBV as shown in **Figure 6** is composed of a relaxed-circular and partially double-stranded DNA (rcDNA) [103]. The long, full-length minus-strand is approximately 3200 nucleotides (nt) in length and has a protein (the viral polymerase) covalently bound to its 5' terminus. The plus-strand, which varies between 1700 and 2800 nt in length depending on species and subtype, has a capped oligoribonucleotide at its 5' end. The plus-strand maintains genome circularity by a cohesive overlap across the 5' and 3' termini of the minus-strand .

The HBV genome includes four open reading frames (ORFs) that encode at least seven translation products through the use of varying in-frame initiation codons. These translation products include three surface antigens (HBsAg): the envelope glycoproteins preS1, preS2, and S; core (C) and e antigens (HBcAg and HBeAg); viral polymerase (Pol); and the X protein (HBx). The complete *PreS/S* gene, a part of the *PreC/C* gene and the *X* gene overlap with the *Pol* gene which leads to the compact organization of the genome. Also, promoters, enhancers and other regulatory genes are all located in the genes that can code for the structural proteins. Therefore, nucleic acid mutations in one of the HBV genes could affect the functions of more than one gene [104].

The genome is also replete with important cis-elements required for the regulation of HBV gene expression and replication. These include viral promoters, enhancers and signal regions. The 5' terminus of both strands contain regions of short (11 nucleotide) repeats, DR1 and DR2, which are essential for priming the synthesis of their respective DNA strands during replication. Not surprisingly, HBV's compact coding organization wastes no sequence space, with every nucleotide falling within a transcribed region and with 50% of translated sequences present in more than one ORF.

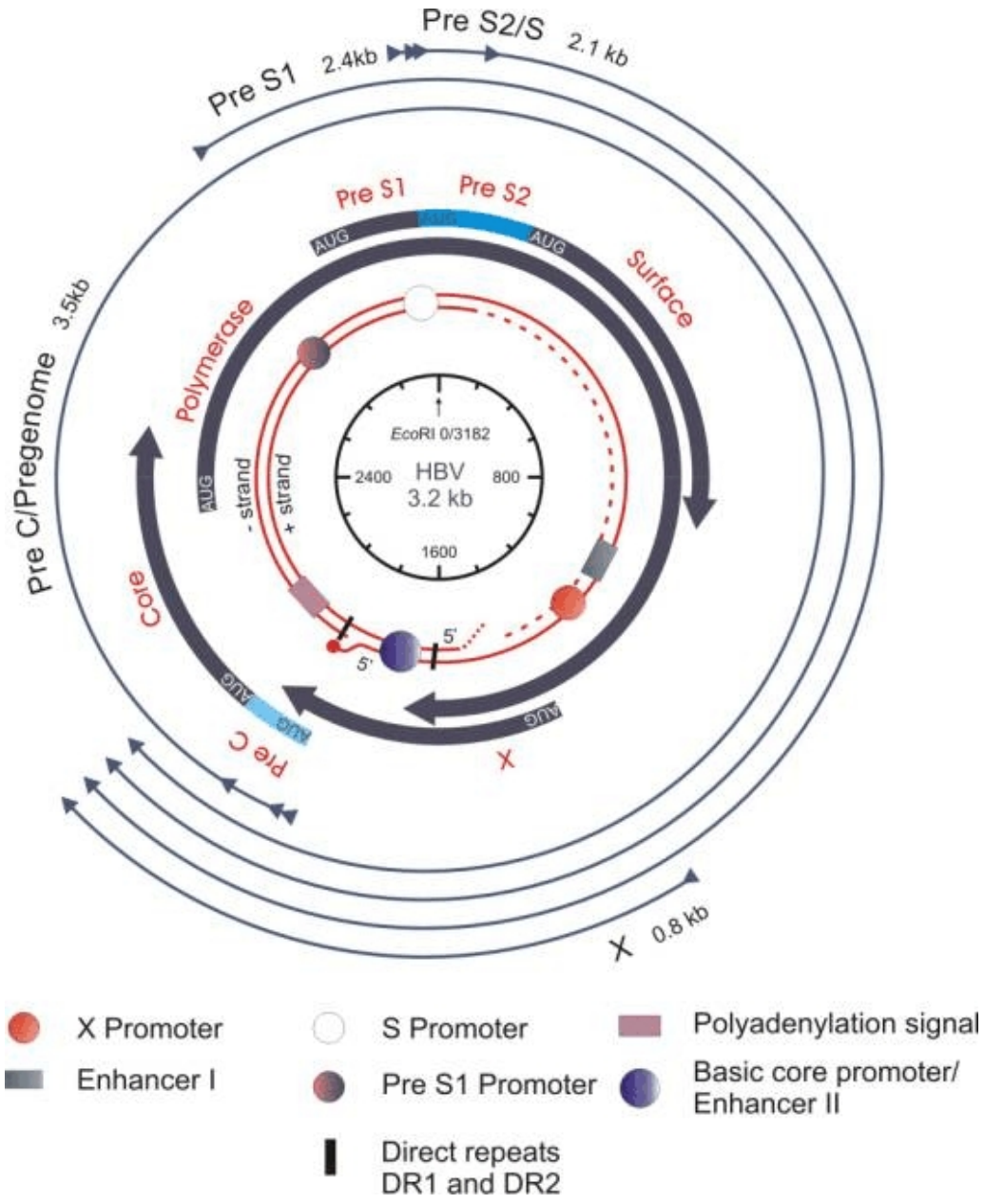

Figure 6: Transcriptional and translational organization of the hepatitis B virus genome (strain ayw).

Partially double stranded HBV DNA comprises + and – strands with cohesive complementary 5′ ends. The cis-elements that regulate HBV transcription are represented by the circular and rectangular symbols. The positions of direct repeats DR1 and DR2 are indicated as black rectangles. Immediately surrounding arrows (thick) indicate the viral open reading frames (with initiation codons) that encompass the entire genome. Four outer arrows, that give the 5′ to 3′ polarity, indicate the HBV transcripts. Multiple arrowheads at the 5′ ends of the PreC/Pregenome and PreS2/S transcripts indicate heterogeneous transcription start sites. The common 3′ end of all the HBV transcripts is depicted by the identical termination site and sequences that overlap with the HBx transcript. The nucleotide numbers are indicated by the numbering system of Galibert et al. [105]. Nucleotide positions of the genome are given relative to the single EcoRI restriction site. [106]

II.2.5 HBV transcripts

Covalently closed circular HBV DNA (cccDNA) is the template for both genomic and subgenomic viral transcript mRNAs, which are produced using cellular RNA polymerase II. The 5' termini of the preC/pregenome and preS2/S transcripts both have heterogeneous transcription start sites. The HBx transcript, however, has only one unique start site. The pregenomic RNA (preC/pgRNA) transcript, which is approximately 3500 nt in size, is more than a full genome in length and contains terminal repeats at each end (see **Figure 6**). The preC/pgRNA transcript is unique in that apart from being the genomic template for reverse transcription, it represents the mRNA for production of the core protein as well as the viral polymerase protein. The S transcript is 2400 nt in length whereas the preS2 family of transcripts are roughly 2100 nt in size. The HBx mRNA is a single 900 nt transcript. The presence of a common 3' polyadenylation termination signal results in all transcripts sharing the same 3' terminal sequences [106,107].
Hepatitis B virus can be further classified into eight major genotypes (A to H) based on nucleotide (nt) diversity of ≥8% [108]. These genotypes have a distinct global geographical distribution.

II.2.6 Hepatitis B capsid

The HBV core (C) gene is the most conserved amongst HBV genes, numerous amino acid substitutions were fixed for its most parts. The C gene has two in-frame initiation AUG codons (**Figure 7a**), it is responsible for the appearance of at least four different polypeptides: p25, p22, p21, and p17 [109]. The p25 precore protein, starting at the first ATG codon, becomes targeted by a signal peptide in the preC sequence to a cell secretory pathway, in which a p22 is formed by N-terminal processing. The p22 undergoes further cleavage at the C-terminal region, after position 149, to generate a p17 protein, or HBe protein, which is early secreted from the cell during infection as the HBe antigen [110]. HBe antigen serves as an important serological marker. The predominant p21 core or capsid protein is synthesized from the second ATG of the 183 amino acids long open reading frame (21 kDa), and constitutes a structural component of the HBcAg, or HBV nucleocapsid. The fine structure of HBc particles (**Figure 7b**) was revealed by X-ray crystallography at 3.3 Å resolution [111]. The folding of the protein is characterized by α-helices and the absence of β-sheets. Association of two amphipathic α- helical hairpins results in the formation of a dimer, stabilized through two disulfide bonds [112], with a four-helix bundle as the major central feature.

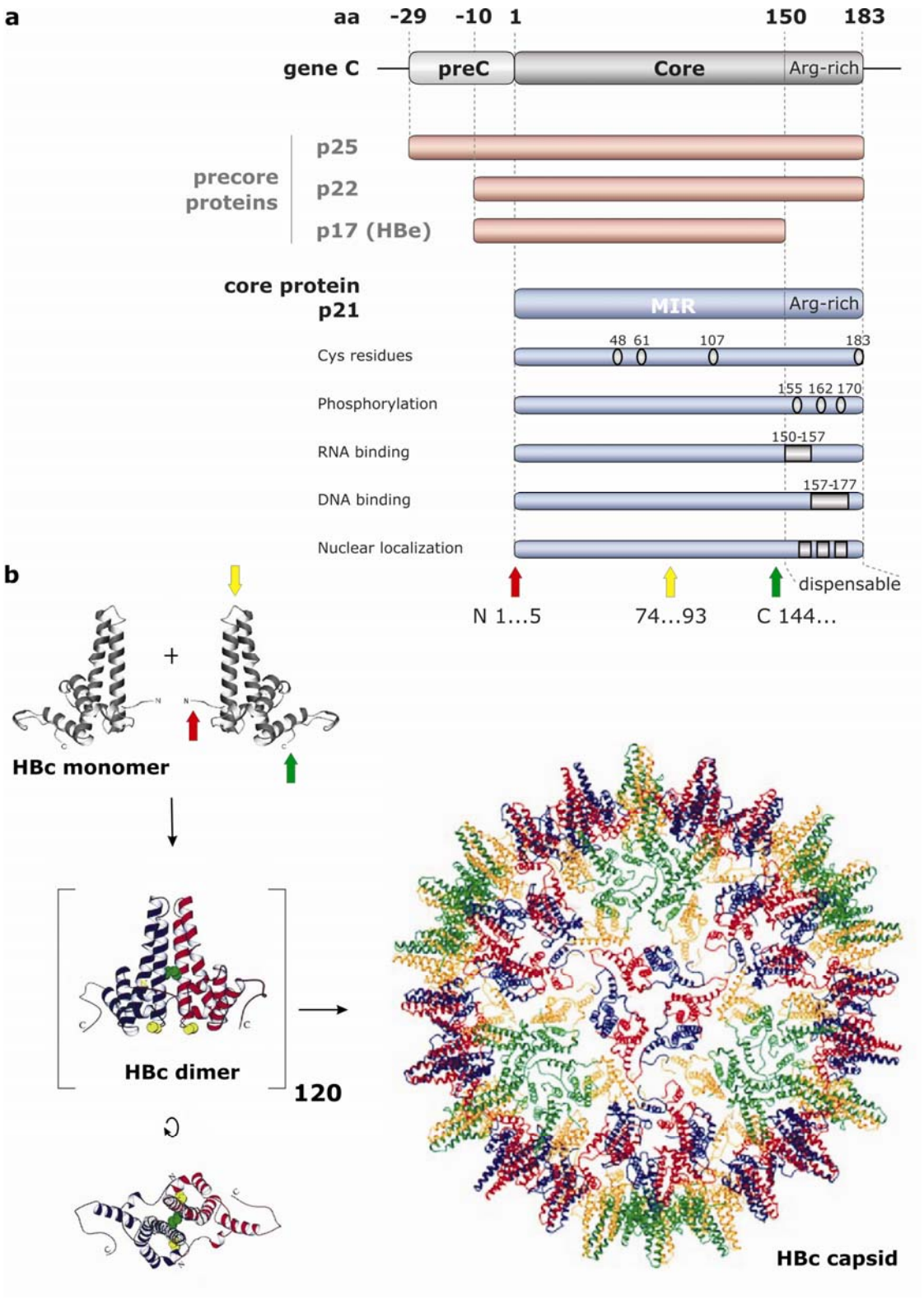

Figure 7: Structure of the hepatitis B capsid

a) Products encoded by the C gene with structural and functional peculiarities of the HBc molecule [113]. Localization of the possible insertion sites for foreign epitopes (colored arrows). **b)** Orthogonal views of the HBc dimer built by two monomers, viewed normal to the local two-fold axis and along the two-fold axis from the outside of the capsid. Cys-61, which forms a disulfide bridge between the two monomers, is shown in green balls, and Cys-48, which does not form disulfides, in yellow balls. Insertion sites for foreign epitopes are marked by colored arrows. T = 4 HBc capsid viewed down an icosahedral three-fold axis [110,111].

The dimers are able to assemble into two types of particles, large and small ones, which are 34 and 30 nm in diameter and correspond to triangulation number *T = 4* and *T = 3* packings. Most HBV capsids (> 90 %) are complexes of 120 dimers (240 HBc protein monomers) arranged with *T = 4* quasi-symmetry. The minority of capsids containing 90 dimers (180 monomers) and arrange with *T = 3* quasi-symmetry (see also **Figure 10b**) [114]. The four-helix bundles protrude, forming spikes. The amino acid stretch 76–81 located at the tips of the spikes presents a central part of the so-called major immunodominant region (MIR) of the HBc particle [110]. The capsid is fenestrated, with large pores with a diameter of 12–15 Å. It has been proposed that these pores are essential to allow nucleotides access to the capsid interior during the DNA synthesis and they are also relevant for the reported accessibility of the C-terminal tail to proteinases and antibodies [111].

It was demonstrated that the C-terminal protamine-like arginine-rich domain of the core protein (aa 150–183, **Figure 8**) is not essential for its self-assembly capability [115]. The C-terminally truncated HBc_{149} particles were indistinguishable from the HBc particles formed by full-length HBc_{183} proteins, as shown by electron cryomicroscopy [116]. However, HBc_{149} particles were less stable, failed to encapsidate nucleic acid and accumulated usually as empty shells, in contrast to the full-length HBc_{183} particles [117]. The C-terminal limit for selfassembly of truncated HBc proteins was mapped experimentally at aa residues 140 [118,119].

Besides capsid-formation, the HBV core protein participates in the viral life cycle and its regulation, including the synthesis of double-stranded DNA as a cofactor of the viral reverse transcriptase-DNA polymerase, viral maturation, recognition of viral envelope proteins and budding from the cell [110]. It appears that HBcAg can recognize specific sites of the envelope proteins S and L [120,121]. Direct electron cryomicroscopic evaluation showed that the peptide which inhibits the binding of the L protein to the HBc particle is located at the tips of the spikes of the latter [122]. Phosphorylation of serine residues by cellular protein kinase [95,123,124] within three repeated SPRRR motifs on the C-terminus of the core protein [125-127] and its role in the maturation of the HBV capsid [94,128,129] reflect the complex function of the HBc particles. According to recent data, phosphorylation of HBc subunits induces a conformational change that exposes the C-terminal sequences, which may protrude through the holes in the capsid wall and become accessible on the surface to serve as a nuclear targeting signal [92,94].

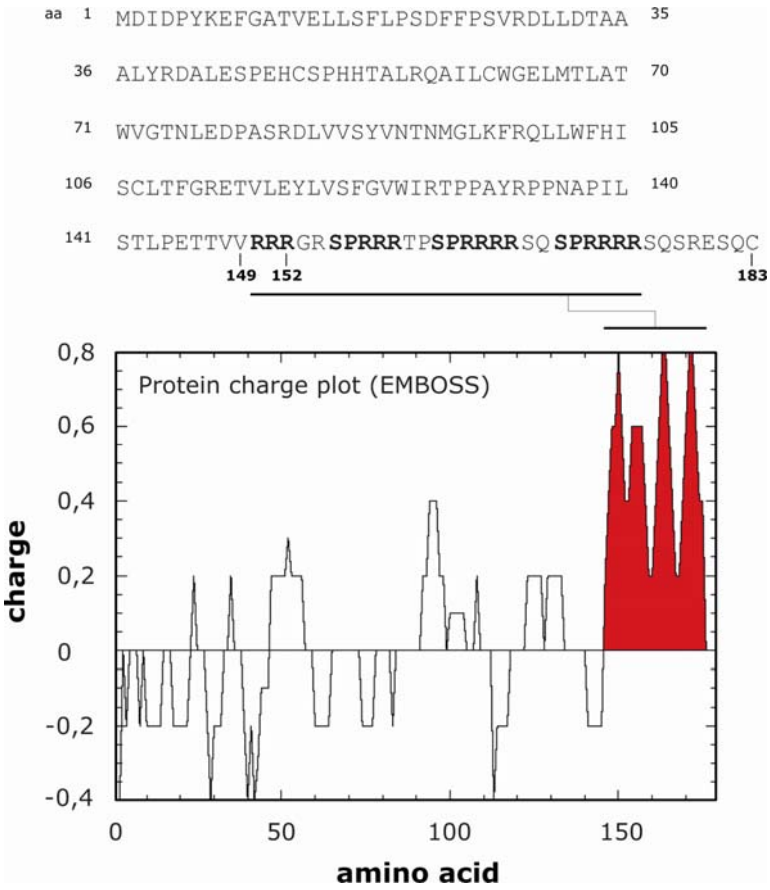

Figure 8: HBV core protein sequence and charge plot
The C-terminus of the 183 aa long wildtype HBV core protein (subtype ayw) contains an Arginine-rich domain (bold letters and red peaks in the charge plot) with within three repeated SPRRR motifs. This domain adds a very positive charged tail to the C-terminus of HBc. The distribution of positively or negatively charged amino acids is given in the EMBOSS charge plot (see also V.2.5.1).

II.2.6.1 Assembly of HBV capsid

In infected human hepatoma cells, the hepatitis B virus capsid assembles around a complex that includes viral pregenomic RNA (pgRNA) and viral reverse transcriptase (HBV polymerase) to form the HBV core (Figure 9). In vitro assembly of empty capsids using the dimeric 149 amino acid long residue assembly domain of the core protein can be induced by high NaCl concentration. Subunit dimers are stable in solution.
Assembly of HBV conforms to thermodynamic and kinetic predictions of the simplest case assembly models. Analysis of assembly kinetics indicated that formation of a trimer of dimers was the rate-limiting step [130]. In the context of the capsid, each dimer interacts with four neighbors (**Figure 10**).

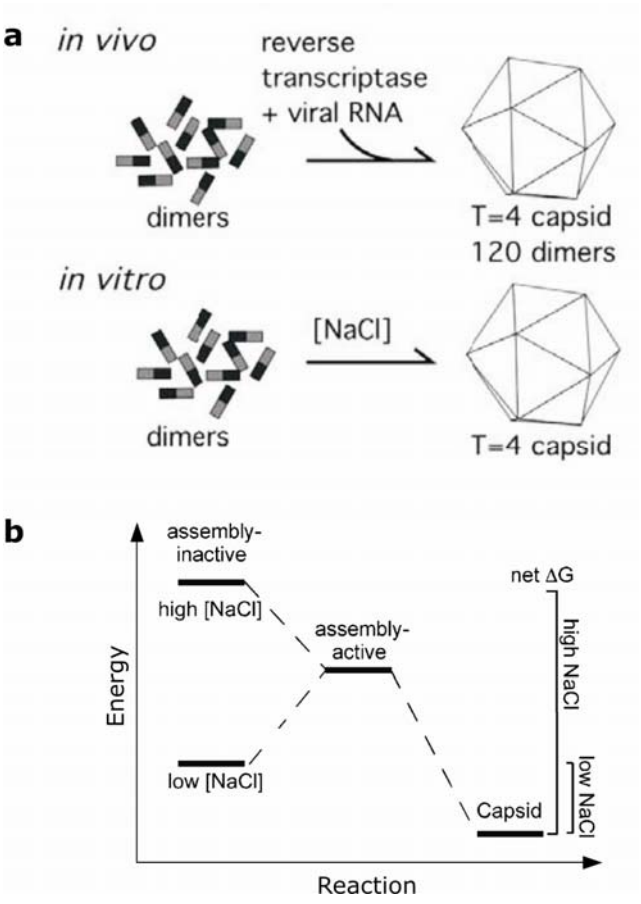

Figure 9: Schematic model of HBV capsid assembly [1]

a) *In vivo* assembly of HBV cores is cytoplasmic: 120 HBc dimers assemble around a complex of viral RNA and reverse transcriptase (HBV polymerase). *In vitro* the kinetics and thermodynamics of assembly is dependent on dimer and NaCl concentration. **b)** Assembly is activated by a conformational change that is induced by high ionic strength. Low ionic strength favors the assembly-inactive state. Dimers must undergo a conformational change prior to or concomitant with assembly, resulting in a weak net energy for capsid formation. In high ionic strength, the assembly-active state is favored, and the net energy for capsid formation is stronger. The energy for polymerization of the assembly-active form is about the same for both conditions [2].

HBV assembly is characterized by positive enthalpy and entropy. The reaction is entropy-driven, consistent with the largely hydrophobic contacts found in the crystal structure. Increasing NaCl concentration raises the magnitude of free energy, enthalpy, and entropy, as if ionic strength were increasing the amount of hydrophobic surface buried by assembly [2].

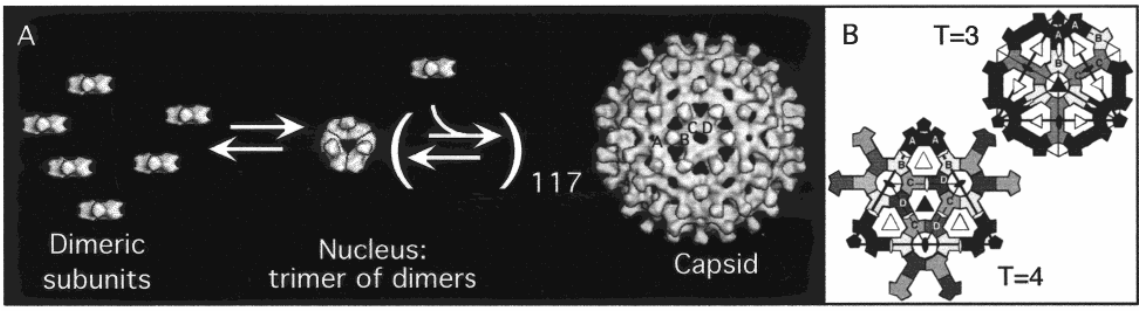

Figure 10: Schema of HBV capsid assembly.

A) Capsid assembly begins with formation of a trimeric nucleus; subsequent elongation reactions add one dimeric subunit at a time. The product of the reaction shown here is a *T* = 4 capsid at 17 Å resolution. **B)** Schematic representation of *T* = 3 and *T* = 4 icosahedral facets, emphasizing the dimeric structure of the subunit. Putative nuclei correspond to *T* = 3(AB-AB-CC), *T* = 4(AB-AB-CD), and *T* = 4(3CD). Icosahedral symmetry operators are shown as solid pentagons (5-fold axes), triangles (3-folds), and ovals (2-folds). Quasi-symmetry axes are shown as white symbols [130].

II.2.7 Surface proteins

A lipid envelope is surrounding the viral core, which is approximately 7 nm thick. It is derived from the host cell membrane and endoplasmic reticulum [131]. Orthohepadnaviruses express three surface proteins embedded in the lipid membrane. The surface gene consists of a single open reading frame divided into three coding regions: preS1, preS2 and S, each starting with an in-frame ATG codon. Through alternate translation initiation at each of the three AUG codons, a large (LHBs; PreS1 + PreS2 + S), a middle (MHBs; PreS2 + S) and a small (SHBs; S) envelope glycoprotein can be synthesized [132]. All three envelope components are glycosylated, type II transmembrane proteins that can form multimers stabilized by disulfide bridges formed by cysteine residues present in the S domain. S, L, and M are all found as components of the 42-nm-diameter infectious viral particles (**Figure 4**) [106]. L and M are present in roughly equal amounts in Dane particles and together constitute approximately 30% of the envelope protein content. S by itself, and together with the larger envelope proteins, also forms filamentous and spherical "surface antigen" particles that are secreted from infected cells in at least 100-fold excess over virions (see **Figure 5**). These spheres and filaments can accumulate to concentrations of several hundred micrograms per milliliter in the blood of HBV-infected patients. Complexes of these particles with their cognate antibodies are probably responsible for the immune complex syndromes that sometimes occur during transient infections [133].

Although much remains to be ascertained, the specific function of the surface proteins appears to be in selectively transporting the nucleocapsid into and out of the host hepatocyte without causing cellular lysis. The preS1 domain appears to have a dual role in HBV biology. PreS1 is both a ligand to core particles during viral envelope assembly (which is why a fraction of L particles display the preS1 domain towards the cytosolic interior), and a substrate to an unidentified host cell receptor during viral infection [134].

The LHBs envelope protein plays the prominent role in cell attachment during infection of HBV and has a regulatory protein function. The preS1 domain comprising the unique N-terminal 120 amino acids bears the prominent cell attachment epitope. This epitope was narrowed down to amino acids 21– 47 of preS1 that was shown to be both required and sufficient to mediate attachment to HepG2 cells. Specific antibodies that were raised against this region block infection [135].

II.2.8 HBV polymerase

The polymerase (P) open reading frame encodes the viral polymerase, which is translated from an internal initiation codon on the viral preC/pgRNA. The *P* gene is transcribed as part of a second cistron of the preC or pgRNA transcript and thus lacks its own direct upstream promoter element. Translation of polymerase requires ribosomal 'scanning' of the preC/pgRNA and is thus produced less efficiently than either eAg or C proteins.

Unlike most DNA viruses, HBV replicates via reverse transcription of an RNA intermediate and is considered to be distantly related to the retroviruses. The process includes polymerization of minus-strand DNA (by RNA-dependent DNA polymerization), degradation of the pregenome from an RNA-DNA heteroduplex (by RNase H activity) as minus-strand DNA synthesis proceeds and synthesis of plus-strand DNA from the minus-strand DNA template (by DNA-dependent DNA polymerization). All enzyme activities responsible for these steps come from the viral polymerase [136].

Polymerase is further endowed with another functional domain: the terminal protein (TP) at its amino terminal. The TP is found covalently linked to the 5' end of HBV minus-strand DNA and is important for the packaging of the RNA pregenome as well as acting as a primer for reverse transcription.

The viral polymerase is also required as a structural component for the packaging of preC/pgRNA into immature core particles and may play a role in orchestrating entry into the host cell nucleus for early replication.

II.2.8.1 Epsilon (ε) signal

The terminal repeats of the pgRNA contain a stem-loop sequence, or epsilon (ε) motif, which plays a key role in HBV DNA encapsidation and reverse transcription. The location of ε was determined by fusing heterologous genes to various regions of the HBV genome and by observing subsequent encapsidation [137,138].

Despite the presence of terminal repeats on the pgRNA, only the 5' ε signal retains functionality resulting in encapsidation of the pgRNA transcript, notwithstanding the fact that all HBV transcripts have the ε coding region at their 3' ends. Analysis of ε shows a series of inverted repeats that fold into a three-dimensional stem-loop structure [139] (**Figure 11**). This stem-loop is conserved among all hepadnaviruses irrespective of differences in the primary sequence [137]. Polymerase recognizes and directly interacts with ε, initiating both encapsidation as well as reverse transcription of the HBV pgRNA [140] (**Figure 12**).

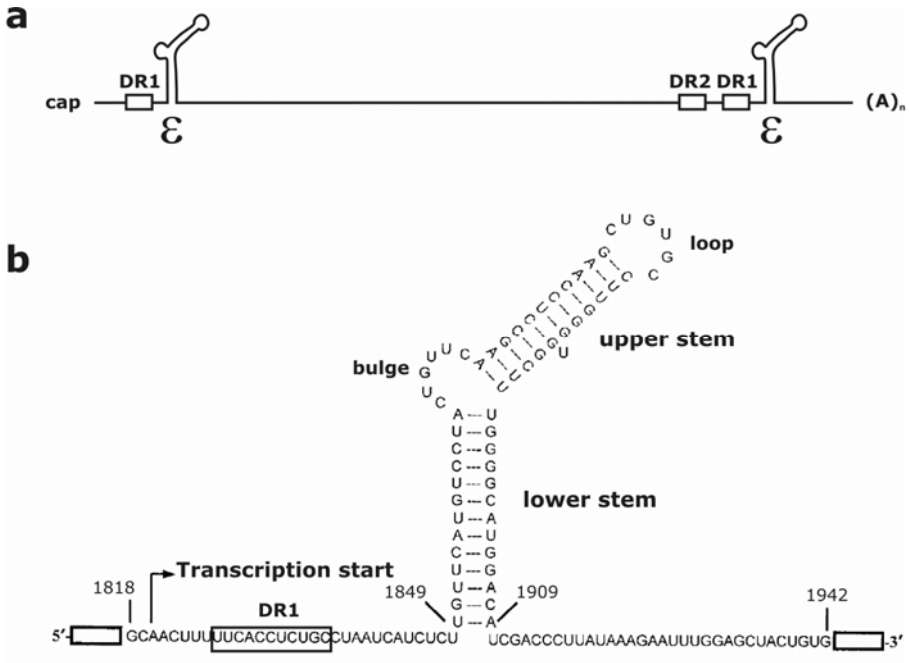

Figure 11: Secondary structure of HBV epsilon signal on pgRNA.
a) Structure of HBV pre genomic RNA. Both the DR1 and the ε sequence are present twice on the pgRNA due to its terminal redundancy. The ε elements are represented by the stem-loop structure. The direct repeat elements, DR1 and DR2, are indicated by open boxes. The sequence at the 5′ end has been enlarged below the pgRNA. The translation initiation codon of the pre-C gene is indicated by AUG. Transcription start sites are indicated by arrows. The nucleotide numbers are indicated by the numbering system of Galibert et al. [105].
b) Nucleotide sequences and predicted structure of the ε element DR1 and the stem-loop structure are indicated. Restriction enzyme sites at both ends of the e insertion cassette are indicated by the open boxes. [139]

A unique polymerase-linked DNA primer consisting of four-nucleotides is synthesized using the bulge region of the stem-loop ε as template. Following primer synthesis, the first of three template switches occurs. The short primer translocates to a homologous four-nucleotide sequence, which is part of a direct repeat (DR1), at the 3′ end of the preC/pgRNA. Subsequently, minus-strand DNA synthesis continues within the immature nucleocapsid [141]. RNase H activity of P then degrades the preC/pgRNA which is hybridized to the minus-strand DNA. Degradation is complete except for a short stretch at the 5′ terminus (16-18 nt) which acts as the primer for synthesis of the complementary plus-strand viral DNA [142].
This primer RNA is translocated to direct repeat 2 (DR2) where the synthesis of plus-strand DNA is initiated. Plus-strand DNA is extended to the physical end of minus-strand DNA using a short terminal repeat (r) followed by a template switch from the 5′- to the 3′-end of the minus-strand DNA [141]. Failure to translocate, referred to as in situ priming, leads to the synthesis of a double-stranded linear form of the viral genome, which is

often observed in tissue culture. A third template switch enables the synthesis of the partially double-stranded DNA genome found in mature, infectious virions [142]. The plus-strand DNA maintains genome circularity through a cohesive overlap across the 5′ and 3′ ends of the minus-strand DNA.

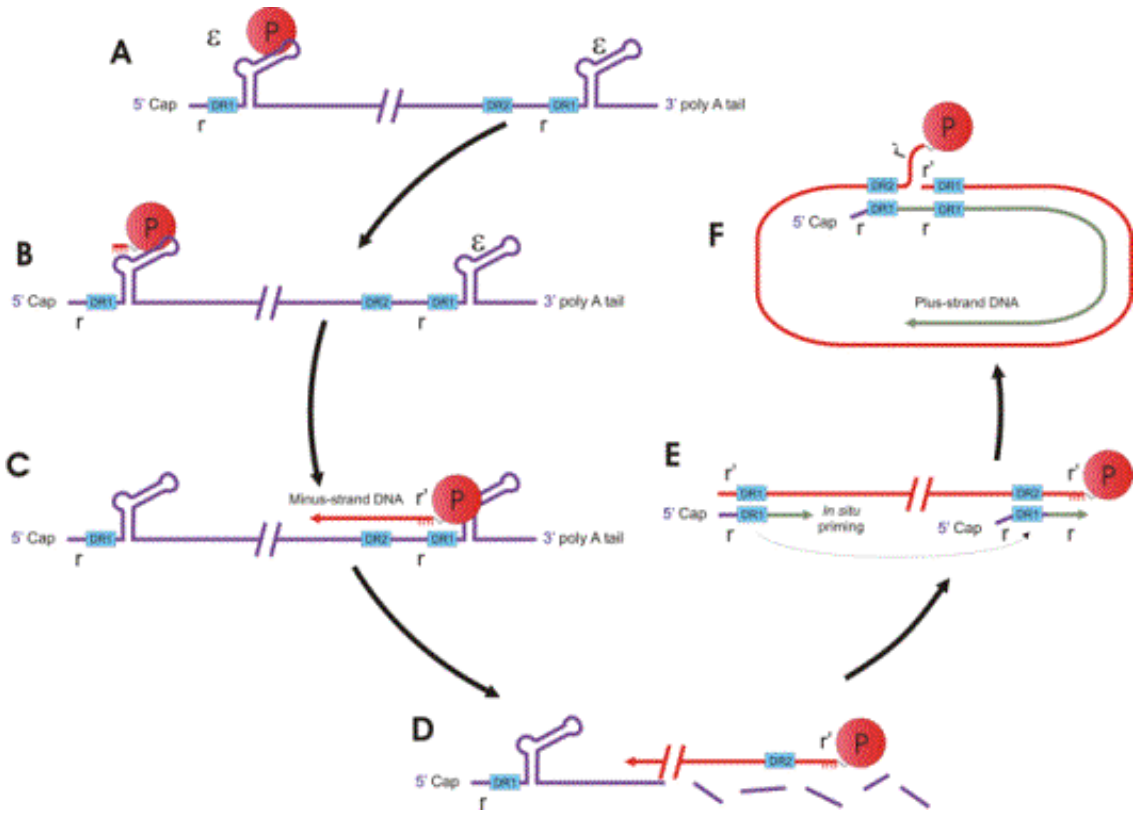

Figure 12: Polymerase recognizes the HBV pgRNA
Essential steps involved in hepadnaviral DNA replication starting from a newly transcribed pgRNA template. **A)** The reverse transcriptase (HBV polymerase, P) binds to the epsilon (ε) stem-loop structure near the 5′ terminus of the preC/pgRNA (blue). **B)** Reverse transcription is by a protein priming mechanism, utilizing a tyrosine located near the amino terminus of the reverse transcriptase itself. **C)** Following the synthesis of four bases in the bulge region of ε, P translocate to the 3′ terminus of the RNA template, where the four bases can anneal with a complimentary sequence. **D)** Elongation of the minus-strand DNA (red) to the 5′ end is concomitant with RNase H degradation of the pregenome sparing the 5′ cap and DR1, which remain hybridized to the minus-strand. **E)** The remaining fragment serves as primer for plus-strand DNA synthesis (green), following its translocation to DR2. Failure to translocate leads to a double-stranded, linear form of the viral genome (a process referred to as in situ priming). **F)** A third translocation occurs when the plus-strand DNA reaches the 5′ terminus of the minus-strand, circularizing the molecule and allowing continued plus-strand elongation. This translocation is facilitated by a short (~9 nt) terminal repeat on the minus strand DNA (r′) [134].

II.2.9 Regulatory protein - HBx

HBx is a small, 17 kDa non-structural protein and although poorly immunogenetic, produces antibodies in sera of infected humans and naturally infected animals. Although the specific function of HBx in natural infection remains elusive, its presence is necessary for the establishment of viral infection in vivo in animals (WHBV) [143].
HBx appears to possess a multitude of activities in vitro. Few of these define a specific role for this protein during viral replication. HBx has been shown to activate the transcription of a variety of viral and host cellular genes [144]. This indiscriminate or 'promiscuous' trans-activation activity appears to be via an indirect process, since HBx, without DNA-binding domains, cannot act alone. Although it appears that trans-activation activity is supplementary to HBx's intended primary function, recent studies suggest a more prominent role for this function in the HBV life cycle as HBx was shown to activate transcription of the core gene in vivo in HBx transgenic mice [145].
Apart from modulating transcription, HBx is a multifunctional viral regulator that adapts cell responses to genotoxic stress (by modulating or affecting DNA repair and apoptosis), protein degradation, signaling pathways, and cell cycle checkpoints [54]. These modulations affect viral replication and viral proliferation, directly or indirectly, and in turn provide insight into the possible role of HBx in the aetiology of HBV-induced carcinogenesis [146]. The overwhelming array of host-cell factors that are associated with HBx (as well as other viral factors) underscores the exquisite sensitivity and specificity which HBV enjoys in its natural cellular environment.

II.3 EARLY STEPS OF VIRAL INFECTION

In general, the early step of virus infection in which the virus enters the cell can be divided into three stages: attachment, fusion, and entry. Enveloped viruses usually entry by an attachment to the host cells, which usually is via the interaction of viral surface protein with the specific receptor on the cell membrane. However, the attachment itself does not necessarily result in viral entry. The fusion of a viral envelope and cell membrane and the following viral genome release finally trigger the viral infection [147]. The viral fusion occurs by one of at least two known mechanisms. The first requires a fusion of the viral envelope with the plasma membrane, leading to the release of the viral nucleocapsid. HIV is an example of a virus that uses this type of mechanism to enter the cell [148]. In the second, for HBV relevant mechanism, an endosomal compartment first takes up the

attached virus. Later, the virus and its capsid harboring the viral genome is released from this endosomal compartment into the cytoplasm by a process that may or may not depend upon a lowering of the pH to activate a virus-mediated fusogenic activity. Beside HBV in some other virus/cell systems, such as those involving influenza virus [149] or paramyxoviruses [150] exogenously added protease generates infective virions from otherwise non-infective ones. The reason is that the proteolysis exposes the viral fusion sequence. The un-coating and the genome release occur instantly after virus fusion [102].

All DNA viruses, except for pox viruses, as well as some viral families that bear RNA genomes, including retroviruses and influenza viruses, replicate in the nucleus. Only here can these viruses gain access to the cellular factors they require for the amplification and transcription of their genome and for the posttranscriptional processing of viral mRNA species. Therefore, these viruses not only have to cross the plasma membrane but must also traverse the cytoplasm and then enter the cell nucleus. Overall, viruses use a "Trojan horse" strategy in which the victim assists the intruder.

II.3.1 Intracellular Transport

After escape from endosomal vesicle, the capsid containing the genome of viruses must be transported to the nucleus. Diffusion in the crowded and highly structured cytosol is not efficient given the large dimensions of most capsids and the long distances they must travel [151,152].

To move inside the cell, incoming viruses often exploit the cytoskeleton and cellular motor proteins. There are two main ways to do this; the viruses can allow endocytic vesicles to ferry them as passive lumenal cargo, or the penetrated capsid can itself interact with the relevant motors [151]. In the latter case, a capsid protein binds and interacts with the cellular factors. Transport to the nucleus generally involves the minus-enddirected microtubule-dependent motor dynein and its adaptor protein, dynactin (**Figure 13**) [148,153]. Actin can also play a role in virus entry [154]. Because it is rich in actin filaments, the cortical cytoskeleton poses a barrier against the inward movement of capsids and viruses that enter directly through the plasma membrane [155]. To overcome the barrier, some viruses, such as SV40, activate tyrosine kinase– induced signaling cascades that lead to the local dissociation of filamentous actin [156,157].

Actin has also been found to promote vesicle budding at the cell surface and to propel virus-containing endocytic vesicles through the cytoplasm [156,158]. The release of baculovirus from endosomes induces

polymerization of actin to one end of the baculovirus capsid, which promotes capsid movement toward the nucleus [159,160].

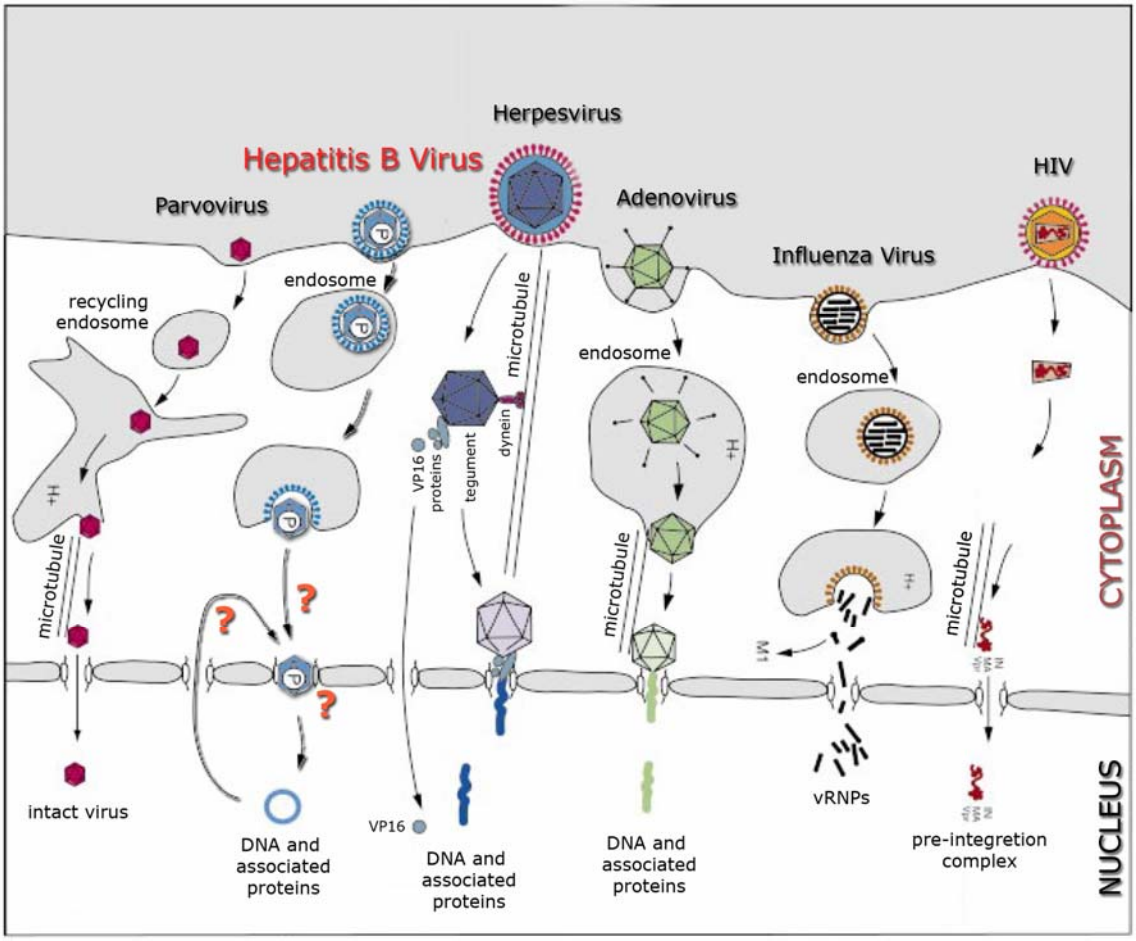

Figure 13: Entry and uncoating mechanisms of selected nuclear-replicating viruses
Viruses can undergo substantial uncoating in the cytoplasm before translocating into the nucleus (HIV and influenza virus). Alternatively, viruses can dock to the NPC and uncoat at the cytoplasmic side of the nuclear membrane (adenovirus and herpesvirus). They may possibly disassemble within the nuclear pore (hepatitis B virus). Finally, viruses can enter the nucleus essentially intact (parvovirus) [21].

II.3.2 Nuclear import

The following transport of the viral genome to nucleus and the start of the virus replication finally complete the early steps of virus life cycle.
The nucleus provides excellent "service" functions for virus replication, ranging from DNA and RNA polymerases to RNA-splicing and -modifying enzymes. However, the nucleus is difficult to enter and exit, and viruses must again rely on cellular mechanisms [21,161].
In interphase cells, the import of virus and viral capsids occurs through the NPCs. For targeting, viruses use nuclear localization signals and cytosolic

import receptors. Human immunodeficient virus type 1 (HIV-1) and adenovirus bind to importin 7, and human papilloma viruses 11 and 45 and influenza virus nucleoproteins are known to bind importins α and β [94,162-164].

In case of the hepatitis B virus, the most direct indication that transport depends on the core protein comes from studies with *E. coli*-expressed isolated core particles that contain the core protein but no other components of the virus. In an *in vitro* assay using digitonin-permeabilized tissue culture cells, these cores bind specifically to the NPCs [94]. Binding is blocked by reagents such as wheat germ agglutinin and anti-NPC antibodies that inhibit NPC binding. Like capsids isolated from virus particles the *E. coli*-grown cores bind importin α (karyopherin α) *in vitro* and depend on importin β (karyopherin β) for NPC attachment. Thus HBV cores behave like endogenous import cargo with a classical NLS. It is important to note that the binding of cores to NPCs in permeabilized cells depends on prior phosphorylation of at least a fraction of the core proteins by a trapped kinase [94]. Phosphorylation occurs at the extreme C terminus, which contains five potential Ser/Thr phosphorylation sites (see **Figure 7**).

Depending on the model system different upper limits in particle diameter for transport trough the NPC were discussed. Early studies in cells injected with various sizes of NLS-coated gold particles show that the NPC can import macromolecules that are up to ~26 nm in diameter [165]. In contrast to that Bustamante et al. showed in a Nucleus-attached patch-clamp assay that colloidal gold particles were not useful probes; they modified NPC gating [166,167] and the physiological size limit is smaller than 30nm. Performing assays with isolated nuclei and permeabilized cells in the presence of cell lysates/extracts Kann et al. suggest that the HBV capsids, with a diameters of 32 or 36 nm are able to enter the nuclear pore complex and therefore the size of NPC has to be greater than 30nm [168,169].

The smallest viruses and capsids, as well as helical capsids in extended form, can probably be imported into the nucleus without disassembly or deformation. An example for that is the Parvovirus (diameter 18 to 24 nm) which was observed imported intact [170].

Larger viruses and capsids must either be deformed or disassemble to allow the genome to pass through the NPC. Adenovirus 2 binds to NUP214/CAN, a protein located at the base of the filaments extending into the cytosol from the nuclear pore [171]. Interaction of the bound virus with histone H1 and importins β and 7 induces disassembly of the virus capsid. The viral DNA thus liberated is imported into the nucleoplasm. The Herpes simplex virus type 1 capsids (diameter 120 nm) also bind to the NPCs in an importin β–

dependent manner, but the DNA is released through one of the icosahedral vertices of the capsid without further capsid disassembly [172]. The empty capsid stays bound to the pore complex for hours after the DNA has been ejected [173]. With the exception of lentiviruses such as HIV-1, retroviruses do not use the NPCs for nuclear entry. Preintegration complexes can only enter the nucleus during mitosis when the nuclear envelope is temporarily absent, limiting their infectivity to dividing cells [157].

II.4 AIM OF THE STUDY

In order to combine the advantages of viral gene transfer (efficacy and protection of the nucleic acid) and of non-viral gene transfer (safety) the aim of this study should be the development of a novel tool in gene transfer technology based on cell permeable HBV-nucleocapsids loaded with non-viral or viral nucleic acids.

To this end the evolutionary developed features of hepatitis B virus capsid should be used to generate VLPs. To bypass the endosomal uptake and to enable the translocation of HBV-derived VLPs into a variety of different cell types it is necessary to create cell permeable capsids. To achieve cell permeability a novel cell permeable peptide (TLM) [52] should be fused to the core protein in a way that it is exposed on the surface of the assembled capsid. The cell permeability, mediated by the TLM, should allow the receptor and endocytosis independent high efficient transfer of TLM capsids across the cell membrane right into the cytoplasm of target cells.

HBV core particles represent a very well characterized model system [111,122] and can be produced by eukaryotic as well as by bacterial expression systems in sufficient amounts.

To generate TLM capsids it is required to identify the optimal position for fusion of the TLM in the three dimensional context of assembled capsids. Furthermore the insertion of TLM should not affect the capacity of recombinant core proteins to assemble to robust VLPs. To test the ability and efficiency of TLM capsids to translocate into the cytoplasm without affecting the integrity of cells, confocal laser scanning microscopy should be performed.

A crucial prerequisite for gene transfer based on cell permeable VLPs is to develop a protocol for packaging nucleic acid (containing a reporter gene) into the TLM core particle. Like in the viral system the genetic material is packaged into the capsid and by this protected against degradation. For this purpose both an *in vitro* de- and re-assembly of capsids in the presence of nucleic acid as well as an *in vivo* packaging system according to the life cycle of HBV should be performed. After having developed an efficient

protocol for packaging nucleic acid into TLM capsids, a gene transfer in hepatoma cell lines and finally in primary human hepatocytes is intended.

Having established a novel gene transfer system based on cell permeable HBV-derived nucleocapsids this system could be very helpful for investigations concerning the post-endosomal steps of HBV infection. In the absence of an efficient *in vivo* infection system, very little information exists regarding the post-entry intracellular trafficking of the HBV nucleocapsid and transport of the viral genome towards the nucleus. Using cell-permeable HBV nucleocapsids it should be possible to establish a physiological model system which allows a detailed analysis of intracellular trafficking and interaction of these viral particles without affecting integrity of the cell.