TOPIC HIGHLIGHT

Dieter Glebe, PhD, *Series Editor*

Hepatitis B virus replication

Juergen Beck, Michael Nassal

Juergen Beck, Michael Nassal, Department of Internal Medicine II / Molecular Biology, University Hospital Freiburg, Germany
Correspondence to: Dr. Michael Nassal, Department of Internal Medicine II/Molecular Biology, University Hospital Freiburg, Hugstetter Street 55, D-79106 Freiburg,
Germany. nassal2@ukl.uni-freiburg.de
Telephone: +49-761-2703507 Fax: +49-761-2703507

Abstract

Hepadnaviruses, including human hepatitis B virus (HBV), replicate through reverse transcription of an RNA intermediate, the pregenomic RNA (pgRNA). Despite this kinship to retroviruses, there are fundamental differences beyond the fact that hepadnavirions contain DNA instead of RNA. Most peculiar is the initiation of reverse transcription: it occurs by protein-priming, is strictly committed to using an RNA hairpin on the pgRNA, ε, as template, and depends on cellular chaperones; moreover, proper replication can apparently occur only in the specialized environment of intact nucleocapsids. This complexity has hampered an in-depth mechanistic understanding. The recent successful reconstitution in the test tube of active replication initiation complexes from purified components, for duck HBV (DHBV), now allows for the analysis of the biochemistry of hepadnaviral replication at the molecular level. Here we review the current state of knowledge at all steps of the hepadnaviral genome replication cycle, with emphasis on new insights that turned up by the use of such cell-free systems. At this time, they can, unfortunately, not be complemented by three-dimensional structural information on the involved components. However, at least for the ε RNA element such information is emerging, raising expectations that combining biophysics with biochemistry and genetics will soon provide a powerful integrated approach for solving the many outstanding questions. The ultimate, though most challenging goal, will be to visualize the hepadnaviral reverse transcriptase in the act of synthesizing DNA, which will also have strong implications for drug development.

http://www.wjgnet.com/1007-9327/13/48.asp

INTRODUCTION

A hallmark of hepatitis B virus (HBV) replication is protein-primed reverse transcription, related to, but mechanistically distinct from, retroviral replication. This review will focus on the various genome transformations that occur during the replication cycle, with emphasis on the cis-elements on the one hand, and the trans-acting factors known or thought to be involved on the other. A general outline of the chain of events during hepadnaviral replication has been established, mainly using transfection of cloned HBV DNA into a few suitable cell lines, and has been the subject of several reviews[1-6]. Much less clear than what happens is, however, how these various steps are achieved and regulated.

Several features of HBV make resolving such mechanistic questions exquisitely difficult: Foremost, until recently, no *in vitro* systems were available to reconstitute individual replication steps under controlled conditions; conversely, there are no simple infection systems to follow the consequences of mutation-induced *in vitro* phenotypes in the context of authentic virus replication. Secondly, for various of the viral components and mechanisms no precedents exist such that drawing conclusions *in silico*, or by analogy to experimentally more tractable systems, is of limited value. Finally, there is a general lack of structural information on the involved viral factors, in particular on P protein, the viral reverse transcriptase. Fortunately, duck HBV (DHBV), the type member of the *avihepadnaviridae*, provides a valuable model system[7]. Although it differs from HBV in some details, the general features of genome replication are highly conserved; in fact, that hepadnaviruses replicate through reverse transcription has been established with DHBV[8]. Beyond allowing feasible *in vitro* and *in vivo* infection studies, the crucial initiation step of DHBV reverse transcription has recently been
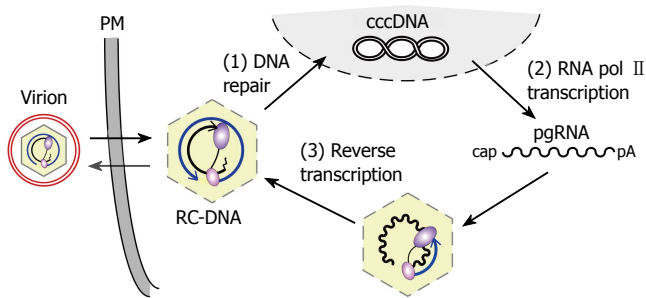
Figure 1 Replication cycle of the hepadnaviral genome. Enveloped virions infect the cell, releasing RC-DNA containing nucleocapsids into the cytoplasm. RC-DNA is transported to the nucleus, and repaired to form cccDNA (1). Transcription of cccDNA by RNA polymerase Ⅱ (2) produces, amongst other transcripts (not shown), pgRNA. pgRNA is encapsidated, together with P protein, and reverse transcribed inside the nucleocapsid (3). (+)-DNA synthesis from the (-)-DNA template generates new RC-DNA. New cycles lead to intracellular cccDNA amplification; alternatively, the RC-DNA containing nucleocapsids are enveloped and released as virions. PM, plasma membrane.

Figure 2 HBV genome organization. The partially double-stranded, circular RC-DNA is indicated by thick black lines, with P covalently linked to the 5′ end of the (-)-DNA, and the RNA primer (zigzag line) at the 5′ end of (+)-DNA. The dashed part symbolizes the heterogeneous lengths of the (+)-strands. DR1 and DR2 are the direct repeats. The outer circle symbolizes the terminally redundant pgRNA with ε close to the 5′ end, and the poly-A tail at the 3′ end. The precore mRNA is nearly identical, except it starts slightly upstream. The relative positions of the open reading frames for core (C), P, preS/S, and X are shown inside. TP, Terminal protein domain of P.

reconstituted in the test tube. On many occasions, we will therefore refer to data obtained with this model virus, and add what is available for human HBV.

OVERVIEW OVER THE HEPADNAVIRAL GENOME REPLICATION CYCLE

Replication of the hepadnaviral genome can broadly be divided into three phases (Figure 1): (1) Infectious virions contain in their inner icosahedral core the genome as a partially double-stranded, circular but not covalently closed DNA of about 3.2 kb in length (relaxed circular, or RC-DNA); (2) upon infection, the RC-DNA is converted, inside the host cell nucleus, into a plasmid-like covalently closed circular DNA (cccDNA); (3) from the cccDNA, several genomic and subgenomic RNAs are transcribed by cellular RNA polymerase Ⅱ; of these, the pregenomic RNA (pgRNA) is selectively packaged into progeny capsids and is reverse transcribed by the co-packaged P protein into new RC-DNA genomes. Matured RC-DNA containing-but not immature RNA containing-nucleocapsids can be used for intracellular cccDNA amplification, or be enveloped and released from the cell as progeny virions. Below we discuss these genome conversions, with emphasis on the reverse transcription step, and particularly its unique initiation mechanism.

RC-DNA TO cccDNA CONVERSION

Persistent viral infections require that the viral genome be present in the infected cell in a stable form that is not lost during cell division, and which therefore can be used for the continuous production of progeny genomes. Many DNA virus genomes harbor replication origins allowing them to directly exploit the cellular replication machinery for amplification; retroviruses integrate a terminally duplicated linear version of their DNA genome into the host genome, such that it is replicated along with the chromosomes. For hepadnaviruses, the genome persists, instead, as a nuclear, episomal covalently closed
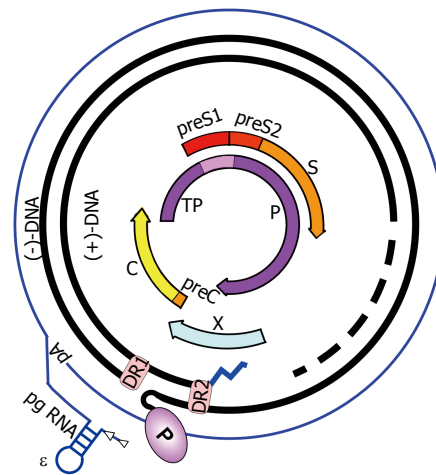
circle, i.e. the cccDNA. The circular form obviates the need for terminal redundancy in that, on the circle, the core promoter/enhancer Ⅱ is placed in front of the start sites for the genomic RNAs; conversely, in typical HBV expression vectors the cloned hepadnaviral DNA is interrupted by plasmid sequences, such that the control regions need to be duplicated[9].

Distinct features of the RC-DNA (Figure 2) are (1), only the (-)-DNA strand (with opposite polarity to the mRNAs) is complete whereas the (+)-strands comprise a cohort of less than full-length molecules; (2), the 5′ end of the (-)-DNA is covalently linked to P protein; (3) the 5′ end of the (+)-strand consists of an RNA oligonucleotide, derived from the pgRNA, which served as the primer for (+)-strand synthesis. For cccDNA formation, all these modifications need to be removed, and both strands need to be covalently ligated.

How this is achieved is not well understood, in particular because unambiguous cccDNA detection in the presence of excess RC-DNA is not technically trivial, not even by PCR approaches[10]. Also, whenever overlength HBV constructs are involved, e.g. upon transduction of cells with HBV carrying adeno-[11] or baculoviruses[12], caution is indicated because homologous recombination could provide for a virus replication independent mechanism of cccDNA formation.

Direct infection avoids this problem but as yet only limited data are available. Earlier evidence obtained with the DHBV-primary duck hepatocyte (PDH) infection system suggested that the activity of the viral P protein is not required for cccDNA generation[13]; however, more recent data using infection with HBV of primary tupaia hepatocytes[14,15] indicate that reverse transcriptase inhibitors can strongly reduce, though not completely block, cccDNA formation[16]. This would be in line with a

role for P protein in the process, probably the completion of the (+)-strands. However, the other steps towards cccDNA are likely to require cellular activities, as suggested by the apparent lack of cccDNA, despite formation of infectious virions, in non-authentic host cells such as hepatocytes from HBV transgenic mice[17]. However, low level cccDNA generation in mice has been reported in certain experimental settings[18,19].

Naturally hepadnavirus infected hepatocytes contain up to 50 or more copies of cccDNA[20], probably in the form of histone-containing minichromosomes[21]. This amplification occurs intracellularly[22,23] in that progeny RC-DNA genomes, like the initially infecting genome, undergo nuclear import and cccDNA conversion (Figure 1). Together with the long half-life (30 to 60 d for DHBV[20,24]), this ensures that cccDNA is not lost during cell division, and even persists during effective antiviral therapy[25]. Nuclear transport, and enveloped virion formation appear to be competing events, such that late in infection, when sufficient amounts of envelope proteins become available, further cccDNA amplification ceases[26]. A recent analysis on the single cell level revealed that the cccDNA copy number in DHBV infected PDH is not uniform and fluctuates over time; about 90% of the cells contained between 1 and 17 copies, the rest more than 17 copies[27,28]. Cells with only one copy may allow segregation of daughter cells containing no cccDNA. This could explain the apparent, though probably less than complete[29], disappearance of cccDNA during spontaneous clearance of HBV infection. However, therapeutic cccDNA elimination from the chronically infected liver remains a major issue even with the latest generation antivirals[30,31]; further investigations into the mechanism of cccDNA formation, and possibly breakdown, are clearly warranted.

FROM cccDNA TO PREGENOMIC RNA

All known hepadnaviral RNAs, i.e. the subgenomic RNAs as well as the greater-than-genome length pgRNA and precore RNA, are transcribed by cellular RNA polymerase Ⅱ (the enzyme responsible for cellular mRNA synthesis) using cccDNA as the template. All contain 5′ cap structures, all are 3′ terminally poly-adenylated at a common site, and all serve as mRNAs for viral gene products. Spliced transcripts do exist, and even can be packaged into progeny virions[32,33], yet their functional role is still obscure although for DHBV splice site mutants appear to have defects[34]. The extent of splicing appears to be controlled in DHBV by a long-range RNA secondary structure[35], and in HBV by the post-transcriptional regulatory element (PRE)[36].

The transcript relevant for virus replication is the pregenomic RNA (pgRNA), encompassing the entire genome length plus a terminal reduncancy of, in HBV, about 120 nt that contains a second copy each of the direct repeat 1 (DR1) and the ε signal, plus the poly-A tail (Figure 3A). The pgRNA starts immediately after the precore initiator codon. Its first essential role is that as mRNA for the core protein and the reverse transcriptase;
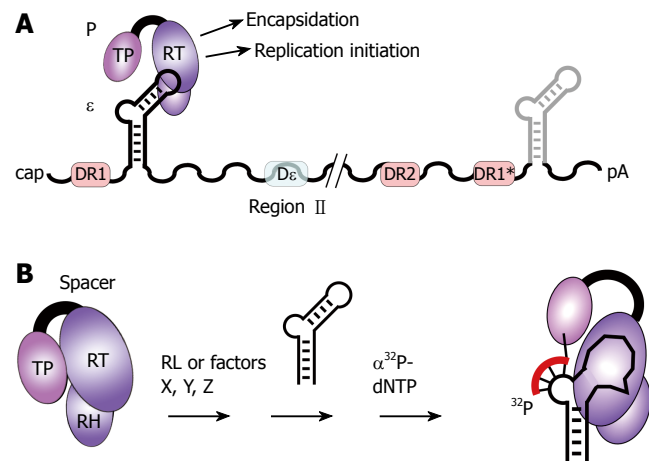

Figure 3 The pgRNA as substrate for P protein. **A**: pgRNA organization. The pgRNA is shown with some major cis-elements, i.e. ε (hairpin structure), DR1, DR2, and DR1*; DHBV, but not HBV, requires for encapsidation an additional cis-element (Dε region Ⅱ). Binding of P protein to 5′ ε, but not 3′ ε, initiates pgRNA encapsidation and replication; **B**: In vitro priming assay. P protein with its Terminal protein (TP), reverse transcriptase (RT), and RNase H (RH) domains can be activated by reticulocyte lysate (RL), or individual factors (X, Y, Z), to bind ε; ε may be supplied as small RNA covering just the hairpin structure. Upon addition of α32P-dNTPs, P uses ε as template to copy 3 to 4 nt from the ε bulge region; by the covalent linkage of the 5′ nt to a tyrosine residue in TP, P protein becomes radioactively labeled, providing a sensitive assay for activity. In vitro priming does thus far not work with human HBV P protein and ε RNA.

unlike retroviral Gag-Pol proteins, P is expressed as a separate polypeptide by an unconventional mechanism[37]. Secondly, pgRNA is the template for generation of new DNA genomes by reverse transcription. The 5′ terminally extended precore RNA contains the initiator codon of the preC region and gives rise to the 25 kDa precore precursor protein of secreted 17 kDa HBeAg. It is unsuited as a pregenome, and is excluded from participating in replication on the level of encapsidation[38].

Of note, RNA polymerase Ⅱ transcription could contribute, in addition to reverse transcription, to hepadnaviral genome variability. Its extent is, however, unclear because the error rate of RNA polymerase Ⅱ is not firmly established and strongly affected by some of the subunits in the holo-enzyme complex[39,40].

CAPSID-ASSISTED REVERSE TRANSCRIPTION OF pgRNA

The next crucial step in hepadnaviral replication is the specific packaging of pgRNA, plus the reverse transcriptase, into newly forming capsids. Key actors are cis-elements on the pgRNA, most notably the encapsidation signal ε, and P protein which binds to ε. This interaction, in a still poorly understood fashion, mediates recruitment of core protein dimers and thus leads to packaging of the pgRNA-P complex. Remarkably, the precore RNA is not packaged although it contains all of the sequence comprising the pgRNA. Likely, active translation from the upstream precore ATG through the ε sequence prevents the P-ε interaction[38]. This implies, in turn, that P binding to ε on the pgRNA interferes with

translation of the core ORF, and evidence supporting this view has been forwarded for DHBV[41].

Once pgRNA and P protein are being encapsidated a second key function of the P-ε interaction is brought to bear, namely the initiation of reverse transcription. At this stage the first DNA nucleotide (nt) is covalently linked to P protein, extended into a complete (-) strand DNA, and (+) strand DNA synthesis ensues, giving rise to a new molecule of RC-DNA; the various immature DNA forms *in statu nascendi* are termed replicative intermediates, visible as a broad multiple band pattern in Southern blots from intracellular DNA.

Completion of reverse transcription before leaving the cell marks another fundamental difference to retroviruses which, except for foamyviruses[42], are secreted as virions containing two copies of RNA; an evolutionary rationale may be that DNA synthesis in the infected cell allows hepadnaviruses the intracellular genome amplification cycle as an alternative to proviral integration for stable virus genome propagation. Yet another difference is that hepadnaviral reverse transcription occurs largely, if not exclusively, in intact nucleocapsids rather than in retrovirus-like reverse transcription complexes which typically lack a continuous core shell[43,44].

Although compartmentalization of the genome amplification process is emerging as a common theme of several classes of RNA and reverse transcribing viruses[45], the hepadnaviral strategy appears as an extreme variation, considering the space restrictions imposed on the replicating complex inside the geometrically defined capsid lumen. Either the RT must slide along the entire 3 kb template, or the template must be pulled through the RT´s active site; at the same time the nucleic acid is most likely in contact with the Arg-rich C termini of the core protein subunits. In fact, capsids from core protein variants lacking part of this region still package pgRNA but appear unable to produce full-length RC-DNA; instead they preferentially reverse transcribe the fraction of spliced genomic RNAs[46,47]. Phosphorylation/dephosphorylation events at the S and T residues in the nucleic acid binding domain clearly accompany genome maturation[48], and mutations affecting the core phosphorylation status[46,49] can influence DNA synthesis in various ways. Outside capsids, or in the absence of core protein, apparently no full-length DNA can be formed.

Together these observations support the view of the capsid as a dynamic replication machine. A recent cryo electron microscopic comparison between recombinant, bacterial RNA containing capsids and authentic genome harboring nucleocapsids indeed showed some structural differences[50]; a complementary study on mutant recombinant cores revealed an enormous flexibility of the capsid structure[51]. Switching between different structural states could well be involved in supporting the progress of DNA synthesis.

The heterogeneous lengths of the (+)-strand DNAs generated by capsid-assisted reverse transcription may result from a non-identical supply of dNTPs inside individual nucleocapsids at the moment of their enclosure by the dNTP-impermeable envelope. This predicts that intracellular cores produced in the absence of envelopment should contain further extended (+)-DNAs. Alternatively, space restrictions in the capsid lumen could prevent (+)-strand DNA completion; in this view, further (+)-strand elongation after infection of a new cell might destabilize the nucleocapsid and thus be involved in genome uncoating.

CIS-ELEMENTS AND TRANSACTING FACTORS ESSENTIAL FOR HEPADNA-VIRAL REPLICATION

The absolute requirements for replication are a template nucleic acid, plus an enzyme that is able to read the template information and use it for synthesis of a complementary nucleic acid. Clearly, these basic components are the pgRNA (and later the (-)-strand DNA) and P protein. However, generation of a functional genome also depends critically on precise start and end points, provided by cis-elements on the template (Figure 3A); a further specialty of P protein is its strict dependence, for activity, on cellular factors, namely heat shock proteins (Hsp´s) or chaperones (see below).

Most initial knowledge on the cis- and trans-factors involved in hepadnaviral replication was derived by transient transfection of mutant viral genomes into stable hepatoma cell lines. This system faithfully mimics several of the authentic replication steps, however its complexity precluded elucidation of many mechanistic details. The recent establishment of *in vitro* systems, culminating in the complete *in vitro* reconstitution of active DHBV replication initiation complexes from purified components[52-54], overcame these restrictions (see below). It should be noted, however, that such minimal systems have their own limitations; for instance, the crucial role of the proper capsid environment for RC-DNA formation has not yet been modeled *in vitro*, and HBV P protein has thus far proven refractory to *in vitro* reconstitution of DNA synthesis activity. Hence which of the two approaches is more useful depends on the question addressed, and where ever possible they should be combined.

STRUCTURE OF THE RNA ENCAPSIDATION SIGNAL ε

The best understood cis-element on the hepdnaviral pgRNA is ε (Figure 4), a stem-loop structure initially defined as the sequence from the 5´ end of HBV pgRNA that P-dependently mediated encapsidation of pgRNA, and also of heterologous transcripts to which it was fused[55]; later, the P-ε interaction was found to constitute the first step in initiation of reverse transcription[56-58]; hence ε also acts as the replication origin (Figure 3A).

The hairpin structure of ε (Figure 4A) was confirmed by secondary structure analyses[59,60], and its importance was established by following the effects of site-directed mutants on the packaging efficacy in transfected cells[59-63]. Furthermore, the ε sequence is highly conserved in other mammalian hepadnaviruses, as well as between different
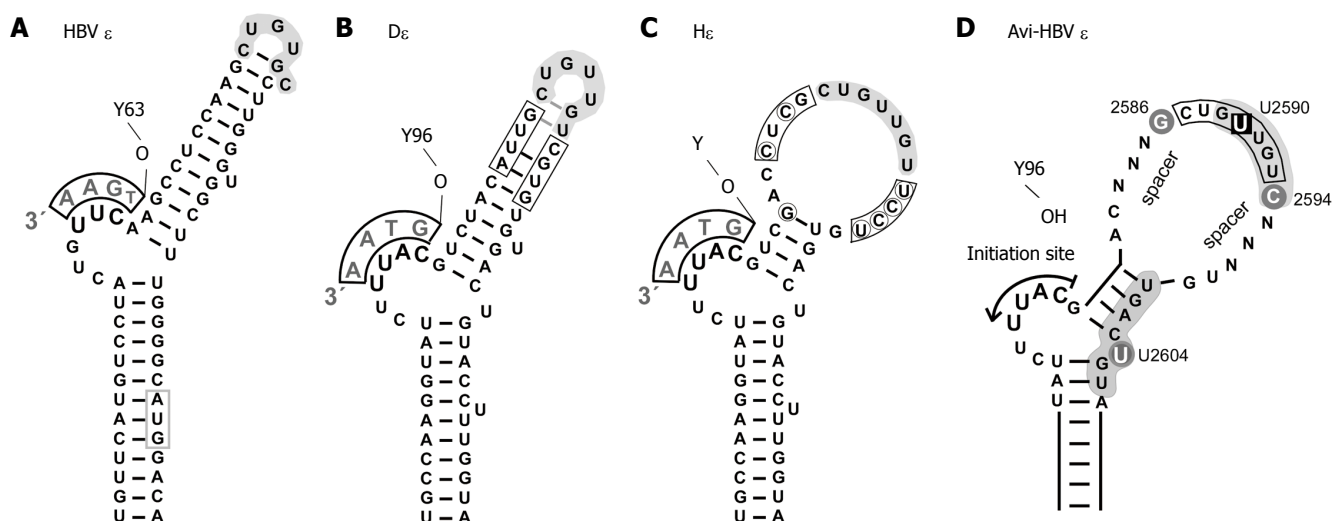
Figure 4 Secondary structures of hepadnaviral ε signals. **A**: HBV ε. The entire hairpin, including in the upper stem region, is stably base-paired; formation of a stable tri-loop, as indicated, is confirmed by direct NMR analysis[68]. The conventional annotation for the loop sequence is high-lighted by grey shading. The bulge-templated DNA oligonucleotide and the priming Y63 are indicated; **B**: DHBV ε. The overall 2D structure is similar to that of HBV ε, including a largely base-paired upper stem[69], as confirmed by preliminary NMR data. The boxed positions were randomized in a SELEX approach, and P binding individuals were selected from the corresponding RNA library[71]; **C**: HHBV ε (Hε). Hε shows substantial sequence variation to **D**: (encircled nt), leading to a largely open upper stem structure; **D**: Generalized secondary structure of an avihepadnaviral ε signal. The scheme summarizes major determinants for productive interaction with P protein[71]. Grey background indicates nt positions that are probably in contact with protein. The bulge and its immediate vicinity, particularly the tip of the lower stem, are essential for P binding whereas the loop appears critically involved in the transition to a productive initiation complex[109]. U2590 > C or U2604 > G mutations abrogate, or strongly reduce, P binding whereas G2586 and C2594 do not affect binding but are important for priming. The major role of the nt termed N may be to provide a proper spacing to the bulge element; their sequence is not important as long as formation of highly stable base-pairing is prevented.

HBV isolates[64,65]. An illustrating example are the HBV precore variants in which HBeAg synthesis is prevented by stop mutations in the ε overlapping precore region. The only mutations causing this phenotype found in nature are those which maintain a stable ε secondary structure[66,67].

It should be noted that RNA secondary structure (2D) analysis provides a mere description of the base-pairing pattern. Hence a true mechanistic understanding requires knowledge of the three-dimensional (3D) structure. The structure of the HBV ε upper stem, recently solved by nuclear magnetic resonance (NMR) techniques[68], revealed that the apical loop consists actually of only 3 nt (as in Figure 4A). This analysis is currently extended to determining the 3D structure of the entire stem-loop, which seems to form a nearly contiguous double-helix (S. Flodell, M. Petersen, F. Girard, J. Zdunek, K. Kidd-Ljunggren, J. Schleucher and S. Wijmenga; submitted); of particular interest will be the structure of the bulge region which is the template for the first few nt of (-)-DNA (see below). Thermodynamic calculations as well as experimental melting curves indicate that the entire HBV ε structure, including the upper stem, is highly stable. Notably, however, in DHBV rearranging this structure is necessary for the RNA to act as a template (see below). Determining the ε RNA structure in the complex with P protein is therefore the ultimate, yet demanding, goal.

Despite limited sequence homology to HBV ε, DHBV ε (Dε) has a similar secondary structure[69] with a bulge and an apical loop (Figure 4B), which suggested that this structure is a common trait of all hepadnaviruses. Surprisingly, the corresponding ε signal (Hε) from the related heron HBV (HHBV) has much reduced base-pairing in the upper stem region (Figure 4C), yet it functionally interacts with DHBV P protein *in vitro* whereas HBV ε does not[70].

Selection, from a library of RNAs with partially randomized upper stems for individuals able to bind to *in vitro* translated DHBV P protein (see Cell-free reconstitution of hepadnaviral replication initiation) revealed the absence of base-pairing in the upper stem region as a common theme (Figure 4D)[71]. Some of the selected P-binding RNAs supported *in vitro* priming while others did not, confirming that a productive interaction, leading to DNA synthesis, requires more than mere physical binding (see below). Hence for avihepadnaviral ε signals an open upper stem structure is beneficial for both physical and productive binding to P; in fact, deliberate Dε stabilization strongly reduces P binding[71]. This is one line of evidence that structural reshaping of the upper stem is a crucial event for initiation of DNA synthesis. Independent support comes from preliminary NMR and melting curve data for wt Dε RNA, according to which the upper stem is the least stable region (F. Girard, O. Ottink, M. Tessari and S. Wijmenga, to be submitted). This marked difference to the highly stable HBV ε structure may be related to the *in vitro* inactivity of the HBV P protein[72].

The functional consequences of mutations affecting ε and other recognized cis-elements are discussed in more detail below. Of note is, however, whereas the HBV ε stem-loop alone is sufficient to mediate encapsidation of heterologous RNAs[62,73], DHBV pgRNA encapsidation requires additional elements. "Region Ⅱ"[74], several hundred nt downstream of 5′ ε, and the intervening sequence[75] may be operative *via* long-range RNA
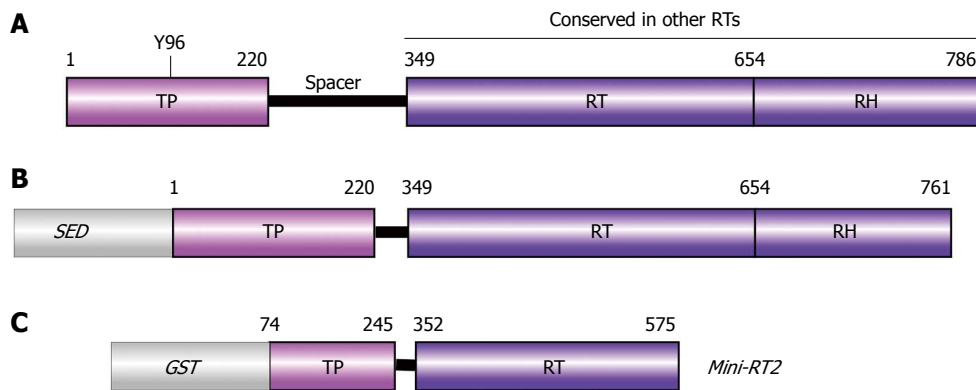
Figure 5 Domain structure of P protein. **A**: Authentic DHBV P protein. Numbers are aa positions for DHBV P protein. The priming Tyr residue Y96 is indicated; **B**: Typical recombinant P protein construct. For solubility, a heterologous solubility enhancing domain (*SED*) such as NusA, GrpE, or *GST* is required, and a short stretch of C terminal aa must be removed. Deletion of the spacer has no negative effects on *in vitro* activity; **C**: Mini-RT2. This heavily truncated recombinant DHBV P protein requires mild detergent, but no chaperones for priming activity.

interactions[76]; why a similar element is dispensable for HBV pgRNA packaging is unclear.

Even HBV ε, however, does not act as a completely autonomous encapsidation element. The 3′ copy of ε cannot substitute for 5′ ε in the context of otherwise authentic pgRNA[77] (although its DHBV counterpart is useable for DHBV P interaction in *in vitro* translation;[78]). Furthermore, 5′ ε mediates encapsidation of heterologous RNAs only up to a limited distance from the 5′ end (about 65 nt), and seems to require the 5′ cap structure[79]. Hence the 5′ cap and factors bound to it appear to have a role in the process, possibly in concert with the 3′ poly-A tail and its associated cellular proteins[80]; this may explain why attempts to reconstitute encapsidation by simultaneous *in vitro* translation of P and core protein from uncapped, non-poly-adenylated RNAs in reticulocyte lysate (RL) were thus far unsuccessful.

An unexplored issue is whether the ε hairpin, or other base-paired regions on the pgRNA, would be substrates for cellular dsRNA recognizing systems such as Toll-like receptor 3 (TLR3) or retinoic acid inducible gene I (RIG-I) which play important roles in innate immune responses against viral infection[81,82], or for enzymes involved in processing of cellular hairpin RNAs such as the microRNA precursors[83]. Direct screens identified a 65 kDa nuclear protein[84] of unknown sequence and, more recently, a novel large RNA-binding ubiquitin ligase, hRUL138[85], as potential cellular ε RNA interaction partners. However, their roles in the viral life-cycle are not known. Also, any cellular ε-binding factor would have to cope with the physical sequestration of pgRNA into nucleocapsids.

P PROTEIN STRUCTURE

The P ORF, covering nearly 80% of the hepadnaviral genome (Figure 2), has a coding capacity of about 830 aa for HBV, and about 790 aa for DHBV (Figure 5). There is no indication for downstream processing of the primary translation products. In further contrast to retroviruses, hepadnaviral virions contain probably just one P protein molecule per particle[86], in line with the covalent linkage to the genome. In transfected cells, P appears to be produced in excess, such that most molecules are not capsid associated[87-89]; they have a short half-life, and are bound to large cytoplasmic structures; their function, if any, is

unclear.

Bioinformatic[90] and genetic analyses showed the presence in all P proteins, of two conserved domains, namely the polymerase/reverse transcriptase (RT) domain, and the C-terminal RNase H (RH) domain. Both are necessary as structural components for pgRNA encapsidation[73]. An absolutely hepadnavirus-specific feature is, however, the Terminal Protein (TP) domain at the N terminus, separated from the RT domain by a highly variable, and dispensable, spacer[91]. TP was first identified as the (-)-DNA linked protein[92] and later was shown to provide a specific Y residue to which the first nt of the (-)-DNA becomes covalently linked (Y96 in DHBV TP,[93,94]; Y63 in HBV TP[95]).

At present, no direct structural information on any hepadnaviral P protein is available although homology-based models for the RT[96] and RH domain[97] have been proposed. The RT model is in accord with drug resistance data, and it is supported by mutational analysis of the putative dNTP pocket of DHBV P protein where a single aromatic residue (F451) was shown to have a homologous role in dNTP versus rNTP discrimination as Y115 in HIV-1 RT[98]; replacement of the bulky F451 by smaller residues conferred to the protein a low but clearly detectable RNA polymerase activity. However, outside the active site the accuracy of the modeled structure is unknown. Hence direct structure determination of the RT and RH domains remains a major objective.

This holds even more for the TP domain, which shares no significant sequence similarity to any other protein in the data base, not even to the few other terminal proteins involved in viral genome replication, such as the VPg in picornaviruses[99], or the terminal proteins of adenovirus[100] and the bacteriophage φ29[101]; moreover, those TP proteins are not covalently linked to their polymerases.

Structure determination requires a source for sufficient amounts of pure, soluble protein which proved to be immensely difficult for P protein. Eventually, this problem was partly overcome by slight modifications in the primary sequence of DHBV P and particularly by adding solubility mediating fusion partners such as GrpE, NusA, or GST[53,102,103]. However, although such fusions (Figure 5B) display activity (see below) they appear to be present as "soluble aggregates" which are unsuited for crystallization.

Particularly TP, when expressed in *E. coli*, is completely insoluble on its own, mostly due to a hydrophobic region

in the C terminal part. By selection from a pool of TP variants with random mutations we could isolate several TP variants, harboring fewer hydrophobic residues in this region, as monodispersely soluble proteins (J. Beck and M. Nassal, unpublished data). Interestingly, the same region was recently implicated to contain a molecular contact site, possibly for the RT domain[104], as suggested by the ability of separate TP and RT/RH domains to trans-complement each other[53,95,105]. Since replacing several conserved hydrophobic regions at a time may affect TP function the challenge will be to find, for crystallization, mutants that combine solubility with functional activity. However, being at the heart of hepadnaviral replication, solving the structure of TP is one of the big current challenges.

CELL-FREE RECONSTITUTION OF HEPADNAVIRAL REPLICATION INITIATION

Cell-free systems are inherently much more manipulatable than intact cells. The first such system to investigate the mechanism of hepadnaviral reverse transcription was based on the observation that DHBV P protein, *in vitro* translated in rabbit RL, became radioactively labeled when the translation reaction was supplied with α^{32}P-dNTPs[78] - as expected if the initial step of reverse transcription, i.e. covalent attachment of the first nucleotide to P protein, had occurred ("*in vitro* priming"; Figure 3B). In fact, the 3´ copy of Dε present on the P protein mRNA was shown to be the template for limited elongation; the role of ε, though only the 5´ copy, as authentic genome replication origin was confirmed for DHBV[56,57], and finally also HBV[58].

The *in vitro* translation system has been used extensively to functionally analyze P protein as well as Dε mutants, particularly because Dε can be added as a separate short RNA covering just the hairpin structure ("trans-priming"). Human HBV P protein and ε, however, show no enzymatic activity in this setting.

THE P-ε COMPLEX: DETERMINANTS FOR BINDING, PRIMING AND ENCAPSIDATION

A first - though not the only-requirement for a productive P-ε interaction is specific binding. In many RNA-protein interactions, structural diversity and hence specificity is achieved by deviations from a fully base-paired double-helical structure[106], e.g. by interspersed single-stranded bulges and loops. Indeed, the Dε bulge structure (but not its actual sequence, unless it affects structure) is absolutely necessary for P binding. Mutants in the upper stem (see Figure 4B) which favor stable non-bulged structures do not bind to P[70]. Similarly critical is the sequence and structure at the junction between the lower stem and the bulge: Mutation of the unpaired U2604 opposite the bulge to G substantially reduced binding (and nearly abolished priming)[71]. At the tip base-pair of the lower stem, base identity of G2605, but not base-pairing itself, is important; in addition, most ribose residues in Dε could be replaced by deoxyribose residues, except in the two top base pairs

of the lower stem and in the bulge residue templating the first nt of (-)-DNA[107]. Hence this small region contains an essential base- and backbone dependent determinant for P-interaction (Figure 4D), likely because it forms, together with the bulge, a distinct three-dimensional recognition surface for P protein.

In the apical loop, various mutations had no drastic effects on P binding, suggesting the loop is not principally required for complex formation (though probably involved in forming a priming competent structure; see below). An exception is U2590 replacement of which by C abrogated binding[70,108]. However, deletion of this residue did not affect binding, and the negative effect of the U2590C mutation was partially rescued by an additional G to A mutation at the neighboring position 2589[61]. A direct 3D structure comparison between wild-type Dε, currently underway (S.S. Wijmenga, personal comm.), and the U2590C mutant RNA may help to explain this complex phenotype.

The existence of Dε RNA variants which bind P but have a much reduced or no template activity indicated that a productive P-ε interaction requires more than binding[61,70,71]. 2D structure comparisons of free versus P bound RNAs provide compelling evidence that the ability of an RNA to undergo a specific conformational shift is such a decisive additional feature[109]. Several base-paired nt in the upper stem of Dε become highly accessible to nucleases once bound to P; similar changes are not seen with priming-deficient variants. Hence it is likely that the RNA, after an initial binding step, must experience an induced-fit alteration into a new structure, and that only this is usable as a template. Non-productive complexes, by contrast, appear to be trapped at the initial binding stage. Foot-printing analysis further revealed that in productive, but not non-productive, complexes the 3' half of the loop plus 3´ adjacent nt are protein bound, as are the unpaired U opposite the bulge and the nt at the tip of the lower stem (Figure 4D). Hence although the loop nucleotides may not be strictly required for initial binding, they probably provide a protein binding site that becomes crucial in the transition to a priming-active complex.

Also P protein undergoes structural alterations in this process. Proteolysis of *in vitro* translated DHBV P protein yielded a distinct proteolytic fragment only in the presence of priming-competent but not priming-inactive RNA variants[108,110]. Hence RNA and P protein mutually alter each others conformation, likely to properly arrange the ε template region and the priming Y residue of TP in the active site of the RT domain.

Notably, of the various P binding RNAs only those that are priming-active also support pgRNA encapsidation. Hence the abilities to initiate reverse transcription, to package pgRNA, and to adopt a distinct RNA-P protein complex conformation appear strictly coupled. In effect, this represents a quality control mechanism ensuring that only RNAs suitable as templates for reverse transcription are packaged. The current methods to analyze these replication-relevant structural changes have a very limited resolution; ultimately, biophysical examination of a priming-active complex will be required, and its
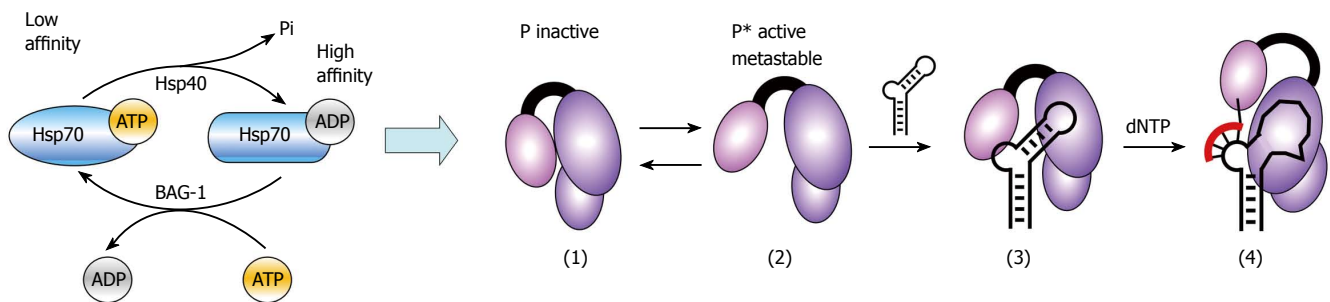
Figure 6 Model for Hsp40/Hsp70 mediated in vitro activation of P protein. The low ATPase activity of Hsp70 is stimulated by Hsp40, yielding high substrate affinity Hsp70/ADP. A new cycle of substrate release and folding requires exchange of the ADP for new ATP which is stimulated by nucleotide exchange factors such as BAG-1. This ATP-dependent Hsp70 cycling applies to the chaperone's global folding activities but likely also to P activation: in the inactive state of P (1), the ε RNA binding pocket is inaccessible; Hsp40/Hsp70 activation creates active P* (2) which is able to bind ε RNA (3); P* is metastable, and decays to the inactive state (1) within minutes. Maintaining a steady-state level of P* requires constant re-activation by Hsp40/Hsp70 and thus a continuous supply of fresh ATP. Complexes containing priming-competent ε RNAs (3) undergo induced-fit rearrangements in the RNA and the protein, enabling them to initiate DNA synthesis upon dNTP addition (4). Several Dε variants bind P but do not act as template; most likely they are trapped at stage (3). The same may hold for human HBV P protein complexes with HBV ε RNA.

comparison with non-productive complexes should be most revealing. However, several significant obstacles will have to be overcome for this approach, not the least being the strict dependence of DHBV P protein activity on additional cellular factors (see below).

This holds even more for active initiation complexes of human HBV. Insect cell expressed HBV P exerted a low but clearly detectable polymerase activity, with part of the DNA products covalently linked to TP as expected from authentic initiation[111]. The system revealed some important features, such as the priming role of Y63 in HBV TP[95], trans-complementation between TP and RT/RH domains[95], and even some nucleocapsid formation by co-expression of core protein[112]. Disturbingly, however, various of these activities were not strictly ε dependent, and they could not be reconstituted after (partial) purification of the P protein. Whether insect cells contain some RNA that can substitute for ε has not been clarified. *In vitro* translated HBV P protein shows no priming activity in reticulocyte lysate and not even in lysates from Huh7 cells (J. Beck and M. Nassal, unpublished data) which support replication when transfected with HBV. Recently, however, Hu et al[72] reported a partial but important progress in that they were able to demonstrate specific binding-though not priming-for HBV P and ε in an *in vitro* reconstitution system (see next paragraph).

IN VITRO RECONSTITUTION OF REPLICATION INITIATION FROM PURIFIED COMPONENTS

One aspect for *in vitro* translation of P protein was, at the time, the lack of alternative ways to recombinantly produce the protein. Surprisingly, wheat germ extract versus rabbit RL translated DHBV P protein had a much reduced priming activity, suggesting the mammalian RL provided additional essential factors[113]. These turned out to be cellular chaperones, which are abundantly present in RL. Apart from Hsp60 chaperonins[114] (GroEL in bacteria), Hsp70 (DnaK in bacteria) and Hsp90 constitute the major chaperone systems. Hsp70 assists folding of many newly

synthesized polypeptides, refolding of misfolded proteins resulting, e.g., from heat-shock, and protein translocation through membranes[115-117]. Hsp90 has also broad but more specialized folding functions[118], usually if not exclusively in concert with Hsp70; the two systems are linked *via* the Hsp70/Hsp90 organizing protein Hop. By analogy to the activation of nuclear hormone receptors[119] it was proposed that Hsp90 plus its small co-chaperone p23 are the essential factors for P protein activation[113,120]. However, the complex overall composition and the high chaperone content of RL compared to the minute amounts of P protein precluded a clear-cut distinction.

This was overcome when DHBV P protein became accessible in larger quantities by expressing, in *E. coli*, slightly modified variants[53,102], in particular as fusions with solubility enhancing heterologous domains (Figure 5B). Such recombinant P proteins could now be added to the RL system and showed activity. Finally, because the translation function of RL was not required any longer, it became possible to systematically analyze the chaperone requirements for P protein activation with purified individual components.

In this way we could demonstrate that DHBV P can efficiently be activated *in vitro* by Hsp40 and Hsp70 plus ATP as an energy source, without the need for Hsp90 or other cofactors[54]. The primary reaction product is an RNA binding-competent form of P protein (P*) that decays quickly in the absence of ε RNA but, in its presence, accumulates in an initiation-competent form (Figure 6).

Maintaining P in its activated P* form requires a constant supply of ATP[54], and the same holds for the general folding activity of Hsp70[117]. This suggested that P activation represents a special form of Hsp70 mediated folding in which the chaperone, rather than helping P from an unfolded into a stable folded state, affects the equilibrium between the inactive P ground state and the metastable activated P* form. Of note, bacterial DnaK has been shown to interact, physiologically, with a few folded, as opposed to misfolded, proteins such as the transcription factor σ32[121,122].

Hsp70 chaperoning is a cyclic ATP-driven process[117].

Hsp70 binds ATP and in this form exerts low affinity to folding substrates (Figure 6, left). ATP hydrolysis then generates the high affinity Hsp70/ADP form. In the presence of substrate, the weak Hsp70 ATPase activity is stimulated up to 1000 fold by Hsp40 and related J-domain proteins (named after the prototypic *E. coli* Hsp40, DnaJ[123]), explaining the important role of Hsp40. Initiating a new folding cycle requires reconversion of Hsp70/ADP into Hsp70/ATP, i.e. replacement of the bound ADP by fresh ATP; otherwise Hsp70 would be trapped and not be available for acting on new substrate molecules. Spontaneous nt exchange is slow but strongly enhanced by nucleotide exchange factors (NEFs). This predicted that addition of BAG-1, an established NEF of Hsp70[124], to the minimal *in vitro* reconstitution system should enhance the formation of P* molecules. Indeed, we observed a strong BAG-1 dependent increase in priming-active P molecules, but not with a BAG-1 mutant unable to interact with Hsp70. Hence ADP-ATP exchange on Hsp70 is the rate limiting step in the *in vitro* priming reaction (M. Stahl, M. Retzlaff, M. Nassal and J. Beck, submitted). A working model for the Hsp40/Hsp70 activation of DHBV P protein is shown in Figure 6.

Though these data are clear-cut, Hu and colleagues, using a similar experimental set-up though a different (i.e. GST) DHBV P fusion protein, reported that in their *in vitro* system P activation was strictly dependent on the additional presence of Hsp90 and the Hsp70/Hsp90 adaptor protein Hop, with the Hsp90 co-chaperone p23 enhancing the reaction rate[52].

In our system, Hop and Hsp90 do have a stimulatory-but not an essential-effect (in a similar range as BAG-1), particularly at low P protein concentrations. In addition, the specific nature of the Hsp40 used can affect activation efficiency. Eukaryotic cells contain numerous Hsp40-like (or J-domain) proteins[125]; all contain the about 70 aa long J-domain which mediates interaction with Hsp70. Hsp70 activation of P protein appears to proceed selectively with the Hdj1 variant of Hsp40 but not Hdj2 or its yeast homolog Ydj1. With Ydj1, Hop and Hsp90 may become essential for detectable activation. Why different Hsp40s have different effects is currently unclear because all Hsp40s can stimulate the Hsp70 ATPase activity; however, there are different ways of how a substrate can enter the folding complex: either it is bound directly by Hsp70, or it is presented to Hsp70 by the Hsp40-like protein[123]. Therefore, the various additional domains present in different J-domain proteins could effect the formation of P* *via* different pathways. Hence in summary, P activation *in vitro* is fundamentally dependent on Hsp70/Hsp40 but can be enhanced by additional factors, including Hop/Hsp90.

Interestingly, the strict chaperone dependence of DHBV P protein activity was relieved by an extensive C terminal truncation that removed the entire RH domain and some 75 aa from the RT domain; the only requirement for priming activity (though restricted to the very first DNA primer nt) of this truncated Mini-RT 2 protein (Figure 5C) was the presence of mild detergent[126]. This suggests that in full-length P protein C terminal parts

somehow block the ε RNA binding site, and that this occlusion is removed by the chaperone action[102,127].

Where exactly the chaperones bind to P and which conformational rearrangements they induce is unclear. Using *in vitro* translated DHBV P protein in RL Tavis *et al*[110] noted the ε-dependent generation of a papain- and trypsin-resistant fragment covering mainly the RT domain. Taking advantage of the simple, defined composition of the *in vitro* reconstitution system we could directly investigate the effects of Hsp40 and Hsp70 on P protein conformation. Limited V8 protease digestion revealed specific chaperone- and ATP-dependent cleavages in the C terminal part of TP (between aa 164 and 199). Hence this TP region is inaccessible in non-activated P protein but becomes exposed in P* (M. Stahl, J. Beck and M. Nassal, unpublished data). The functional relevance of this conformational alteration is supported by its correlation with the ability of P to bind ε RNA, and the shielding of the same region from protease attack as long as the Dε RNA is bound; furthermore, the larger fragment encompasses two residues, K182 and R183, which are essential for RNA binding. These data support the model shown in Figure 6. Inclusion of Hop and Hsp90 in this assay should now allow one to monitor whether these, or additional factors, have differential effects on P conformation, or whether they mainly stabilize the changes already established by Hsp40/Hsp70.

Extending the scope of the *in vitro* reconstitution system, Hu and colleagues recently demonstrated that specific binding to ε by the human HBV P protein also appears to be controlled by cellular chaperones[72,128]. On the ε RNA side, the region immediately surrounding the bulge was essential but the apical loop was not. This was surprising given that the loop is essential for pgRNA encapsidation and initiation of replication in intact cells. However, in light of the above described DHBV data, it may just be a drastic manifestation that mere physical binding is not sufficient for a productive interaction; this is further supported in that only about half the P protein sequence was required for RNA binding, with even the catalytic YMDD motif being dispensable. Apparently, the crucial second step, which in Dε involves rearranging the upper stem and the loop, does not occur in the reconstituted HBV P protein-ε complexes, possibly due to the high stability of this region in HBV ε. Finding conditions under which HBV P protein exerts authentic ε-dependent *in vitro* priming-activity remains therefore another major challenge in HBV biology.

IMPORTANCE OF CHAPERONES FOR HEPADNAVIRUS REPLICATION IN INTACT CELLS

Both the Hsp70 and the Hsp90 chaperoning activities are subject to regulation by a multitude of co-chaperones[117,129]. Hence which of the *in vitro* reconstitutable processes is the relevant one inside cells is not trivial to address, in particular because the chaperones are crucial for very many fundamental and regulatory cellular processes. For
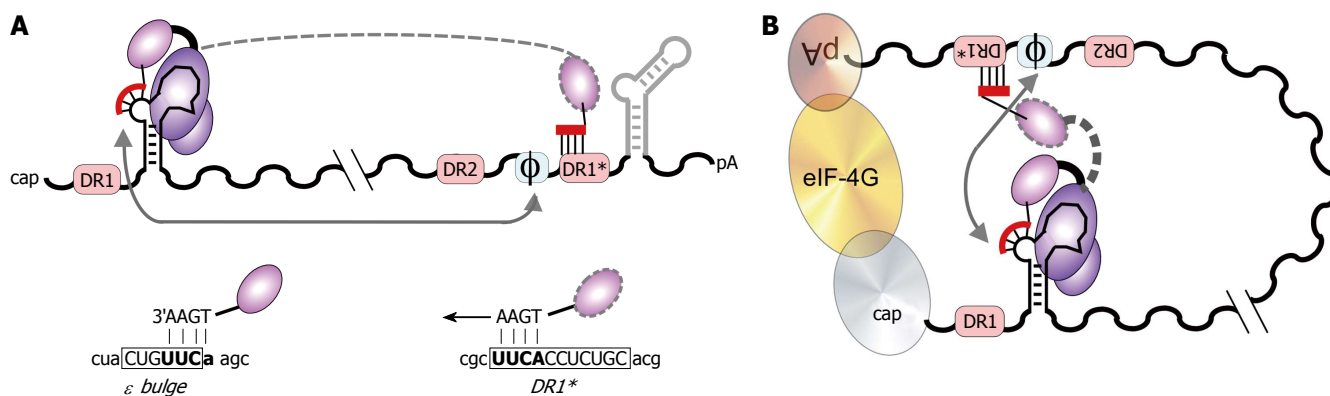
Figure 7 DNA primer translocation (first template switch). P copies 3 to 4 nt from the 5′ ε bulge, yielding the TP-linked DNA oligonucleotide which is translocated to the complementary motif in the 3′ proximal DR1*. **A**: Linear representation. DR1* is nearly 3 kb apart from 5′ ε in primary sequence, and numerous other UUCA motifs are not used as acceptors. φ denotes a newly identified cis-element with partial sequence complementarity to the 5′ half of ε. 3′ ε (light grey) is dispensable. The HBV specific sequences in the ε bulge, and in DR1* are shown below in capitals, flanking sequences in lower case; **B**: Models for juxtaposition of 5′ ε and DR1*. A general mechanism would be closed loop formation[80] of the pgRNA by cap-binding and poly-A binding factors (ovals), e.g. *via* elongation initiation factor 4G (eIF-4G; large oval). More specifically, ε might base-pair with φ[146,147], as indicated by the grey arrows. In such an arrangement, a small movement, rather than a big jump, of TP with the bound DNA primer (dashed outline) would suffice for specific translocation to DR1*.

instance, geldanamycin (GA), an inhibitor of Hsp90, strongly interfered with DHBV replication in transfected cells at a concentration of 10 ng/mL whereas significant reduction of the *in vitro* priming activity required much higher concentrations[113]. Therefore, the antiviral effect in cells might be indirect, perhaps *via* one of the cellular kinases that are regulated by Hsp90[130]. Such effects on the cell make it also difficult to imagine that chaperone inhibitors could therapeutically be used against HBV infection without causing severe adverse effects. That GA analogs, nonetheless, exert tumor-specific therapeutic value[131] is probably due to the selective presence, in cancer cells, of Hsp90 in a high GA affinity state[132]. Whether this also holds for HBV infected cells is not known.

Various additional chaperones have, in part indirectly, been implicated in affecting P protein, for instance Hsp60[133-135] and the Hsp90 family member GRP94[136]; however, GRP94 is an endoplasmic reticulum (ER) resident protein for which a role in P protein activation is difficult to imagine. Overexpression of p50/cdc37, a co-chaperone of Hsp90 involved in activation of several signal transduction kinases, stimulated DHBV replication in transfected cells and a dominant negative mutant of p50/cdc37 inhibited DHBV P protein priming *in vitro*[137]. Again, the physiological relevance of these observations remains to be confirmed.

One possible solution would be to map the contact sites of the various chaperones on P protein using biochemical and biophysical methods, and then to generate P mutants with specific interaction defects. Monitoring their replication phenotypes in transfected, or better in infected, cells should then allow to narrow down which of the various reported interactions are truly significant for virus propagation *in vivo*.

Though we have stressed the uniquely strict chaperone-dependence of hepadnaviral P protein activation, there is accumulating evidence that other polymerases are also affected by chaperones. Probably the closest analogy exists to telomerase, the cellular reverse transcriptase that maintains chromosome end integrity[138,139]. However, also

the DNA polymerase of Herpes simplex virus[140], and the RNA polymerases of flock house virus[141] and influenza virus appear to require chaperone assistance[142,143]. Closely watching progress in those areas might also provide clues as to the mechanism of chaperone-assisted hepadnavirus replication; however, for polymerases without a similarly sophisticated protein-primed initiation mode the chaperones could act at rather different levels.

DNA PRIMER TRANSLOCATION AND (-)-DNA COMPLETION

In contrast to initiation, the subsequent steps for RC-DNA formation are currently not amenable to *in vitro* analysis, and they appear intimately related to the proper environment of intact nucleocapsids. Hence reverse genetics is still the most rewarding approach to address these equally puzzling and complex events.

The initial model of hepadnaviral (-)-DNA formation assumed that synthesis would start inside the 3′ copy of DR1 (DR1*), for HBV at the motif 5′ UUCA. As discussed above, the complementary sequence 3′ AAGT at the 5′ end of (-)-DNA is instead copied from the UUCA motif in the ε bulge (Figure 7A). Hence the oligonucleotide bound to TP must specifically be translocated to the 3′ DR1*, nearly 3 kb apart from 5′ ε. Given that there are about 20 further UUCA motifs on the pgRNA, and that even fewer than 4 nt of identity between the template region in ε and the target site in DR1* are sufficient for specific transfer[58], additional elements ensuring proper translocation must be operating. One model is that DR1* and 5′ ε are brought into close proximity (Figure 7B). A general mechanism would be closed-loop formation of the pgRNA *via* cellular proteins such as elongation initiation factor 4G (eIF-4G) which links 5′ cap and 3′ poly-A binding proteins[80]. A recent more specific model is a long-range RNA interaction between ε and a new cis-element ("φ" or "β5") slightly upstream of DR1* that is involved in proper (-)-DNA synthesis[144-146]; it contains a sequence that
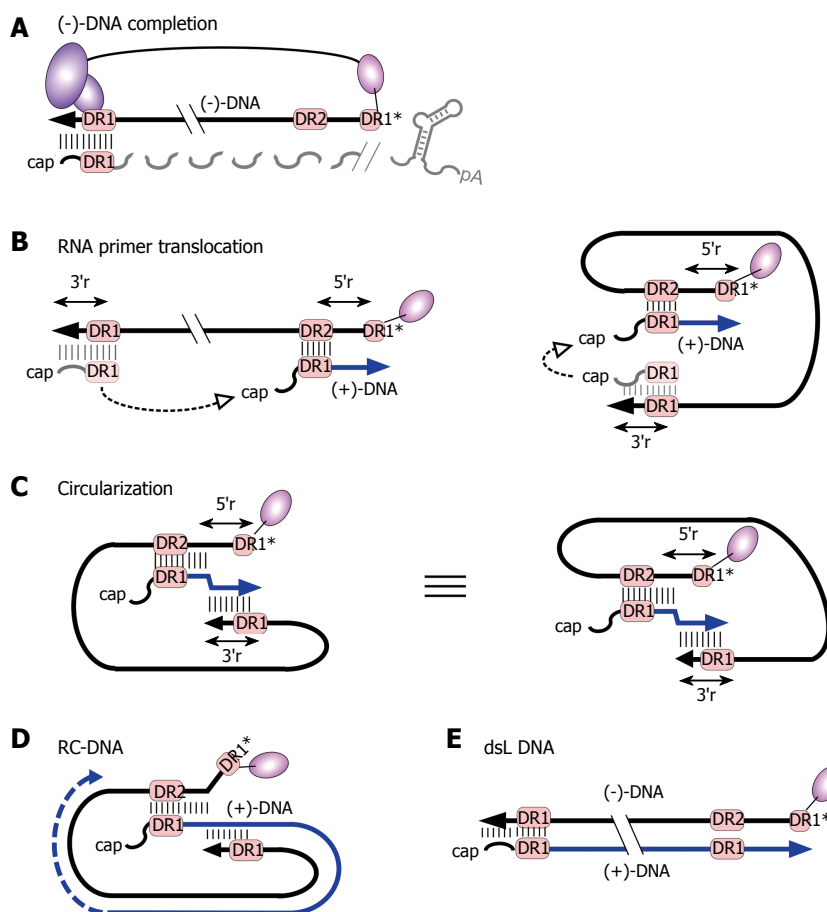
Figure 8 RC-DNA formation. **A**: (-)-DNA completion. The DNA primer, still linked to TP, is extended from DR1* to the 5′ end of pgRNA. The RNA is simultaneously degraded by the RH domain, except for its capped 5′ terminal region including 5′ DR1; the fate of the poly-adenylated 3′ end is unclear; **B**: RNA primer translocation (second template switch). The RNA primer translocates to DR2, and is extended to the 5′ end of (-)-DNA. 3′ r and 5′ r denote an about 10 nt redundancy on the (-)-DNA. As above, several cis-elements appear to promote close proximity of the DR1 donor and the DR2 acceptor, as schematically indicated in the right hand figure; **C**: Circularization (third template switch). Having copied 5′ r, the growing 3′ end of the (+)-DNA switches to 3′ r on the (-)-DNA, enabling further elongation. This reaction must involve juxtaposition of 5′ r and 3′ r. For easier comprehension, the switch is also depicted on the basis of the representation shown on the right of Figure 8B; both are topologically equivalent; **D**: RC-DNA. Extension on the (-)-DNA template creates a set of (+)-DNA strands of various length; **E**: Double-stranded linear (dsL) DNA. This minor DNA form originates when the RNA primer, having failed to translocate to DR2, is extended from its original position ("in situ priming").

is partially complementary to the 5′ half of ε. Mutations affecting the base-pairing potential between the two sequences can, indeed, influence the efficiency of (-)-DNA synthesis[147]. Exactly how an ε-φ interaction would aid in primer translocation is unclear because base-pairing would simultaneously affect the proper ε structure, implying some temporal regulation. Similar φ elements have been proposed to be present in the other hepadnaviruses[146]; for DHBV we noted, however, that Dε mutants with reduced potential base-pairing to the supposed φ element showed no obvious (-)-DNA synthesis defects ([71] and K. Dallmeier, B. Schmid and M. Nassal, unpublished data).

The primer translocation process must also involve remodeling of the P-ε complex; the priming Y residue of TP must give way to the growing DNA oligonucleotide, then ε must be replaced by DR1* as the template. Hence P, like other protein-priming polymerases, must have distinct initiation and elongation modes. The recently solved structure of the bacteriophage φ29 DNA-dependent DNA polymerase with its (separately expressed) terminal protein[101] gives an impression of the dynamic changes that have to occur. φ29 uses for protein priming an S residue in one of three distinct TP domains. In the initiation mode, the priming domain mimics the primer-template nucleic acid that occupies the same site during elongation. The initiation reaction continues until 6 to 10 nt have been attached; at this time the priming domain is pushed out of the polymerase´s active site, marking the transition to elongation mode. Similar events must happen in P protein. Notably, initiation of φ29 replication does not start at the very terminal template nt; rather the (identical) penultimate

nt is copied. Next, the copied nt slips back to the terminal template position, and the penultimate nt is copied again; adenovirus uses a similar mechanism[100]. Interestingly, such slip-back and re-copying can also occur with HBV P protein on mutant ε templates[58]. A full understanding, however, will obligatorily require high resolution structural data on the hepadnaviral replication complex.

Notably, the single-stranded (-)-DNA is an intermediate that, as in retroviral replication, could be a target for cytidine deamination, and consequently inhibition of replication, by APOBEC enzymes[148]. APOBEC3G, one of several family members, has indeed been reported to interfere with HBV replication[149], however, the mechanism does not involve editing[150]. Also, expression of the corresponding APOBEC mRNA is low in hepatic cells, although it might be inducible[151]. Further research is needed to clarify whether any of the APOBEC enzymes is genuinely involved in the innate response against HBV infection. At any rate, sequestration into intact nucleocapsids of this potentially vulnerable single-stranded DNA replication intermediate may be an efficient counter-defense of HBV.

(+)-STRAND DNA SYNTHESIS AND CIRCULARIZATION

The end product of (-)-DNA synthesis is a unit length DNA copy of the pgRNA from its 5′ end to, in HBV, the UUCA motif in the 3′ DR1* (Figure 8A); hence it contains a small, about 10 nt, terminal redundancy ("r"). Most of

the pgRNA template is degraded concomitantly to (-)-DNA synthesis by the RNase H domain of P. The fate of the non-copied 3´ end of the pgRNA from DR1* to the polyA tail is not exactly known. It seems to be underrepresented in the packaged RNA, both in DHBV[152], and in HBV with mutations in the nucleic acid binding domain of the core protein[46]. Whether some of these 3´ ends are never completely encapsidated, or whether they dissociate out of the capsid, has not yet conclusively been shown.

Well established is, however, that the 5´ terminal about 15 to 18 nt of the pgRNA including the 5´ DR1 sequence are spared from degradation, probably because the active site of RH is halted at this distance when the RT domain reaches the template 5´ end[153]. This capped 5´ RNA oligo is essential as primer for (+)-DNA synthesis. Extension of the RNA from its original position ("in situ priming") gives rise to a double-stranded linear (dsL) DNA (Figure 8E) which occurs to a small percentage in all hepadnaviruses[154]. Lacking the core promoter/enhancer for pgRNA translation upstream of the pgRNA start site it is unsuited for virus propagation but may be of pathogenic potential[155]. For RC-DNA formation, the RNA primer must be transferred to the 3´ proximal DR2 (Figure 8B); pgRNAs with improper 5´ ends, not containing the DR1 sequence in the RNA primer, fail to undergo this essential second template switch. Why the RNA primer predominantly jumps to DR2 although its complementarity to the initial site is larger (more than 15 versus 11 or 12 nt) is not obvious. An interesting explanation has been proposed for DHBV, according to which burying part of the 5´ DR1 sequence in a competing intramolecular hairpin structure effectively shortens the sequence available for hybridization with the RNA primer[156].

From its new location on DR2 the RNA primer is extended towards the P bound 5´ end of the (-)-DNA, including the 5´ r redundancy. Further elongation requires a third template switch, i.e. circularization (Figure 8C). In effect, the growing (+)-DNA end is transferred from 5´ r to 3´ r on the (-)-DNA template from where it can further be extended to yield RC-DNA (Figure 8D).

Though sequence identity between 5´ r and 3´ r is important, additional cis-elements are again required to ensure efficient RNA primer translocation and circularization. Using quantitative genetic techniques, the Loeb laboratory has defined, mostly in DHBV, several such cis-elements on the (-)-DNA[144,156-161], e.g. 3E, M, and 5E, which are located at both termini and in the middle of pgRNA. Collectively, these data provide evidence that these cis-elements, *via* long range base-pairing, allow for a close juxtaposition of the corresponding donor and acceptor sites, and thus facilitate the proper template switches (Figure 8B, right panel). How this is achieved inside the replicating nucleocapsid is not easily envisaged. Also, we noted that chimeric heron-duck HBVs are less sensitive than DHBV to reduced base-pairing in some of these elements (K. Dallmeier and M. Nassal, unpublished). There is evidence, however, that similar potentially base-pairing cis-elements are present in human HBV[144, 162] and assist in circularization (E. Lewellyn and D.D. Loeb, pers. comm.).

Hence intramolecular base-pairing is probably an important mechanism that ensures proper shaping of the viral genome for the various template switches that eventually yield RC-DNA. More details than is possible to show here can be found in informative schematic representations in references[144,160].

CONCLUSIONS AND PERSPECTIVES

Since the original discovery that hepadnaviruses replicate through reverse transcription numerous novel and unique aspects of the replication process have been unraveled by using reverse genetics in transfected cells. Indisputably, this system will continue to be an important tool in future HBV research; in some cases because even a cancer-derived stable cell line mimics many aspects of the complex natural environment for viral replication, allowing to look into the interplay between virus replication and cellular networks by genomics and/or proteomics; in others because we have currently no alternative due to the complexity of the interactions between the viral components themselves. An example is the control of the various template switches during RC-DNA generation by long-range nucleic acid interactions which apparently can only occur, and be monitored, in the context of assembled nucleocapsids.

For a mechanistic understanding, however, cell-free systems are indispensable. Two striking examples for their power are the discovery of the replication origin function of ε and the chaperone-dependence of P protein activation; both would have gone unnoticed for a long time without the *in vitro* translation system. To disentangle these multifactorial processes on the molecular level, *in vitro* reconstitution from purified components is the approach of choice. Despite the bewildering complexity of the chaperoning systems, it will now be possible, starting from relatively simply composed systems, to systematically add-in additional factors and monitor their contributions to P protein activation; this should, *inter alia*, help answer the pressing question why human HBV P protein shows no enzymatic activity under conditions where its DHBV counterpart is active. An *in vitro* activity assay for HBV P protein would also be an important screening tool for better HBV antivirals. Similarly rewarding should be the development of *in vitro* systems that address other steps of the viral replication cycle, foremost perhaps inclusion of the core protein into the now available replication initiation systems.

Ultimately, structural biophysics needs to enter the field. At this time the HBV core protein, without its nucleic acid binding domain, is the only hepadnavirus component for which a high resolution structure is available[163]. Fortunately, 3D structural analyses of the ε element are well underway, and at least individual domains of the P protein may become amenable to direct structure investigations. The most exciting, though also most challenging aim given its multifactorial composition, will be to obtain high resolution data for the P protein-ε RNA complex, caught in the act of DNA synthesis. However, the relevance of any such biophysical and biochemical

data will have to be corroborated in the context of the complete virus replication cycle, whenever possible in an *in vivo* setting.

Finally, although the complex interactions between viral, and viral and cellular, components may seem of foremost interest for basic HBV biology, they also hold the keys for novel, and more efficacious therapies of chronic hepatitis B. The few currently approved anti-HBV drugs are all nucleos(t)idic inhibitors of the reverse transcriptase. Their long-term efficacy is limited by the virtually unavoidable emergence of resistant HBV variants[30]; any new compounds with the same mechanism of action will also face this problem, which will be aggravated by cross-resistance, particularly against different analogs of the same natural nucleotide. Each of the steps in the HBV replication cycle that is now being elucidated in molecular detail provides new, unconventional targets for interference. Replication initiation alone depends on many specific interactions, including the TP and RT/RH domains of P protein, ε RNA, and different chaperones. Blocking any of these interactions, e.g. by small molecules binding to, and altering the structure of ε or preventing chaperone binding to P, would abolish reverse transcription by mechanisms entirely different from that of nucleoside analogs. Knowledge of the 3D structure of P would allow to design better nucleosidic as well as non-nucleoside inhibitors, and additional opportunities will certainly arise once the process of cccDNA formation is better understood. Especially in combination with conventional antivirals, such potential new drugs should greatly increase the chances for curing, rather than just controlling, chronic hepatitis B.

REFERENCES

1   **Nassal M**, Schaller H. Hepatitis B virus replication. *Trends Microbiol* 1993; **1**: 221-228

2   **Nassal M**, Schaller H. Hepatitis B virus replication--an update. *J Viral Hepat* 1996; **3**: 217-226

3   **Nassal M**. Hepatitis B virus replication: novel roles for virus-host interactions. *Intervirology* 1999; **42**: 100-116

4   **Nassal M**. Macromolecular interactions in hepatitis B virus replication and particle formation. In: Cann AJ, editor. Frontiers in Molecular Biology: DNA virus replication. Oxford: Oxford University Press, 2000: 1-40

5   **Ganem D**, Schneider R. Hepadnaviridae: The Viruses and Their Replication. In: Knipe DM, Howley PM, editors. Fields Virology. 4th ed. Philadelphia: Lippincott Williams & Wilkins, 2001: 2923-2969

6   **Seeger C**, Mason WS. Hepatitis B virus biology. *Microbiol Mol Biol Rev* 2000; **64**: 51-68

7   **Schultz U**, Grgacic E, Nassal M. Duck hepatitis B virus: an invaluable model system for HBV infection. *Adv Virus Res* 2004; **63**: 1-70

8   **Summers J**, Mason WS. Replication of the genome of a hepatitis B--like virus by reverse transcription of an RNA intermediate. *Cell* 1982; **29**: 403-415

9   **Protzer U**, Nassal M, Chiang PW, Kirschfink M, Schaller H. Interferon gene transfer by a hepatitis B virus vector efficiently suppresses wild-type virus infection. *Proc Natl Acad Sci USA* 1999; **96**: 10818-10823

10   **Sun D**, Nassal M. Stable HepG2- and Huh7-based human hepatoma cell lines for efficient regulated expression of infectious hepatitis B virus. *J Hepatol* 2006; **45**: 636-645

11   **Ren S**, Nassal M. Hepatitis B virus (HBV) virion and covalently closed circular DNA formation in primary tupaia hepatocytes and human hepatoma cell lines upon HBV genome transduction with replication-defective adenovirus vectors. *J Virol* 2001; **75**: 1104-1116

12   **Isom HC**, Abdelhamed AM, Bilello JP, Miller TG. Baculovirus-mediated gene transfer for the study of hepatitis B virus. *Methods Mol Med* 2004; **96**: 219-237

13   **Köck J**, Schlicht HJ. Analysis of the earliest steps of hepadnavirus replication: genome repair after infectious entry into hepatocytes does not depend on viral polymerase activity. *J Virol* 1993; **67**: 4867-4874

14   **Köck J**, Nassal M, MacNelly S, Baumert TF, Blum HE, von Weizsäcker F. Efficient infection of primary tupaia hepatocytes with purified human and woolly monkey hepatitis B virus. *J Virol* 2001; **75**: 5084-5089

15   **von Weizsäcker F**, Köck J, MacNelly S, Ren S, Blum HE, Nassal M. The tupaia model for the study of hepatitis B virus: direct infection and HBV genome transduction of primary tupaia hepatocytes. *Methods Mol Med* 2004; **96**: 153-161

16   **Köck J**, Baumert TF, Delaney WE, Blum HE, von Weizsäcker F. Inhibitory effect of adefovir and lamivudine on the initiation of hepatitis B virus infection in primary tupaia hepatocytes. *Hepatology* 2003; **38**: 1410-1418

17   **Guidotti LG**, Matzke B, Schaller H, Chisari FV. High-level hepatitis B virus replication in transgenic mice. *J Virol* 1995; **69**: 6158-6169

18   **Raney AK**, Eggers CM, Kline EF, Guidotti LG, Pontoglio M, Yaniv M, McLachlan A. Nuclear covalently closed circular viral genomic DNA in the liver of hepatocyte nuclear factor 1 alpha-null hepatitis B virus transgenic mice. *J Virol* 2001; **75**: 2900-2911

19   **Takehara T**, Suzuki T, Ohkawa K, Hosui A, Jinushi M, Miyagi T, Tatsumi T, Kanazawa Y, Hayashi N. Viral covalently closed circular DNA in a non-transgenic mouse model for chronic hepatitis B virus replication. *J Hepatol* 2006; **44**: 267-274

20   **Zhu Y**, Yamamoto T, Cullen J, Saputelli J, Aldrich CE, Miller DS, Litwin S, Furman PA, Jilbert AR, Mason WS. Kinetics of hepadnavirus loss from the liver during inhibition of viral DNA synthesis. *J Virol* 2001; **75**: 311-322

21   **Bock CT**, Schwinn S, Locarnini S, Fyfe J, Manns MP, Trautwein C, Zentgraf H. Structural organization of the hepatitis B virus minichromosome. *J Mol Biol* 2001; **307**: 183-196

22   **Tuttleman JS**, Pourcel C, Summers J. Formation of the pool of covalently closed circular viral DNA in hepadnavirus-infected cells. *Cell* 1986; **47**: 451-460

23   **Wu TT**, Coates L, Aldrich CE, Summers J, Mason WS. In hepatocytes infected with duck hepatitis B virus, the template for viral RNA synthesis is amplified by an intracellular pathway. *Virology* 1990; **175**: 255-261

24   **Addison WR**, Walters KA, Wong WW, Wilson JS, Madej D, Jewell LD, Tyrrell DL. Half-life of the duck hepatitis B virus covalently closed circular DNA pool in vivo following inhibition of viral replication. *J Virol* 2002; **76**: 6356-6363

25   **Ganem D**, Prince AM. Hepatitis B virus infection--natural history and clinical consequences. *N Engl J Med* 2004; **350**: 1118-1129

26   **Summers J**, Smith PM, Horwich AL. Hepadnavirus envelope proteins regulate covalently closed circular DNA amplification. *J Virol* 1990; **64**: 2819-2824

27   **Zhang YY**, Zhang BH, Theele D, Litwin S, Toll E, Summers J. Single-cell analysis of covalently closed circular DNA copy numbers in a hepadnavirus-infected liver. *Proc Natl Acad Sci USA* 2003; **100**: 12372-12377

28   **Zhang YY**, Theele DP, Summers J. Age-related differences in amplification of covalently closed circular DNA at early times after duck hepatitis B virus infection of ducks. *J Virol* 2005; **79**: 9896-9903

29   **Le Mire MF**, Miller DS, Foster WK, Burrell CJ, Jilbert AR. Covalently closed circular DNA is the predominant form of duck hepatitis B virus DNA that persists following transient infection. *J Virol* 2005; **79**: 12242-12252

30   **Zoulim F**. Antiviral therapy of chronic hepatitis B. *Antiviral Res* 2006; **71**: 206-215

31   **Zoulim F**. Entecavir: a new treatment option for chronic hepatitis B. *J Clin Virol* 2006; **36**: 8-12

32   **Terré S**, Petit MA, Bréchot C. Defective hepatitis B virus particles are generated by packaging and reverse transcription of spliced viral RNAs in vivo. *J Virol* 1991; **65**: 5539-5543

33   **Sommer G**, van Bömmel F, Will H. Genotype-specific synthesis and secretion of spliced hepatitis B virus genomes in hepatoma cells. *Virology* 2000; **271**: 371-381

34   **Obert S**, Zachmann-Brand B, Deindl E, Tucker W, Bartenschlager R, Schaller H. A splice hepadnavirus RNA that is essential for virus replication. *EMBO J* 1996; **15**: 2565-2574

35   **Loeb DD**, Mack AA, Tian R. A secondary structure that contains the 5′ and 3′ splice sites suppresses splicing of duck hepatitis B virus pregenomic RNA. *J Virol* 2002; **76**: 10195-10202

36   **Heise T**, Sommer G, Reumann K, Meyer I, Will H, Schaal H. The hepatitis B virus PRE contains a splicing regulatory element. *Nucleic Acids Res* 2006; **34**: 353-363

37   **Sen N**, Cao F, Tavis JE. Translation of duck hepatitis B virus reverse transcriptase by ribosomal shunting. *J Virol* 2004; **78**: 11751-11757

38   **Nassal M**, Junker-Niepmann M, Schaller H. Translational inactivation of RNA function: discrimination against a subset of genomic transcripts during HBV nucleocapsid assembly. *Cell* 1990; **63**: 1357-1363

39   **Gnatt AL**, Cramer P, Fu J, Bushnell DA, Kornberg RD. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. *Science* 2001; **292**: 1876-1882

40   **Nesser NK**, Peterson DO, Hawley DK. RNA polymerase II subunit Rpb9 is important for transcriptional fidelity in vivo. *Proc Natl Acad Sci USA* 2006; **103**: 3268-3273

41   **Cao F**, Tavis JE. Suppression of mRNA accumulation by the duck hepatitis B virus reverse transcriptase. *Virology* 2006; **350**: 475-483

42   **Delelis O**, Lehmann-Che J, Saïb A. Foamy viruses--a world apart. *Curr Opin Microbiol* 2004; **7**: 400-406

43   **Fassati A**, Goff SP. Characterization of intracellular reverse transcription complexes of human immunodeficiency virus type 1. *J Virol* 2001; **75**: 3626-3635

44   **Iordanskiy S**, Berro R, Altieri M, Kashanchi F, Bukrinsky M. Intracytoplasmic maturation of the human immunodeficiency virus type 1 reverse transcription complexes determines their capacity to integrate into chromatin. *Retrovirology* 2006; **3**: 4

45   **Ahlquist P**. Parallels among positive-strand RNA viruses, reverse-transcribing viruses and double-stranded RNA viruses. *Nat Rev Microbiol* 2006; **4**: 371-382

46   **Köck J**, Nassal M, Deres K, Blum HE, von Weizsäcker F. Hepatitis B virus nucleocapsids formed by carboxy-terminally mutated core proteins contain spliced viral genomes but lack full-size DNA. *J Virol* 2004; **78**: 13812-13818

47   **Le Pogam S**, Chua PK, Newman M, Shih C. Exposure of RNA templates and encapsidation of spliced viral RNA are influenced by the arginine-rich domain of human hepatitis B virus core antigen (HBcAg 165-173). *J Virol* 2005; **79**: 1871-1887

48   **Perlman DH**, Berg EA, O'connor PB, Costello CE, Hu J. Reverse transcription-associated dephosphorylation of hepadnavirus nucleocapsids. *Proc Natl Acad Sci USA* 2005; **102**: 9020-9025

49   **Melegari M**, Wolf SK, Schneider RJ. Hepatitis B virus DNA replication is coordinated by core protein serine phosphorylation and HBx expression. *J Virol* 2005; **79**: 9810-9820

50   **Roseman AM**, Berriman JA, Wynne SA, Butler PJ, Crowther RA. A structural model for maturation of the hepatitis B virus core. *Proc Natl Acad Sci USA* 2005; **102**: 15821-15826

51   **Böttcher B**, Vogel M, Ploss M, Nassal M. High plasticity of the hepatitis B virus capsid revealed by conformational stress. *J Mol Biol* 2006; **356**: 812-822

52   **Hu J**, Toft D, Anselmo D, Wang X. In vitro reconstitution of functional hepadnavirus reverse transcriptase with cellular chaperone proteins. *J Virol* 2002; **76**: 269-279

53   **Beck J**, Nassal M. Reconstitution of a functional duck hepatitis B virus replication initiation complex from separate reverse transcriptase domains expressed in Escherichia coli. *J Virol* 2001; **75**: 7410-7419

54   **Beck J**, Nassal M. Efficient Hsp90-independent in vitro activation by Hsc70 and Hsp40 of duck hepatitis B virus reverse transcriptase, an assumed Hsp90 client protein. *J Biol Chem* 2003; **278**: 36128-36138

55   **Junker-Niepmann M**, Bartenschlager R, Schaller H. A short cis-acting sequence is required for hepatitis B virus pregenome encapsidation and sufficient for packaging of foreign RNA. *EMBO J* 1990; **9**: 3389-3396

56   **Wang GH**, Seeger C. Novel mechanism for reverse transcription in hepatitis B viruses. *J Virol* 1993; **67**: 6507-6512

57   **Tavis JE**, Perri S, Ganem D. Hepadnavirus reverse transcription initiates within the stem-loop of the RNA packaging signal and employs a novel strand transfer. *J Virol* 1994; **68**: 3536-3543

58   **Nassal M**, Rieger A. A bulged region of the hepatitis B virus RNA encapsidation signal contains the replication origin for discontinuous first-strand DNA synthesis. *J Virol* 1996; **70**: 2764-2773

59   **Knaus T**, Nassal M. The encapsidation signal on the hepatitis B virus RNA pregenome forms a stem-loop structure that is critical for its function. *Nucleic Acids Res* 1993; **21**: 3967-3975

60   **Pollack JR**, Ganem D. An RNA stem-loop structure directs hepatitis B virus genomic RNA encapsidation. *J Virol* 1993; **67**: 3254-3263

61   **Pollack JR**, Ganem D. Site-specific RNA binding by a hepatitis B virus reverse transcriptase initiates two distinct reactions: RNA packaging and DNA synthesis. *J Virol* 1994; **68**: 5579-5587

62   **Rieger A**, Nassal M. Distinct requirements for primary sequence in the 5′- and 3′-part of a bulge in the hepatitis B virus RNA encapsidation signal revealed by a combined in vivo selection/in vitro amplification system. *Nucleic Acids Res* 1995; **23**: 3909-3915

63   **Fallows DA**, Goff SP. Mutations in the epsilon sequences of human hepatitis B virus affect both RNA encapsidation and reverse transcription. *J Virol* 1995; **69**: 3067-3073

64   **Laskus T**, Rakela J, Persing DH. The stem-loop structure of the cis-encapsidation signal is highly conserved in naturally occurring hepatitis B virus variants. *Virology* 1994; **200**: 809-812

65   **Lok AS**, Akarca U, Greene S. Mutations in the pre-core region of hepatitis B virus serve to enhance the stability of the secondary structure of the pre-genome encapsidation signal. *Proc Natl Acad Sci USA* 1994; **91**: 4077-4081

66   **Kramvis A**, Kew MC. Structure and function of the encapsidation signal of hepadnaviridae. *J Viral Hepat* 1998; **5**: 357-367

67   **Guarnieri M**, Kim KH, Bang G, Li J, Zhou Y, Tang X, Wands J, Tong S. Point mutations upstream of hepatitis B virus core gene affect DNA replication at the step of core protein expression. *J Virol* 2006; **80**: 587-595

68   **Flodell S**, Schleucher J, Cromsigt J, Ippel H, Kidd-Ljunggren K, Wijmenga S. The apical stem-loop of the hepatitis B virus encapsidation signal folds into a stable tri-loop with two underlying pyrimidine bulges. *Nucleic Acids Res* 2002; **30**: 4803-4811

69   **Beck J**, Bartos H, Nassal M. Experimental confirmation of a hepatitis B virus (HBV) epsilon-like bulge-and-loop structure in avian HBV RNA encapsidation signals. *Virology* 1997; **227**: 500-504

70   **Beck J**, Nassal M. Sequence- and structure-specific determinants in the interaction between the RNA encapsidation signal and reverse transcriptase of avian hepatitis B viruses. *J Virol* 1997; **71**: 4971-4980

71   **Hu K**, Beck J, Nassal M. SELEX-derived aptamers of the duck hepatitis B virus RNA encapsidation signal distinguish critical

and non-critical residues for productive initiation of reverse transcription. *Nucleic Acids Res* 2004; **32**: 4377-4389

72 **Hu J**, Boyer M. Hepatitis B virus reverse transcriptase and epsilon RNA sequences required for specific interaction in vitro. *J Virol* 2006; **80**: 2141-2150

73 **Bartenschlager R**, Junker-Niepmann M, Schaller H. The P gene product of hepatitis B virus is required as a structural component for genomic RNA encapsidation. *J Virol* 1990; **64**: 5324-5332

74 **Calvert J**, Summers J. Two regions of an avian hepadnavirus RNA pregenome are required in cis for encapsidation. *J Virol* 1994; **68**: 2084-2090

75 **Ostrow KM**, Loeb DD. Characterization of the cis-acting contributions to avian hepadnavirus RNA encapsidation. *J Virol* 2002; **76**: 9087-9095

76 **Ostrow KM**, Loeb DD. Chimeras of duck and heron hepatitis B viruses provide evidence for functional interactions between viral components of pregenomic RNA encapsidation. *J Virol* 2004; **78**: 8780-8787

77 **Rieger A**, Nassal M. Specific hepatitis B virus minus-strand DNA synthesis requires only the 5′ encapsidation signal and the 3′-proximal direct repeat DR1. *J Virol* 1996; **70**: 585-589

78 **Wang GH**, Seeger C. The reverse transcriptase of hepatitis B virus acts as a protein primer for viral DNA synthesis. *Cell* 1992; **71**: 663-670

79 **Jeong JK**, Yoon GS, Ryu WS. Evidence that the 5′-end cap structure is essential for encapsidation of hepatitis B virus pregenomic RNA. *J Virol* 2000; **74**: 5502-5508

80 **Mangus DA**, Evans MC, Jacobson A. Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol* 2003; **4**: 223

81 **Wagner H**, Bauer S. All is not Toll: new pathways in DNA recognition. *J Exp Med* 2006; **203**: 265-268

82 **Meylan E**, Tschopp J. Toll-like receptors and RNA helicases: two parallel ways to trigger antiviral responses. *Mol Cell* 2006; **22**: 561-569

83 **Cullen BR**. Transcription and processing of human microRNA precursors. *Mol Cell* 2004; **16**: 861-865

84 **Perri S**, Ganem D. A host factor that binds near the termini of hepatitis B virus pregenomic RNA. *J Virol* 1996; **70**: 6803-6809

85 **Kreft SG**, Nassal M. hRUL138, a novel human RNA-binding RING-H2 ubiquitin-protein ligase. *J Cell Sci* 2003; **116**: 605-616

86 **Bartenschlager R**, Schaller H. Hepadnaviral assembly is initiated by polymerase binding to the encapsidation signal in the viral RNA genome. *EMBO J* 1992; **11**: 3413-3420

87 **Yao E**, Schaller H, Tavis JE. The duck hepatitis B virus polymerase and core proteins accumulate in different patterns from their common mRNA. *Virology* 2003; **311**: 81-88

88 **Yao E**, Tavis JE. Kinetics of synthesis and turnover of the duck hepatitis B virus reverse transcriptase. *J Biol Chem* 2003; **278**: 1201-1205

89 **Cao F**, Tavis JE. Detection and characterization of cytoplasmic hepatitis B virus reverse transcriptase. *J Gen Virol* 2004; **85**: 3353-3360

90 **Xiong Y**, Eickbush TH. Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J* 1990; **9**: 3353-3362

91 **Radziwill G**, Tucker W, Schaller H. Mutational analysis of the hepatitis B virus P gene product: domain structure and RNase H activity. *J Virol* 1990; **64**: 613-620

92 **Bartenschlager R**, Schaller H. The amino-terminal domain of the hepadnaviral P-gene encodes the terminal protein (genome-linked protein) believed to prime reverse transcription. *EMBO J* 1988; **7**: 4185-4192

93 **Weber M**, Bronsema V, Bartos H, Bosserhoff A, Bartenschlager R, Schaller H. Hepadnavirus P protein utilizes a tyrosine residue in the TP domain to prime reverse transcription. *J Virol* 1994; **68**: 2994-2999

94 **Zoulim F**, Seeger C. Reverse transcription in hepatitis B viruses is primed by a tyrosine residue of the polymerase. *J Virol* 1994; **68**: 6-13

95 **Lanford RE**, Notvall L, Lee H, Beames B. Transcomplementation of nucleotide priming and reverse transcription between independently expressed TP and RT domains of the hepatitis B virus reverse transcriptase. *J Virol* 1997; **71**: 2996-3004

96 **Das K**, Xiong X, Yang H, Westland CE, Gibbs CS, Sarafianos SG, Arnold E. Molecular modeling and biochemical characterization reveal the mechanism of hepatitis B virus polymerase resistance to lamivudine (3TC) and emtricitabine (FTC). *J Virol* 2001; **75**: 4771-4779

97 **Allen MI**, Deslauriers M, Andrews CW, Tipples GA, Walters KA, Tyrrell DL, Brown N, Condreay LD. Identification and characterization of mutations in hepatitis B virus resistant to lamivudine. Lamivudine Clinical Investigation Group. *Hepatology* 1998; **27**: 1670-1677

98 **Beck J**, Vogel M, Nassal M. dNTP versus NTP discrimination by phenylalanine 451 in duck hepatitis B virus P protein indicates a common structure of the dNTP-binding pocket with other reverse transcriptases. *Nucleic Acids Res* 2002; **30**: 1679-1687

99 **Ferrer-Orta C**, Arias A, Escarmís C, Verdaguer N. A comparison of viral RNA-dependent RNA polymerases. *Curr Opin Struct Biol* 2006; **16**: 27-34

100 **de Jong RN**, van der Vliet PC, Brenkman AB. Adenovirus DNA replication: protein priming, jumping back and the role of the DNA binding protein DBP. *Curr Top Microbiol Immunol* 2003; **272**: 187-211

101 **Kamtekar S**, Berman AJ, Wang J, Lázaro JM, de Vega M, Blanco L, Salas M, Steitz TA. The phi29 DNA polymerase: protein-primer structure suggests a model for the initiation to elongation transition. *EMBO J* 2006; **25**: 1335-1343

102 **Hu J**, Anselmo D. In vitro reconstitution of a functional duck hepatitis B virus reverse transcriptase: posttranslational activation by Hsp90. *J Virol* 2000; **74**: 11447-11455

103 **Beck J**, Nassal M. In vitro reconstitution of epsilon-dependent duck hepatitis B virus replication initiation. *Methods Mol Med* 2004; **95**: 315-325

104 **Cao F**, Badtke MP, Metzger LM, Yao E, Adeyemo B, Gong Y, Tavis JE. Identification of an essential molecular contact point on the duck hepatitis B virus reverse transcriptase. *J Virol* 2005; **79**: 10164-10170

105 **Lanford RE**, Kim YH, Lee H, Notvall L, Beames B. Mapping of the hepatitis B virus reverse transcriptase TP and RT domains by transcomplementation for nucleotide priming and by protein-protein interaction. *J Virol* 1999; **73**: 1885-1893

106 **Chen Y**, Varani G. Protein families and RNA recognition. *FEBS J* 2005; **272**: 2088-2097

107 **Schaaf SG**, Beck J, Nassal M. A small 2′-OH- and base-dependent recognition element downstream of the initiation site in the RNA encapsidation signal is essential for hepatitis B virus replication initiation. *J Biol Chem* 1999; **274**: 37787-37794

108 **Tavis JE**, Ganem D. Evidence for activation of the hepatitis B virus polymerase by binding of its RNA template. *J Virol* 1996; **70**: 5741-5750

109 **Beck J**, Nassal M. Formation of a functional hepatitis B virus replication initiation complex involves a major structural alteration in the RNA template. *Mol Cell Biol* 1998; **18**: 6265-6272

110 **Tavis JE**, Massey B, Gong Y. The duck hepatitis B virus polymerase is activated by its RNA packaging signal, epsilon. *J Virol* 1998; **72**: 5789-5796

111 **Lanford RE**, Notvall L, Beames B. Nucleotide priming and reverse transcriptase activity of hepatitis B virus polymerase expressed in insect cells. *J Virol* 1995; **69**: 4431-4439

112 **Seifer M**, Hamatake R, Bifano M, Standring DN. Generation of replication-competent hepatitis B virus nucleocapsids in insect cells. *J Virol* 1998; **72**: 2765-2776

113 **Hu J**, Seeger C. Hsp90 is required for the activity of a hepatitis B virus reverse transcriptase. *Proc Natl Acad Sci USA* 1996; **93**: 1060-1064

114 **Tang YC**, Chang HC, Roeben A, Wischnewski D, Wischnewski

N, Kerner MJ, Hartl FU, Hayer-Hartl M. Structural features of the GroEL-GroES nano-cage required for rapid folding of encapsulated protein. *Cell* 2006; **125**: 903-914

115   **Hartl FU**, Hayer-Hartl M. Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* 2002; **295**: 1852-1858

116   **Young JC**, Agashe VR, Siegers K, Hartl FU. Pathways of chaperone-mediated protein folding in the cytosol. *Nat Rev Mol Cell Biol* 2004; **5**: 781-791

117   **Mayer MP**, Bukau B. Hsp70 chaperones: cellular functions and molecular mechanism. *Cell Mol Life Sci* 2005; **62**: 670-684

118   **Young JC**, Moarefi I, Hartl FU. Hsp90: a specialized but essential protein-folding tool. *J Cell Biol* 2001; **154**: 267-273

119   **Pratt WB**, Galigniana MD, Morishima Y, Murphy PJ. Role of molecular chaperones in steroid receptor action. *Essays Biochem* 2004; **40**: 41-58

120   **Hu J**, Toft DO, Seeger C. Hepadnavirus assembly and reverse transcription require a multi-component chaperone complex which is incorporated into nucleocapsids. *EMBO J* 1997; **16**: 59-68

121   **Guisbert E**, Herman C, Lu CZ, Gross CA. A chaperone network controls the heat shock response in E. coli. *Genes Dev* 2004; **18**: 2812-2821

122   **Takaya A**, Matsui M, Tomoyasu T, Kaya M, Yamamoto T. The DnaK chaperone machinery converts the native FlhD2C2 hetero-tetramer into a functional transcriptional regulator of flagellar regulon expression in Salmonella. *Mol Microbiol* 2006; **59**: 1327-1340

123   **Hennessy F**, Nicoll WS, Zimmermann R, Cheetham ME, Blatch GL. Not all J domains are created equal: implications for the specificity of Hsp40-Hsp70 interactions. *Protein Sci* 2005; **14**: 1697-1709

124   **Brehmer D**, Rüdiger S, Gässler CS, Klostermeier D, Packschies L, Reinstein J, Mayer MP, Bukau B. Tuning of chaperone activity of Hsp70 proteins by modulation of nucleotide exchange. *Nat Struct Biol* 2001; **8**: 427-432

125   **Walsh P**, Bursać D, Law YC, Cyr D, Lithgow T. The J-protein family: modulating protein assembly, disassembly and translocation. *EMBO Rep* 2004; **5**: 567-571

126   **Wang X**, Qian X, Guo HC, Hu J. Heat shock protein 90-independent activation of truncated hepadnavirus reverse transcriptase. *J Virol* 2003; **77**: 4471-4480

127   **Cho G**, Suh SW, Jung G. HBV polymerase interacts independently with N-terminal and C-terminal fragments of Hsp90beta. *Biochem Biophys Res Commun* 2000; **274**: 203-211

128   **Hu J**, Flores D, Toft D, Wang X, Nguyen D. Requirement of heat shock protein 90 for human hepatitis B virus reverse transcriptase function. *J Virol* 2004; **78**: 13122-13131

129   **Bracher A**, Hartl FU. Hsp90 structure: when two ends meet. *Nat Struct Mol Biol* 2006; **13**: 478-480

130   **Pratt WB**, Toft DO. Regulation of signaling protein function and trafficking by the hsp90/hsp70-based chaperone machinery. *Exp Biol Med (Maywood)* 2003; **228**: 111-133

131   **Workman P**. Altered states: selectively drugging the Hsp90 cancer chaperone. *Trends Mol Med* 2004; **10**: 47-51

132   **Kamal A**, Thao L, Sensintaffar J, Zhang L, Boehm MF, Fritz LC, Burrows FJ. A high-affinity conformation of Hsp90 confers tumour selectivity on Hsp90 inhibitors. *Nature* 2003; **425**: 407-410

133   **Park SG**, Jung G. Human hepatitis B virus polymerase interacts with the molecular chaperonin Hsp60. *J Virol* 2001; **75**: 6962-6968

134   **Park SG**, Lim SO, Jung G. Binding site analysis of human HBV pol for molecular chaperonin, hsp60. *Virology* 2002; **298**: 116-123

135   **Park SG**, Lee SM, Jung G. Antisense oligodeoxynucleotides targeted against molecular chaperonin Hsp60 block human hepatitis B virus replication. *J Biol Chem* 2003; **278**: 39851-39857

136   **Kim SS**, Shin HJ, Cho YH, Rho HM. Expression of stable hepatitis B viral polymerase associated with GRP94 in E. coli. *Arch Virol* 2000; **145**: 1305-1320

137   **Wang X**, Grammatikakis N, Hu J. Role of p50/CDC37 in hepadnavirus assembly and replication. *J Biol Chem* 2002; **277**: 24361-24367

138   **Holt SE**, Aisner DL, Baur J, Tesmer VM, Dy M, Ouellette M, Trager JB, Morin GB, Toft DO, Shay JW, Wright WE, White MA. Functional requirement of p23 and Hsp90 in telomerase complexes. *Genes Dev* 1999; **13**: 817-826

139   **Keppler BR**, Grady AT, Jarstfer MB. The biochemical role of the heat shock protein 90 chaperone complex in establishing human telomerase activity. *J Biol Chem* 2006; **281**: 19840-19848

140   **Burch AD**, Weller SK. Herpes simplex virus type 1 DNA polymerase requires the mammalian chaperone hsp90 for proper localization to the nucleus. *J Virol* 2005; **79**: 10740-10749

141   **Kampmueller KM**, Miller DJ. The cellular chaperone heat shock protein 90 facilitates Flock House virus RNA replication in Drosophila cells. *J Virol* 2005; **79**: 6827-6837

142   **Momose F**, Naito T, Yano K, Sugimoto S, Morikawa Y, Nagata K. Identification of Hsp90 as a stimulatory host factor involved in influenza virus RNA synthesis. *J Biol Chem* 2002; **277**: 45306-45314

143   **Deng T**, Sharps J, Fodor E, Brownlee GG. In vitro assembly of PB2 with a PB1-PA dimer supports a new model of assembly of influenza A virus polymerase subunits into a functional trimeric complex. *J Virol* 2005; **79**: 8669-8674

144   **Liu N**, Ji L, Maguire ML, Loeb DD. cis-Acting sequences that contribute to the synthesis of relaxed-circular DNA of human hepatitis B virus. *J Virol* 2004; **78**: 642-649

145   **Shin MK**, Lee J, Ryu WS. A novel cis-acting element facilitates minus-strand DNA synthesis during reverse transcription of the hepatitis B virus genome. *J Virol* 2004; **78**: 6252-6262

146   **Tang H**, McLachlan A. A pregenomic RNA sequence adjacent to DR1 and complementary to epsilon influences hepatitis B virus replication efficiency. *Virology* 2002; **303**: 199-210

147   **Abraham TM**, Loeb DD. Base pairing between the 5′ half of epsilon and a cis-acting sequence, phi, makes a contribution to the synthesis of minus-strand DNA for human hepatitis B virus. *J Virol* 2006; **80**: 4380-4387

148   **Harris RS**, Liddament MT. Retroviral restriction by APOBEC proteins. *Nat Rev Immunol* 2004; **4**: 868-877

149   **Turelli P**, Mangeat B, Jost S, Vianin S, Trono D. Inhibition of hepatitis B virus replication by APOBEC3G. *Science* 2004; **303**: 1829

150   **Rösler C**, Köck J, Kann M, Malim MH, Blum HE, Baumert TF, von Weizsäcker F. APOBEC-mediated interference with hepadnavirus production. *Hepatology* 2005; **42**: 301-309

151   **Bonvin M**, Achermann F, Greeve I, Stroka D, Keogh A, Inderbitzin D, Candinas D, Sommer P, Wain-Hobson S, Vartanian JP, Greeve J. Interferon-inducible expression of APOBEC3 editing enzymes in human hepatocytes and inhibition of hepatitis B virus replication. *Hepatology* 2006; **43**: 1364-1374

152   **Ostrow KM**, Loeb DD. Underrepresentation of the 3′ region of the capsid pregenomic RNA of duck hepatitis B virus. *J Virol* 2004; **78**: 2179-2186

153   **Loeb DD**, Hirsch RC, 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* 1991; **10**: 3533-3540

154   **Staprans S**, Loeb DD, Ganem D. Mutations affecting hepadnavirus plus-strand DNA synthesis dissociate primer cleavage from translocation and reveal the origin of linear viral DNA. *J Virol* 1991; **65**: 1255-1262

155   **Bill CA**, Summers J. Genomic DNA double-strand breaks are targets for hepadnaviral DNA integration. *Proc Natl Acad Sci USA* 2004; **101**: 11135-11140

156   **Habig JW**, Loeb DD. The conformation of the 3′ end of the minus-strand DNA makes multiple contributions to template switches during plus-strand DNA synthesis of duck hepatitis B virus. *J Virol* 2003; **77**: 12401-12411

157   **Havert MB**, Loeb DD. cis-Acting sequences in addition to donor and acceptor sites are required for template switching during synthesis of plus-strand DNA for duck hepatitis B virus. *J Virol* 1997; **71**: 5336-5344

158 **Havert MB**, Ji L, Loeb DD. Analysis of duck hepatitis B virus reverse transcription indicates a common mechanism for the two template switches during plus-strand DNA synthesis. *J Virol* 2002; **76**: 2763-2769

159 **Habig JW,** Loeb DD. Template switches during plus-strand DNA synthesis of duck hepatitis B virus are influenced by the base composition of the minus-strand terminal redundancy. *J Virol* 2003; **77**: 12412-12420

160 **Liu N**, Tian R, Loeb DD. Base pairing among three cis-acting sequences contributes to template switching during hepadnavirus reverse transcription. *Proc Natl Acad Sci USA* 2003; **100**: 1984-1989

161 **Mueller-Hill K**, Loeb DD. cis-Acting sequences 5E, M, and 3E interact to contribute to primer translocation and circularization during reverse transcription of avian hepadnavirus DNA. *J Virol* 2002; **76**: 4260-4266

162 **Lee J**, Shin MK, Lee HJ, Yoon G, Ryu WS. Three novel cis-acting elements required for efficient plus-strand DNA synthesis of the hepatitis B virus genome. *J Virol* 2004; **78**: 7455-7464

163 **Wynne SA**, Crowther RA, Leslie AG. The crystal structure of the human hepatitis B virus capsid. *Mol Cell* 1999; **3**: 771-780