



Mutational profiling of the variability of individual amino acid positions in the hepatitis B virus matrix domain

Beate Schittl, Volker Bruss*

Institute of Virology, Helmholtz Zentrum München, Ingolstädter Landstraße 1, D-85764 Neuherberg, Germany

ARTICLE INFO

Article history:

Received 23 January 2014

Returned to author for revisions

12 February 2014

Accepted 22 April 2014

Available online 14 May 2014

Keywords:

Hepatitis B virus

Envelopment

Large surface protein

Matrix domain

ABSTRACT

The hepatitis B virus (HBV) is formed by budding. A stretch of 22 amino acids (aa) (matrix domain, MD, R103 – S124) in the large envelope protein L is crucial for virion formation and probably establishes contact to the nucleocapsid. Here, we assess the impact of sequence variations at numerous individual aa positions within the MD on virion formation. We generated panels of L mutants covering all 19 possible aa for 11 positions and tested the capacity of these mutants to rescue virus production by an L-defective HBV genome. At four positions (L112, R113, P117, W122), any replacement of the wild type (WT) aa reduced virus assembly to undetectable levels. Virus production was strongly diminished by substitutions at five other positions (R103, T106, S115, H116, A119). Only two tested positions (D114, Q118) tolerated several substitutions. The restricted positions may represent promising targets for the development of novel antiviral strategies.

© 2014 Elsevier Inc. All rights reserved.

Introduction

The hepatitis B virus (HBV) is an enveloped DNA virus and the prototype of the family *Hepadnaviridae*. Despite the existence of an efficient vaccine, this virus still chronically infects 360 million people worldwide and causes high morbidity and mortality (Chan and Sung, 2006). The therapy is limited to immune stimulation by interferon and long-term treatment with nucleoside/nucleotide analogs that block the viral reverse transcription. However, the success of the therapy is compromised by the appearance of resistant virus variants (Gish et al., 2012). Alternative strategies for inhibiting virus replication are desirable to expand therapeutic options. Since HBV encodes only one protein with enzymatic activity, the P protein with reverse transcriptase/DNA polymerase and RNaseH activity, protein–protein interactions during cell entry or virion assembly may represent interesting targets for therapeutic intervention.

During viral replication (Beck and Nassal, 2007), an icosahedral nucleocapsid with a diameter of 30 nm is formed in the cytosol of the infected cell. The nucleocapsid is enveloped during budding through an intracellular membrane containing the three viral membrane proteins L (large), M (middle), and S (small membrane protein) (Bruss, 2007). The L and the S proteins are required for virion formation, while the M protein is dispensable. The three envelope proteins are translation products of a single unspliced

open reading frame. The L protein is encoded by the entire reading frame, while the translation of the M and S proteins starts from two internal initiation sites at codon 120 and 175, respectively (for genotype A used in this work). The 119 aa long region unique to L is referred to as the preS1 domain, the 55 aa long stretch present in L and M is named preS2, and the 226 aa long sequence of S is present in all three proteins.

The HBV envelope proteins also form subviral lipoprotein particles (SVP) of spherical or filamentous shape with a diameter of 22 nm without the participation of the capsid. These particles appear in the luminal compartment of the secretory organelles and are released in large amounts from infected cells. The S domains traverse the membrane four times with the N and C termini oriented away from the cytosol. One distinctive feature of the L protein is that it forms two different transmembrane topologies: in the e-preS form its preS domains are located at the luminal side of the internal membrane and in the i-preS form the preS domains are at the cytoplasmic side (in i-preS, probably, the first hydrophobic domain in the S part is not traversing the membrane) (Bruss et al., 1994; Ostapchuk et al., 1994; Prange and Strecker, 1995).

Several experimental observations are consistent with the model that a stretch of 22 aa between R103 and S124 of the i-preS form of L, referred to as the matrix domain (MD), contacts the HBV nucleocapsid during envelopment. (i) The formation of the i-preS conformation can be prevented by fusing an N-terminal secretion signal to L which cotranslationally translocates preS across the membrane of the endoplasmic reticulum. This alters the i-preS conformation of L to its e-preS conformation and blocks

* Corresponding author. Tel.: +49 89 3187 3021; fax: +49 89 3187 3329.

E-mail address: volker.bruss@helmholtz-muenchen.de (V. Bruss).

nucleocapsid envelopment (Bruss and Vieluf, 1995). (ii) N-terminal truncations of L up to G102 were compatible with nucleocapsid envelopment and release of virions (Bruss and Thomssen, 1994). (iii) Substitutions of stretches of 4–8 aa between aa 125 and aa 171 (Bruss, 1997) as well as deletions between aa 132 and aa 234 allowed the release of enveloped nucleocapsids (Kluge et al., 2005; Ni et al., 2010). (iv) Six out of seven double point mutations in the MD strongly blocked virion formation (Bruss, 1997).

Candidates for MD binding sites on the capsid have been mapped to the crystal structure of the particle (Wynne et al., 1999) by screening point mutations of the capsid protein C (Pairan and Bruss, 2009; Ponsel and Bruss, 2003), the single viral protein forming the isosahedral capsid with $T=4$ or $T=3$ symmetry. Mutations in two narrow areas at the base of the 2.5 nm long spikes protruding from the capsid block capsid envelopment. It is conceivable that the MD reaches down to one or both of these sites and mediates the contact between capsid and envelope. In favor of this view is that the distance between the MD and the membrane must not be shorter than 26 aa (Kluge et al., 2005).

We asked what degree of sequence variability is possible in individual amino acid positions of the MD without disturbing capsid envelopment. To approach this question we mutagenized 11 individual positions of the MD to all 19 non-WT aa and tested pools of mutants for complementation of an L negative HBV genome. We found that some residues were totally invariable: replacement of the WT aa by any other aa at these positions reduced virion formation to undetectable levels.

Results

Virion quantification in the supernatant of transiently transfected cells

We wanted to examine the effect of mutations in the matrix domain of the large HBV envelope protein on virion formation by expressing the virus in Huh7 cells transiently transfected with an HBV genomic plasmid construct (Fig. 1B). For quantitative analysis of hepatitis B virions in the supernatant of transfected cells, we measured viral genomic DNA by quantitative polymerase chain reaction (qPCR). However, such a measurement is not trivial because the plasmid used for transfection usually produces a strong background. In addition the cells release substantial amounts of naked capsids containing viral genomes which also disturb the determination of virion titers by qPCR.

To ensure that the readout of the qPCR corresponds to the amount of viral genomes in virions we combined three procedures as described in the Material and methods section (virus quantification assay): (i) to exclude naked capsids, virions were immunoprecipitated from the culture supernatant with antibodies against the major envelope proteins S (anti-HBs) prior to viral DNA preparation. To control the effectiveness of this step a sample from cells transfected only with plasmid pSVHBV1.1LE⁻ carrying the envelope-negative HBV genome was processed in parallel. These cells produced no virions but naked capsids in amounts comparable to cells expressing the WT viral genome (data not shown). (ii) Before viral DNA preparation, the immunoprecipitate was subjected to DNase treatment to remove unpackaged DNA. (iii) Primers (forward primer nt 1784 – nt 1805, reverse primer nt 2005 – nt 1985) were used that bind to regions only 200 bp apart in the original viral genome (Fig. 1A), while the binding sites for these primers are separated by 3.5 kb in the genomic HBV plasmid pSVHBV1.1LE⁻ used for transfection (Fig. 1B). The qPCR was performed with very short extension times to favor the generation of products initiated from viral genomes over products generated from the transfected plasmids.

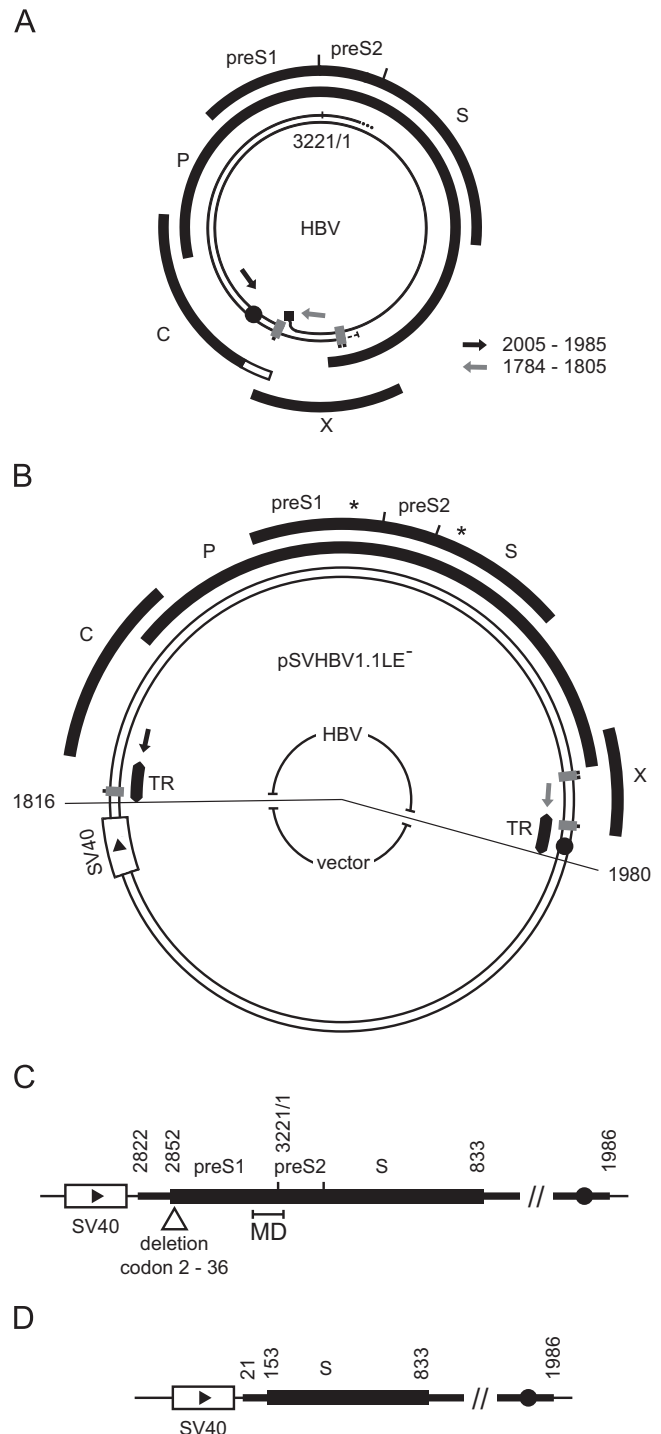


Fig. 1. Plasmid maps. (A) Map of the HBV genome. The inner two circles represent the complete (–) strand covalently linked at its 5' end to the P protein (black square) and the incomplete (+) strand, respectively. The black disk labels the position of the polyA signal. The gray rectangles together with one or two dots denote direct repeats I and II, respectively. The two arrows show the positions of the sense (gray) and antisense (black) primers used for qPCR. The outer circle segments represent open reading frames. (B) Map of plasmid pSVHBV1.1LE⁻. The plasmid carries a 1.1 mer terminal redundant (TR) copy of the HBV genome. Transcription of the pregenomic RNA is initiated at the SV40 early promoter (SV40). The sites for hybridization of the two primers shown in panel (A) are indicated. The two asterisks indicate the position of two nonsense mutations in the preS1 and S open reading frame, respectively, which are both silent in the P reading frame. (C) Expression vector for the synthesis of all three HBV envelope proteins. Codons 2 to 36 of preS1 were deleted (triangle). (D) Expression vector for the synthesis of the S protein. All numbers refer to nucleotides in the HBV genome sequence.

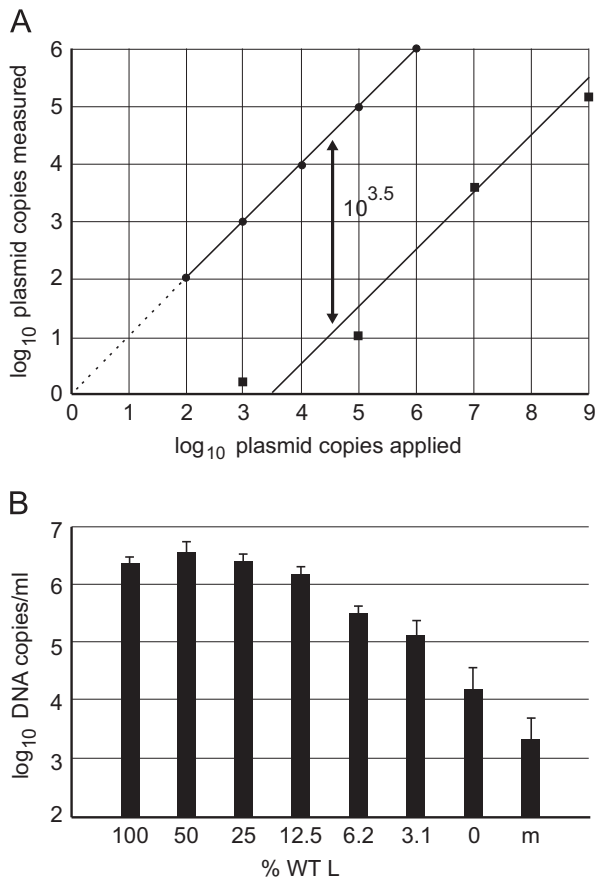


Fig. 2. Virion detection in the supernatant of transiently transfected cells. (A) Different amounts of plasmid pHBV1.5 carrying a continuous stretch of the HBV sequence 1784–2005 (dots) and of plasmid pSVHBV1.1LE⁻ (squares) as indicated on the abscissa were quantified using the qPCR described in [Materials and methods](#) section. Plasmid pHBV1.5 (as a surrogate for the HBV genome) was quantified as expected with high fidelity. For the HBV genomic plasmid pSVHBV1.1LE⁻ used for transfections on the other hand an amount approximately 3000-fold lower than applied was measured. (B) A mixture of the plasmid for HBV surface protein expression (WT with respect to the MD sequence) ([Fig. 1C](#)) and the identical plasmid I1 carrying a substitution in MD rendering the L protein non-functional in nucleocapsid envelopment were cotransfected in the virus quantification assay. m: Transfection with the envelope-negative genomic construct ([Fig. 1B](#)) alone. The standard deviation is indicated, $n=3$.

The selectivity of this qPCR for viral genomes relative to the plasmid used for transfection was demonstrated as follows: a plasmid (pHBV1.5) containing the uninterrupted HBV target sequence from nt 1784 to nt 2005 was used as a surrogate for the genome in virions. This plasmid was quantified as expected with high sensitivity by the qPCR. However, the measured amounts for the genomic plasmid pSVHBV1.1LE⁻ were approximately 3000 fold ($10^{3.5}$) lower than the amounts applied to the assay ([Fig. 2A](#)).

For phenotypic characterization of matrix domain mutations we cotransfected a viral genome unable to direct the synthesis of functional envelope proteins due to two stop mutations in the corresponding open reading frame (pSVHBV1.1LE⁻) with an expression vector (pSV45–57) for the missing envelope proteins L, M, and S carrying the mutation of the matrix domain ([Fig. 1C](#)). Codons 2–36 of the preS1 region were deleted in this plasmid because they are dispensable for virion formation ([Bruss and Thomssen, 1994](#)) but their presence reduces the amount of secreted virions in a dose dependant fashion when the L protein is overexpressed relative to S and M protein ([Persing et al., 1986](#)). In addition, the ratio of L to S protein generated by plasmid pSV45–57 was unfavorable for optimal virion production. Therefore, the

expression plasmid pSVBX24H for the synthesis of the small envelope protein S ([Fig. 1A](#)) was cotransfected to adjust the L: S ratio to an optimal level.

Since we planned to use pools of MD mutants in the virion detection assay, we were concerned that non-functional mutants in the pool could suppress a functional mutant in a transdominant negative way. To address this issue we cotransfected different ratios of a plasmid directing the synthesis of N-terminal truncated wild type (WT) L (pSV45–57) with a plasmid encoding the non-functional, transdominant negative MD mutant I1 containing a partial substitution of the MD ([Bruss, 1997](#)) and used the culture supernatant in the virion quantification assay ([Fig. 2B](#)). The WT alone produced approximately $10^{6.3}$ virions/ml culture supernatant. The envelope-negative HBV genome without cotransfection of envelope expressing plasmids is unable to direct virion formation ([Bruss and Ganem, 1991](#)) but generated a signal which would correspond to approximately $10^{3.3}$ virions/ml ([Fig. 2B](#), column labeled ‘m’). This signal was considered as the background of the assay. When the negative MD mutant was expressed in the absence of WT L approximately $10^{4.2}$ virions/ml were found (approximately 1% of the WT signal). Apparently, the MD mutant I1 was not totally blocked in nucleocapsid envelopment but was 100 times less effective than the WT. Former characterizations of this mutant were done by a radioactive endogenous polymerase assay ([Bruss, 1997](#)) which was less sensitive. Surprisingly, a 1:1 mixture of WT and mutant consistently gave slightly larger amounts of virions than the WT alone which cannot be explained by us. Cotransfection of as low as 12.5% WT L and 87.5% mutant L resulted in virion amounts comparable to the transfection with 100% WT L. Even 3.1% WT L/96.9% mutant L was sufficient to achieve a signal 10-fold stronger than the signal from the transfection with 0% WT L/100% mutant L. We concluded that it would be possible and more efficient to test pools of 19 mutants (all possible non-WT aa at a given position). If just one mutant in the pool (approximately 5%) would be fully functional a signal at least 10-fold over background should be achieved.

Testing of the impact of sequence variations at individual positions of the matrix domain with pools of mutants

Eleven positions (R103, T106, L112, R113, D114, S115, H116, P117, Q118, A119, and W122) of the 22 aa long matrix domain were chosen, and pools of mutants representing all possible amino acid substitutions were generated for each position. We arbitrarily excluded the proline rich region (107PISPP) and a few positions at both ends of the MD from this analysis. The resulting 11 pools of plasmids were tested in the virion quantification assay ([Fig. 3A](#)) described above. Four pools (L112, R113, P117, W122) resulted in signals corresponding to only 1–2% of the signal generated by WT L protein. We conclude that the WT aa were required at these positions for efficient virion formation and that any other aa drastically reduced the level of virion formation. Two pools (D114 and Q118) showed relatively high signals around 50–60% of WT suggesting that several different aa at these positions were compatible with a functional matrix domain. The remaining 5 pools (R103, T106, S115, H116, and A119) gave signals around 10% of WT. Such a value would be expected if e.g. just one non-WT aa of the pool would confer full matrix functionality and the other 18 aa would result in non-functional L proteins.

Testing of the impact of sequence variations at some positions of the matrix domain with individual mutants

In-depth analysis of the impact of individual amino acid substitutions on virus production was performed for three positions of the MD (W122, D114, and Q118) ([Fig. 3B](#)). The results of

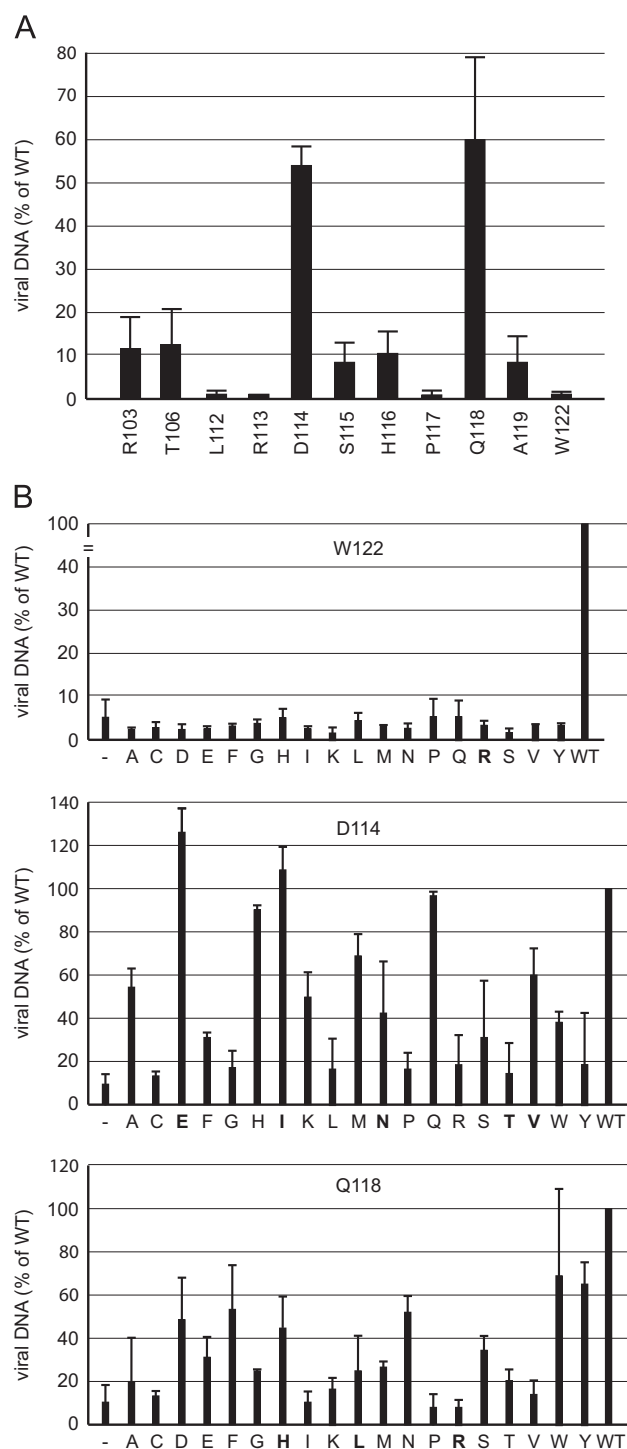


Fig. 3. Influence of MD mutations on virion formation. (A) Eleven pools containing all 19 possible point mutations at the indicated MD positions were separately used in the virus quantification assay. One hundred % signal on the ordinate was reached when L protein with WT MD was cotransfected. (B) Single mutants from pools W122, D114 and Q118 were cotransfected in the virus quantification assay. The standard deviation is indicated, $n=3$.

the tests with the pooled mutants (Fig. 3A) led us to predict that all mutants of W122 would be non-functional whereas various amino acid substitutions at position 114 and 118 would retain functionality. These expectations were confirmed: all tested 18 non-WT aa at position 114 resulted in an L protein defective in nucleocapsid envelopment. (The W122T mutant is not included because it turned out later in the project that this mutant carried a

frameshift mutation in L.) On the other hand the substitution of D114 by several aa like E, H, I, M, Q allowed virion formation to levels above 60% of the WT signal. Also the exchange of residue Q118 by F, N, W, and Y was compatible with at least 50% virion production relative to WT.

Stable expression of L point mutants

We then evaluated whether MD substitution mutants are efficiently expressed and are capable of co-assembling with M and S proteins within intracellular membrane. To this end we investigated the incorporation of mutants into subviral particles released from cells transfected with the L expression vector alone (Fig. 1C) by western blotting (Fig. 4). The incorporation of WT amounts of the L mutant into secreted subviral particles would indicate protein stability and coassembly competence with M and S protein close to WT. We were not able to test all 201 mutants but we tested all 18 mutants of position W122 (Fig. 4A) and all substitutions of the other 10 MD positions with phenylalanine (Fig. 4B). Since the L background carried a deletion of aa 2–36, the truncated L derivatives ran very close to M on the SDS-PAGE but could still be identified. All L proteins with W122 mutations were detected at similar levels as the WT L protein in both lysates and supernatants of cells. Also, all phenylalanine mutations in 10 MD positions allowed efficient incorporation of L into subviral particles. The S protein levels could be used as internal controls for transfection efficiency and gel loading. The amounts of M varied relative to S, apparently some mutations influenced M expression. Overall, the result showed that L protein stability and coassembly with HBV envelope proteins was not sensitive to point mutations in the MD. Therefore, it is likely that a negative phenotype of an MD mutant in the virion quantification assay indicates a defect in nucleocapsid envelopment.

Variability of the preS sequences retrieved from databases

A large number of complete HBV genome sequences carrying a preS1/preS2/s open reading frame with correct length were available from the NCBI database (3579 unique entries). We compared the preS sequences and counted the occurrence of all 20 aa at each position (Table Appendix A, Supplementary material). In Fig. 5 the number of different aa appearing at each aa position between aa 70 and aa 141 are shown. Amino acids appearing only once were ignored to reduce the risk that non-functional genomes were included which e.g. were packaged into virions by transcomplementation of a coinfecting WT virus.

The analysis revealed the following: the region can be divided into 4 areas: an N-terminal area from position 70 to 83 where, on the average, 3.5 different aa were found more than once per position; a central left area between position 84 and 91 with 8.3 different aa per position; a central right area from position 92 to 124 with again lower variability (4.0 different aa per position); and a C-terminal area with slightly higher variability from position 125 to 141 (5.7 different aa per position). The MD maps to the C-terminal part of the central right area (position 103 to 124).

The two positions characterized as being variable (D114 and Q118) showed different features in the sequence comparison: D114 was also relatively weakly conserved with only 82% of all sequences carrying aspartate at this position. And five other aa were found more than once at position 114. The second variable position Q118 was relatively conserved among the 3579 sequences, 99% of all sequences carried a glutamine at this position. Only 10 × H, 6 × L, and 2 × R could be found beside the consensus aa, and these residues are not among those allowing the highest virion formation levels in vitro (F, N, W, Y).

Inspection of the four positions of the MD characterized in this work as being crucial for virion formation (Fig. 3A) showed that these positions were also conserved among the stored sequences.

However, at residue 112 a glycine is present in 879 sequences instead of the consensus aa leucine. We separately tested the L112G mutation and found very low virion formation (twice over background, data not shown). Apparently, this variation results in low functionality in our HBV isolate.

Discussion

In the present study we analyzed the effect of point mutations in the matrix domain of the HBV L protein on virion assembly. Surprisingly, in four out of eleven tested positions of the MD the exchange of the WT aa by any other aa destroyed detectable virion formation. This demonstrates a remarkable invariability in this region and fits to a similar invariability of the corresponding matrix binding domain (MBD) on the HBV capsid (Pairan and Bruss, 2009), where the exchange of individual aa on the capsid surface to a large number of non-WT residues also blocked envelopment.

It is not known how many L proteins are involved in or required for the envelopment of a nucleocapsid. The virion envelope probably contains between 360 and 400 surface proteins (Dryden et al., 2006; Seitz et al., 2007). The molar ratio of L:M:S is approximately 1:1:4 (Heermann et al., 1984). Therefore roughly 60–70 L proteins are present in virions. It is possible that all L chains are in the i-preS form during budding or that already half of the molecules switched their topology to the e-preS form. In this case only 30–35 L molecules have their preS domains oriented towards the capsid. When L mutant I1 impaired in nucleocapsid envelopment was mixed with a WT L protein as low as 12.5% WT was sufficient to allow almost WT levels of virion formation (Fig. 2A). This would correspond to just 4 to 9 functional L proteins per virion envelope. However, it is not known whether non-functional L proteins are excluded from the budding process leading to an accumulation of functional L at the site of capsid envelopment.

The variability of certain MD residues characterized in this work often matched the conservation of MD positions in HBV genomes isolated from patients (Fig. 5). However, for some positions this comparison gave puzzling results. E.g. L112 did not tolerate any mutation in the pool (Fig. 3A), and the individual L112G mutant resulted in very low virus production. However, 804 sequences from the database carried a glycine at this position. Why glycine was not tolerated in our HBV isolate could not be answered. And for position Q118 several other aa resulted in approximately 50% of WT virus production in vitro (Fig. 3B). However, this residue is relatively conserved in the database. Possibly, changes in the overlapping open reading frame for the viral polymerase by mutations in this codon limit its variability (Jones et al., 2014). In our assay system the L protein was provided in trans and mutations in the MD had no influence on the polymerase. It is also possible that residue Q118 is conserved due to another pivotal function of L that is not relevant for in vitro virion formation. A further possibility is that the underlying DNA

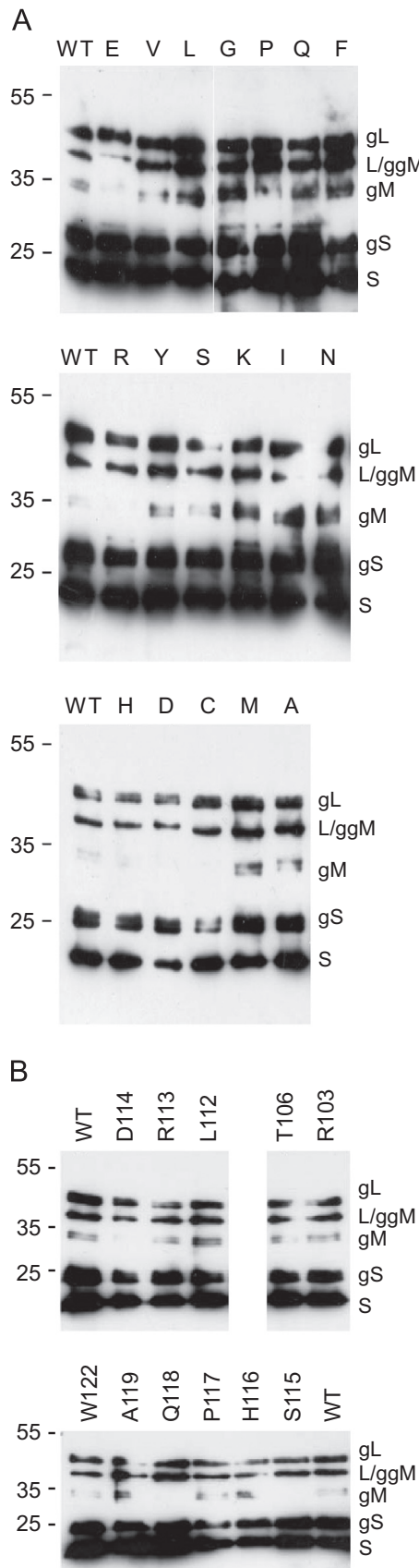


Fig. 4. Detection of L mutants in cell culture supernatants. (A) The expression vector for HBV envelope protein synthesis (Fig. 1C) carrying a mutation of preS codon 122 for replacement of the WT tryptophan residue by the indicated aa was used for transfection of Huh7 cells. Cell culture supernatants were harvested, the HBV envelope proteins were separated by polyacrylamide gel electrophoresis and visualized by western blotting with an anti-HBs antibody. (B) Like in (A) but the indicated codons were replaced by codons for tryptophan. S: unglycosylated S (24 kDa), gS: N-glycosylated S (27 kDa), gM: single N-glycosylated M (33 kDa), ggM: double N-glycosylated M (36 kDa), L: unglycosylated L (here 36 kDa due to the N-terminal deletion of aa 2–36), gL: N-glycosylated L (here 39 kDa). Numbers on the left indicate the position and molecular weight of marker proteins.

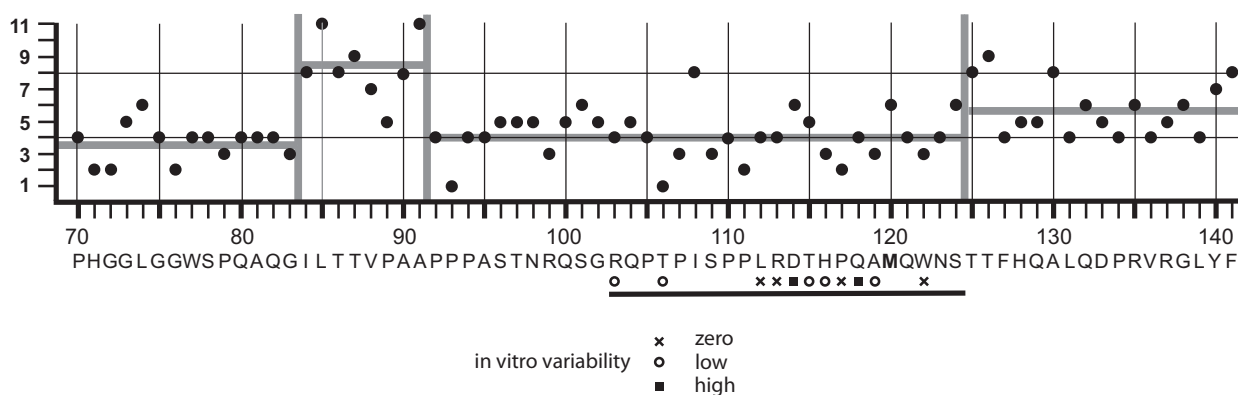


Fig. 5. Variability in the preS domain among sequences from the NCBI database. The consensus sequence from 3579 HBV genomes retrieved from the NCBI database between the preS positions 70 and 141 is indicated at the bottom. The MD is underlined. The isolate used in this work has the following differences in this region relative to the consensus sequence: L74I, G75L, P89S, A90T, A91I, T115S. The residues analyzed in this work (Fig. 3A) are indicated by crosses (only the WT aa allowed capsid envelopment), circles (the corresponding pools of L mutants resulted in low virus production), and squares (the corresponding pools of L mutants resulted in high virus production). Numbers on the ordinate show how many different aa could be found in more than one sequence from the database at the position indicated on the abscissa. The gray vertical bars divide the region into four parts displaying different levels of sequence variability. The gray horizontal bars show the average value of all data points in the corresponding region.

sequence is required for M protein expression. The M protein is not required for virion morphogenesis and a specific function of M in the viral lifecycle is not known, however, the M protein is conserved in all mammalian hepadnaviruses.

Structural proteins assemble during virion morphogenesis in the virus producing cell but they have to disassemble in the newly infected cell in order to release the viral genome. The correct execution of both processes requires a delicate equilibrium of the binding forces between viral structural components (Ceres and Zlotnick, 2002; Tan et al., 2013), and mutations in viral proteins are deleterious if they destroy the balance between these opposing processes. This study suggests that the MD in the L protein has a very limited potential to mutate without losing function. The invariant amino acid positions might represent convenient targets to interfere with virus replication by small molecules.

Materials and methods

Plasmids

The plasmid pSVHBV1.1LE⁻ (Fig. 1B) carries a terminal redundant 1.1-mer copy of an HBV genome, genotype A (Valenzuela et al., 1980), starting with nucleotide 1816 and terminating with nucleotide 1980 (numbering of the 3221 bp long HBV DNA sequence starts with the deoxycytidin residues of the unique EcoRI site in the preS2 region). The redundant region has a length of 165 nucleotides. Codon 52 of the preS1 and codon 15 of the S open reading frame were mutated to nonsense codons (TGG to TAG and TTA to TGA, respectively) to block the expression of functional envelope proteins. Both mutations are silent in the overlapping open reading frame for the viral DNA polymerase. Transcription initiated from the Simian Virus 40 early promoter directly upstream of the HBV sequence and terminated by the HBV polyA signal in the 3' redundancy generates an RNA corresponding to the viral pregenomic RNA. The plasmid was generated by a deletion in the 3' redundancy of plasmid pSVHBV1.5LE⁻ (Kluge et al., 2005). Plasmid pSVHBV1.5LE⁻ is named pHBV1.5 in the current work.

Plasmid pSV45-57 (Fig. 1C) directs the expression of an L protein with a deletion of aa 2–aa 36 of the preS1 region and WT M and S proteins. Plasmid I1 is identical to plasmid pSV45-57 but carries a substitution in the region encoding the MD. The corresponding L mutant has a substitution of the peptide N99 to

Q104 by the aa sequence PASGIRIPE (Bruss, 1997). Plasmid pSVBX24H (Fig. 1D) directs the expression of the S protein.

Mutagenesis

For the introduction of all 19 non-WT aa at a fixed position of the matrix domain plasmid pSV45-57 was used together with a mixture of 19 oligonucleotides representing the desired 19 point mutations in one reaction with the TransformerTM Site-Directed Mutagenesis Kit (Clontech, Saint-Germain-en-Laye, France). Sequencing of 70–80 molecular clones usually resulted in the identification of 12–17 different mutants. Missing mutants were generated by using individual mutant oligonucleotides for the same procedure.

Cell culture, transfection, harvest

Huh7 cells grown in 6-well dishes were transiently transfected with a total amount of 1 µg of plasmid using Eugene 6, Eugene HD, or X-tremeGene HP (Roche, Rotkreuz, Switzerland) according to the manufacturer's instructions. Cotransfections were performed with an equal molar ratio of plasmids. After transfection, the culture supernatant was removed, the cells were washed once with 2 ml of PBS and 2 ml of fresh medium was added to the cells. Five days post transfection the culture supernatant was harvested and centrifuged for 10 min at 13,000 rpm. From the supernatant 750 µl was used for immunoprecipitation or 40 µl for western blotting. The cells on the dish were washed with 1 ml of PBS and lysed by adding 0.5 ml of lysis buffer (50 mM Tris-HCl pH 7.5, 100 mM NaCl, 20 mM ethylenediaminetetraacetate, 0.5% (v/v) Nonidet P40) and incubated on ice for 10 min. The cell lysate was collected and spun at 13,000 rpm for 10 min. Forty microlitre of the supernatant was used for western blotting.

Virus quantification assay

To immunoprecipitate virions from cell culture supernatants of transfected HuH7 cells, 25 µl of protein G Plus agarose beads (Santa Cruz Biotechnologies) were washed three times with PBS and incubated with 1 µl of sheep polyclonal anti-HBs serum (gift from Wolfram Gerlich, Giessen) in an appropriate volume of PBS on a rotator for 6 h at 4 °C. After another washing step with PBS, 750 µl of cell culture supernatant was added and binding of the antigen to the anti-HBs occurred during overnight rotation at 4 °C.

The beads were washed three times with PBS and the amount of residual plasmid DNA from the transfection was further diminished by a treatment with DNase (Qiagen, Hamburg, Germany). For this purpose, the beads were suspended in 87 µl of PBS, 10 µl of buffer supplied by the manufacturer and 2.5 µl of DNase (30 Kunitz units) and incubated at room temperature for 30 min. To inactivate the DNase and to release the HBV genomes, the beads were finally resuspended in 45 µl of PBS, 50 µl of 2 × proteinase K buffer (20 mM Tris-Cl pH 7.5, 20 mM EDTA, 2% (w/v) SDS) and 50 µg of proteinase K (Applchem, Darmstadt, Germany), and incubated at 56 °C for 3 h. After three phenol/chloroform extractions the viral DNA was ethanol precipitated and the pellet was dissolved in 30 µl of water.

The extracted viral DNA was used for quantification of the HBV genome by a quantitative PCR using the forward primer 5'GGCA-TAAATTGGTCTGCGCACC hybridizing to the region between nt 1784 and nt 1805 on the HBV genome and the reverse primer 5'GCTGAGGCGGTGTCTAGGAGA hybridizing to the region between nt 2005 and nt 1985 on the HBV genome. Ten microlitre of SYBRGreen I Master (Roche, Penzberg, Germany), 1 µl each of a 20 µM solution of the two primers, and 8 µl of either template DNA or water were mixed and treated in a Roche LightCycler 480 II using the LightCycler 480 software release 1.5.0 SP4 (version 1.5.0.39). The temperature program was 5 min 95 °C/40 × (20 s 95 °C/10 s 60 °C/10 s 72 °C).

Western blotting

Forty microlitre of the cleared cell culture supernatant or lysate of transfected HuH7 cells were denatured and reduced with sodium dodecylsulfate and dithiothreitol and used for electrophoresis through 12% polyacrylamide gels. The proteins in the gel were transferred onto a nitrocellulose membrane (Trans-Blot Pure Nitrocellulose Membrane, 0.45 µm, Biorad, Munich, Germany) by semi-dry blotting. Envelope proteins were detected with the mouse monoclonal antibody HB1 against a linear epitope (aa 120–125) in the S protein (courtesy provided by Aurelia Zvirbliene, Vilnius University, Lithuania).

Comparison of HBV sequences

DNA sequences were retrieved from the NCBI nucleotide database by searching for “hepatitis B virus complete genome”. The data set contained 5088 entries in December 2013. However, numerous items did not represent complete HBV genomes. Therefore, this set was first filtered based on the sequence length (3221 nt for genotype A, 3215 nt for genotypes B, C, F, and H, 3182 nt for genotype D, 3212 nt for genotype E, and 3248 nt for genotype G) resulting in 3768 entries. Next, 49 redundant sequences were deleted. From these genomes the preS1/preS2/S open reading frame was translated starting with the methionine at position 12 for genotype A and the homologous positions in the other genotypes. Entries that did not translate to a 389 aa long peptide were deleted leaving 3579 unique full length sequences. The occurrence of all 20 aa was counted for each position of the protein (Table Appendix A, Supplementary material). The number of different aa found more than once at each position between 70P and 141F was determined (Fig. 5). All extractions from the retrieved data set and calculations were done by self-written computer programs. These programs will be made available on request.

Acknowledgments

We thank Ke Zhang (Institute for Virology, Technical University of Munich, Germany) for his help to develop the qPCR assay, Aurelia Zvirbliene, Vilnius University, Lithuania for providing the monoclonal anti-HBs antibody HB1, and Ruth Brack-Werner (Helmholtz-Zentrum München) for critical reading of the manuscript.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.virol.2014.04.030>.

References

- Beck, J., Nassal, M., 2007. Hepatitis B virus replication. *World J. Gastroenterol.* 13, 48–64.
- Bruss, V., 1997. A short linear sequence in the pre-S domain of the large hepatitis B virus envelope protein required for virion formation. *J. Virol.* 71, 9350–9357.
- Bruss, V., 2007. Hepatitis B virus morphogenesis. *World J. Gastroenterol.* 13, 65–73.
- Bruss, V., Ganem, D., 1991. The role of envelope proteins in hepatitis B virus assembly. *Proc. Natl. Acad. Sci. USA* 88, 1059–1063.
- Bruss, V., Thomssen, R., 1994. Mapping a region of the large envelope protein required for hepatitis B virion maturation. *J. Virol.* 68, 1643–1650.
- Bruss, V., Vieluf, K., 1995. Functions of the internal pre-S domain of the large surface protein in hepatitis B virus particle morphogenesis. *J. Virol.* 69, 6652–6657.
- Bruss, V., Lu, X., Thomssen, R., Gerlich, W.H., 1994. Post-translational alterations in transmembrane topology of the hepatitis B virus large envelope protein. *EMBO J.* 13, 2273–2279.
- Ceres, P., Zlotnick, A., 2002. Weak protein–protein interactions are sufficient to drive assembly of hepatitis B virus capsids. *Biochemistry* 41, 11525–11531.
- Chan, H.L., Sung, J.J., 2006. Hepatocellular carcinoma and hepatitis B virus. *Semin. Liver Dis.* 26, 153–161.
- Dryden, K.A., Wieland, S.F., Whitten-Bauer, C., Gerin, J.L., Chisari, F.V., Yeager, M., 2006. Native hepatitis B virions and capsids visualized by electron cryomicroscopy. *Mol. Cell* 22, 843–850.
- Gish, R., Jia, J.D., Locarnini, S., Zoulim, F., 2012. Selection of chronic hepatitis B therapy with high barrier to resistance. *Lancet Infect. Dis.* 12, 341–353.
- Heermann, K.H., Goldmann, U., Schwartz, W., Seyffarth, T., Baumgarten, H., Gerlich, W.H., 1984. Large surface proteins of hepatitis B virus containing the pre-s sequence. *J. Virol.* 52, 396–402.
- Jones, S.A., Clark, D.N., Coa, F., Tavis, J.E., H.J., 2014. Comparative analysis of hepatitis B virus polymerase sequences required for viral RNA binding, RNA packaging, and protein priming. *J. Virol.* 88, 1564–1572.
- Kluge, B., Schlager, M., Pairan, A., Bruss, V., 2005. Determination of the minimal distance between the matrix and transmembrane domains of the large hepatitis B virus envelope protein. *J. Virol.* 79, 7918–7921.
- Ni, Y., Sonnabend, J., Seitz, S., Urban, S., 2010. The pre-s2 domain of the hepatitis B virus is dispensable for infectivity but serves a spacer function for L-protein-connected virus assembly. *J. Virol.* 84, 3879–3888.
- Ostapchuk, P., Hearing, P., Ganem, D., 1994. A dramatic shift in the transmembrane topology of a viral envelope glycoprotein accompanies hepatitis B viral morphogenesis. *EMBO J.* 13, 1048–1057.
- Pairan, A., Bruss, V., 2009. Functional surfaces of the hepatitis B virus capsid. *J. Virol.* 83, 11616–11623.
- Persing, D.H., Varmus, H.E., Ganem, D., 1986. Inhibition of secretion of hepatitis B surface antigen by a related presurface polypeptide. *Science* 234, 1388–1391.
- Ponsel, D., Bruss, V., 2003. Mapping of amino acid side chains on the surface of hepatitis B virus capsids required for envelopment and virion formation. *J. Virol.* 77, 416–422.
- Prange, R., Streeck, R.E., 1995. Novel transmembrane topology of the hepatitis B virus envelope proteins. *EMBO J.* 14, 247–256.
- Seitz, S., Urban, S., Antoni, C., Bottcher, B., 2007. Cryo-electron microscopy of hepatitis B virions reveals variability in envelope capsid interactions. *EMBO J.* 26, 4160–4167.
- Tan, Z., Maguire, M.L., Loeb, D.D., Zlotnick, A., 2013. Genetically altering the thermodynamics and kinetics of hepatitis B virus capsid assembly has profound effects on virus replication in cell culture. *J. Virol.* 87, 3208–3216.
- Valenzuela, P., M. Quiroga, J. Zaldivar, R. Gray, W. Rutter, 1980. The nucleotide sequence of the hepatitis B viral genome and the identification of the major viral genes. In: *Proceedings of the UCLA Symposium on Molecular and Cellular Biology*, vol. 18, pp. 57–70.
- Wynne, S.A., Crowther, R.A., Leslie, A.G., 1999. The crystal structure of the human hepatitis B virus capsid. *Mol. Cell* 3, 771–780.