Labeling of hepatitis B virus middle envelope protein with enhanced green fluorescent protein

PAULA E. FLORIAN1*, RAMONA N. GALANTONU2*, CATALIN LAZAR2, SIMONA RUTA3,4, ANCA ROSEANU1 and NORICA NICHITA2

1Ligand-Receptor Interactions Department, Institute of Biochemistry of the Romanian Academy, Bucharest 060031, Romania
2Viral Glycoproteins Department, Institute of Biochemistry of the Romanian Academy, Bucharest 060031, Romania
3Department of Virology, Carol Davila University of Medicine and Pharmacy, Bucharest 020022, Romania
4Emergent Diseases Department, Stefan S. Nicolau Institute of Virology, Bucharest 030304, Romania.

Corresponding author: Norica Nichita, tel: 004021 223 90 69 fax: 004021 223 90 68, nichita70@yahoo.co.uk nichita@biochim.ro
*These authors contributed equally to this work
#Present address: University of Nice Sophia Antipolis, 650 Route Des Colles, 06560 Valbonne Sophia Antipolis, France

Abstract

Hepatitis B virus (HBV) is a DNA enveloped virus, which infects human liver leading to severe pathologies. The HBV genome is packed in a nucleocapsid, surrounded by the viral envelope containing three proteins: large (L), medium (M) and small (S), with different functions in the viral life-cycle. Functional labeling of HBV particle for imagistic studies has failed so far, despite the use of various strategies. Here we aimed to investigate the possibility of fluorescently labeling the M protein, the only envelope component which is dispensable for HBV infection, despite being incorporated into mature virions. M protein was N-terminally fused with the enhanced green fluorescent protein (EGFP) and the recombinant protein was characterized biochemically and functionally. The results showed that EGFP:M displayed the typical glycosylation and dimerization pattern, characteristic to the wild-type counterpart. Microscopy analysis revealed that EGFP:M partially shared the same intracellular location as the other envelope proteins. The protein was secretion-competent when expressed on its own; however, the process occurred at significantly lower rates compared to the wild-type protein. Trans-complementation assays have shown that EGFP:M is incorporated during viral assembly when expressed in the context of the full length HBV genome, only in the absence of the wild-type M protein.

Keywords: Hepatitis B virus, enhanced green fluorescent protein, labeling, viral envelope

Introduction

Human hepatitis B virus (HBV) infection is a global health problem affecting about 350 millions of people who are prone to develop hepatitis, cirrhosis or liver cancer [1, 2]. Current treatment is often inefficient and patients develop resistance due to mutations occurring in the viral polymerase following prolonged therapy [3, 4]. Therefore, there is a rising interest in dealing with therapy resistance and controlling the disease in long term.
HBV belongs to the Hepadnaviridae family. Its 3.2 kb DNA is packed in a nucleocapsid surrounded by the viral envelope, which contains three proteins: large (L), medium (M) and small (S). It has been proposed recently that HBV infection is initiated by virus binding to a hepatocyte-specific membrane receptor, the sodium taurocholate co-transporting polypeptide (NTCP) [5]. Other cellular factors such as clathrin heavy chain, clathrin adaptor protein AP-2 and caveolin-1 play a role in HBV infection [6, 7]. The envelope proteins share a common S domain and play different functions in the virus life-cycle. The S protein is important in subviral particles (SVPs) and virion formation and secretion [8]. Interestingly, the M protein, which contains the S domain and a preS2 extension, appears to be dispensable for these processes, as well as infection. The L protein, which includes the M sequence and the preS1 extension, is important for nucleocapsid envelopment during virion assembly [9] and plays a key role in infection [5, 10].

While many aspects of HBV replication were clarified in the last decades, some important details of attachment, internalization [11-13] and intracellular trafficking leading to productive infection into the host cell have only recently been revealed [14]. Developing new antiviral agents to improve the current standard of care of HBV-infected patients is tightly linked to understanding the molecular mechanism of the viral infection. In this respect, one approach to address HBV internalization into host cells and the intracellular trafficking leading to infection is to monitor a labeled HBV virus particle, using either microscopy or more quantitative assays. Previous reports have described different strategies in their attempt to label HBV: the S and the core protein (HBcAg) were chosen as targets for the green fluorescent protein (GFP) [15] or biarsenical labeling, respectively [16]. To increase the efficiency of the labeling process, both attempts have targeted abundant proteins present in the mature virion, with crucial functions in its assembly and infectivity. However, it was not clear whether the changes induced at the level of the structural proteins could affect the infection properties of the virus.

The purpose of this study was to investigate the potential use of the M protein as a tag acceptor for labeling of the HBV envelope. Our hypothesis was that modification of the only envelope protein that was shown to be dispensable for HBV infection [17] may have less impact on the virus functionality. The genetic fusion between the M envelope protein and a fluorescent marker, such as the GFP, would enable tracking the viral protein routes and the early stages of the viral life cycle.

**Materials and Methods**

*Cells.* HEK 293T human embryonic kidney and Huh 7 human hepatoma cells (European Collection of Animal Cell Culture, Porton Down, UK) were routinely maintained in RPMI medium supplemented with 10% fetal bovine serum, 50 U/ml penicillin, 50 µg/ml streptomycin, 2 mM Glutamax (GIBCO, Life Science) and maintained in the incubator at 37°C in 5% CO2 atmosphere.

*Plasmids and transfection.* Plasmids pCiS, pCiM or pCiL coding for S, M and L proteins, were described previously [18]. Plasmids pTriEx HBV1.1 containing 1.1 units of the whole HBV genome, supporting viral replication, assembly and secretion of fully infectious virions and pTriEx HBVΔM lacking expression of the M protein were also used in secretion assays [19]. The plasmid pEGFP.M expressing the M envelope protein N-terminally fused with the enhanced GFP (EGFP), under the control of the CMV promoter was obtained by cloning the EGFP cDNA in the Nhe I sites of the pCiM vector. For transfection experiments,
HEK 293T and Huh7 cells were seeded at a confluence of 80% in 6 wells plates and the next day were transfected with 3µg of plasmid DNA using Lipofectamine 2000 (Invitrogen). After 48h of transfection, cells and supernatants were collected.

**SDS-PAGE and Western blotting.** Transfected or mock transfected (control) cells were lysed in a buffer containing 10 mM Tris-HCl (pH 7.5), 150 mM NaCl, 2 mM EDTA, 0.5% TritonX-100 and a mixture of protease inhibitors (Sigma-Aldrich). Equal amounts of total proteins were migrated on SDS PAGE under either reducing –R (5 mM DTT) or non-reducing-NR (no reducing agent) conditions before blotting on nitrocellulose membranes using a semi-dry bloter (Millipore). The blots were incubated with rabbit anti-S (Europa Bioproducts, dilution 1/2000) antibodies (Abs), followed by anti-rabbit (Santa Cruz, dilution 1/20000) Abs conjugated with horseradish peroxidase. The proteins were detected using an enhanced chemiluminescence (ECL) system (Santa Cruz).

**Pulse-labeling and immunoprecipitation.** Transfected cell monolayers were washed once with phosphate-buffered saline (PBS) and then starved in cysteine-free medium. After 1h, the cells were pulse-labeled with 100 µCi/ml of 35Smethionine-35S(Trans (Trans 35S-label, 1,100 Ci/mmol; ICN Flow) for 15 min at 37°C. After 3 h of chase, the cells and supernatants were harvested. Cell lysates and medium were further incubated with Protein A-Sepharose and anti-S Abs (diluted 1:2000) overnight at 4°C. Labelled proteins were separated by SDS-PAGE and analyzed by autoradiography.

**Quantification of secreted HBsAg by ELISA.** Cells seeded in 75 cm² flasks were transfected for 24h as indicated in the figures and the supernatants were 100-fold ultracentrifuge-concentrated. Resulting samples were analyzed for the amount of secreted HBsAg, using the Monolisa HBsAg Ultra Kit (BiO-Rad). The results were obtained as ratios of signal to cut-off value.

**Immunofluorescence assay.** Cells seeded on glass cover-slips in 12 wells plates were transfected the next day for 24h, fixed with paraformaldehyde and permeabilized with 0.2% Triton-X100. For internal antigens detections, cells were incubated with rabbit anti-S, rabbit anti-preS1(dilution 1:5000, Santa Cruz) and anti-protein disulfide isomerase (PDI, dilution 1:500, Santa Cruz) Abs prior to incubation with either anti-rabbit AlexaFluor 488-conjugated or anti- AlexaFlour 594-conjugated (dilutions 1:400, Molecular Probes). DAPI (4’6’-diamidino-2-phenylindole) was used to stain the nuclei. The samples were mounted with ProLong Gold-antifade reagent (Invitrogen) and visualized under a Zeiss LSM710 laser scanning confocal microscope. Images were processed with AxioVision Rel 4.8 software.

**Spectrofluorimetry.** Cells seeded in 75 cm² flasks were transfected with pCiEGFP.M, pTriEx HBV1.1, and pTriExHBVΔM plasmids and the amounts of secreted fluorescent SVps was quantified by spectrofluorimetry (Jasco FP-6500; 490 nm excitation/ 509 nm emission wavelengths). Supernatants of cells transfected with the pCi plasmid were used as blanks.

**Results and Discussion**

The study describes the cloning and the functional characterization of the recombinant EGFP:M protein. These investigations addressed the folding, N-glycosylation status and stability of the protein, the intracellular localization and the ability to form secretion-competent SVps. Previous reports have suggested that the N-terminus of the S envelope protein is more permissive for the GFP insertion than the C-terminus [15]. Based on this data, the EGFP sequence was fused to the N-terminus of the HBV M envelope protein.

Due to its properties to generate a highly visible, efficiently emitting internal fluorophore, the GFP has been widely used for visualization of protein expression and
intracellular localization and has also been used to successfully tag other viruses. Thus, there have been studies reporting production of recombinant viruses in which structural proteins labeled with GFP were incorporated (vesicular stomatitis virus, herpes simplex virus type I, coronavirus, Bunyamwera virus and HIV [20-24].

**EGFP:M recombinant protein expression and glycosylation**

The S domain of all three envelope proteins is partially glycosylated at Asn-146. The preS2 domain of M protein is fully glycosylated at Asn-4 (Fig. 1A), however, this second glycosylation site is not occupied in the L protein. The N-glycosylation status is particularly relevant for the M protein since the calnexin-assisted folding and secretion of virions depend on the presence of the N-linked glycan at Asn-4 and its correct trimming [25-29].

To analyze the expression and glycosylation pattern of the EGFP:M protein, HEK cells were transfected with pEGFP.M. Twenty four hours after transfection, cells were harvested, lysed and immunoblotted with antibodies specific for the S domain of the envelope proteins. As expected, the wild-type M was expressed as a mixture of fully- (ggM) and partially glycosylated (gpM) polypeptide (Fig. 1B). In addition to these forms, a doublet with a lower electrophoretic mobility was observed in pEGFP.M-transfected cells, in agreement with the molecular weight of the fusion protein, confirming correct cloning and expression of the EGFP-tagged protein (Fig 1B). According to the molecular weights the proteins within the slower migrating doublet were identified as the glycoforms ggEGFP:M and gpEGFP:M, respectively. Since the fused EGFP-M sequence contains two internal translation initiations ATG codons, the pCiEGFP.M expression vector produces small quantities of M and S proteins, in addition to the main EGFP:M fusion protein (Fig. 1B).

**Figure 1.** Expression of the EGFP:M recombinant protein. **A.** Schematic representation of the middle (M) HBV envelope glycoprotein. N-glycosylation sites (NG) are indicated: Ψ – partially occupied NG site in S domain and Ψ’ – totally occupied NG site in preS2. **B.** HEK cells were transfected with either pCiM or increasing amounts of pCiEGFP.M plasmids and analyzed by Western blotting with anti-S Abs under reducing conditions.
These results show that the recombinant EGFP:M protein is stable and importantly, it maintains the glycosylation pattern characteristic of the M wild-type. The wild-type M protein contains four hydrophobic transmembrane segments common to the S domain and both, the N-terminal pre-S2 sequence and the C-terminal end are oriented towards the ER lumen [30]. Interestingly, analysis of the topology of the GFP:S fusion protein has suggested a dual conformation, with the GFP domain displayed both, towards the ER lumen and the cytoplasm [15]. Formation of such a mixed topology is rare among membrane proteins but is a prominent feature of the L envelope protein of HBV [30].

In our study, EGFP:M likely adopts a similar conformation with the M wild-type, considering the glycosylation status of the proteins as a marker for the ER exposure of the pre-S2 domain.

*Folding of the recombinant EGFP:M protein*

Shortly after synthesis and acquirement of the native conformation within the ER, the viral envelope proteins begin the assembly process by forming homo- and heterodimers linked by disulfide bridges, without preference for a specific association [31, 32]. Thus, dimerization may be considered an indicator of appropriate folding of the corresponding polypeptides. To determine the oligomerization pattern of the EGFP:M protein, lysates of cells transfected with pEGFP:M were subjected to SDS-PAGE under non-reducing conditions. Cells expressing both, the wild-type M and the EGFP-fused M proteins were also included as control in this analysis.

![Figure 2. Dimerization of the wild-type and EGFP:M recombinant protein. HEK cells were co-transfected with pCiEGFP.M and increasing amount of pCiM plasmids. After 24 h cell lysates were analyzed for the presence of M monomers or dimers by Western Blotting under non-reducing conditions, using the anti-preS2 Abs.](image)

As shown in Fig. 2, EGFP:M is able to form homodimers when expressed on its own, as well as heterodimers with the M wild-type. As the dimerization of the envelope proteins is controlled by the S domain, this result suggests that most likely, the S domain of the M protein and not the EGFP peptide ensures the interaction within the homo- and heterodimers. This is an important observation showing that EGFP:M is a functional partner for dimerization with the wild-type counterpart and therefore, it is likely to be incorporated into the viral envelope, during virion assembly. A possible dimerization EGFP-EGFP is very improbable; although homodimers of GFP were reported in solutions and crystals, the associations are very weak [33].
Intracellular localization of the EGFP:M recombinant protein

We further wanted to visualize the intracellular distribution of the recombinant EGFP:M protein in cells of hepatic origin, such as the Huh7 cell line. To determine the intracellular localization of the recombinant protein the protein disulfide isomerase (PDI), a well-established ER-resident protein and a marker for this organelle, was also considered in this study.

As seen in Fig. 3 (first row) and as expected, the wild-type M and PDI co-localize strongly. When expressed alone, the EGFP:M protein shows a perinuclear localization and a micropunctate pattern of distribution resembling vesicles only partially co-localizing with ER compartment (Fig. 3, second row). It was previously shown that the GFP:S recombinant protein may aggregate if incorrectly folded and transported for degradation [15]. The vesicular localization of the EGFP:M recombinant protein is an indication that the protein could have, at least in part, the same fate. Interestingly, when expressed in the context of the whole HBV genome, an important co-localization of the recombinant EGFP:M protein with the wild-type L protein was observed, indicating that the two viral proteins share the same intracellular distribution (Fig. 3, third row).

Figure 3. Intracellular localization of EGFP:M recombinant protein. Huh7 cells were transfected with pCiM (upper panels), pCiEGFP.M (middle panels) or co-transfected with pCiEGFP.M and pTriEx HBV 1.1 plasmids (lower panels). Cells were fixed, permeabilized and immunostained with specific Abs for M, L and PDI. Merged images containing DAPI staining for nuclei visualization are shown. Scale bar: 20 µm.

Secretion of the EGFP:M recombinant protein

Having established that EGFP:M is a stable, assembly competent protein with a similar folding as the wild-type counterpart and, an intracellular distribution co-localizing at least partially with the wild-type HBV envelope proteins, it was important to determine its efficiency of secretion. To this end, metabolically-labeled HEK cells transfected with
pCiEGFP.M and their supernatants were analyzed by immunoprecipitation using antibodies against the S domain. The results shown in Fig. 4 A demonstrate again the intracellular expression of the recombinant protein.

**Figure 4.** Secretion of recombinant EGFP:M protein in HEK cells. Intracellular expression (A) and secreted (B) EGFP:M recombinant protein in medium of the transfected cells, detected by pulse labeling and immunoprecipitation with anti-S antibody.

Although a fraction of the EGFP:M protein was secreted into the medium (Fig. 4 B), an important amount of the protein was still present within the cells, despite the 3 hour-chase, during which time most of the synthesized protein is expected to exit the cells. Thus, it is possible that a pool of the expressed protein is retained in the cells and targeted to lysosomes for further degradation.

Interestingly, a previous report has shown that the recombinant protein GFP:S was secreted only if co-expressed with S wild-type. Moreover, the GFP:S protein had all the functions warranted by inter-protein interactions with wild-type chains [15]. By contrast, the EGFP:M protein is secreted in the absence of other envelope proteins. These observations suggest that the EGFP:M protein is more stable and can better tolerate the addition of a rather large sequence at its N-terminal end.

To further investigate the efficiency of EGFP:M secretion and the ability of the fluorescent protein to be incorporated together with the wild-type proteins into HBV SVPs and virions a trans-complementation assay was performed. The fluorescent protein was supplied in cells expressing the wild-type envelope proteins, by co-transfection of pEGFP.M with either pTriExHBV (encoding for all viral proteins) or pTriexHBVΔM (lacking M protein expression due to mutation of the internal ATG codon driving M biosynthesis). The secretion of SVPs from these cells was measured by an HBsAg-specific ELISA, while fluorescence of the HBV virions and SVPs was quantitatively determined by spectrofluorimetry. As shown in Fig. 5, secretion of HBsAg from cells expressing EGFP:M alone is significantly reduced compared to the cells expressing full-length, wild-type HBV. Co-expression of the fluorescent chimera in the context of the wild-type HBV proteins, either in the presence (pTriEx HBV) or absence (pTriEx HBVΔM) of the M protein also results in a slight inhibition of HBsAg secretion, suggesting an intracellular retention effect. Notably, an important increase of HBV particles fluorescence was detected when cells were co-transfected with pEGFP.M and pTriEx HBVΔM. This indicates that EGFP:M is able to “replace” the wild-type M protein during HBV assembly and becomes incorporated, albeit with low efficiency, into SVPs/virions. However, the fluorescently labeled M protein is not able to compete with the wild-type counterpart for virus assembly, the latter being preferred.
Figure 5. Secretion of EGFP:M–containing HBV particles. Huh7 cells were transfected with pTriEx HBV 1.1 or pTriEx HBVΔM, in the presence or absence of pEGFP.M. Cell supernatants were analyzed for HBsAg secretion by ELISA (black box) or EGFP fluorescence by spectrofluorimetry (gray box).

Conclusion

The HBV M protein was successfully labeled with EGFP at its N-terminal end. The resulting protein is stable, oligomerization- and secretion-competent. In addition, the fluorescent protein has a similar glycosylation pattern to the wild-type counterpart and colocalizes, at least partially, with wild-type envelope proteins. However, the efficiency of EGFP:M secretion is significantly reduced when the recombinant protein is expressed on its own. Trans-complementation experiments have indicated that the fluorescent protein is incorporated into HBV particles, when the wild-type M protein is absent. Future work will address the exact nature of these particles and the efficiency of incorporation of the recombinant M protein into the virion while preserving its infectivity properties. Thus, unlike the S envelope protein which is more abundant but strictly required for a functional HBV particle, the M protein may be a better viral target for HBV labeling.

Acknowledgments

This paper is supported by the Sectoral Operational Programme Human Resources Development (SOP HRD), financed from the European Social Fund and by the Romanian Government under the contract number POSDRU/159/1.5/S/137390 and by Romanian Academy Project 3 of the Institute of Biochemistry

References


