



Phosphorylation of the Hepatitis B Virus Large Envelope Protein

Marie-Laure Fogeron^{1†}, Lauriane Lecoq^{1†}, Laura Cole¹, Roland Montserret¹, Guillaume David¹, Adeline Page², Frédéric Delolme², Michael Nassal³ and Anja Böckmann^{1*}

¹Molecular Microbiology and Structural Biochemistry, Labex Ecofect, UMR 5086 CNRS/Université de Lyon, Lyon, France,

²Protein Science Facility, SFR BioSciences CNRS UAR 3444, Inserm US8, UCBL, ENS de Lyon, Lyon, France, ³Department of Medicine II / Molecular Biology, Medical Center, University Hospital Freiburg, University of Freiburg, Freiburg im Breisgau, Germany

We here establish the phosphorylation sites in the human hepatitis B virus (HBV) large envelope protein (L). L is involved in several functionally important interactions in the viral life cycle, including with the HBV cellular receptor, HBV capsid, Hsc70 chaperone, and cellular membranes during fusion. We have recently shown that cell-free synthesis of the homologous L protein of duck HBV in wheat germ extract results in very similar phosphorylation events to those previously observed in animal cells. Here, we used mass spectrometry and NMR to establish the phosphorylation patterns of human HBV L protein produced by both *in vitro* cell-free synthesis and in *E. coli* with the co-expression of the human MAPK14 kinase. While in the avian virus the phosphorylation of L has been shown to be dispensable for infectivity, the identified locations in the human virus protein, both in the PreS1 and PreS2 domains, raise the intriguing possibility that they might play a functional role, since they are found at strategic sites predicted to be involved in L interactions. This would warrant the further investigation of a possible function in virion formation or cell entry.

Keywords: hepatitis B, L HBsAg, phosphorylation, NMR, cell-free (CF) protein synthesis, mass spectrometry

OPEN ACCESS

Edited by:

Piero Andrea Temussi,
University of Naples Federico II, Italy

Reviewed by:

Giovanna Musco,
San Raffaele Hospital (IRCCS), Italy

Richard J. Bingham,
University of Huddersfield,
United Kingdom

*Correspondence:

Anja Böckmann
a.boeckmann@ibcp.fr

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Structural Biology,
a section of the journal
Frontiers in Molecular Biosciences

Received: 24 November 2021

Accepted: 22 December 2021

Published: 23 February 2022

Citation:

Fogeron M-L, Lecoq L, Cole L,
Montserret R, David G, Page A,
Delolme F, Nassal M and Böckmann A
(2022) Phosphorylation of the Hepatitis
B Virus Large Envelope Protein.
Front. Mol. Biosci. 8:821755.
doi: 10.3389/fmolb.2021.821755

INTRODUCTION

The three hepatitis B virus (HBV) envelope proteins L, M, and S [large, middle and small hepatitis B surface antigen (HBsAg)] form the viral envelope. S is an integral membrane protein for which structural models predict four transmembrane-spanning helices, interrupted by an N-proximal cytosolic loop and, after the second helix, the antigenic loop, which presents a complex yet poorly defined structure stabilized by multiple disulfide bridges. M and L share their C-termini with S but carry in addition the PreS1 plus PreS2 (for L) and PreS2 (for M) domains. PreS1 and PreS2 together, collectively termed PreS, are suspected to represent an intrinsically disordered protein domain. This was experimentally supported for the avian (duck) homolog, DHBV PreS (Urban et al., 2000), while the corresponding data for PreS from human HBV lacked.

The PreS part of L plays a central role in a variety of interactions with viral and cellular proteins. The very N-terminal residues of PreS represent a first important site, with the first 48 amino acids involved in binding to the HBV receptor (Glebe et al., 2005; Gripon et al., 2005), the sodium taurocholate co-transporting polypeptide (NTCP) (Yan et al., 2012). N-terminal myristoylation is central in this context (Persing et al., 1987). Further down the PreS sequence, several hydrophobic stretches can be identified that are concentrated between residues 50–70. Most models favor that different residues distributed along preS are responsible for membrane interaction, hinting at a

possible fusion mechanism (Núñez et al., 2009; Delgado et al., 2015), but a recent proposal locates the fusion peptide exactly to this hydrophobic region (Pérez-Vargas et al., 2021). A third region of interest is localized at the PreS1/PreS2 border, comprising residues between approximately amino acids 90–120, which is believed to be involved in interactions with the viral capsid during particle formation (Gudima et al., 2007; Xi et al., 2021). Finally, another site in PreS1 is involved in the interaction with the Hsc70 chaperone, reported to be a determinant in the i-PreS orientation observed in immature viral particles (Prange et al., 1999; Prange, 2012). In sum, several specific sites in preS have been identified to be involved in the important functional interactions of L.

Post-translational modifications have been shown to play a central role in the HBV life cycle (Heger-Stevec et al., 2018; Yang, 2018). The phosphorylation of PreS has been thoroughly investigated in the DHBV L variant (Grgacic and Anderson, 1994; Grgacic et al., 1998). Metabolic labeling with ³²P and digestion with residue-specific phosphatases (Grgacic and Anderson, 1994) have identified several phosphorylation sites, whose functional importance has, in part, been assessed by the mutational analysis of Ser and Thr residues (Grgacic et al., 1998). Mutations mimicking absent or constitutive PreS phosphorylation had no detectable effect on DHBV replication (Grgacic et al., 1998), although DHBV L phosphorylation was found important in host cell–virus cross talk (Rothmann et al., 1998).

We have recently found, during the wheat-germ cell-free protein synthesis of DHBs proteins for structural studies, that the DHBV L protein is phosphorylated in this system (David et al., 2018; David et al., 2019). While it was known that the cell-free extract indeed contains kinases, allowing post-translational phosphorylation (Badillo et al., 2017), our analysis of the DHBV L protein confirmed that the phosphorylation sites are the same as in animal cells (Grgacic et al., 1998) and identified two additional sites (David et al., 2019).

As an important step in the ongoing studies in our group on HBV preS–capsid interactions, we here investigate preS phosphorylation, in the protein, from the human virus, as obtained in recombinant systems, and apply a similar analysis to the HBV L protein and PreS fragments thereof. Based on the results, we designed a bacterial kinase co-expression system able to phosphorylate PreS, which we show to reproduce the results obtained in the acellular system. In addition, we used NMR to structurally analyze the PreS protein and to confirm major phosphorylation sites by NMR chemical-shifts. Our approach confirmed the disordered nature of the protein, and identified four consensus sites as well as several additional potential PreS phosphorylation sites.

MATERIALS AND METHODS

Plasmids

For cell-free expression, PreS1, PreS, and full-length HBV L sequences were amplified by PCR from the HBV isolate H2815 (genotype D5, GenBank Accession Number

KP322603.1), and cloned into a pEU-E01-MCS vector (CellFree Sciences, Matsuyama, Japan). A *Strep*-tag II, shortly named “tag” in the following, was fused either to the N- or C-terminal end for purification (Schmidt and Skerra, 2007), resulting in the four constructs PreS1_{tag}, tagPreS, PreS_{tag}, and tagL. The plasmids were amplified in *Escherichia coli* TOP10 cells (Life Technologies, Carlsbad, CA, USA). DNA was isolated using a NucleoBond Xtra Maxi kit (Macherey-Nagel). Plasmids were further purified by phenol/chloroform extraction, according to CellFree Sciences recommendations.

For the bacterial expression of PreS (^{E.coli}PreS_{tag}), cDNA-encoding PreS was cloned into a pRSF-T7 vector and into an analogous pRSF-T7 vector carrying a MAPK14 gene under a Tet promoter in order to promote PreS phosphorylation. A solubility-enhancing fusion protein, GB1, with an N-terminal His-Tag, was fused at the N-terminus of PreS and a *Strep*-tag II to its C-terminus. A TEV protease cleavage site was inserted between GB1 and PreS, resulting in the following constructs: pRSF_T7-H6-GB1-ENLYFQG-preS-*Strep*-tag-II and pRSF_Tet-H6MAPK14_T7-H6-GB1-ENLYFQG-preS-*Strep*-tag-II. Plasmids were amplified in *E. coli* TOP10 cells (Life Technologies). DNA was isolated using QIAprep Spin Miniprep Kit (Qiagen, Hilden, Germany).

Wheat Germ Cell-free Expression and Purification of PreS and L

Homemade wheat germ extract was prepared using non-treated durum wheat seeds (Semences du Sud, Vic-Fezensac, France), as described in Fogeron et al. (2017). Protein synthesis was performed with uncoupled transcription and translation. Transcription was performed using 100 µg/ml plasmid, 2.5 mM NTP mix (Promega, Charbonnières-les-Bains, France), 1 U/µl RNase inhibitor (CellFree Sciences, Matsuyama, Japan), and 1 U/µl SP6 RNA polymerase (CellFree Sciences, Matsuyama, Japan) in transcription buffer (CellFree Sciences, Matsuyama, Japan) composed of 80 mM HEPES-KOH pH 7.6, 16 mM magnesium acetate, 10 mM DTT, and 2 mM spermidine in nuclease-free water. The solution was incubated for 6 h at 37°C; the produced mRNA solution was then used directly for translation. Translation was performed using the bilayer method (Takai et al., 2010; Fogeron et al., 2015), either at a small scale with the one well of a 6-well plate (6 ml total reaction volume), or at a large scale with two 6-well plates (2 × 36 ml total reaction volume) in order to obtain a sufficient amount of protein to perform NMR experiments. The feeding buffer composition was 30 mM HEPES-KOH pH 7.6, 100 mM potassium acetate, 2.7 mM magnesium acetate, 16 mM creatine phosphate, 0.4 mM spermidine, 1.2 mM ATP, 0.25 mM GTP, and 4 mM DTT supplemented with 6 mM amino acid mix (0.3 mM average concentration per amino acid). The translation mix was prepared with the mRNA solution, wheat germ extract (250 µl for each well of a 6-well plate), 6 mM amino acid mix, and 40 µg/ml creatine kinase. On the bottom of the each well, the translation mix (518 µl for each well) was then deposited under the feeding buffer (5.5 ml for each well), allowing for the formation of a bilayer. The plate was incubated overnight at 22°C without

shaking. For NMR sample preparation, large-scale production was performed in presence of a mixture of (^{15}N) or (^2H - ^{13}C - ^{15}N)-labeled amino acids (Cambridge Isotope Laboratories) added to the reaction solution and the feeding buffer. A summary of the different sample preparations is given in **Supplementary Table S1**.

For *Strep*-Tactin affinity chromatography, the total cell-free reaction was incubated with homemade benzonase 250 U/ μl (50 μl per well) on a rolling wheel for 30 min at room temperature. This solution was then centrifuged at 20,000 g, 4°C for 30 min. The supernatant obtained was loaded either on a 200- μl (small-scale production) or on two 1-ml (large-scale production) *Strep*-Tactin Superflow[®] gravity flow columns (IBA Lifesciences, Göttingen, Germany). Purification was performed as described previously (Fogeron et al., 2015; Fogeron et al., 2016). The protein of interest was eluted in 100 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, and 2.5 mM D-desthiobiotin (IBA Lifesciences, Göttingen, Germany).

All experiments were assessed using 15% polyacrylamide SDS-PAGE gels. Samples were resuspended in a loading buffer containing 62.5 mM Tris-HCl pH 6.8, 10% glycerol (v/v), 2% SDS (w/v), 5% β -mercaptoethanol (v/v), and 0.01% bromophenol blue (w/v).

Bacterial Expression and Purification of *E. coli* PreS_{tag}

E. coli BL21 (DE3) or *E. coli* BL21*CP (DE3) cells were transformed with the plasmids for PreS expression or PreS plus MAPK14 co-expression, respectively, and grown at 37°C, either in LB (Lysogeny Broth) for purification setup and mass spectrometry or in M9 minimal medium containing 2 g/L of ^{13}C -labeled glucose and ^{15}N -labeled ammonium chloride for NMR spectroscopy. T7 promoter-controlled protein expression was induced at an OD_{600nm} of 1.2 using 1 mM of isopropyl- β -D-1-thiogalactopyranoside (IPTG) for 17 h at 30°C and cells were harvested (6,000 g, 20 min, 4°C). Purification procedures were the same for labeled and unlabeled samples. Cells were resuspended in 4 ml/g of cell pellet lysis buffer (20 mM NaPO₄, pH 7.5, 500 mM NaCl, 40 mM imidazole) supplemented with EDTA-free protease inhibitor (Roche). Cell lysis was performed by incubating with 1 mg/ml of lysozyme (Sigma) for 50 min at 4°C under rotation, and nucleic acids were digested with Benzonase nuclease 250 U/ μl (6 μl /L of culture) supplemented with 2 mM MgCl₂, for 30 min at room temperature. Cellular membranes were broken by passing three times through a microfluidizer (Microfluidics M-110P) at 15,000 psi. Soluble proteins were isolated by centrifugation (25,000 g, 30 min, 4°C) and filtrated using a 0.45 μm filter, before being loaded into a 5 ml HisTrap (GE Healthcare) affinity column, connected to a Biorad NGC chromatography system. The HisTrap column was washed with lysis buffer until A_{280nm} returned to baseline, and His-tagged proteins were eluted through a one-step elution with elution buffer (20 mM NaPO₄ pH 7.5, 500 mM NaCl, 1 M imidazole). The eluted proteins were dialyzed using a 3.5 kDa cut-off membrane (Spectrum labs) in TEV protease reaction buffer (50 mM Tris-HCl pH 7.5, 1 mM EDTA, 5 mM DTT), at 4°C

under slow stirring. GB1 fusion proteins were cleaved using homemade TEV protease 10 U/ μl (0.2 ml or 2,000 U per milligram of PreS protein) by incubating overnight at 4°C under rotation. Cleaved PreS protein was recovered using *Strep*-Tactin resin (IBA Lifesciences), in batch mode, following the manufacturer's recommendations. EDTA-free protease inhibitor (Roche) 1X was added to the eluted protein.

PreS was dialyzed using a 3.5 kDa cut-off membrane (spectrum labs) overnight in final NMR buffer (20 mM NaPO₄ pH 6, 50 mM NaCl). Protein concentration was determined using a Nanodrop instrument (Thermo Fisher) and the absorbance at 280 nm. Subsequently, the protein solution was concentrated by immersing the dialysis bag containing the protein and protease inhibitor into Sephadex G-25 powder at 4°C. Concentration was followed by weighing the dialysis bag before adding the powder and every 12–20 h thereafter. This approach was used to avoid PreS sticking to cellulose membranes used in concentrators as Amicon (Merck) or Vivaspinn (Sartorius). The protein concentration was measured by Nanodrop before storage at –80°C. Protein solutions in all purification steps were analyzed using 15% polyacrylamide SDS-PAGE gels. The typical yields of PreS and PreS-MAPK14 co-expressed protein were around 30 mg in LB and 10 mg per liter of M9 medium culture.

Solution-State NMR Spectroscopy

Isotopically labeled samples were dialyzed in 20 mM HEPES-KOH pH 7.5 containing 50 mM NaCl for ^{15}N - and ^2H - ^{13}C - ^{15}N -PreS_{1tag}, in 20–50 mM phosphate buffer pH 6.0 containing 50 mM NaCl for ^2H - ^{13}C - ^{15}N -PreS_{tag} and tagPreS produced in cell-free, and ^{13}C - ^{15}N -*E. coli* PreS_{tag} proteins \pm MAPK14 produced in the bacterial expression system. A pH of 6.0 could not be used for PreS_{1tag} due to its close theoretical isoelectric point of 6.37, which could result in protein aggregation. D₂O was added to a final volume of 7%, and protein concentration was quantified by NanoDrop. Concentrations were estimated for cell-free samples at 150 μM for PreS_{1tag}, 60 μM for PreS_{tag} and 120 μM for tagPreS, and 50 μM for both bacterial samples MAPK14-*E. coli* PreS_{tag} and *E. coli* PreS_{tag}. About 1 μl of 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) was added to each sample for chemical-shift referencing. NMR experiments were recorded at 298 K (25°C) on Bruker Avance II spectrometers operating at 600 MHz (tagPreS and *E. coli* PreS_{tag} samples) and 950 MHz (PreS_{1tag} and PreS_{tag}). Backbone resonances were assigned using two-dimensional (2D) BEST-TROSY spectra and three-dimensional (3D) BEST-TROSY versions of HNCA, HNcaCO, as well as HNCACB_2H, HNcoCACB_2H, and HncaCO_2H optimized for deuterated proteins (cell-free samples) (Solyom et al., 2013) and HNCACB, HNcoCACB, and HncaCO for non-deuterated proteins (*E. coli* PreS_{tag} samples). Pulse sequences were installed on the 600 MHz using the NMRlib tool (Favier and Brutscher, 2019). For details on NMR samples, NMR experiments, and assignment, see **Supplementary Table S2**. All PreS samples started to show degradation products in the NMR spectra after 1 week, even in the presence of a protease inhibitor. NMR data were processed using TopSpin 4.0.7 (Bruker) and analyzed with CcpNmr Analysis 2.4.2 (Vranken et al., 2005; Stevens et al., 2011).

For the secondary chemical shifts, the C α and C β chemical shifts of each residue for PreS1 and both PreS constructs were compared to their random coil shift taken from Wang and Jardetzky (2002). The difference between Δ C α and Δ C β were calculated, where positive values indicate the tendency to form an α -helix and negative values indicate the tendency to form a β -strand.

MALDI-TOF Mass Spectrometry

Mass spectra were acquired with a Voyager-DE PRO (Sciex, Framingham, MA, USA) equipped with a nitrogen laser emitting at 337 nm. Ions were accelerated to a final potential of 20 kV, and the mass spectrum was the sum of 300 laser shots. An external mass calibration was used (a mixture of peptides from the Sequazyme™ standards kit, AB Sciex). The analysis was performed in linear mode (instrumental mass accuracy is 0.05%). Samples were prepared by diluting 10-fold the protein solution (0.2 mg/ml) in the matrix sinapinic acid (Sigma-Aldrich, St. Louis, MI, USA), used without further purification and dissolved in 0.1%TFA/acetonitrile (70/30 v/v). About 1 μ l of the mixture was deposited onto the MALDI sample plate and let dried to complete co-crystallization.

Nano LC-MS/MS Analysis

The solutions of all purified PreS1 and PreS samples were digested overnight at 37°C with 1/100 the amount (w/w) of trypsin (trypsin porcine; Promega, Charbonnières-les-Bains, France). $_{tag}L$ protein solution was reduced with 5 mM TCEP for 45 min at 57°C, then alkylated with 10 mM iodoacetamide for 30 min in the dark at room temperature and under agitation (850 rpm) and digested overnight at 25°C with chymotrypsin 1/100 ratio (bovine pancreas chymotrypsin; Promega, Charbonnières-les-Bains, France). Peptide digests were desalted using C18 spin columns (Thermo Scientific). Peptides were dried in a speed-vac and suspended in 0.1% HCOOH. Samples were analyzed using an Ultimate 3,000 nano-RSLC (Thermo Scientific) coupled on line with a Q-Exactive HF mass spectrometer *via* a nano-electrospray ionization source in positive ionization mode (Thermo Scientific). Peptide mixtures were loaded on a C18 Acclaim PepMap100 trap-column and then separated on a C18 Acclaim Pepmap100 nano-column 50 cm \times 75 μ m i. d, 2 μ m, 100 Å (Thermo Scientific) with a 60 min linear gradient from 3.2 to 40% buffer 0.1% formic acid in © at a flow rate of 300 nl/min. Samples were analyzed using TOP20 HCD; mass data were acquired in a data-dependent strategy, selecting the fragmentation events based on the 20 most abundant precursor ions in the survey scan (375–1,600 Th). The resolution of the survey scan was 60,000 at m/z 200 Th. The ion target values for the survey scans in the Orbitrap and the MS² mode were set to 3E6 and 1E5, respectively, and the maximum injection time was set to 60 ms for both scan modes. The parameters for acquiring HCD MS/MS spectra were set to a collision energy of 27 and an isolation width of 2 m/z. The precursors with unknown charge state or a charge state of 1 were excluded. The peptides selected for MS/MS acquisition were then placed on an exclusion list for 20 s using the dynamic exclusion mode to limit duplicate spectra. Peptides were

identified by database searching using Sequest HT and MS Amanda with Proteome Discoverer 2.2 software (Thermo Scientific) against the HBV sequence. Precursor mass tolerance was set at 10 ppm, and fragment mass tolerance was set at 0.02 Da, and up to two missed cleavages were allowed. Oxidation (M), acetylation (protein N-terminus, K) and phosphorylation (S, T, Y) were set as variable modification and carbamidomethylation (C) as fixed modification. Peptides were filtered with a fixed-value PSM validator and rank 1. Phosphorylation sites were then manually validated.

Circular Dichroism

Far UV circular dichroism (CD) spectra were recorded on a Chirascan spectrometer (Applied Photophysics, Leatherhead, United Kingdom) calibrated with 1S-(+)-10-camphorsulfonic acid. Measurements were carried out at room temperature in a 0.1 cm path length quartz cuvette (Hellma). Spectra were measured in a 180–260 nm wavelength range with an increment of 0.2 nm, band pass of 0.5 nm, and integration time of 1 s. Spectra were processed, baseline-corrected, and smoothed with the Chirascan software. $_{tag}PreS$ was in 50 mM phosphate buffer pH 6.5 at a concentration of 9.5 μ M.

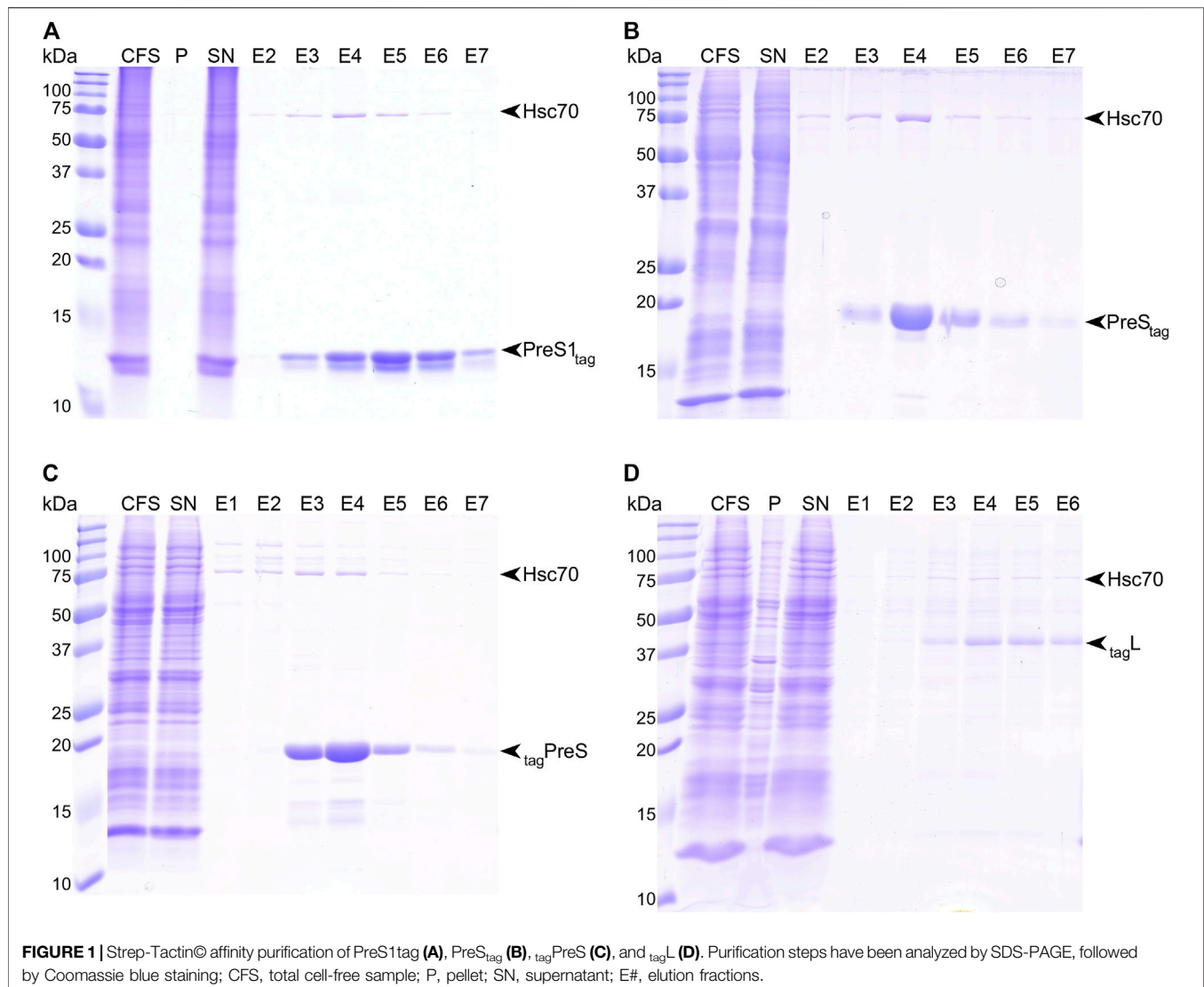
Data Availability

The ¹³C and ¹⁵N backbone chemical shifts of PreS have been deposited in the BioMagResBank (<http://www.bmrb.wisc.edu/>) under accession code 51186.

RESULTS

Cell-free Synthesis and Purification of HBV L and PreS Peptides

We synthesized full-length HBV L, as well as PreS1 and PreS fragments thereof, using wheat-germ cell-free protein synthesis (WG-CFPS) (Sawasaki et al., 2002; Takai et al., 2010; Fogeron et al., 2017; Fogeron et al., 2021). All cell-free synthesized proteins carried a Strep-tag II for affinity purification, for L at the N-terminus and for PreS1 $_{tag}$ at the C-terminus; we produced both versions for PreS, named $_{tag}PreS$ and PreS $_{tag}$. **Figure 1A** shows the results of the WG-CFPS of PreS1 $_{tag}$, followed by affinity purification, as analyzed by SDS-PAGE, followed by Coomassie blue staining. The protein was fully soluble even in the absence of detergent, as no PreS1 was detected in the pellet fraction. The band corresponding to PreS1 $_{tag}$ is clearly visible on the gel (yields are given in **Supplementary Table S1**), and interestingly, two separate bands with different intensities are observed for all elution fractions. We recently reported a similar observation for the duck HBV L protein, where it resulted from alternative translation initiation in addition to phosphorylation (David et al., 2019). Both PreS $_{tag}$ and $_{tag}PreS$ have also been successfully purified by affinity chromatography (**Figure 1B** and **Figure 1C**, respectively). Final yields (**Supplementary Table S1**) might be slightly overestimated as some degradation is visible on the gels, indicating that these constructs might not be stable on the long term unless protease inhibitors are present. A CD spectrum was recorded on the purified protein



(**Supplementary Figure S1**), which already indicates that PreS is likely unstructured.

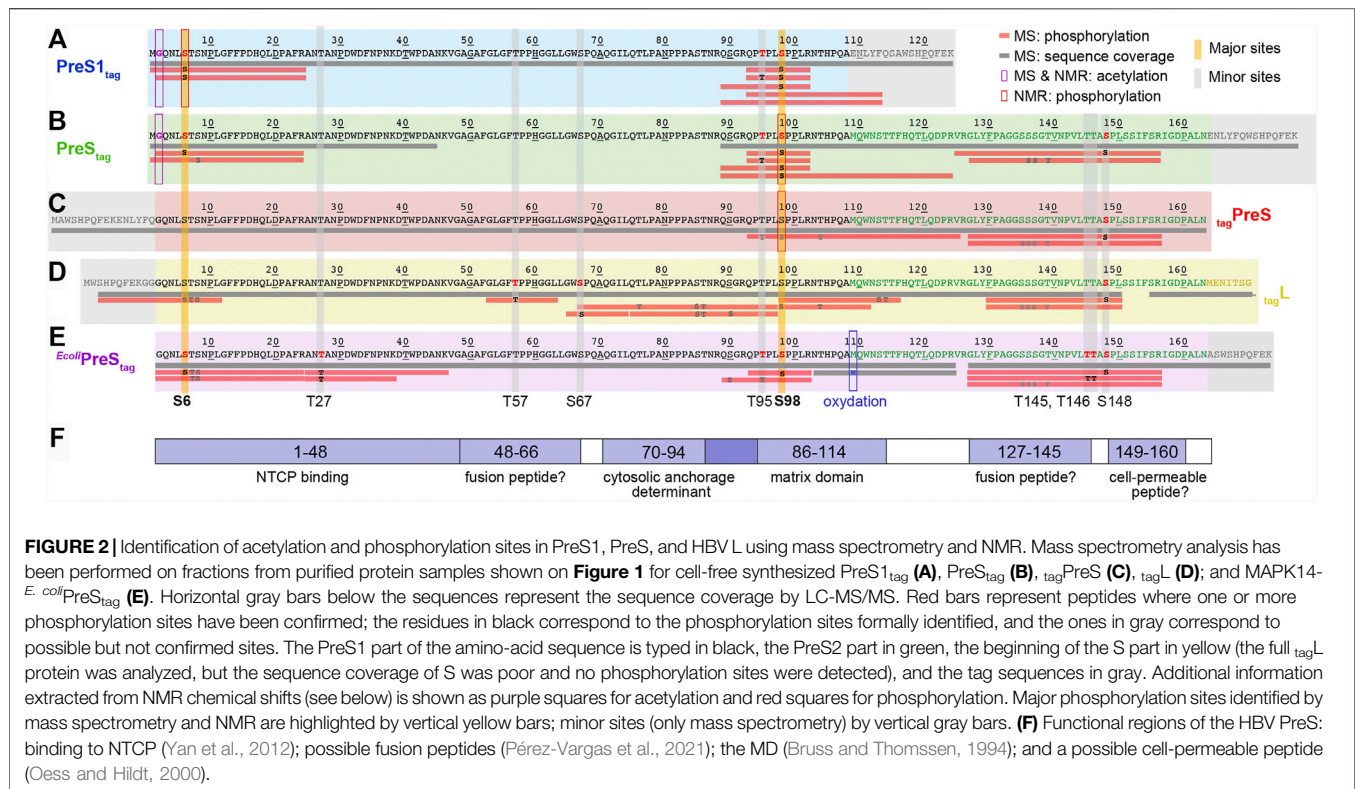
As a membrane protein, the full-length construct tag^L required the presence of detergent to be synthesized in a solubilized form for further purification. We have identified Brij-58 as a suitable detergent for the soluble production of this construct, with a concentration of 0.05% being optimal. Affinity purification was successfully performed in the presence of 0.1% n-dodecyl-β-D-maltoside (DDM) (**Figure 1D**). While L_{tag} could be expressed in a soluble form as well, the purification of this construct was unsuccessful, so it was not included in our analyses.

One can note that, for all constructs, in addition to the two bands corresponding to PreS1_{tag}, a third band eluted from the column with an apparent molecular weight below 75 kDa, which corresponds to Hsc70, as identified by the mass spectrometry of the band cut from the gel. Hsc70 is known to be an interactant of PreS in cells, and our work thus identifies at least one binding site to be localized in PreS1, in line with previous studies that assigned it to residues 70–107 of PreS (Prange et al., 1999).

In summary, all PreS-containing constructs expressed well using WG-CFPS, with yields between 0.3 and 1.7 mg protein per milliliter of wheat-germ extract (**Supplementary Table S1**), and in a soluble form, which was obtained for L through addition of detergent to the reaction. All proteins could be purified to high homogeneity, using in some cases detergent, resulting in higher purity.

Identification of PreS Phosphorylation Sites Using Mass Spectrometry

As multiple bands observed by SDS-PAGE (typically for PreS, **Figure 1A**) can be an indication of a post-translational modification such as phosphorylation (Ubersax et al., 2003) we set out to assess the modifications using mass spectrometry, in line with previous work on the DHBV L protein (David et al., 2019) and also HBV core (Heger-Stevic et al., 2018). First, the amount of post-translational modifications of PreS1_{tag} were evaluated while analyzing the total mass of the



protein by MALDI-TOF mass spectrometry. The region of interest in **Supplementary Figure S2** showed four peaks, all corresponding to PreS1_{tag}. The first peak from the left corresponds to the protein from which the N-terminal methionine (13,364.0 Da) was stripped. Peaks representing PreS1_{tag} with one additional acetyl group (+42 Da, 13,405.8 Da), with one phosphoryl group (+80 Da, 13,444.2 Da), and with both acetyl and phosphoryl groups (+122 Da, 13,486.3 Da) were observed as well. These signals indicate a single phosphorylation per chain, though not necessarily on the same residue in all proteins in the sample. Major and minor sites can overlap and yield a global sum of one site.

In order to localize the modifications on the protein sequence, PreS1_{tag} was further analyzed by LC-MS/MS mass spectrometry (**Supplementary Figure S3**), and one can see that identified peptides completely covered the sequence of PreS1_{tag} (**Figure 2A**). Phosphorylation in PreS1_{tag} was unambiguously identified for three amino acids, namely, S6, T95, and S98, which are highlighted in red bold font on the sequence. The analysis also confirmed the removal of M1, as well as acetylation on G2 (**Supplementary Figure S4**).

For the PreS_{tag} construct, even if the entire sequence was not covered, the analysis by LC-MS/MS mass spectrometry (**Figure 2B**) revealed phosphorylation on four amino acids: S6, T95, S98, and S148 (in bold red type, **Supplementary Figure S5**). Further unconfirmed sites pointed to S8, and S136/S137/T139. As for PreS1_{tag}, N-terminal methionine processing and G2 acetylation could be identified (data not shown). We also analyzed the complementary construct tagPreS, for which LC-MS/MS mass spectrometry shows full sequence coverage (**Figure 2C**).

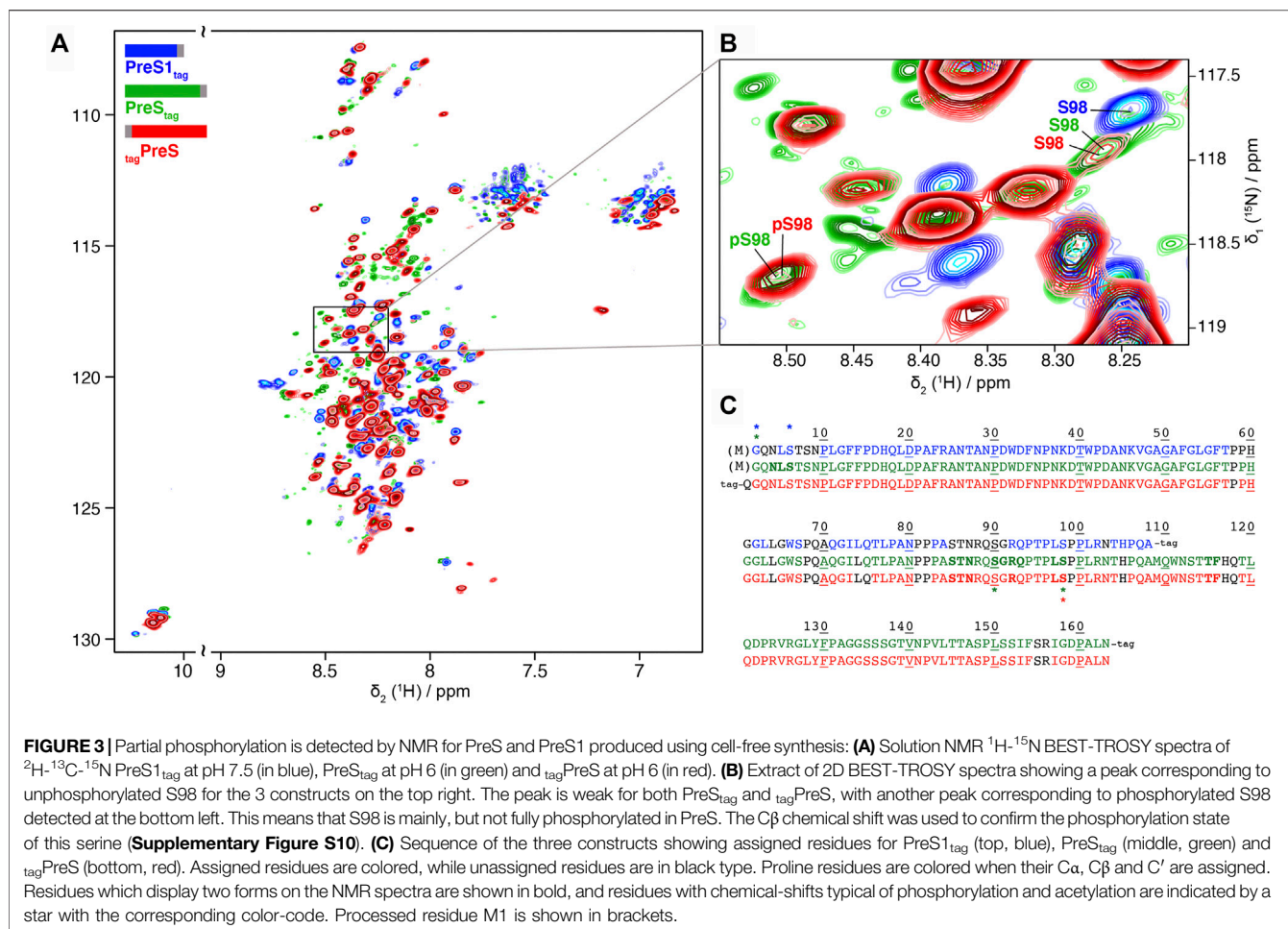
However, phosphorylation was unambiguously identified only for one single amino acid, namely, S148 (**Supplementary Figure S6**). Further ambiguous sites are proposed for T95/S98/T104, and also S135/S136/S137/T139. This indicates that the N-terminal tag interferes with phosphorylation in the N-terminal portion.

The LC-MS/MS mass spectra of HBV L (**Figure 2D**) show that the region between amino acids 50–70 is best covered in L. Three phosphorylation sites could be unambiguously identified in L by LC-MS/MS: T57, S67, and S148 (**Supplementary Figure S7**). Several ambiguous sites, including S6/T7/S8 and T76/S85/T86, were revealed as well. As the phosphorylation of S6 was clearly identified for the isolated PreS1_{tag} and PreS_{tag} forms as described above, the phosphorylation of HBV L thus most probably also occurs on this residue, and not on T7 or S8.

When combining the results obtained for L and its fragments as produced by WG-CFPS, LC-MS/MS mass spectrometry thus identified S6, T57, S67, T95, S98, and S148 as phosphorylated, and highlights other possible but unconfirmed sites. At the same time, the MALDI-TOF analyses of PreS1 suggested that only one major site exists (S6, T95, S98) and that other sites are minor. These findings are summarized in **Figures 2A–D**, where horizontal bars represent the peptides that could be analyzed, with those for which phosphorylation was unambiguously confirmed in red.

Identification of Phosphorylation Sites by NMR Spectroscopy.

We produced PreS1_{tag}, PreS_{tag}, and also tagPreS on a large scale and uniformly ²H-¹³C-¹⁵N labeled for NMR studies, with yields



between 0.3 and 0.6 mg protein per ml WGE (**Supplementary Table S1**). Solution NMR experiments were recorded on PreS1_{tag}, PreS_{tag} and tagPreS, and 2D BEST-TROSY spectra are shown in **Figure 3A**, and an extract in **Figure 3B**. The peak pattern reveals a narrow chemical shift dispersion, with $^1\text{H}_\text{N}$ resonances observed between 7.5 and 8.5 ppm, revealing with atomic detail the intrinsically disordered nature of all three PreS fragments. Backbone resonances were assigned using a combination of 3D NMR spectra (Solyom et al., 2013). Assigned residues are shown on the sequences of the three constructs in **Figure 3C**, and assignment statistics are given in **Supplementary Table S2**. A 2D-HN assigned spectrum is shown in **Supplementary Figure S8**. Secondary chemical shifts derived from the sequential assignments reveal that the proteins do not display any partial secondary structures (**Supplementary Figure S9**). HBV PreS is thus, as DHBV PreS, an intrinsically disordered protein.

NMR chemical shifts are sensitive to phosphorylation events, as addition of this group typically strongly affects the electronic environment of the neighboring spins. We thus listed random-coil chemical shifts of phosphorylated Ser and Thr residues taken from Hendus-Altanburger et al. (2019) (**Supplementary Table S3** and **Supplementary Table S4**), and compared them to all assigned chemical shifts of Ser and Thr residues in the three

samples. This allowed to clearly identify the phosphorylation of S6 in PreS1_{tag}, and partial phosphorylation of S98 in both PreS_{tag} and tagPreS constructs. S6 and S98 thus correspond to major phosphorylation sites in the WG-CFPS system. S98 is, however, only partially phosphorylated, as can be seen from the presence of a second, weaker signal for this residue at a chemical shift indicative for the non-modified residue, as shown in **Figure 3B**. The ratios of the peaks allowed to estimate that around 50% of S98 is phosphorylated in PreS_{tag} and 65% in tagPreS. The observed peak splitting could possibly also be due to cis-trans isomerization of neighboring proline residues (Hull and Kricheldorf, 1980); still, the chemical shift of the C β is typical for a phosphorylated residue (**Supplementary Figure S10**), and thus clearly points to phosphorylation.

The NMR spectra reveal phosphorylation only for a subset of the residues identified by mass spectrometry as being phosphorylated. This means that several residues, in the different constructs, are only phosphorylated on a subset of proteins in the sample, which is below the detection limit of NMR (about 10%). They thus correspond to minor sites, in agreement with the PreS1_{tag} MALDI-TOF analysis that pointed to a single major phosphorylation site in this construct (**Supplementary Figure S1**), which thus must be S6 in PreS1_{tag} according to the NMR analysis.

Interestingly, the analysis of the NMR spectra revealed signals corresponding to phosphorylation of an unconfirmed site, S90 in PreS_{tag} (**Supplementary Table S3**). This site however has not been observed in the two other PreS constructs, and its identity thus remains to be confirmed.

In addition, NMR corroborates removal of the N-terminal methionine and the acetylation of G2 for both PreS_{1tag} and PreS_{tag}, as shown in the 3D extracts of the HNcoCAB in **Supplementary Figure S11**. Indeed, a peak at 24.6 ppm displays a ¹³C chemical shift typical of an acetyl glycine, while this is not observed in N-terminally tagged_{tag}PreS and_{tag}L, where no M1 processing and acetylation can take place due to the tag.

Taken together, the NMR data show that among the phosphorylation sites identified by mass spectrometry, S6 and S98 represent major phosphorylation sites, while T27, T57, S67, T95, T145, T146, and S148 must be minor sites. This is summarized on **Figure 2** by the yellow vertical bars, annotated in bold type for major sites, and the gray bars for minor sites.

Design of an *E. coli* Expression System for Phosphorylated PreS

With major phosphorylation sites being identified according to **Figure 2**, we predicted, combining information from several web programs (Li et al., 2018; Wang et al., 2020) (<http://gps.biocuckoo.org/links.php>), the kind of kinase that could generate such a pattern, in order to design a recombinant system to produce phosphorylated PreS in larger amounts for structural and interaction studies. cdk5 and MAPK14 were the best hits, and since cdk5 is reputed to be difficult to produce in bacteria, we included MAPK14 in the *E. coli* co-expression system. The design of the plasmid was based on a previous one developed for the HBV core protein (Heger-Stevic et al., 2018), and included a cleavable GB1 fusion protein in N-terminus. Cleavage results in PreS with a C-terminal *Strep*-tag II as shown in **Supplementary Figure S12**. We expressed PreS in bacteria with and without MAPK14 co-expression and analyzed the obtained protein with mass spectrometry for phosphorylation. The total mass of the *E. coli* PreS_{tag} co-expressed with MAPK14 was analyzed by MALDI-TOF mass spectrometry, revealing the presence of heterogeneous phosphorylation, with up to five cumulative phosphorylated sites (**Supplementary Figure S13**). The identified phosphorylation sites by LC-MS/MS mass spectrometry are shown in **Figure 2E**, with confirmed sites at S6, T27, T95, S98, T145, T146, and S148 (**Supplementary Figure S14**), with T95 possibly also assigned to S90. Also, similarly as for L_{tag}, T7 and S8 have been proposed as alternatives to S6; yet with S6 clearly confirmed for PreS_{1tag} and PreS_{tag}, T7 and S8 are unlikely options. Four sites overlap with the previously identified phosphorylation sites using WG-CFPS, namely, S6, T95, S98, and S148. Surprisingly, while mass spectrometry allowed to evidence phosphorylation in *E. coli* PreS_{tag}, it could not be detected by NMR as shown in **Supplementary Figure S15**. Most likely, phosphorylation in the current recombinant system is not quantitative as seen in the mass spectrum from **Supplementary Figure S13** where the major species remains the

unphosphorylated protein, which probably places the phosphorylated residues below the detection level of 2D and 3D NMR spectra. This highlights the interesting ability of the cell-free system to efficiently induce phosphorylation by endogenous kinases.

DISCUSSION

We have shown that the HBV L protein is phosphorylated at several sites *in vitro* when synthesized both in a wheat-germ cell-free system or by bacterial co-expression with MAPK14. Proteomics studies (Mak et al., 2006) on wheat germs have highlighted the presence of several kinases, including also serine/threonine kinases; a study of kinases present in different organs of the wheat plant has revealed that germs contain a variety of kinases as well (Wei and Li, 2019). The phosphorylation of DHBV L in WG-CFPS has been observed to mainly occur at sites, followed by a proline residue, however, with exceptions (David et al., 2019). This is also the case for HBV L, where five sites are followed by Pro. Interestingly, as also for DHBV L, the N-terminal site (S6 in HBV L and S8 in DHBV L, respectively) is not preceding a Pro. Most residues for which phosphorylation could be confirmed are highly conserved in the sequence, and also most subsequent Pro residues (with the exception of 149), as shown in **Supplementary Figure S15**. An exception is S27, which is in about 20% of cases a Thr. We identified a total of nine phosphorylation sites, with several observed in different constructs by mass spectrometry and two confirmed by NMR as being major.

All identified sites are located in or just next to regions closely linked to different PreS functions [reviewed recently in Sun et al. (2021)] as shown in **Figure 2F**. Indeed, S6 is located in the interaction sequence of L with the cellular HBV receptor NTCP (Gripon et al., 2005; Yan et al., 2012). Since S6 is close to the central myristoylation site, its phosphorylation clearly has the potential to impact NTCP binding. T27 is located inside the identified NTCP-binding peptide as well. T57, S67, and T145 are located in or just next to the putative fusion peptides that have been proposed recently using a combined computational and experimental approach (Pérez-Vargas et al., 2021). Residues A70–P94 have been described to contain the amino-acid stretch that determines the cytosolic anchorage of PreS, presumably through interaction with the cognate heat shock protein Hsc70 (Prange et al., 1999). T95 and S98 are located in the so-called matrix domain (MD), which comprises the stretch of amino acids T86–T114 in the large envelope protein L. MD is the presumed interaction site of PreS with the core protein of the HBV capsid, central in the process of envelopment (Poisson et al., 1997; Le Pogam and Shih, 2002). This region is crucial for virion formation, and is believed to establish contact to the nucleocapsid, since truncations up to G92 were compatible with envelopment (Bruss and Thomssen, 1994), and several point mutations within the MD-blocked virion formation (Bruss, 1997). It has been shown also that a peptide comprising this domain interacts with the core particle (Poisson et al., 1997). Phosphorylation on T95 and S98 is centrally located in this

domain and has the potential to change the required interaction interfaces, resulting in productive envelopment. S148 is located just downstream from the proposed amphipathic PreS2 translocation motif (Oess and Hildt, 2000; Sun et al., 2021). It is intriguing that in, or just next to, each proposed functional sites, phosphorylation is observed in the present experiments, pointing to a possible role thereof in the regulation of PreS function. No role for phosphorylation has yet been identified for the phosphorylation sites in DHBV PreS (Grgacic and Anderson, 1994; Grgacic et al., 1998). However, DHBV differs from HBV in fundamental aspects, including by a much-larger core protein (Makbul et al., 2020) and the lack of an HBx-like transactivator that is crucial in HBV infection (Slagle and Bouchard, 2018). Hence, a functional relevance of PreS phosphorylation in the human virus would not be surprising.

CONCLUSION

We here reported phosphorylation on nine different sites of the PreS domain of the human HBV L envelope protein, enabled by endogenous kinases in the wheat germ extract used for cell-free protein synthesis, or in *E. coli* by co-expression with the MAPK14 kinase. We identified two major phosphorylation sites, on S6 and S98, and seven minor ones, using a combination of mass spectrometry and NMR. We found phosphorylation to occur in all major functional regions of PreS, which raises the possibility that phosphorylation is involved in the regulation of these functions in the human virus protein, although this has not been identified in the avian virus homolog. We however found also that full phosphorylation is provided in neither recombinant system, which positions phosphorylation mimics by S/T to E mutations as the best strategy to explore the impact of phosphorylation on PreS interactions in structural studies.

REFERENCES

- Badillo, A., Receveur-Brechot, V., Sarrazin, S., Cantrelle, F.-X., Delolme, F., Fogeron, M.-L., et al. (2017). Overall Structural Model of NS5A Protein from Hepatitis C Virus and Modulation by Mutations Confering Resistance of Virus Replication to Cyclosporin a. *Biochemistry* 56, 3029–3048. doi:10.1021/acs.biochem.7b00212
- 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. doi:10.1128/jvi.71.12.9350-9357.1997
- Bruss, V., and Thomssen, R. (1994). Mapping a Region of the Large Envelope Protein Required for Hepatitis B Virion Maturation. *J. Virol.* 68, 1643–1650. doi:10.1128/jvi.68.3.1643-1650.1994
- David, G., Fogeron, M.-L., Montserret, R., Lecoq, L., Page, A., Delolme, F., et al. (2019). Phosphorylation and Alternative Translation on Wheat Germ Cell-free Protein Synthesis of the DHBV Large Envelope Protein. *Front. Mol. Biosci.* 6, 138. doi:10.3389/fmolb.2019.00138
- David, G., Fogeron, M. L., Schledorn, M., Montserret, R., Haselmann, U., Penzel, S., et al. (2018). Structural Studies of Self-Assembled Subviral Particles: Combining Cell-Free Expression with 110 kHz MAS NMR Spectroscopy. *Angew. Chem. Int. Ed.* 57, 4787–4791. doi:10.1002/anie.201712091
- Delgado, C. L., Núñez, E., Yélamos, B., Gómez-Gutiérrez, J., Peterson, D. L., and Gavilanes, F. (2015). Study of the Putative Fusion Regions of the PreS Domain

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://bmr.io>, 51186.

AUTHOR CONTRIBUTIONS

AB, M-LF, and LL conceived and planned the experiments. M-LF, RM, and GD carried out cell-free syntheses, LC carried out bacterial expression, and MN designed bacterial expression vectors. LL carried out NMR analyses. AP and FD carried out mass spectrometry analyses. M-LF, LL, MN, and AB interpreted results and wrote the manuscript, with input from all authors.

ACKNOWLEDGMENTS

Financial support from the ANRS (ECTZ71388), the CNRS (CNRS-Momentum 2018) and the IR-RMN-THC Fr3050 CNRS for conducting the research is gratefully acknowledged. The authors acknowledge the financial support from ITMO Cancer AVIESAN (Alliance Nationale pour les Sciences de la Vie et de la Santé, National Alliance for Life Sciences and Health) within the framework of the cancer plan for Orbitrap mass spectrometer funding.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2021.821755/full#supplementary-material>

- of Hepatitis B Virus. *Biochim. Biophys. Acta (Bba) - Biomembranes* 1848, 895–906. doi:10.1016/j.bbmem.2014.12.020
- Favier, A., and Brutscher, B. (2019). NMRlib: User-Friendly Pulse Sequence Tools for Bruker NMR Spectrometers. *J. Biomol. NMR* 73, 199–211. doi:10.1007/s10858-019-00249-1
- Fogeron, M.-L., Badillo, A., Penin, F., and Böckmann, A. (2017). Wheat Germ Cell-free Overexpression for the Production of Membrane Proteins. *Methods Mol. Biol.* 1635, 91–108. doi:10.1007/978-1-4939-7151-0_5
- Fogeron, M.-L., Jirasko, V., Penzel, S., Paul, D., Montserret, R., Danis, C., et al. (2016). Cell-free Expression, Purification, and Membrane Reconstitution for NMR Studies of the Nonstructural Protein 4B from Hepatitis C Virus. *J. Biomol. NMR* 65, 87–98. doi:10.1007/s10858-016-0040-2
- Fogeron, M.-L., Lecoq, L., Cole, L., Harbers, M., and Böckmann, A. (2021). Easy Synthesis of Complex Biomolecular Assemblies: Wheat Germ Cell-free Protein Expression in Structural Biology. *Front. Mol. Biosci.* 8, 639587. doi:10.3389/fmolb.2021.639587
- Fogeron, M.-L., Paul, D., Jirasko, V., Montserret, R., Lacabanne, D., Molle, J., et al. (2015). Functional Expression, Purification, Characterization, and Membrane Reconstitution of Non-structural Protein 2 from Hepatitis C Virus. *Protein Expr. Purif.* 116, 1–6. doi:10.1016/j.pep.2015.08.027
- Glebe, D., Urban, S., Knoop, E. V., Çağ, N., Krass, P., Grün, S., et al. (2005). Mapping of the Hepatitis B Virus Attachment Site by Use of Infection-Inhibiting preS1 Lipopeptides and tupaia Hepatocytes. *Gastroenterology* 129, 234–245. doi:10.1053/j.gastro.2005.03.090

- Grgacic, E. V., and Anderson, D. A. (1994). The Large Surface Protein of Duck Hepatitis B Virus Is Phosphorylated in the Pre-S Domain. *J. Virol.* 68, 7344–7350. doi:10.1128/jvi.68.11.7344-7350.1994
- Grgacic, E. V., Gazina, E. V., Lin, B., Anderson, D. A., and Snooks, M. J. (1998). Normal Phosphorylation of Duck Hepatitis B Virus L Protein Is Dispensable for Infectivity. *J. Gen. Virol.* 79 (Pt 11), 2743–2751. doi:10.1099/0022-1317-79-11-2743
- Gripon, P., Cannie, I., and Urban, S. (2005). Efficient Inhibition of Hepatitis B Virus Infection by Acylated Peptides Derived from the Large Viral Surface Protein. *J. Virol.* 79, 1613–1622. doi:10.1128/JVI.79.3.1613-1622.2005
- Gudima, S., Meier, A., Dunbrack, R., Taylor, J., and Bruss, V. (2007). Two Potentially Important Elements of the Hepatitis B Virus Large Envelope Protein Are Dispensable for the Infectivity of Hepatitis Delta Virus. *J. Virol.* 81, 4343–4347. doi:10.1128/JVI.02478-06
- Heger-Stevic, J., Zimmermann, P., Lecoq, L., Böttcher, B., and Nassal, M. (2018). Hepatitis B Virus Core Protein Phosphorylation: Identification of the SRPK1 Target Sites and Impact of Their Occupancy on RNA Binding and Capsid Structure. *Plos Pathog.* 14, e1007488. doi:10.1371/journal.ppat.1007488
- Hendus-Altenburger, R., Fernandes, C. B., Bugge, K., Kunze, M. B. A., Boomsmma, W., and Kragelund, B. B. (2019). Random Coil Chemical Shifts for Serine, Threonine and Tyrosine Phosphorylation over a Broad pH Range. *J. Biomol. NMR* 73, 713–725. doi:10.1007/s10858-019-00283-z
- Hull, W. E., and Kricheldorf, H. R. (1980). 15N-nmr Spectroscopy. 20.Cis/trans Isomerism and Neighboring Residue Effects of Proline-Containing peptides/Cis/trans Isomerism and Neighboring Residue Effects of Proline-Containing Peptides. *Biopolymers* 19, 1103–1122. doi:10.1002/bip.1980.360190602
- Le Pogam, S., and Shih, C. (2002). Influence of a Putative Intermolecular Interaction between Core and the Pre-S1 Domain of the Large Envelope Protein on Hepatitis B Virus Secretion. *J. Virol.* 76, 6510–6517. doi:10.1128/JVI.76.13.6510-6517.2002
- Li, F., Li, C., Marquez-Lago, T. T., Leier, A., Akutsu, T., Purcell, A. W., et al. (2018). Quokka: a Comprehensive Tool for Rapid and Accurate Prediction of Kinase Family-specific Phosphorylation Sites in the Human Proteome. *Bioinformatics* 34, 4223–4231. doi:10.1093/bioinformatics/bty522
- Mak, Y., Skylas, D. J., Willows, R., Connolly, A., Cordwell, S. J., Wrigley, C. W., et al. (2006). A Proteomic Approach to the Identification and Characterisation of Protein Composition in Wheat Germ. *Funct. Integr. Genomics* 6, 322–337. doi:10.1007/s10142-005-0018-8
- Makbul, C., Nassal, M., and Böttcher, B. (2020). Slowly Folding Surface Extension in the Prototypic Avian Hepatitis B Virus Capsid Governs Stability. *eLife* 9, e57277. doi:10.7554/eLife.57277
- Núñez, E., Yélamos, B., Delgado, C., Gómez-Gutiérrez, J., Peterson, D. L., and Gavilanes, F. (2009). Interaction of preS Domains of Hepatitis B Virus with Phospholipid Vesicles. *Biochim. Biophys. Acta (Bba) - Biomembranes* 1788, 417–424. doi:10.1016/j.bbmem.2008.10.014
- Oess, S., and Hildt, E. (2000). Novel Cell Permeable Motif Derived from the PreS2-Domain of Hepatitis-B Virus Surface Antigens. *Gene Ther.* 7, 750–758. doi:10.1038/sj.gt.3301154
- Pérez-Vargas, J., Teppa, E., Amirache, F., Boson, B., Pereira de Oliveira, R., Combet, C., et al. (2021). A Fusion Peptide in preS1 and the Human Protein Disulfide Isomerase ERp57 Are Involved in Hepatitis B Virus Membrane Fusion Process. *eLife* 10, e64507. doi:10.7554/eLife.64507
- Persing, D. H., Varmus, H. E., and Ganem, D. (1987). The preS1 Protein of Hepatitis B Virus Is Acylated at its Amino Terminus with Myristic Acid. *J. Virol.* 61, 1672–1677. doi:10.1128/JVI.61.5.1672-1677.1987
- Poisson, F., Severac, A., Hourieux, C., Goudeau, A., and Roingeard, P. (1997). Both pre-S1 and S Domains of Hepatitis B Virus Envelope Proteins Interact with the Core Particle. *Virology* 228, 115–120. doi:10.1006/viro.1996.8367
- Prange, R. (2012). Host Factors Involved in Hepatitis B Virus Maturation, Assembly, and Egress. *Med. Microbiol. Immunol.* 201, 449–461. doi:10.1007/s00430-012-0267-9
- Prange, R., Werr, M., and Löffler-Mary, H. (1999). Chaperones Involved in Hepatitis B Virus Morphogenesis. *Biol. Chem.* 380, 305–314. doi:10.1515/BC.1999.042
- Rothmann, K., Schnölzer, M., Radziwill, G., Hildt, E., Moelling, K., and Schaller, H. (1998). Host Cell-Virus Cross Talk: Phosphorylation of a Hepatitis B Virus Envelope Protein Mediates Intracellular Signaling. *J. Virol.* 72, 10138–10147. doi:10.1128/jvi.72.12.10138-10147.1998
- Sawasaki, T., Ogasawara, T., Morishita, R., and Endo, Y. (2002). A Cell-free Protein Synthesis System for High-Throughput Proteomics. *Proc. Natl. Acad. Sci.* 99, 14652–14657. doi:10.1073/pnas.232580399
- Schmidt, T. G., and Skerra, A. (2007). The Strep-Tag System for One-step Purification and High-Affinity Detection or Capturing of Proteins. *Nat. Protoc.* 2, 1528–1535. doi:10.1038/nprot.2007.209
- Slagle, B. L., and Bouchard, M. J. (2018). Role of HBx in Hepatitis B Virus Persistence and its Therapeutic Implications. *Curr. Opin. Virol.* 30, 32–38. doi:10.1016/j.coviro.2018.01.007
- Solyom, Z., Schwarten, M., Geist, L., Konrat, R., Willbold, D., and Brutscher, B. (2013). BEST-TROSY Experiments for Time-Efficient Sequential Resonance Assignment of Large Disordered Proteins. *J. Biomol. NMR* 55, 311–321. doi:10.1007/s10858-013-9715-0
- Stevens, T. J., Fogh, R. H., Boucher, W., Higman, V. A., Eisenmenger, F., Bardiaux, B., et al. (2011). A Software Framework for Analysing Solid-State MAS NMR Data. *J. Biomol. NMR* 51, 437–447. doi:10.1007/s10858-011-9569-2
- Sun, H., Chang, L., Yan, Y., and Wang, L. (2021). Hepatitis B Virus pre-S Region: Clinical Implications and Applications. *Rev. Med. Virol.* 31, e2201. doi:10.1002/rmv.2201
- Takai, K., Sawasaki, T., and Endo, Y. (2010). Practical Cell-free Protein Synthesis System Using Purified Wheat Embryos. *Nat. Protoc.* 5, 227–238. doi:10.1038/nprot.2009.207
- Ubersax, J. A., Woodbury, E. L., Quang, P. N., Paraz, M., Blethrow, J. D., Shah, K., et al. (2003). Targets of the Cyclin-dependent Kinase Cdk1. *Nature* 425, 859–864. doi:10.1038/nature02062
- Urban, S., Schwarz, C., Marx, U. C., Zentgraf, H., Schaller, H., and Multhaup, G. (2000). Receptor Recognition by a Hepatitis B Virus Reveals a Novel Mode of High Affinity Virus-Receptor Interaction. *EMBO J.* 19, 1217–1227. doi:10.1093/emboj/19.6.1217
- Vranken, W. F., Boucher, W., Stevens, T. J., Fogh, R. H., Pajon, A., Llinas, M., et al. (2005). The CCPN Data Model for NMR Spectroscopy: Development of a Software Pipeline. *Proteins* 59, 687–696. doi:10.1002/prot.20449
- Wang, C., Xu, H., Lin, S., Deng, W., Zhou, J., Zhang, Y., et al. (2020). GPS 5.0: An Update on the Prediction of Kinase-specific Phosphorylation Sites in Proteins. *Genomics, Proteomics & Bioinformatics* 18, 72–80. doi:10.1016/j.gpb.2020.01.001
- Wang, Y., and Jardetzky, O. (2002). Probability-based Protein Secondary Structure Identification Using Combined NMR Chemical-Shift Data. *Protein Sci.* 11, 852–861. doi:10.1110/ps.3180102
- Wei, K., and Li, Y. (2019). Functional Genomics of the Protein Kinase Superfamily from Wheat. *Mol. Breed.* 39, 1–23. doi:10.1007/s11032-019-1045-9
- Xi, J., Liu, H., and Hu, J. (2021). Regulation of Hepatitis B Virus Virion Release and Envelopment Timing by Nucleocapsid and Envelope Interactions. *J. Virol.* JVI0130521. doi:10.1128/JVI.01305-21
- Yan, H., Zhong, G., Xu, G., He, W., Jing, Z., Gao, Z., et al. (2012). Sodium Taurocholate Cotransporting Polypeptide Is a Functional Receptor for Human Hepatitis B and D Virus. *eLife* 1, e00049. doi:10.7554/eLife.00049
- Yang, F. (2018). Post-translational Modification Control of HBV Biological Processes. *Front. Microbiol.* 9, 2661. doi:10.3389/fmicb.2018.02661

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Fogeron, Lecoq, Cole, Montserret, David, Page, Delolme, Nassal and Böckmann. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.