UNIL | Université de Lausanne

Master Thesis in Medical Studies

Biochemical Characterization of Hepatitis E Virus ORF3 Protein

Student
Maxime Matter

Tutor
Prof. Darius Moradpour, MD
Service of Gastroenterology and Hepatology, CHUV

Co-tutor
PD Dr. Jérôme Gouttenoire, PhD
Service of Gastroenterology and Hepatology, CHUV

Expert
Prof. Stefan Kunz, PhD
Institute of Microbiology, CHUV

Faculty of Biology & Medicine, University of Lausanne
December 2016

Table of contents

Table of contents 1

Abstract 2

1. Introduction 3

1.1 Hepatitis E virus 3

1.2 Classification and genetic heterogeneity 3

1.3 Route of transmission 4

1.4 Clinical presentation of hepatitis E 5

Acute Hepatitis 5

Chronic hepatitis E 6

Extrahepatic manifestations 6

1.5 Diagnosis and treatment 6

1.6 HEV genetic organisation 7

1.7 Viral life cycle 7

1.8 ORF3 protein, current knowledge 8

2. Aim of the study 10

3. Materials and methods 11

3.1 Reagents 11

3.2 Cell culture 11

3.3 Transfection 11

3.4 SDS-PAGE and western blot 12

3.5 Immunoprecipitation 13

3.6 Coomassie blue staining 13

3.7 Deglycosylation assay 13

4. Results 15

4.1 ORF3 has a different apparent molecular weight when expressed in mammalian cells *vs* cell-free expression systems 16

4.2 Co-expression of the ORF2 capsid protein does not influence ORF3 apparent molecular weight 17

4.3 Establishment of an immunoprecipitation protocol to purify an ORF3-FLAG protein expressed in mammalian cells 18

4.4 Investigation of potential post-translational modifications of ORF3 protein 20

5. Discussion 22

6. Conclusion 26

7. Acknowledgments 27

8. References 27

Abstract

Background and Aim: Hepatitis E virus (HEV) infection is the most common cause of acute hepatitis worldwide. This faecal-orally transmitted virus is responsible for epidemics and sporadic cases of hepatitis E in emerging countries but has now been recognized as a zoonotic infection in developed countries through consumption of contaminated food. HEV is a small icosahedral, non-enveloped virus with a positive-strand RNA genome of 7.2 kb encoding three open reading frames (ORFs). ORF1 encodes the replicase, ORF2 the viral capsid and ORF3 a small, hitherto poorly characterized protein. To date, ORF3 protein is believed to play an important role in particle secretion. The aim of this study was to further characterize the ORF3 protein by biochemical approaches.

Methods: Various plasmids allowing the expression of wild-type and mutant ORF3 protein were transfected into mammalian cells. Cell free expression systems were employed to translate recombinant ORF3 protein in vitro. Protein lysates were analysed by SDS-PAGE and western blot (WB) analyses. An immunoprecipitation protocol using FLAG® superscript M2 magnetic beads was developed to purify ORF3-FLAG protein for mass spectrometry (MS) analysis.

Results: We observed that ORF3 protein displays different and higher apparent molecular weight (MW) when expressed in mammalian cells as compared to cell free expression system. Furthermore, the observed higher MW was not due to potential phosphorylation at Ser 70. In addition, co-expression of the ORF2 capsid protein did not influence the apparent MW of ORF3 protein. We successfully optimized an immunoprecipitation protocol to efficiently and specifically recover a C-terminally FLAG-tagged ORF3 protein expressed in mammalian cells. Despite efforts to increase the amount of protein recovered for MS analysis, the latter did not reveal any substantial modification in ORF3 protein. Finally, a more targeted approach did not reveal any glycosylation of ORF3 protein.

Conclusion: Our findings reveal that HEV ORF3 protein has a higher MW than expected and suggest that it is likely post-translationally modified. Therefore, it opens new perspectives in the understanding of this viral protein as well as of the HEV life cycle.

Key words: HEV - post-translational modification - signal peptide - viral Hepatitis

1. Introduction

1.1 Hepatitis E virus

The hepatitis E virus (HEV) is the most recently described of the five hepatotropic viruses, identified by molecular cloning in the early 1990s [1]. HEV was previously considered to be endemic in emerging countries and responsible for several epidemics as well as sporadic cases of hepatitis. In developed countries the cases of HEV infection were initially attributed to travellers returning from endemic areas [2]. However, a substantial number of locally acquired cases of infection has been diagnosed and appears to be increasing over the last few years in developed countries with an estimated seroprevalence of 7 % to 21 % [3]. In developing countries, the virus is responsible for more than 50% of acute hepatitis although high variations between the different regions are observed [4]. Finally, it is estimated that 20.1 million people have been infected with HEV genotypes 1 and 2 in 2005, in nine different regions representing 71% of the world population [5]. Therefore, it is now well established that this virus is present worldwide as a common cause of acute hepatitis and becomes a public health concern.

1.2 Classification and genetic heterogeneity

Hepatitis E virus belongs to the *Hepeviridae* family which includes the mammalian HEV infecting humans and animals but also avian, cutthroat trout or bats viruses [6, 7]. HEV which infects humans is classified into 4 major genotypes with several subtypes. All these genotypes are represented in a single serotype. These different genotypes show a different geographical world distribution (Figure 1).
Genotype 1 (gt 1) is mainly present in Asia and Africa where it is highly endemic, and also found in travellers from non-endemic regions. Genotype 2 (gt 2) is present in Mexico, Nigeria and Chad. Both gt 1and gt 2 only infect humans and have been identified in many outbreaks and sporadic cases among developing countries, gt 1 being responsible for the majority of these cases [8, 9]. Moreover, their prevalence for 2005 in Africa and Asia is approximately 20.1 million incident HEV infections with 3.4 million cases of symptomatic disease, 70,000 deaths, and 3,000 stillbirths [5, 10]. Genotype 3 (gt 3) is the most frequent type with a worldwide distribution including USA, North America, European countries, New-Zealand, China, and Japan. Genotype 4 (gt 4) is predominant in China, India, Indonesia, Vietnam, Taiwan and Japan. Both gt 3 and gt 4 infect humans but also many different animal species particularly pigs, swine, deer and cattle constituting a reservoir for HEV with a possible transmission to human by crossing species barrier. Hepatitis E is thus defined as a zoonotic disease in developed countries where gt 3 is predominant [4]. These two genotypes are mainly known to cause sporadic cases and rarely epidemic but seem to be less pathogenic than gt 1 and 2 in humans [2, 11].

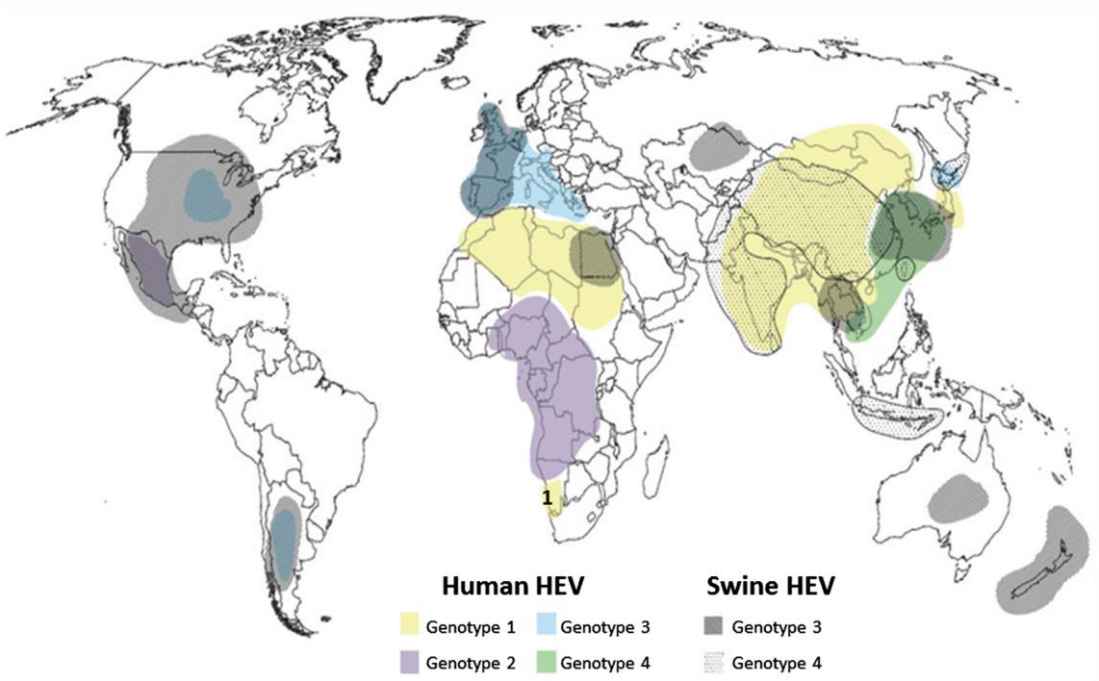

Figure 1. Worldwide distribution of HEV genotypes in human and swine [4].

1.3 Route of transmission

The way humans and/or animals are infected with HEV differs significantly between the different geographic regions in the world (endemic/non-endemic areas) and can be presented either as epidemic and sporadic events.

In developing countries, faecal-oral transmission is mainly due to the consumption of contaminated water with animals and human sewage which occur specially during floods and torrential rain associated with poor sanitation and hygiene. This way of contamination is responsible for epidemic that occurred in these regions where HEV is endemic. It represents the most common route of transmission of HEV especially for gt 1 and 2 as they are limited to humans [2, 4, 5].

In industrialized countries, hepatitis E is considered as a zoonotic disease with infected pigs as main reservoir. Food-borne contamination is predominant in these regions by eating raw or undercooked meat from infected animals and is mainly responsible for the transmission of gt 3 and 4. Fruits and vegetables washed with contaminated water and seafood have also been reported as vector of HEV in humans as well [11, 4]. Other transmission routes that are less frequently implicated include parenteral transmission *via* transfusion of infected blood products or materno-fetal transmission (*in utero* contamination). Finally, direct contacts between infected pigs or humans have been observed but are very uncommon [4]. Nevertheless, it appears that the precise way of acquisition of HEV could not be determined precisely in many cases from non-endemic regions and sporadic cases in endemic countries [10].

1.4 Clinical presentation of hepatitis E

Acute Hepatitis

Clinical presentations of acute hepatitis E are very similar in endemic and non-endemic countries, ranging from subclinical forms to fulminant liver failure. Acute hepatitis E is a self-limited disease, with a spontaneous recovery occurring generally after 4 to 7 weeks [4, 8, 12]. Shortly after the infection HEV RNA is found in the serum and in stools. The incubation period lasts from 2 to 10 weeks, followed by the elevation of transaminase (ALT) together with the onset of symptoms and finally by the apparition of IgM and IgG antibodies persisting for 3 to 12 months and for years respectively [2, 4]. The duration of viremia is variable (Fig. 2). Although 67-98% of the HEV infections are asymptomatic, jaundice is the main symptom reported [4]. The clinical phase associated with the elevation of transaminase (ALT) is also characterized by various non-specific symptoms including fever, flu-like symptoms, joint and abdominal pain, vomiting, anorexia, weight loss and myalgia [13, 14].
In endemic regions, where gt 1 and 2 are predominant, HEV can affect all age groups but some populations are particularly at risk such as young adults (15-45 years), patients with pre-existing chronic liver disease and pregnant women. The latter are associated with a higher risk to develop fulminant hepatitis (acute liver failure) with a mortality rate of approximatively 20-30% mostly during the third trimester [5, 12]. It is associated with an elevated risk of mother and fetus death, abortion, prematurity or low birthweight [2, 15, 16]. However, HEV gt 3, the most frequent in industrialized countries, may have a predilection for middle aged and elderly persons with a predominance for men specially those affected by coexistent disease conditions, including chronic liver disease [2].

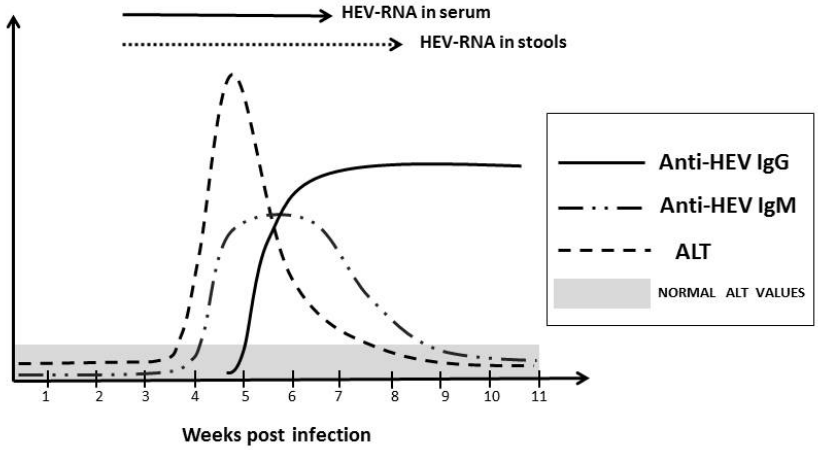

Figure 2. Clinical course of acute hepatitis E [8].

Of note, this disease is often misdiagnosed with other aetiologies of acute hepatitis for instance HAV, HBV, EBV, drug-induced liver injury, autoimmune or alcoholic related hepatitis. This could explain the high variability and underestimated prevalence allocated to HEV specially in developed countries [12].

Chronic hepatitis E

Chronic hepatitis E is defined by the persistence of HEV-RNA in serum and/or stools, IgM antibodies and elevated liver enzymes for more than 6 months [4]. To date, it has only been documented with HEV gt 3, 4 and 7 (also known as camel HEV) in immunocompromised patients such as solid-organ transplant recipients, post-chemotherapy, haematological malignancies and HIV-infected patients [2, 4, 12]. The prevalence of chronic disease in immunosuppressed patients is of approximately 1% in European countries however nearly 60% of them may develop liver disease [12]. The clinical feature among these patients is usually pauci-symptomatic only one third of them expressing symptoms and is associated with an increased risk to develop liver fibrosis and also cirrhosis if not treated early on [9].

Extrahepatic manifestations

Diverse extrahepatic manifestations have been described to be associated to HEV infection, most often with HEV gt 3.
Neurological complications have been described such as Guillain-Barré and Turner syndromes, Bell's palsy, acute transversal myelitis, inflammatory polyradiculitis, encephalitis, acute meningo-encephalitis, myopathy, neuralgic amyotrophy and peripheral neuropathy [8, 9]. These manifestations usually appear concurrently with the increase of IgM antibodies in serum and concern approximately 5% of the infected individuals, half of them being chronic hepatitis E patients [17].
Pancreatitis is another complication that may occurs few weeks after the onset of jaundice but spontaneously recovers in most of the time [18]. Some immune-mediated haematological diseases have also been described among these thrombocytopenia, and haemolytic anemia. Other immunological disorders have been reported such as membrano-proliferative or membranous glomerulonephritis and Henoch-Schonlein purpura [4, 8]. Finally, rash and arthralgia could also be observed in infected patients with HEV [19].

1.5 Diagnosis and treatment

Hepatitis E is diagnosed either indirectly by dosing the IgM and/or IgG anti-HEV antibodies in the serum or directly with the detection of HEV-RNA in serum and/or stool by PCR. The latter technique is the most accurate to confirm the disease and may allow theoretically the detection of the virus up to two weeks after appearance of the clinical symptoms [4].
Although, two capsid-based vaccines have been developed and have proven efficacy, they are licensed and available only in China [23].
Thought most of patients spontaneously recover from this infection, no treatment is usually required. Currently no direct antiviral agent against HEV is available but treatment with ribavirin as well as pegylated interferon-α have demonstrated a substantial efficacy of about 80% in chronic hepatitis E [20, 21]. Moreover, reduction of immunosuppressing drugs (e.g. Tacrolimus) is an alternative which proved efficacy in HEV-infected transplant recipients [9, 22].

1.6 HEV genetic organisation

HEV is a small icosahedral (30-33nm in diameter) non-enveloped virus with a positive single-stranded RNA genome of 7.2kb. The genome is composed of two non-coding regions at both extremities (NCR), a 5'end with a 7-methylguanine cap, a polyadenylated 3'end and three open reading frames (ORFs) that encode at least 3 viral proteins [24, 25]. ORF1 encodes a methyltransferase, a Y-domain, a so-called papain-like cysteine protease (PCP) domain, a Macro domain (also named X), a helicase (Hel) and a RNA-dependant RNA polymerase (RdRp) involved in the viral replication. ORF1 also contains a proline-rich hypervariable region (HVR) between the *C* terminus of PCP and the *N* terminus of the Macro domain [10]. ORF2 at the 3'end of the genome, encodes a 660 amino acids (aa) capsid protein that oligomerizes to encapsidate the genome and form the virion [10]. The capsid protein possesses three glycosylation sites which play an important role in the formation of the viral particle and in the interaction with host cells [24]. ORF3 overlaps the ORF2 capsid gene and is translated from a subgenomic RNA into a small protein whose precise functions and three-dimensional structure still remain unclear.

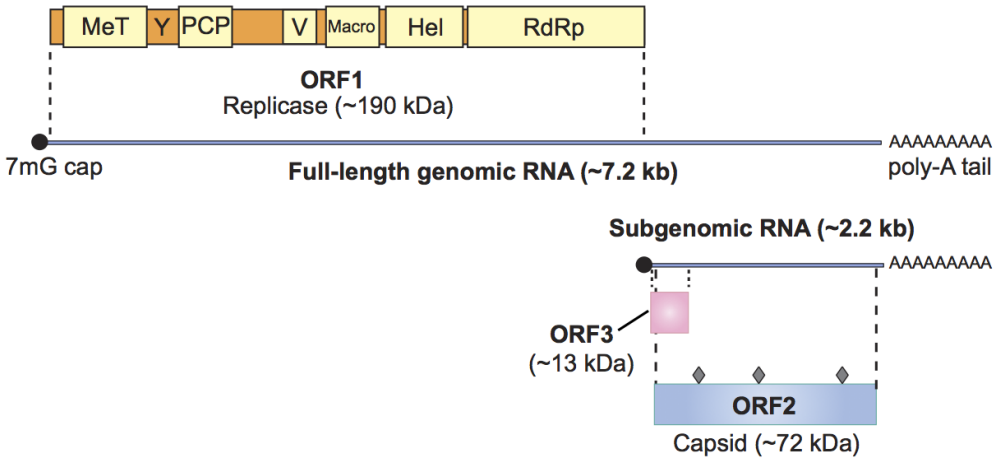

Figure 3. HEV genome organisation [26]. The 7.2 kb positive-strand RNA genome has a 5' 7-methylguanylate cap and a 3' polyadenylated tail. It comprises three ORFs. ORF1 encodes a polyprotein of about 190 kDa. ORF2 and ORF3 are translated from a 2.2 kb subgenomic RNA generated during viral replication.

1.7 Viral life cycle

The life cycle of HEV is still poorly understood mainly due to several technical hurdles that are hopefully being overcome in the near future thanks to newly developed HEV infectious clone [27].
First, it has been shown that viral particle attachment is mediated by the interaction of the capsid protein ORF2 with two types of molecules, e.g. heat shock cognate protein 70 (HSC70) and heparin sulfate proteoglycans (HSPGs) both located on the cell surface [4]. The receptor(s) triggering the entry of the virus is still unknown, however this step involved the clathrin-mediated endocytosis. Upon entry, the virus releases the RNA genome which is then uncoated and serve as template for the translation of the ORF1 protein. This step is initiated by the recruitment of the 40S ribosomal unit by the 7-methylguanine cap of the RNA genome [24]. A negative single-stranded RNA intermediate is then synthetised by the viral RdRp harboured by ORF1. This intermediate RNA species serves as a template for *de novo* synthesis of full-length RNA genomes as well as of subgenomic RNA which is

further translated in ORF2 and ORF3 proteins [4]. Finally, the ORF2 protein is likely involved in the packaging of the viral genome into the newly formed particles. ORF3 is believed to play a crucial role in the secretion of infectious virion.

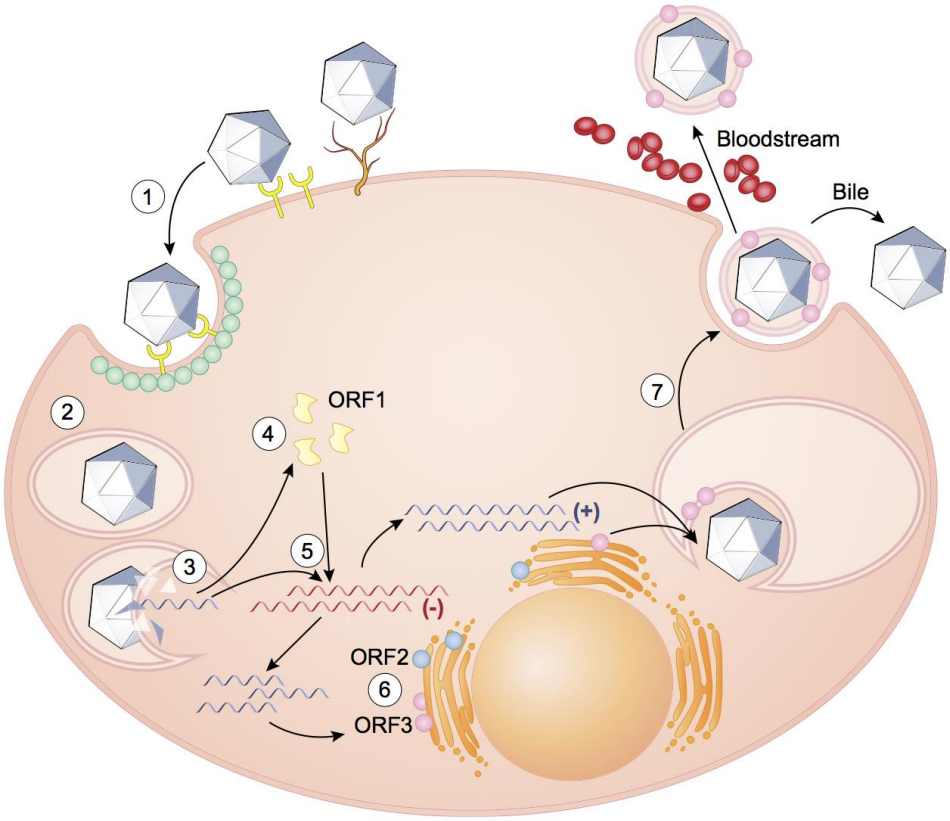

Figure 4. HEV life cycle [26]. The HEV life cycle includes the following steps: 1) viral attachment to heparan sulfate proteoglycans and entry through as yet unidentified receptor(s); 2) clathrin-mediated endocytosis; 3) release of the viral positive-strand RNA genome into the cytosol; 4) translation to yield the ORF1 protein; 5) replication through a negative-strand RNA intermediate and synthesis of full-length as well as 2,2 kb subgenomic RNAs; 6) translation of the subgenomic RNA to yield the ORF2 and ORF3 proteins; and 7) packaging, assembly and release of the newly formed virus.

1.8 ORF3 protein, current knowledge

The hepatitis E virus ORF3 protein is believed to be a 13 kDa-protein composed of two N-terminal hydrophobic domains D1 and D2 and two proline-rich domains P1 and P2 [28]. Furthermore, it has been described that ORF3 protein can be phosphorylated which in turn regulates its interaction with the ORF2 capsid protein [29]. The phosphorylation site has been mapped at a serine residue at aa position 70 or 71 depending on the HEV genotype, respectively gt 3 and gt 1. This interaction may play a regulatory role in the assembly of viral particles [29]. Moreover, it has been proposed that ORF3 associates with the cytoskeleton of the host cells through an interaction with its hydrophobic domain D1 and binds to the microtubules and a MAPK phosphatase as well [30-32]. The D2 domain might interact with hemopexin and thereby altering the cellular iron homeostasis [33]. Moreover, ORF3 may interact with diverse host proteins such as bikunin, α1-microglobulin or fibrinogen Bβ chain and may stabilize the hypoxia-inducible factor 1 (HIF-1) which in turn should upregulate enzymes belonging to the glycolytic pathway [34-37]. Different studies

also suggest that ORF3 may promote cell survival and proliferation through various mechanisms especially by regulating cellular pathways such as Src, MAPK, mitochondrial apoptosis, STAT-3 or c-Met [32,38,39,40,41]. All these observations await a validation in a proper infectious model system.

Although ORF3 appears to be dispensable for HEV RNA replication, it may be important for both infection and virion shedding from host cells [42-45]. Of note, anti-ORF3 antibodies are able to react with HEV recovered from cell culture supernatants as well as from serum of infected patients in contrast to stool samples. Thus, it is postulated that ORF3 protein may be present on the surface of nascent HEV virions likely on the quasi-enveloped HEV [46]. Indeed, HEV has been described as being quasi-enveloped by a cellular membrane when secreted in cell culture or in bloodstream (Fig. 4) [46]. One of the most validated interaction partner of ORF3 is the tumor susceptibility gene 101 (Tsg101). Tsg101 is believed to bind at the PSAP motif of ORF3 and this interaction promotes the formation of the endosomal sorting complexes required for transport (ESCRT). Tsg101 as well as two vacuolar protein-sorting proteins (Vsp4A and Vsp4B) have been suggested to be involved in the secretion of HEV particles through the multivesicular body pathway [47-49].

Taken together these findings suggest that HEV ORF3 protein is involved in the viral release of the infected cells through its interaction with Tsg101. In addition, ORF3 protein interacts with several host proteins which in turn may also contribute to the HEV replication.

2. Aim of the study

The main objective of this project is to characterize further the poorly characterized HEV viral protein ORF3 by biochemical approaches. It involves the expression of ORF3 and various previously designed mutants in mammalian cells, the establishment of an immunopurification protocol for the ORF3-FLAG protein to analyse it by mass spectrometry. In addition, the potential presence of post-translational modifications of ORF3 protein, such as glycosylation and phosphorylation, shall be explored.

3. Materials and methods

3.1 Reagents

Plasmid DNA constructs used in the study were previously prepared in the laboratory. These plasmids consist in pcDNA3.1-based vector for expression of gt 3 ORF3 (Kernow-C1 strain) and mutants in mammalian cells. Briefly, the plasmids used were: (1) pCMVORF3wt, allowing full-length ORF3 protein expression; (2) pCMVORF3$_{S70A}$, harbouring an alanine substitution on the putative ORF3 phosphorylation site described at aa position 70; (3) pCMVORF3$_{spED}$, harbouring two substitutions in the putative signal peptide sequence to invalidate its functionality; (4) pCMVORF3$_{\Delta SP}$, expressing an ORF3 protein truncated of the first 18 aa; (5) pCMVGFP, a GFP expression vector; (6) pTMORF3, allowing the T7-driven expression of ORF3; (7) pTMORF3-FLAG, allowing the T7-driven expression of ORF3-FLAG; (8) pTMORF2/3, allowing the T7-driven expression of ORF2 and ORF3; (9) pTMGFP, allowing the T7-driven expression of GFP.
Antibodies used in this study are listed below in Table 1.

Target	Type of antibody	Dilution
HEV ORF3	Rabbit polyclonal	1:1000
GFP (JL8)	Mouse monoclonal	1:5000
HCV NS5A (11H)	Mouse monoclonal	1:100
FLAG tag (FLAG®M2)	Mouse monoclonal	1:1000
Mouse IgG	Sheep, HRP conjugated	1:5000
Rabbit IgG	Swine, HRP conjugated	1:5000

Table 1. Antibodies used in this study.

3.2 Cell culture

U-2 OS human osteosarcoma cell line has been employed in this study for transient transfection of pCMV-based plasmids. Human hepatoma Huh-7-derived cells, such as S10.3-T7bsd and H7-T7-IZ stably expressing the T7 RNA polymerase have been used for transfection of pTM-based plasmids. Both cell lines have been cultured either in 10-cm dishes (10 ml) or 6-well plate (2 ml) with DMEM® (Dulbecco's Modified Eagle's Medium) culture medium complemented with 10% FBS (fetal bovine serum) and 0.5% Gentamicin at 37°C in a 5% CO_2-saturated incubator. Cells were manually counted using a KOVA counting slide. Cells were seeded to 400’000 cells per well in a 6-well plate prior to transfection.

3.3 Transfection

All cell lines were transfected following the same procedure. Briefly, each transfection mix was prepared from two solutions using the reduced serum media Opti-MEM® (Life Technologies) (HEPES buffer, 2,400 mg/L sodium bicarbonate, hypoxanthine, thymidine, sodium pyruvate, L-glutamine, trace elements, growth factors, and phenol

red reduced to 1.1 mg/L) and the polymeric reagent PEI (polyethyleneimine, Sigma-Aldrich) according to Table 2, below. A ratio of 1:3 between plasmid DNA and PEI is respected. The exact volume of plasmid DNA is calculated according to their respective concentrations. When prepared, the solutions A and B of each condition were mixed and incubated at room temperature (RT) for 15 min. before transferred onto cells dropwise. The cells are cultured 24 or 48h until harvesting. After medium removal and one wash with PBS (phosphate buffer saline), cells were lysed with 80 µl of Laemmli sample buffer for Western blot (5 mM Tris Cl pH 6.8, glycerol 6.5%, SDS 0.2%, 50 mM DTT, bromophenol blue 0.01%). Protein lysates were heated at 95°C for 5 minutes and then conserved at -20°C until being analysed by Western-blot.

Plasmid	Solution A		Solution B	
	DNA (3 µg)	Opti-MEM	PEI	Opti-MEM
pCMVORF3$_{WT}$	3.6 µl	100 µl	9 µl	100 µl
pCMVORF3$_{\Delta sp}$	3.1 µl	100 µl	9 µl	100 µl
pCMVORF3$_{spED}$	3 µl	100 µl	9 µl	100 µl
pCMVORF3$_{S70A}$	1.2 µl	100 µl	9 µl	100 µl
pCMVGFP	3 µl	100 µl	9 µl	100 µl
pTMGFP	2.5 µl	100 µl	9 µl	100 µl
pTMORF3-FLAG	3.75 µl	100 µl	9 µl	100 µl
pTMORF2/3	2.5 µl	100 µl	9 µl	100 µl

Table 2. Composition of the transfection mix prepared for our study.

3.4 SDS-PAGE and western blot

For an optimal separation of the different proteins, a 17% SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis) were employed. After preparation of the 17% polyacrylamide gel, 15 µl of the previously prepared protein lysates were loaded on it together with 5 µl of protein marker V (Peqlab). Migration operates in running buffer 1X (200 ml of prepared solution (15 g Tris base, 72 g Glycine, 5 g SDS, H_2O to 1 L), 800 ml H_2O) in two steps beginning with 100 V for 10 min. (migration of the samples in stacking gel) then 200 V for 50 min. The proteins were then transferred onto a PVDF (polyvinylidene difluoride) membrane by 80 V during 1h30 in transfer buffer 1 X (100 ml of prepared solution (15 g Tris base, 72 g Glycine, H_2O to 2 L), 160 ml ethanol, 740 ml deoxygenated H_2O). The membrane is saturated with blocking buffer (3% dry milk, 0.05% Tween-20) for 1h before being incubated with the first antibody for 1h or overnight. Serial washing with PBS-T (PBS; 0.1% Tween-20) to elude the unbound antibodies, the 2nd antibody is incubated with the membrane for 1h. A final washing step with PBS-T (3 times) was preceding the incubation of the membrane with ECL reagent (enhanced chemiluminescence) before exposing a film.

3.5 Immunoprecipitation

Immunoprecipitation (IP) has been used to purify and load a large quantity of ORF3 onto a SDS-PAGE gel for further mass spectrometry analysis. To this end, transfected cells cultured in two 10-cm dishes (10 ml) have been required. The cells were collected with 750 µl of PBS/well and centrifuge at 3'000 rpm for 3 min. Then, the cells were lysed with TBS (Tris-buffered saline) supplemented with 1% detergent Triton 100 x and protease inhibitors and then incubated on ice for 15 min. After centrifugation at 13'200 rpm for 10 min, the cleared protein lysate is incubated with FLAG®M2 magnetic beads. Fifty microliters of the lysate were kept frozen as input control. In the meantime, 50 µl of FLAG®M2 magnetic beads suspension was transferred in 1.5 ml tube. The beads were washed 3 X with TBS (Tris-buffered saline) using a magnet device. The washed beads were then incubated with the protein lysate for 1h at RT on a rolling wheel. After incubation, the lysate was removed and kept frozen as "flow through" sample. The beads were subsequently washed 3 X in TBS to eliminate the remaining unbound proteins. Finally, the beads with bound ORF3 were incubated with 50µl of Laemmli sample buffer to lyse all bound proteins. The samples were then kept frozen until WB analysis.

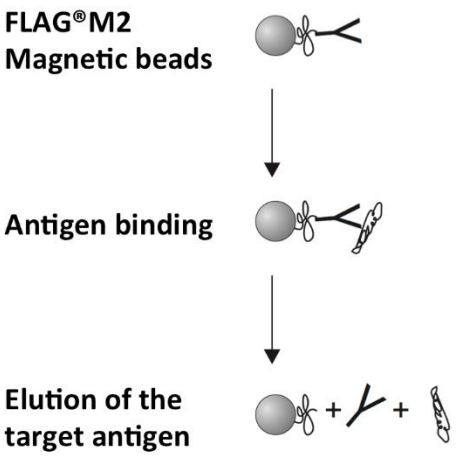

Figure 5. Immunoprecipitation procedure using magnetic beads. Magnetic beads bearing FLAG®M2 monoclonal antibody were incubated with protein lysates of cell expressing ORF3-FLAG. After washing steps, the ORF3-FLAG is being eluted by addition of Laemmli buffer.

3.6 Coomassie blue staining

Coomassie blue staining was used to detect the purified protein after IP of ORF3-FLAG followed by SDS-PAGE. Briefly, after running, the acrylamide gel was incubated with 10 ml of Coomassie blue solution (0.25% Coomassie blue, 40% methanol, 10% acetic acid) for 1h at room temperature. Then, the gel was washed out with 10 ml of destaining buffer (10% acetic acid, 40% methanol) for 10 min. and then a second time with an overnight incubation.

3.7 Deglycosylation assay

Transfected cells were first scrapped in 500 µl PBS (per well) after 3 X PBS washes. Cells were centrifuged at 3'000 rpm for 3 min. and the cell pellet is subjected to 150 µl of RIPA (radioimmunoprecipitation assay) lysis buffer (25 mM Tris HCl pH 7.6; 150 mM NaCl; 1% NP-40; 1% sodium deoxycholate; 0.1% SDS) supplemented with

protease inhibitor cocktail (Complete, Roche). After 15 min. incubation at 4°C, the protein lysate was centrifuged at 13’200 rpm for 10 min. and the cleared lysate was kept. Two different tubes were prepared with 18 µl of protein lysate and 2 µl of Glycoprotein denaturing buffer 10 X (NEB). Both tubes were incubated at 99°C for 10 min. One tube served as negative control, without enzyme, and the other is complemented with 1 µl of the deglycosidase PNGase F (NEB) and then incubated for 1h at 37°C. Finally, Laemmli sample buffer was added to all tubes to stop the reaction before WB analysis.

4. Results

First of all, sequence analyses of the ORF3 protein have been performed with Clustal W alignment followed by the use of TMPred and SignalP algorithms to predict for the presence of, respectively, transmembrane passages and signal peptide. The sequence alignment between HEV gt 1 and gt 3 revealed a high degree of aa conservation in the primary sequence. Furthermore, TMPred identified two transmembrane domains while SignalP found a potential signal peptide located at the first 18 aa (Figure 1A).

In our study, we used several ORF3 constructs, derived from HEV gt 3 (Kernow-C1 strain) that were expressed in various mammalian cell lines as well as two different *in vitro* cell-free expression systems obtained from wheat germ embryos and rabbit reticulocyte lysates. The native form of the ORF3 protein has been compared to a form with two mutations located in the putative signal peptide domain (L7E/F10D) and a truncated form which is depleted by this domain. Another form with a mutation of the serine 70 with an alanine is employed to assess the presence or not of the phosphorylation site of the protein. A FLAG-tagged ORF3 protein is also used later to carry out the immunoprecipitation procedure and mass spectrometry analysis.

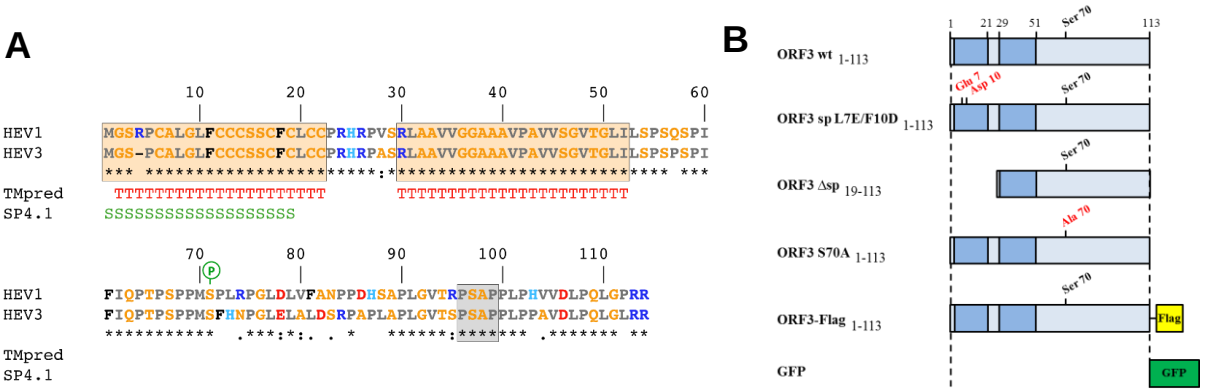

Figure 6. Sequence alignment of ORF3 protein and schematic overview of the constructs used in our study. (**A**) Amino acid sequence alignment of ORF3 protein from HEV1 (Sar55 strain, GenBank accession number AF444002) and HEV3 (Kernow-C1 strain, GenBank accession number JQ679013) obtained with ClustalW. Transmembrane domain (T, red) prediction was performed using TMpred and signal peptide (S, green) prediction was performed using SignalP 4.1. The PSAP motif is boxed in grey and the phosphorylated serine residue is highlighted by a green Ⓟ. (**B**) Different DNA constructs used in this study. ORF3wt correspond to the full-length ORF3 sequence. The ORF3spL7E7F10D construct correspond to an ORF3 mutant where the signal peptide properties of the *N* terminus are no longer retained thanks to the substitution of aa 7 and 10 of ORF3 by a Glu and an Asp residue, respectively. ORF3Δsp correspond to a deletion construct where the first 19 aa have been deleted. The ORF3-S70A is an alanine substitution mutant where the potential phosphorylation site has been abolished. ORF3-FLAG construct is yielding to the expression of ORF3 with a FLAG tag fused at its C-terminal end to be used for ORF3 purification. GFP is a green fluorescent protein expression vector.

4.1 ORF3 has a different apparent molecular weight when expressed in mammalian cells *vs* cell-free expression systems

To express and analyse the ORF3 protein constructs we first used mammalian cells derived from a human osteosarcoma cell line which were cultured and transfected according to the protocols previously described. Rabbit reticulocyte and wheat germ embryo cell-free systems were also employed to express proteins *in vitro*. After transfection, the protein lysates are then separated onto a 17% SDS-PAGE which appears to give a good resolution at low molecular weight (MW) such as the expected MW of ORF3 protein. We also performed other experiments using a more concentrated gel composition such as 19% with the same protocol settings but it did not improve the accuracy of the results and we obtained even poorer outcomes (data not shown). We used in our experiments the expression of the green fluorescent protein (GFP) to assess the quality of cell transfection as a negative control for WB analysis.

First, we noticed that the ORF3 protein produced in both cell-free systems migrates at the same level corresponding to a molecular weight of about 11-12 kDa. Wild-type (WT) ORF3 protein produced in U-2 OS cells migrates at a higher MW, about 14 kDa, than cell free-expressed ORF3. Furthermore, several different MWs are observed (Fig. 2A and B). Interestingly, these different forms of ORF3 are also present at 48h post-transfection but with a predominant form located at a lower MW around 11-12kDa. This observation suggests that the post-transfection time is a parameter that plays an important role in the proportion of the different forms.

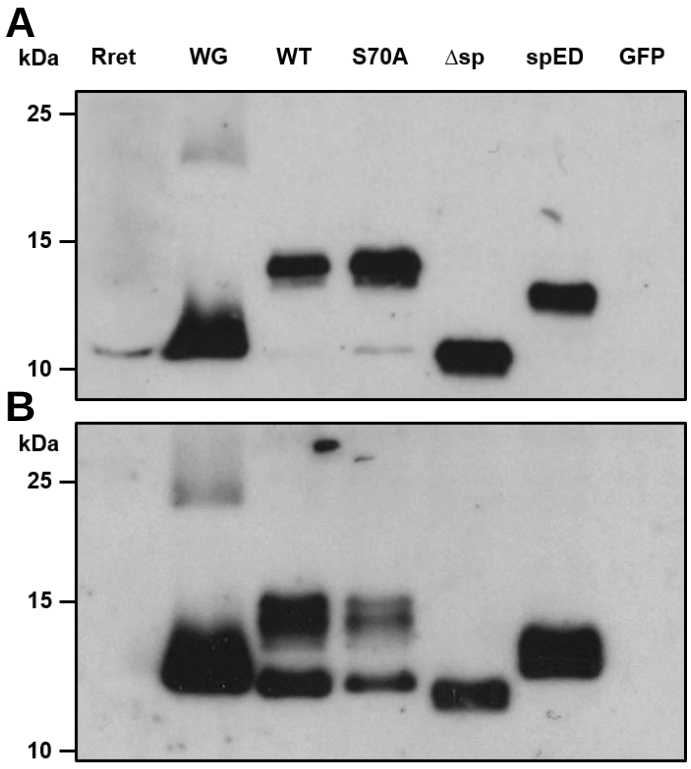

Figure 7. Western blot analysis of the expression of ORF3 constructs in mammalian cells. ORF3 constructs were transfected in U-2 OS cells and protein lysates were analysed by SDS-PAGE followed by western blot analysis, using rabbit polyclonal anti-ORF3 antiserum, 24h (**A**) or 48h (**B**) post-transfection. Rret, in vitro translated ORF3 using rabbit reticulocyte lysate; WG, cell free expressed ORF3 using wheat germ embryos extract; WT, S70A, Δsp, spED and GFP, expression of the corresponding constructs referred in Fig. 6 in U-2 OS cells.

The first 18 aa of the ORF3 protein are predicted to function as a signal peptide. When using a construct in which this region is deleted (sample "Δsp"), likely mimicking a cleavage of the putative signal peptide, the protein migrates at a much lower position than the WT ORF3. However, no significant difference is observed with

a longer transfection time. Compared to “WG” and “Rret” *in vitro* samples, this construct migrates almost at the same location but slightly underneath them.
A construct with a leucine residue substituted by a glutamic acid residue at aa position 7 and a phenylalanine substituted by an aspartic acid residue at aa position 10 is also used to assess the potential presence of the signal peptide. When tested with the signal peptide prediction algorithm SignalP, the ORF3 sequence harbouring these 2 substitutions do not predict signal peptide any longer. Therefore, this mutant is believed to be not sensitive to signal peptide cleavage, if any. Surprisingly, $ORF3_{spED}$ migrated at an intermediate MW which is below the highest MW of $ORF3_{wt}$.
All these findings do not support the initial hypothesis of a cleaved signal peptide but rather suggest that ORF3 undergoes another type of modification when expressed in mammalian cells.
In the meantime, we also wanted to confirm the presence of the phosphorylation of ORF3 previously described in the literature, i.e. at aa position 70 in gt 3. To this aim, we used a construct with an alanine substitution at aa position 70, namely $ORF3_{S70A}$, in order to prevent the phosphorylation of ORF3. Interestingly, $ORF3_{S70A}$ showed the same migration pattern than $ORF3_{wt}$. Thus, a possible phosphorylation at Ser 70 of ORF3 does not explain the high MW observed for ORF3 in WB.

4.2 Co-expression of the ORF2 capsid protein does not influence ORF3 apparent molecular weight

ORF2 and ORF3 are closely related since they are translated from the same subgenomic RNA and that they are believed to interact each other likely during virion assembly and/or secretion. Given that ORF3 protein was successfully expressed in the previous experiments, we wanted to investigate whether ORF2 influence the observed changes in the apparent molecular weight of ORF3 protein.
To this end, we employed a construct that allows both ORF2 and ORF3 expression, e.g. pTMORF2/3. This construct required the usage of cell line stably expressing the T7 RNA polymerase e.g. S10-3_T7bsd or H7-T7-IZ. We compared the expression of the simultaneous ORF3 and ORF2 expression with the single expression of ORF3 either with pCMV or pTM construct. GFP expression was used as negative control for ORF3 WB detection and as transfection efficiency control. Twenty-four hours post-transfection, we observed that ORF3 protein expressed in the different plasmids (pCMV *vs* pTM) showed the same pattern in WB with a main species at a MW of about 14-15 kDa and a minor species of 12 kDa. Interestingly, the co-expression of ORF2, did not change the WB pattern of ORF3 protein. It suggests that, in the present system, the co-expression of ORF2 capsid protein has no visible effect on the apparent molecular weight of ORF3 protein and likely on potential post-translational modifications.

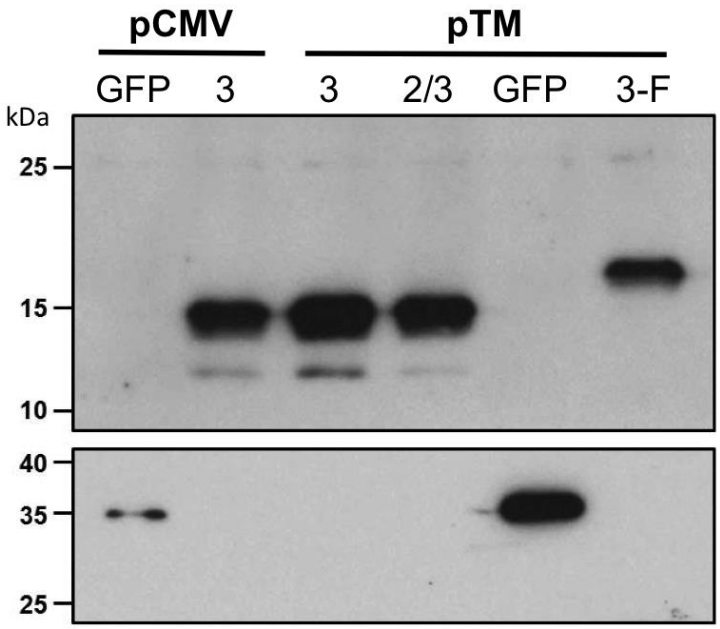

Figure 8. Expression of ORF3 together with ORF2 or fused with a C-terminal FLAG tag. H7-T7-IZ cells transfected with either pCMV or pTM expression constructs were harvested 24h post-transfection and submitted to SDS-PAGE followed by WB analysis using anti-ORF3 (upper panel) or anti-GFP antibodies (lower panel). GFP, green fluorescent protein; 3, ORF3 expression construct; 2/3, ORF2/ORF3 expression construct; 3-F, ORF3-FLAG expression construct.

In this experiment, we also tried to assess whether the FLAG-tagged ORF3 protein was behaving as the untagged, e.g. with a higher apparent molecular weight. As expected, the ORF3-FLAG protein has a MW of about 16-17 kDa consistent with a potential post-translational modification. Therefore, this result confirmed us that we can use a FLAG-tagged ORF3 protein as a model to identify the potential modification using a mass spectrometry analysis. Finally, we noticed that the quality of expression of ORF3 protein in hepatoma cells was comparable to that in osteosarcoma cells.

4.3 Establishment of an immunoprecipitation protocol to purify an ORF3-FLAG protein expressed in mammalian cells

To further characterize ORF3 protein, we wanted to analyse its composition and chemical structure by using mass spectrometry (MS), a highly sensitive technique with a good resolution. To this end, we established a procedure to immunopurify the FLAG-tagged ORF3 before MS analysis.
First, two 10-cm dishes of S10.3_T7bsd cells were transfected either with pTMORF3 (as untagged control) or with pTMORF3-FLAG construct and were harvested 24h later by scrapping in PBS. It enabled us to avoid the use of trypsin at this step which is a potential source of enzymatic proteolysis. Cells were lysed in a 1% Triton X-100 containing buffer. In parallel, we also tried the n-Dodecyl-β-D-maltoside (DDM 1%) as detergent but it did not give satisfactory outcomes. In brief, protein lysate was cleared by centrifugation (13'200 rpm, 10 min.) to separate the solubilized proteins from cell debris and nucleic acids (D). The cleared lysate corresponds to the "Input" sample and was incubated with anti-FLAG®M2 magnetic beads which have been chosen for simplicity of manipulation. In order to improve the ORF3-FLAG recovery and to minimize the elution volume, incubation time as well as the elution time and volume have been adjusted. Elution was performed by addition of Laemmli buffer to the beads (sample "IP"). The different collected fractions along the IP procedure have been loaded and separated onto SDS-PAGE to assess the quality of our procedure.

First, the “D“ sample likely containing all the cell debris and nucleic acids allowed us to evaluate the quality of the cell lysis. Conversely, the “Input“ sample should contain the solubilized ORF3 (tagged or untagged). As shown in Figure 9, we did not detect ORF3 in the “D“ samples while ORF3 was well-detectable in “Input“ samples. It indicated that the cell lysis and removal of cell debris and nucleic acids by centrifugation was efficient enough.
The efficiency of the FLAG IP protocol is given by the analysis of the samples after IP. While ORF3 was not detected in IP when untagged, a robust signal was detected in the IP sample for the ORF3-FLAG. These results indicated that the FLAG immunoprecipitation worked efficiently to recover FLAG-tagged ORF3 and was specific enough. Indeed, since we did not detect ORF3 (untagged) in the IP sample showed that our procedure was enough stringent to clear potential contaminants and pulled down only FLAG-tagged products.

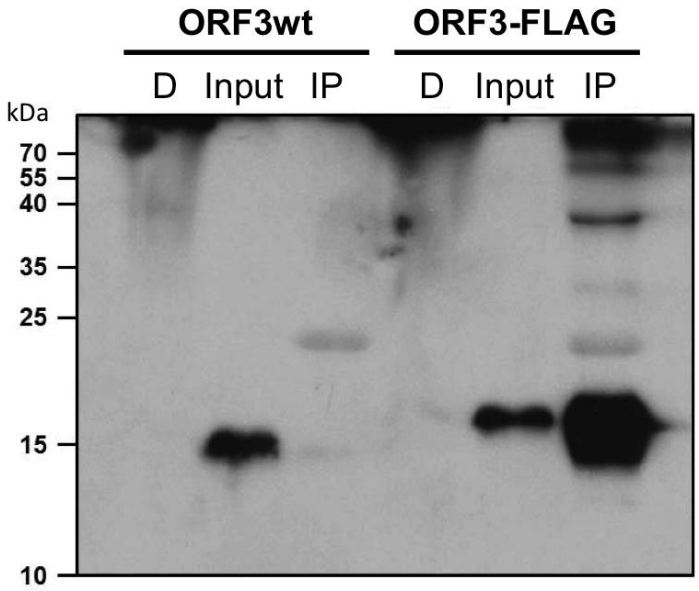

Figure 9. FLAG immuno-purification of ORF3-FLAG expressed in human cells. U-2 OS cells expressing either ORF3wt or ORF3-FLAG were harvested 24h post-transfection and submitted to the IP protocol. D, cell debris of the ORF3wt or ORF3-FLAG samples; Input, pre-IP sample; IP, sample after FLAG IP.

Further attempts to purify larger amounts of ORF3-FLAG to be able to detect it on Coomassie gel were done (starting with 20x10 cm-dishes) but were unsuccessful due to limitations in key steps such as cell lysis as well as protein separation onto SDS-PAGE. Therefore, we were not able to extract enough quantity of protein for mass spectrometry analysis in the timeframe of the Master work. However, the protocol has been further elaborated in the laboratory in my absence. Subsequently, they have been able to achieve a first mass spectrometry analysis of ORF3-FLAG protein expressed in mammalian cells as well as of ORF3 protein produced in wheat germ cell-free expression system (Fig. 10). MS analysis implies a trypsin digestion step to generate peptides (cleavage after Arg or Lys residues) that are subsequently detected by the mass spectrometer. Therefore, the expected results should inform on the quantity and the quality of the different peptides detected for each recombinant protein analysed. Here, we analysed the cell-free produced ORF3 protein using the wheat germ embryo extracts. This sample served a control for an unprocessed and non-modified protein since no such modification could happen in the cell-free system.

Figure 10. Mass spectrometry analysis of recombinant ORF3 produced in cell-free system and mammalian cells. (A) MS analysis of ORF3 protein produced in wheat germ cell-free system (WG). **(B)** MS analysis of ORF3-FLAG recombinant protein purified from U-2 OS cells lysates. Only 44% of the entire sequence have been identified and analysed.

As revealed by the results shown in Figure 10A, peptides covering 92% of ORF3 aa sequence have been successfully identified. Moreover, ORF3-FLAG purified from mammalian cell lysates showed a poorer coverage with only 2 peptides detected and with information missing for the *N* terminus of the protein. This result is disappointing because we therefore failed to show if ORF3 expressed in mammalian cells is processed or not by signal peptidases and if post-translational modifications occur in the non-covered regions. We could explain the results by a lower amount of the protein analysed by MS compared to the cell-free expressed ORF3 and by intrinsic reasons like a lower propensity to detect the N-terminal peptide.

4.4 Investigation of potential post-translational modifications of ORF3 protein

In addition to the unbiased approach by MS analysis, we also wanted to use a more targeted approach to investigate the potential presence of glycosylation and/or phosphorylation, which are two of the most frequent post-translational protein modifications.
The osteosarcoma cell line U-2 OS has been employed as cellular background for these experiments. Transfection with a GFP expression plasmid was used as quality control of the transfection and a GFP-fused HCV NS5B construct (the last 26 aa of

NS5B, C26) harbouring a synthetic glycosylation tag was used as positive control for glycosylation [50]. We used the PNGase F as deglycosylation enzyme.

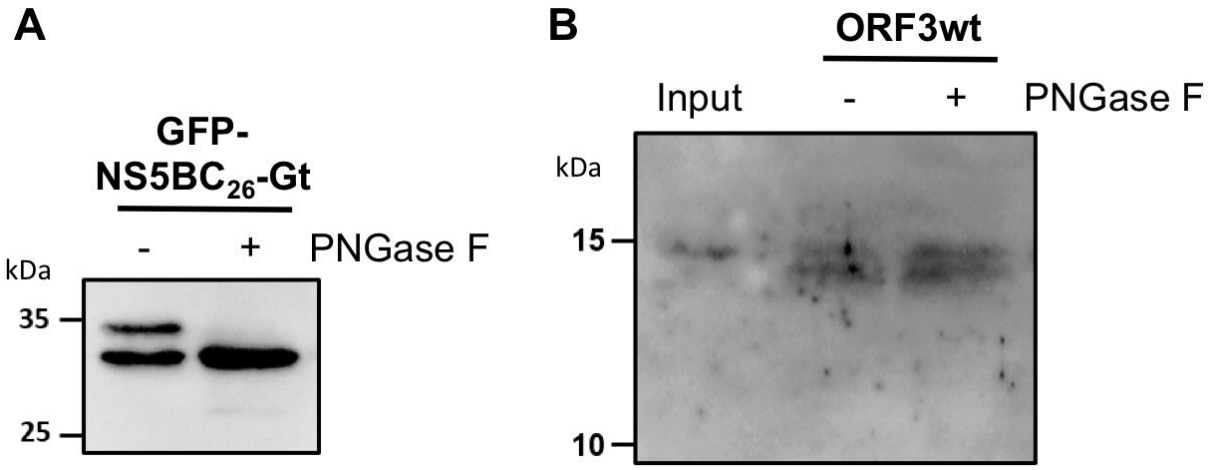

Figure 11. ORF3 protein is likely not glycosylated. (**A**) U-2 OS cells transfected with constructs pCMVGFP-NS5BC26-gt were harvested 24h post-transfection. Cell lysates were subjected (+) or not (-) to treatment with the deglycosidase PNGase F and further analysed by SDS-PAGE followed by WB analysis using anti-GFP antibody. Treatment of GFP-NS5BC26-gt by PNGase F served as a positive control for deglycosylation. (**B**) U-2 OS cells transfected with constructs pCMVORF3wt were harvested 24h post-transfection. Cell lysates prior to treatment (Input) or after the deglycosylation procedure in presence (+) or not (-) of PNGase F were analysed by SDS-PAGE followed by WB analysis using anti-ORF3 antibody.

Our positive control for deglycosylation procedure, i.e. GFP-NS5B C26-gt, shows that the higher MW band was disappearing upon PNGase F treatment (Fig. 11A). Therefore, our experimental settings were working well and could be applied to ORF3. By treating with PNGase F cell lysates prepared from pCMVORF3wt-transfected cells, we did not observe any change with the untreated control. ORF3 was detected as two bands in WB but no change in the ratio of these band was observed. Hence, the post-translational modification of ORF3 is not due to a glycosylation of the protein.

5. Discussion

The function(s) as well as the protein characterization of the HEV ORF3 are still poorly understood yet. In this Master work, we have been able to express the HEV ORF3 protein in different mammalian cell lines and in a cell-free system as well as to establish a suitable immunoprecipitation protocol for mass spectrometry analysis of HEV ORF3 protein. Our findings contribute to a better understanding of HEV ORF3 protein.

Several attempts to optimize the migration and separation of ORF3 protein, including optimization of acrylamide gel concentration, i.e. 15%, 17% and 19%, yielded to the routine use of 17% SDS-PAGE. To our surprise, ORF3 protein displays several forms of different MW ranging from 12 kDa to about 15 kDa. We observed mainly 3 bands, one at 12 kDa and two as a doublet between 14 and 15 kDa. To our knowledge, it is the first time that a report shows ORF3 having different apparent MWs. The theoretical MW of ORF3 is 11.5 kDa which may correspond to the lower band observed on the WB. This hypothesis is further strengthened by the detection of a single band at the same MW for ORF3 produced in cell free expression systems such as rabbit reticulocyte lysates or wheat germ embryo extracts. It should be noted that signal peptide cleavage does not occur in these systems unless microsomal membranes are added [51]. It means that the resulting ORF3 protein detected on Western-blot represents the full-length protein without any cleavage and post-translational modifications even if an active signal peptide is present. Because protein prediction algorithm SignalP indicated a potential signal peptide at the *N* terminus of ORF3, we wanted to investigate whether a signal peptide was indeed present. The peptide signal is defined as a short sequence of 16 to 30 aa which is usually located at the N-terminal extremity of the newly synthesized proteins. Its main function is to guide the proteins from the translational machinery through its recognition by the signal recognition particle to the endoplasmic reticulum for further processing as secreted or membrane-associated proteins. The signal peptide is usually cleaved but not always by a cellular signal peptidase [52]. In addition, HEV ORF2 which is the viral capsid believed to interact with ORF3, possesses a N-terminal signal peptide [53]. The interaction between the 2 viral proteins suggests that ORF3 may also share the same route of secretion and may possess a signal peptide. Therefore, to address this issue we used various ORF3 expression constructs including some where the predicted signal peptide was either deleted (ORF3Δsp) or mutated (ORF3 L7E/F10D). The deletion construct resulted in a single band on WB which is slightly below the 12 kDa-band observed with ORF3wt. However, the theoretical MW of ORF3Δsp is lower than 10 kDa which is obviously not what we observed here. Therefore, we hypothesized that ORF3 does not possess a cleavable signal peptide and/or possesses additional post-translational modifications. The expression of the construct with the two substitutions within the potential signal peptide (ORF3 L7E/F10D) resulted in a doublet band which was at intermediate MWs around 13 kDa. As we observed a lower MW with the mutant, it would further suggest that no cleavage of the *N* terminus occurs with the wild type ORF3 but rather suggest that the mutated region may be involved in post-translational modifications.

Altogether, our observations suggest that the various forms observed for ORF3 likely result from post-translationally modifications. Interestingly, when cell lysate was prepared 48h post-transfection rather than 24h, we observed an accumulation of the low MW ORF3 form. A possible explanation would be the saturation of cellular pathways involved in the ORF3 modification. These results have been reproduced in different cell lines such as the

human osteosarcoma U-2 OS and hepatoma cell lines S10-3 and H7-T7-IZ, both derived from Huh-7 cells. These results significantly differ from previously reported observations which suggest a unique MW for ORF3 of 13.5 kDa [53]. This reported MW is slightly below what we observed for the modified ORF3 and significantly different from the theoretical one. Interestingly, an apparent MW of ORF3 higher than the theoretical one has already been observed *in vivo* in HEV-infected mice [54]. Therefore, our *in vitro*-observation appears relevant to the conditions of infection *in vivo* and justify the use of our *in vitro*-model (expression of ORF3 in mammalian cells) to further characterize the ORF3 protein modifications.

Glycosylation and phosphorylation are ones of the most frequent protein post-translational modifications. We thus decided to use a targeted approach with deglycosylation and dephosphorylation procedures to investigate their potential impact on the ORF3 protein. First of all, we performed the deglycosylation assay using PNGase F enzyme on ORF3 protein. PNGase F is a N-glycosidase F that cleaves internal glycosylation residues such as N-acetylglucosamines (GlcNAc), asparagine residues of high mannose, hybrid and complex oligosaccharides as well N-linked glycoproteins. While our positive control, a GFP-fused HCV NS5B construct harbouring a glycosylation acceptor tag at the C terminus, showed a clear MW shift down upon treatment with PNGase F, we have been unable to observe any change for ORF3 in WB. This result is supported by the absence of glycosylation prediction for ORF3 using bioinformatic tools such as PROSITE software. In addition to glycosylation we have also explored potential phosphorylation of ORF3. It has been shown earlier that ORF3 was phosphorylated on Ser 70 [30], however, as discussed here, we did not see any difference in WB between the wild-type and S70A ORF3 proteins. Given the number of threonine and serine residues in ORF3, it is likely possible that phosphorylation occurs at other sites. It is why we tried to use a dephosphorylation protocol to see whether the higher MW bands we observed for ORF3 were corresponding to phosphorylated forms. To this aim, we employed the Lambda protein phosphatase, however we were unable to present coherent data in this manuscript. Indeed, we did not see any changes in ORF3 MW after treatment with Lambda protein phosphatase but our positive control, the phosphorylated form of HCV NS5A protein, remained also non sensitive to the treatment. This investigation is still worth to pursue with appropriate controls and, furthermore, in relevant system such as cells replicating the full-length genome. Indeed, modification such as phosphorylation are often regulating signalling pathways or transient events. At the virus life cycle scale, it may be involved in RNA replication, encapsidation, viral release or in interaction with host cell responses. To this aim, we would have to further optimize the dephosphorylation protocol as well as the detection sensitivity. Indeed, we have to deal with a low amount of ORF3 in infected cells that we may have to overcome by immunoprecipitation of ORF3.

The ORF2 protein is translated into a precursor that contains a signal peptide of about 22 amino acids which is cleaved for further processing into the mature protein in the endoplasmic reticulum [29,34,53]. Furthermore, it has been demonstrated that the phosphorylated form of the ORF3 protein interacts preferentially with the non-glycosylated form of the ORF2 protein [29]. These informations together with the role of ORF3 in the secretion of HEV suggest that ORF3 protein may closely interfere with the ORF2 capsid protein. Therefore, we tried to co-express both viral proteins together to see whether it influences the post-translational modification of ORF3. Our results, obtained in H7-T7-IZ cells with a T7 RNA polymerase-driven expression of ORF2 and ORF3, did not show any

difference on the apparent MW of ORF3 in WB. Unfortunately, we did not succeed to detect the ORF2 protein within the co-expression sample likely due to technical limitation, i.e. detection sensitivity. This would have been an important control to have and, in addition, it would have told us, the other way around, if ORF3 expression could influence ORF2 protein. Additionally, it would have been judicious to add a sample with ORF2 protein expressed alone.

As there is a great variety of post-translational protein modifications, we thought to design an unbiased proteomics approach to identify the type of modifications harboured by ORF3. To this aim, we need to get a large amount of ORF3 protein to analyse by mass spectrometry. We decided to purify ORF3 expressed in mammalian cells using a FLAG-tagged version of the protein and an immunoprecipitation (IP) approach with magnetic beads. As a pre-requirement, we checked if ORF3-FLAG was also post-translationally modified as it is the case for the untagged ORF3. As revealed on Figure 8, ORF3-FLAG has an apparent molecular weight of about 16-17 kDa which likely indicates that the fused protein is also modified. Indeed, taking into consideration the addition of the FLAG tag to ORF3, the expected MW should be in this range of MW. In our study, we have been able to purify the FLAG-tagged ORF3 protein from transfected cell lysates through the easy and fast immunoprecipitation procedure using anti-FLAG®M2 magnetic beads. Our IP experimental settings allowed a good recovery of ORF3-FLAG. In addition, the FLAG IP showed an excellent specificity since our negative control, which was the untagged ORF3, was not pulled-down in our IP. For the cell lysates preparation, we decided to use the hepatocarcinoma-derived cell line S10-3 T7_bsd to express ORF3 in a cellular environment where HEV is replicating. As previously reported in the result section the adjunction of the non-ionic detergent Triton X-100 in the lysis buffer gave more satisfactory results than with the n-Dodecyl-β-D-maltoside (DDM 1%). Few published studies already employed the immunoprecipitation of HEV ORF3 and used a similar lysis buffer containing Triton X-100, i.e. 20 mM Tris-HCl [pH 7.5], 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 1% Triton X-100, protease inhibitor [36,38]. Other studies used the RIPA buffer (10 mM Tris-HCl [pH 8.0], 140 mM NaCl, 5 mM iodoacetamide, 0.5% Triton X-100, 1% sodium deoxycholate, 0.1% sodium dodecyl sulfate [SDS], 2 mM phenylmethylsulfonyl fluoride) [53] or the GST binding buffer (20 mM Tris [pH 7.9], 180 mM KCl, 0.2 mM EDTA, 5 mM $MgCl_2$, 1 mM phenylmethylsulfonyl fluoride, 0.01% Nonidet P-40, and 1 mM dithiothreitol containing 1 g/ml bovine serum albumin) [29]. These informations from the literature would indicate that a successful immunoprecipitation of ORF3 required stringent detergent conditions.

Finally, further improvement of the IP protocol, especially to scale-up the experiment, led to the purification of ORF3-FLAG in quantity that was enough to extract the corresponding band on Coomassie stained SDS-PAGE and led to the mass spectrometry analysis. ORF3 protein produced in cell free-based wheat germ embryo extract system was also analysed as control. The analysis reported that the first 18 amino acids (aa) of the putative signal peptide was present in the WG sample but remained undetectable in the ORF3-FLAG sample most probably because the corresponding peptide was below the detection threshold. Another explanation could be that this peptide is bearing the post-translational modification thereby changing the intrinsic physico-chemical features of the peptide. Thus, we still cannot definitively conclude on the absence of the N-terminal peptide in the ORF3-FLAG sample.

In our study, both with a mutagenesis approach as well as with the MS analysis, we have been unable to show that ORF3 is phosphorylated at Ser 70. It is an interesting and important question to address for the understanding of HEV ORF3. Indeed, it has been shown that ORF3 can be phosphorylated on this residue to interact with ORF2 [29]. In our hands the co-expression of ORF3 and ORF2 did not lead to any change on the apparent MW of ORF3 in WB. As ORF1 protein and viral RNA genome were missing in this system, it would be now important to investigate it in cells producing infectious HEV. However, the question is difficult to address since we do not detect easily the viral proteins in cells replicating the full-length genome and that we may have to pull-down ORF3 by IP in conditions preserving the phosphorylation. In addition, constructing a virus harbouring a mutation invalidating the phosphorylation on the Ser residue at aa position 70 is challenging since the ORF3 and ORF2 sequences are overlapping in the HEV genome. Indeed, any change in the nucleotide sequence on ORF3 may lead to aa change on the ORF2 sequence as well.

In the meantime, we have noticed that ORF3 protein harbours a high number of cysteine residues at its *N* terminus, 8 Cys out of 21 amino acids. This unusual pattern may suggest that these residues are involved in the observed modifications. Cysteine is an essential neutral aa with a Sulphur atom in its lateral chain which may be involved in disulfide bridge formation upon oxidative conditions. Disulfide bounds are stabilizing protein conformation and/or protein-protein interactions [55]. However, cysteine may also be engaged in post-translational modifications such as S-acyl palmitoylation which consists in the addition of a palmitic acid, a 16 carbons fatty acid chain by a cellular palmitoyltransferase [56]. This hypothesis is currently being pursued in the laboratory by using ^{3}H-palmitate incorporation, site-directed mutagenesis of the ORF3 cysteine residues as well as palmitoylation inhibitors and palmitoyltransferase genes knockdown.

6. Conclusion

Taken together, our results showed that HEV ORF3 expressed in mammalian cell lines has a higher apparent molecular weight than expected and compared to the protein produced in cell free systems. While a signal peptide has been predicted by protein prediction algorithms, the investigations we performed did not support neither the presence of a signal peptide nor its cleavage. In addition, our study suggests that HEV ORF3 protein is post-translationally modified which open new perspectives in the understanding of this viral protein. Viruses, and particularly HEV, have evolved with a limited number of proteins which have then multifunctional properties. Post-translational modifications are offering a great advantage to the protein functional diversity without changing the intrinsic sequence information. The viruses are employing this strategy to their own advantage. Of the best studied modifications of viral proteins, we can mention the phosphorylation and glycosylation but many other do exist. HEV ORF3 protein is a very small protein of 113 aa involved in several protein interactions as well as in trafficking pathways [38,47,49,57]. These informations together with our findings strongly suggest that HEV ORF3 protein also undergo post-translational modifications to efficiently accomplish its numerous functions in the HEV life cycle.

7. Acknowledgments

I would like to sincerely thank Prof. Darius Moradpour for his warm welcome in his laboratory which gave me the opportunity to discover the field of fundamental research as well as his great availability and valuable advice on this research project.
I also would like to thank PD Dr. Jérôme Gouttenoire very much for guiding me throughout the project providing all the crucial theoretical and practical knowledge in a stimulating, fruitful and collegial atmosphere. And last but not least, a special thanks to the whole team of the lab and the IMUL for the numerous shared advice and agreeable moments making this Master Thesis a memorable experience.

8. References

1. Reyes GR, Purdy MA, Kim JP, Luk KC, Young LM, Fry KE, et al. Isolation of a cDNA from the virus responsible for enterically transmitted non-A, non-B hepatitis. Science 1990;247:1335-1339.
2. Aggarwal R, Naik S. Epidemiology of hepatitis E: current status. J Gastroenterol Hepatol 2009;24:1484-1493.
3. Purdy MA, Khudyakov YE. The molecular epidemiology of hepatitis E virus infection. Virus Res 2011;161:31-39.
4. Pérez-Gracia MT, Suay B, Mateos-Lindemann ML. Hepatitis E: an emerging disease. Infect Genet Evol 2014;22:40-59.
5. Rein DB, Stevens GA, Theaker J, Wittenborn JS, Wiersma ST. The global burden of hepatitis E virus genotypes 1 and 2 in 2005. Hepatology. 2012;55:988-997.
6. Drexler JF, Seelen A, Corman VM, Tateno AF, Cottontail V, Zerbinati RM, et al. Bats Worldwide Carry Hepatitis E Virus-Related Viruses That Form a Putative Novel Genus within the Family Hepeviridae. J Virol 2012;86:9134-9147.
7. Kamar N, Bendall R, Legrand-Abravanel F, Xia N-S, Ijaz S, Izopet J, et al. Hepatitis E. Lancet 2012;379:2477-2488.
8. Marano G, Vaglio S, Pupella S, Facco G, Bianchi M, Calizzani G, et al. Hepatitis E: an old infection with new implications. Blood Transfus 2015;13:6-17.
9. Bonnet D, Kamar N, Izopet J, Alric L. L'hépatite virale E : une maladie émergente. La Revue de Médecine Interne. 2012;33:328-334.
10. Kumar S, Subhadra S, Singh B, Panda BK. Hepatitis E virus: the current scenario. International Journal of Infectious Diseases 2013;17:e228-233.
11. Pariente A, Renou C. [Epidemiology of hepatitis E: a (re) emerging disease?]. Presse Med 2015;44:333-338.
12. Scobie L, Dalton HR. Hepatitis E: source and route of infection, clinical manifestations and new developments. J Viral Hepat 2013;20:1-11.
13. Dalton HR, Stableforth W, Thurairajah P, Hazeldine S, Remnarace R, Usama W, et al. Autochthonous hepatitis E in Southwest England: natural history, complications and seasonal variation, and hepatitis E virus IgG seroprevalence in blood donors, the elderly and patients with chronic liver disease. Eur J Gastroenterol Hepatol 2008;20:784-790.
14. Aggarwal R, Aggarwal RA. Hepatitis E: clinical presentation in disease-endemic areas and diagnosis. Semin Liver Dis 2013;33:30-40.
15. Hussaini SH, Skidmore SJ, Richardson P, Sherratt LM, Cooper BT, O'Grady JG. Severe hepatitis E infection during pregnancy. J Viral Hepat 1997;4:51-54.
16. Navaneethan U, Al Mohajer M, Shata MT. Hepatitis E and pregnancy: understanding the pathogenesis. Liver Int 2008;28:1190-1199.
17. Kamar N, Bendall RP, Peron JM, Cintas P, Prudhomme L, Mansuy JM, et al. Hepatitis E virus and neurologic disorders. Emerging Infect Dis. 2011;17:173-179.
18. Deniel C, Coton T, Brardjanian S, Guisset M, Nicand E, Simon F. Acute pancreatitis: a rare

complication of acute hepatitis E. J Clin Virol 2011;51:202-204.
19. Al-Shukri I, Davidson E, Tan A, Smith DB, Wellington L, Johannessen I, et al. Rash and arthralgia caused by hepatitis E. Lancet 2013;382:1856.
20. Kamar N, Abravanel F, Garrouste C, Cardeau-Desangles I, Mansuy JM, Weclawiak H, et al. Three-month pegylated interferon-alpha-2a therapy for chronic hepatitis E virus infection in a haemodialysis patient. Nephrol Dial Transplant 2010;25:2792-2795.
21. Kamar N, Rostaing L, Abravanel F, Garrouste C, Lhomme S, Esposito L, et al. Ribavirin therapy inhibits viral replication on patients with chronic hepatitis e virus infection. Gastroenterology 2010;139:1612-1618.
22. Haagsma EB, van den Berg AP, Porte RJ, Benne CA, Vennema H, Reimerink JHJ, et al. Chronic hepatitis E virus infection in liver transplant recipients. Liver Transpl. 2008;14:547-553.
23. Zhang J, Zhang X-F, Huang S-J, Wu T, Hu Y-M, Wang Z-Z, et al. Long-term efficacy of a hepatitis E vaccine. N Engl J Med 2015;372:914-922.
24. Cao D, Meng X-J. Molecular biology and replication of hepatitis E virus. Emerg Microbes Infect 2012;1:e17.
25. Izopet J, Lhomme S, Abravanel F, Roque A-M, Kamar N. Le virus de l'hépatite E. La Presse Médicale. 2015;44:328-332.
26. Debing Y, Moradpour D, Neyts J, Gouttenoire J. Update on hepatitis E virology: Implications for clinical practice. Journal of Hepatology. 2016;65:200-212.
27. Shukla P, Nguyen HT, Faulk K, Mather K, Torian U, Engle RE, et al. Adaptation of a genotype 3 hepatitis E virus to efficient growth in cell culture depends on an inserted human gene segment acquired by recombination. J Virol 2012;86:5697-5707.
28. Ahmad I, Holla RP, Jameel S. Molecular virology of hepatitis E virus. Virus Res 2011;161:47-58.
29. Tyagi S, Korkaya H, Zafrullah M, Jameel S, Lal SK. The phosphorylated form of the ORF3 protein of hepatitis E virus interacts with its non-glycosylated form of the major capsid protein, ORF2. J Biol Chem 2002;277:22759-22767.
30. Zafrullah M, Ozdener MH, Panda SK, Jameel S. The ORF3 protein of hepatitis E virus is a phosphoprotein that associates with the cytoskeleton. J Virol 1997;71:9045-9053.
31. Kannan H, Fan S, Patel D, Bossis I, Zhang Y-J. The hepatitis E virus open reading frame 3 product interacts with microtubules and interferes with their dynamics. J Virol 2009;83:6375-6382.
32. Kar-Roy A, Korkaya H, Oberoi R, Lal SK, Jameel S. The hepatitis E virus open reading frame 3 protein activates ERK through binding and inhibition of the MAPK phosphatase. J Biol Chem 2004;279:28345-28357.
33. Ratra R, Kar-Roy A, Lal SK. The ORF3 protein of hepatitis E virus interacts with hemopexin by means of its 26 amino acid N-terminal hydrophobic domain II. Biochemistry 2008;47:1957-1969.
34. Tyagi S, Surjit M, Lal SK. The 41-amino-acid C-terminal region of the hepatitis E virus ORF3 protein interacts with bikunin, a kunitz-type serine protease inhibitor. J Virol 2005;79:12081-12087.
35. Tyagi S, Surjit M, Roy AK, Jameel S, Lal SK. The ORF3 protein of hepatitis E virus interacts with liver-specific alpha1-microglobulin and its precursor alpha1-microglobulin/bikunin precursor (AMBP) and expedites their export from the hepatocyte. J Biol Chem 2004;279:29308-29319.
36. Moin SM, Chandra V, Arya R, Jameel S. The hepatitis E virus ORF3 protein stabilizes HIF-1alpha and enhances HIF-1-mediated transcriptional activity through p300/CBP. Cell Microbiol 2009;11:1409-1421.
37. Ratra R, Kar-Roy A, Lal SK. ORF3 protein of hepatitis E virus interacts with the Bbeta chain of fibrinogen resulting in decreased fibrinogen secretion from Huh-7 cells. J Gen Virol 2009;90:1359-1370.
38. Chandra V, Kar-Roy A, Kumari S, Mayor S, Jameel S. The hepatitis E virus ORF3 protein modulates epidermal growth factor receptor trafficking, STAT3 translocation, and the acute-phase response. J Virol 2008;82:7100-7110.
39. Korkaya H, Jameel S, Gupta D, Tyagi S, Kumar R, Zafrullah M, et al. The ORF3 protein of hepatitis E virus binds to Src homology 3 domains and activates MAPK. J Biol Chem

2001;276:42389-42400.
40. Moin SM, Panteva M, Jameel S. The hepatitis E virus Orf3 protein protects cells from mitochondrial depolarization and death. J Biol Chem 2007;282:21124-21133.
41. Chandra V, Kalia M, Hajela K, Jameel S. The ORF3 protein of hepatitis E virus delays degradation of activated growth factor receptors by interacting with CIN85 and blocking formation of the Cbl-CIN85 complex. J Virol 2010;84:3857-3867.
42. Emerson SU, Nguyen H, Torian U, Purcell RH. ORF3 protein of hepatitis E virus is not required for replication, virion assembly, or infection of hepatoma cells in vitro. J Virol 2006;80:10457-10464.
43. Graff J, Nguyen H, Yu C, Elkins WR, St Claire M, Purcell RH, et al. The open reading frame 3 gene of hepatitis E virus contains a cis-reactive element and encodes a protein required for infection of macaques. J Virol 2005;79:6680-6689.
44. Emerson SU, Nguyen HT, Torian U, Burke D, Engle R, Purcell RH. Release of genotype 1 hepatitis E virus from cultured hepatoma and polarized intestinal cells depends on open reading frame 3 protein and requires an intact PXXP motif. J Virol 2010;84:9059-9069.
45. Yamada K, Takahashi M, Hoshino Y, Takahashi H, Ichiyama K, Nagashima S, et al. ORF3 protein of hepatitis E virus is essential for virion release from infected cells. J Gen Virol 2009;90:1880-1891.
46. Takahashi M, Yamada K, Hoshino Y, Takahashi H, Ichiyama K, Tanaka T, et al. Monoclonal antibodies raised against the ORF3 protein of hepatitis E virus (HEV) can capture HEV particles in culture supernatant and serum but not those in feces. Arch Virol 2008;153:1703-1713.
47. Surjit M, Oberoi R, Kumar R, Lal SK. Enhanced alpha1 microglobulin secretion from Hepatitis E virus ORF3-expressing human hepatoma cells is mediated by the tumor susceptibility gene 101. J Biol Chem. 24 mars 2006;281:8135-42.
48. Nagashima S, Takahashi M, Jirintai null, Tanaka T, Yamada K, Nishizawa T, et al. A PSAP motif in the ORF3 protein of hepatitis E virus is necessary for virion release from infected cells. J Gen Virol. févr 2011;92:269-78.
49. Nagashima S, Takahashi M, Jirintai S, Tanaka T, Nishizawa T, Yasuda J, et al. Tumour susceptibility gene 101 and the vacuolar protein sorting pathway are required for the release of hepatitis E virions. J Gen Virol. déc 2011;92:2838-48.
50. Ivashkina N, Wölk B, Lohmann V, Bartenschlager R, Blum HE, Penin F, et al. The hepatitis C virus RNA-dependent RNA polymerase membrane insertion sequence is a transmembrane segment. J Virol. déc 2002;76:13088-93.
51. Promega. Chapter 8: post-translational modifications. In: Promega. In vitro protein expression guide. [Online]. 2005 [quoted 11 dec 2016]. 39-44. Available: https://ch.promega.com/-/media/files/resources/product-guides/in-vitro-protein-expression/ivex_gde.pdf?la=en
52. Kapp K, Schrempf S, Lemberg MK, Dobberstein B. Post-Targeting Functions of Signal Peptides. Landes Bioscience 2013.
53. Jameel S, Zafrullah M, Ozdener MH, Panda SK. Expression in animal cells and characterization of the hepatitis E virus structural proteins. J Virol 1996;70:207-216.
54. Sayed IM, Verhoye L, Cocquerel L, Abravanel F, Foquet L, Montpellier C, et al. Study of hepatitis E virus infection of genotype 1 and 3 in mice with humanised liver. Gut 2016;[Epub ahead of print].
55. Cuq J-L. Biochimie des protéines. [Online]. Université de Montpellier. Sep 2006. [quoted 11 dec 2016]. 1-147. Available: http://mon.univ-montp2.fr/claroline/backends/download.php?url=L0Jpb2NoX1Byb3RfSW50ZWdyZS5wZGY%3D&cidReset=true&cidReq=BIOCHPROT
56. Pouliot J-F, Béliveau R. Modifications post-traductionnelles des protéines par les lipides. M/S. [online]. Jan 1994. [quoted 11 dec 2016]. 10: 65-73. Available: http://www.ipubli.inserm.fr/bitstream/handle/10608/2481/MS_1994_1_65.pdf
57. Chandra V, Holla P, Ghosh D, Chakrabarti D, Padigaru M, Jameel S. The hepatitis E virus ORF3 protein regulates the expression of liver-specific genes by modulating localization of hepatocyte nuclear factor 4. PLoS One 2011;6:e22412.