Aus dem Institut für Virologie
des Fachbereichs Veterinärmedizin
der Freien Universität Berlin

Membrane Topology and Processing of Glycoprotein 3 (GP3) of Porcine Respiratory and Reproductive Syndrome Virus

Inaugural-Dissertation
zur Erlangung des Grades eines
Doctor of Philosophy (PhD) in Biomedical Science
an der
Freien Universität Berlin

vorgelegt von
Minze Zhang
Tierarzt aus Guangzhou, Volksrepublik China

Berlin 2018
Journal-Nr.: 4081

Aus dem Institut für Virologie
des Fachbereichs Veterinärmedizin
der Freien Universität Berlin

Membrane Topology and Processing of Glycoprotein 3 (GP3) of Porcine Respiratory and Reproductive Syndrome Virus

Inaugural-Dissertation
zur Erlangung des Grades eines
Doctor of Philosophy (PhD) in Biomedical Science
an der Freien Universität Berlin

vorgelegt von
Minze Zhang
Tierarzt aus Guangzhou, Volksrepublik China

Berlin 2018
Journal-Nr.: 4081

Gedruckt mit Genehmigung

des Fachbereichs Veterinärmedizin

der Freien Universität Berlin

Dekan: Univ.-Prof. Dr. Jürgen Zentek

Erster Gutachter: PD Dr. Michael Veit

Zweiter Gutachter: Univ.-Prof. Dr. Klaus Osterrieder

Dritter Gutachter: PD Dr. Soroush Sharbati

Deskriptoren (nach CAB-Thesaurus):

Porcine reproductive and respiratory syndrome virus, membranes, glycoproteins, amino acids, SDS-PAGE, fluorescence microscopy, real time PCR

Tag der Promotion: **16.08.2018**

Table of Contents

List of Figures

List of Tables

Abbreviations

BHK-21	Baby hamster kidney cells
C	Carboxyl terminus
cDNA	Complementary DNA
CD163	Scavenger receptor
CHO-K1	Chinese hamster ovary cells
Cys	Cysteine
EAV	Equine arteritis virus
Endo H	Endoglycosidase H
ER	Endoplasmic reticulum
ERGIC	ER-Golgi intermediate compartment
GFP	Green fluorescent protein
GP2	Glycoprotein 2
GP3	Glycoprotein 3
GP4	Glycoprotein 4
GP5	Glycoprotein 5
HR	Hydrophobic region
HCMV	Human cytomegalovirus
kb	Kilo base pair
kDa	Kilo Dalton
LDV	Lactate dehydrogenase-elevating virus
M	Membrane protein
MAb	monoclonal antibody
m.o.i	Multiplicity of infection
mRNA	Messenger RNA
N	Nucleocapsid protein
N	Amino terminus
nsp	Non-structural protein
ORF	Open reading frame
ORF5a	Open reading frame 5a protein
OST	Oligosaccharyl transferase
PAM	Porcine alveolar macrophage
PBS	Phosphate-buffered saline
PDI	Disulphide isomerases

PNGase F	Peptide-n-glycosidase
pp	Polyprotein
PRRSV	Porcine reproductive and respiratory syndrome virus
pSn	Porcine sialoadhesin
PVDV	Polyvinylidene difluoride
RFS	Ribosomal frame shift
SDS	Sodium dodecyl sulfate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
sgRNA	subgenomic RNA
SHFV	Simian hemorrhagic fever virus
SP	Signal peptide
SPase	Signal peptidase
SPC	Signal peptidase complex
SPP	Signal peptide peptidase
SR	Signal recognition particle receptor
SRCR	Scavenger receptor cysteine-rich domain
SRP	Signal recognition particle
TM	Transmembrane
TMD	Transmembrane domain
TMR	Transmembrane region
US11	Unique short 11 protein
UTR	Untranslated region
VLPs	Virus-like particles
WB	Western blot
WPDV	Wobbly possum disease virus
wt	Wild-type
YFP	Yellow fluorescent protein

Chapter 1

General introduction

Porcine reproductive and respiratory syndrome virus (PRRSV), a member of the family *Arteriviridae*, is the causative agent of porcine reproductive and respiratory syndrome (PRRS). It causes the most important infectious disease of pigs, leading to huge financial losses in the pork industry worldwide. In China highly pathogenic strains emerged that even kill 90% of infected pigs. So far, vaccines have failed to eliminate the virus, which is due to the large variation between strains and their ability to escape the immunity of the host.

1.1 Taxonomy and clinics of *Arteriviridae*

Arteriviridae, together with the family *Mesoniviridae*, *Roniviridae* and *Coronaviridae* belong to order *Nidovirales*, which are enveloped, positive sense, single stranded RNA viruses. The family *Arteriviridae* did contain only a single genus, *Arterivirus*, which, besides PRRSV, included equine arteritis virus (EAV) in horse, lactate dehydrogenase-elevating virus in mice (LDV) and simian hemorrhagic fever virus in monkey (SHFV). PRRSV has two distinct genotypes: European type (PRRSV-1) and North American type (PRRSV-2) (Benfield et al., 1992; Bryans et al., 1957; Notkins and Scheele, 1963; Tauraso et al., 1968; Wensvoort et al., 1991). However, several new viruses have been discovered recently: eleven highly divergent simian *Arteriviruses* in diverse African non-human primates, and a novel *Arterivirus,* the wobbly possum disease virus (WPDV) in common brushtails in New Zealand (Dunowska et al., 2012; Lauck et al., 2011). Thus, the *Arterivirus* nomenclature was not in accordance with the most recent version of the International Code of Virus Classification and Nomenclature and a revision of this classification has been proposed. As a result, the genus *Arterivirus* is replaced by five genera: *Equartevirus* (for EAV), *Rodartevirus* (LDV + PRRSV), *Simartevirus* (SHFV + simian *arteriviruses*), *Nesartevirus* (for the *arterivirus* from forest giant pouched rats), and *Dipartevirus* (common brushtail *arterivirus*) (Kuhn et al., 2016).

Clinical features of *Arterivirus* are divergent, but some are similar. For instance, both EAV and PRRSV can cause a systemic and persistent infection with respiratory diseases, and abortion in their host. However, the first two symptoms are rarely observed in infected macaques or mice by SHFV or LDV. Despite a lifelong persistent infection, LDV-infected mice usually have a normal life expectancy and exhibit no clinical syndromes other than the elevated level of enzymes in the blood and subtle

changes in their immune status. SHFV results in acute fatal hemorrhagic fever disease in macaques but long-term persistent infection in other species of African non-human primates, whereas the newly identified WPDV causes neurologic syndromes (Brinton et al., 2015).

1.2 Genome organization and virion structure

The PRRSV genome varies from 14.9 kb to 15.5 kb in length and expresses a range of accessory and structural proteins through two distinct transcription mechanisms. The genomic organization is depicted in Fig. 1.1 (Rascon-Castelo et al., 2015). The PRRSV genome begins with a 5' proximal untranslated region (5'UTR) of 217-222 nucleotides (PRRSV-1) and 188-191 nucleotides (PRRSV-2) in length (Yun and Lee, 2013). Downstream of the 5'UTR are at least 10 open reading frames (ORF). The ORF1a/b share a single translational start site but are augmented by ribosomal frame shift (RFS). The products of ORF1a/b are two large non-structural polyproteins, pp1a and pp1ab. These two polyproteins are processed by four viral proteases to release at least 16 distinct non-structural proteins (nsp): nsp1α, 1β, nsp2-6, nsp7α, 7β, nsp8, the RNA-dependent RNA polymerase (nsp9), a helicase (nsp10), an endonuclease (nsp11) and nsp12. These nsps assemble into membrane-associated complexes, which perform and regulate replication and transcription of the viral genome (Snijder et al., 2013). ORF 2-7 encode structural proteins (GP2, E, GP3, GP4, GP5, ORF5a, M and N), which partially overlap with neighboring genes. The 3' UTR is located directly downstream of ORF7 and consists of approximately 150 nucleotides excluding the polyadenylation site (polyA tail) (Beerens and Snijder, 2007; Choi et al., 2006; Yin et al., 2013).

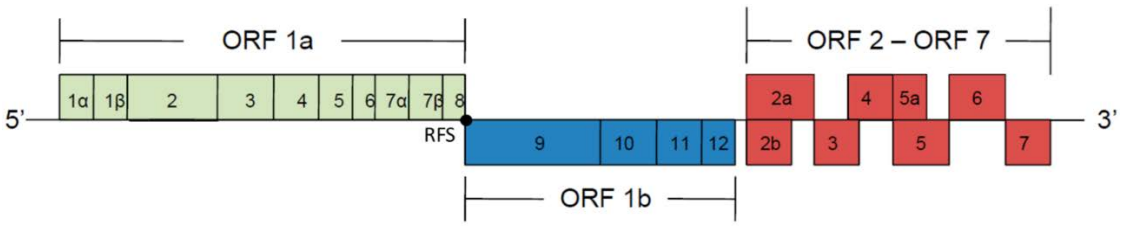

Figure 1.1 Genome organization of PRRSV. The ORF1a and ORF1b encode nonstructural proteins. The 3' end encodes structural proteins (ORF2-ORF7), which are transcribed as six subgenomic RNAs. (●): ribosomal frame shift (RFS) site. Modified from Rascón-Castelo et al., Viruses 2015, 7, 873-886.

PRRSV contains at least seven structural proteins (Fig. 1.2, (Veit et al., 2014)). The nucleocapsid protein N builds the scaffold for the genomic RNA, and several

membrane proteins are incorporated into the viral envelope: the glycoproteins GP2, GP3, GP4 and GP5 as well as the non-glycosylated membrane proteins M and E. GP5 and M form a disulphide-linked heterodimer, the major virion component , whereas GP2/3/4 form a disulphide-linked heterotrimeric complex in virus particles, and are referred along with E as minor structural proteins (Wissink et al., 2005). In addition, a recently discovered membrane-anchored protein ORF5a, encoded by an alternative open reading frame of the subgenomic mRNA encoding also GP5, is incorporated into virus particles, but probably as a very minor component (Johnson et al., 2011). Very recently, nsp2 was also shown to be incorporated into virions of a variety of PRRSV strains (Kappes et al., 2013). This protein is predicted to have several transmembrane-spanning regions and assists in the induction of double membrane vesicles, which are ER-derived membrane scaffolds that form the viral replication and transcription complex (Snijder et al., 2001). The role of nsp2 in virions is not clear yet.

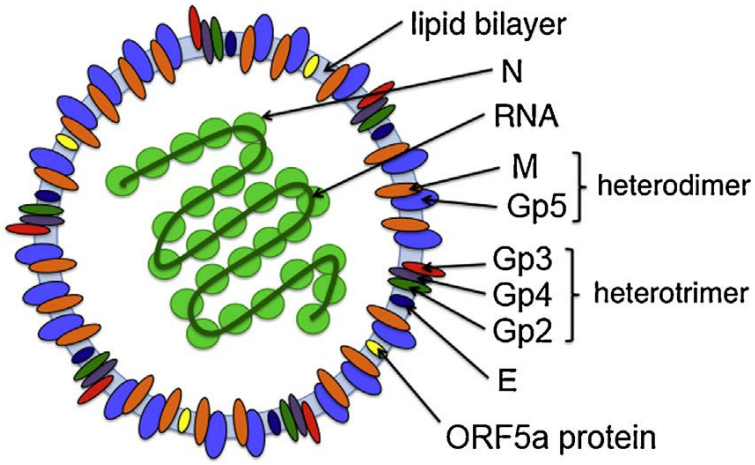

Figure 1.2 Scheme of PRRSV virion. The membrane contains the major GP5/M complex (blue and orange), the minor glycoprotein complex GP2/3/4 (green, red and purple, respectively), the small hydrophobic E protein (dark blue) and the ORF5a protein (yellow). Used with permission from Veit et al., Virus Res. 2014 Dec 19; 194:16-36.

1.3 Cell tropism, receptors and virus entry into host cells

PRRSV exhibits a restricted cell tropism. In vivo, the virus prefers cells of the monocyte/macrophage lineage and infects specific subsets of differentiated porcine macrophages in lungs, lymphoid tissues and placenta (Duan et al., 1997; Kreutz, 1998). In vitro PRRSV infects primary cultures of porcine alveolar macrophage (PAM) and the African green monkey kidney cell line MA-104 and cells derived thereof (Marc-145) (Kim et al., 1993).

A model on PRRSV entry into macrophages has been previously described and often adopted (Fig. 1.3(Van Breedam et al., 2010). PRRSV first attaches to macrophages via heparin sulphate (Delputte et al., 2002), then the virus is internalized via sialoadhesin (pSn, Siglec-1). Interaction of the virus with this receptor involves binding of the viral M/Gp5 complex to the N-terminal part of pSn. The sialic acid-binding domain at the N-terminus of pSn and sialic acids on GP5 on the virion surface are critical for this interaction (Van Breedam et al., 2010). Attachment of the virus to pSn is followed by the uptake of the virus-receptor complex via clathrin-mediated endocytosis (Nauwynck et al., 1999). Upon internalization, the viral genome is released via membrane fusion into the cytoplasm. This last stage of the entry occurs when the virus is present in the early endosome and is critically dependent on acidification of the endosome and on scavenger receptor CD163 (Calvert et al., 2007; Van Gorp et al., 2009). The role of CD163 in genome release or virus disassembly may require interaction with the GP2 and GP4 glycoproteins and relies on a functional CD163 SRCR domain 5 (Das et al., 2010; Van Gorp et al., 2010). In addition, successful infection is dependent on proteases in the endosome, including cathepsin E and a yet unidentified serine protease. The role of the proteases might be to cleave a viral protein and to release the fusion peptide in order to initiate membrane fusion (Harrison, 2008).

Among the described proteins, CD163 is now recognized to be the most essential receptor for PRRSV, as recombinant pigs having a deleted CD163 gene are fully protected from PRRSV infection (Burkard et al., 2017; Calvert et al., 2007; Whitworth et al., 2016). In contrast, deletion of pSn from the pig genome had no effect on PRRSV replication in vivo indicating that it is not essential for attachment or internalization of virus (Harrison, 2008).

Apart from that, other potential cell mediators involved in PRRSV entry were also reported. The nucleocapsid of PRRSV is described to bind to the intermediate filament vimentin, which is suggested to mediate transport of the virus to the cytosol (Kim et al., 2006). CD151 may be involved in fusion of the viral envelope and the endosome, but the precise mechanism is still unknown (Shanmukhappa et al., 2007). pDC-SIGN protein, a porcine C-type lectin, was found to be expressed on the surface of monocyte-derived macrophages and dendritic cells, alveolar macrophages cells. BHK cell expressing pDC-SIGN enhances the transmission of PRRSV to Marc-145 cells in trans (Huang et al., 2009).

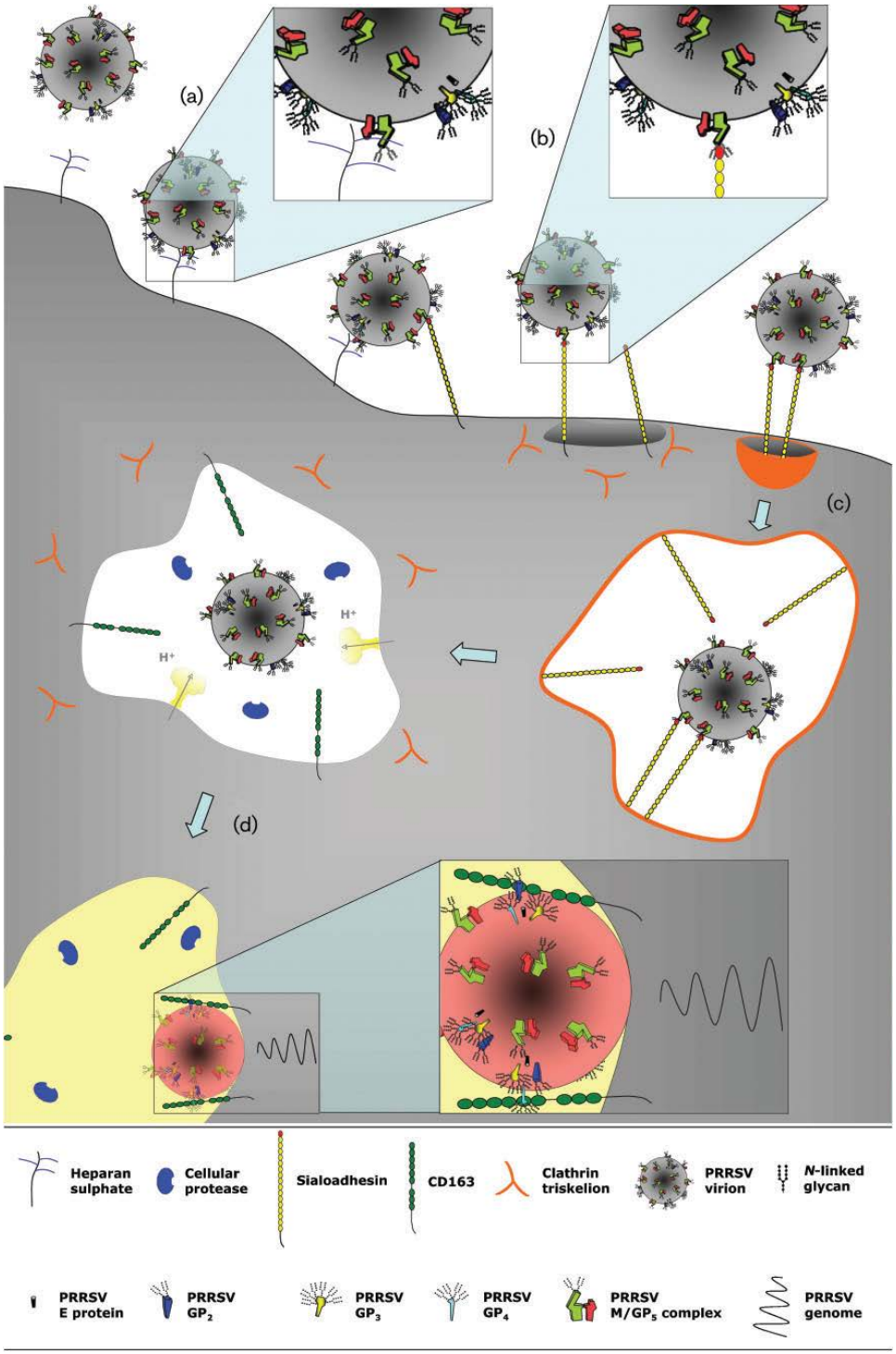

Figure 1.3 Model of PRRSV entry into the porcine macrophage. (a) PRRSV firstly attaches to heparan sulphate present on the macrophage surface. (b) Then the virus binds to the sialoadhesin via M/GP5 complexes present in the viral envelope. (c) Upon attachment to sialoadhesin, the virus-receptor complex is internalized via the clathrin-mediated endocytosis. (d) Upon internalization, the viral genome is released from the early endosome into the cytoplasm of the host cell, where CD163 is essential for this genome release and may exert its function through interaction with GP2 and GP4. In addition, cellular proteases have been

implicated and also a modest pH drop within the early endosome is crucial for viral genome release. Used from Van Breedam et al., J Gen Virol. 2010 Jul; 91(Pt 7):1659-67.

Siglec-10, a member of the same family as sialoadhesin was identified in macrophages and can significantly improve PRRSV infection and production in a CD163-transfected cell line (Xie et al., 2017). In addition, a non-muscle myosin heavy chain 9 (MYH9), a cellular protein in both Marc-145 cells and PAMs, has been recently identified as an important factor for PRRSV infection. MYH9 physically interacts with the PRRSV GP5 protein via its C-terminal domain and confers the improved susceptibility of cells to PRRSV infection if CD163 is co-expressed (Gao et al., 2016).

1.4 The major envelope proteins GP5 and M of PRRSV

The major components of the PRRSV envelope are **GP5** and **M**, which together comprise at least half of the viral proteins and form disulfide-linked heterodimers in the virus (Mardassi et al., 1996; Meulenberg et al., 1995). The glycosylated GP5 of 200 amino acids for PRRSV-2 (201 for PRRSV-1), encoded by ORF5, has a size of about 25 kDa. It possesses a cleavable signal peptide (SP), followed by an ectodomain of roughly 30 amino acids, containing several N-glycosylated sites, two of which (N44 and N51 in PRRSV-2, N46 and N53 in PRRSV-1) are highly conserved between virus strains. The region between residues 60-125 of GP5 comprises a hydrophobic region that includes either one or three transmembrane (TM) helices, followed by a large C-terminal endodomain containing around residues 130 to 200 (Dea et al., 2000).

Additionally, the **ORF5a** is a non-glycosylated protein, encoded by an alternative reading frame of ORF5 that is present in all *Arteriviruses*, and contains 51 residues in PRRSV-2 (52 in PRRSV-1). ORF5a protein is essential for virus replication in EAV and PRRSV (Firth et al., 2011; Johnson et al., 2011; Sun et al., 2013). ORF5a protein of PRRSV-2 possesses two cysteines at positions 29 and 30, which are not essential for virus viability (Sun et al., 2015).

The 16 kDa (174 amino acids in PRRSV-2, 173 for PRRSV-1) non-glycosylated membrane protein M, encoded by ORF6, contains a short 16-residue N-terminal ectodomain followed by three transmembrane (TM) segments and an 84-residue C-terminal endodomain. It is the most conserved structural protein of PRRSV. Disulphide-linked heterodimer are formed by M and GP5 via Cys 9 of M and Cys 48 of GP5 in PRRSV-2 and via Cys 8 and Cys 50 in PRRSV-1 (Dea et al., 2000; Mardassi et al., 1996; Wissink et al., 2005).

The product of ORF7, between 123-128 amino acids N protein, interacts with the viral RNA to form the viral nucleocapsid, and is the most abundant viral protein expressed in infected cells (Dea et al., 2000). N, together with GP5 and M are minimal requirements of virus-like particles formation (VLPs), therefore their main function is assembly and budding.

1.5 The minor envelope proteins GP2, GP3, GP4 and E of PRRSV

The **GP2** glycoprotein, encoded by ORF2a, has 256 amino acids in PRRSV-2 (249 in PRRSV-1). GP2 contains a predicted N-terminal signal peptide of residues 1-40 in PRRSV-2 (1-37 in PRRSV-1), followed by a 168 residues ectodomain, a single transmembrane domain (TMD) and a 20 residues cytoplasmic tail. GP2 has two conserved glycosylation sites, at residues N178 and N184 in PRRSV-2 (N173 and N179 in PRRSV-1), However, the glycosylation was not essential for infectious virus production in both PRRSV-1 and PRRSV-2 (Wei et al., 2012; Wissink et al., 2004). By contrast, another study showed that the N184 glycosylation site in GP2 of PRRSV-2 is necessary for infectious virus production (Das et al., 2011)

The small non-glycosylated **E protein** is encoded from ORF2b, has 70-73 residues, which is fully embedded within ORF2a that encodes GP2 (Wu et al., 2001; Wu et al., 2005). E consists of a single predicted TMD and is thought to form an oligomeric ion channel. This protein is most likely involved in the viral fusion and/or internalization process, as proton channels are required to lower the internal pH of the virus during fusion in the low pH endosomal compartment (Lee and Yoo, 2006).

The **GP4 glycoprotein**, encoded by ORF4, has 178 amino acids in PRRSV-2 (183 in PRRSV-1) (van Nieuwstadt et al., 1996). It contains two main hydrophobic domains: a predicted cleaved signal peptide of 1-21 residue, and a C-terminal hydrophobic region (HR) at residues 156-177 in PRRSV-2 (161-181 for PRRSV-1). GP4 has four potential glycosylation sites at residues 37, 84, 120 and 130. Individual glycosylation site in GP4 is not essential for infectious virus recovery but mutations at least two are fatal (Das et al., 2011). In contrast, another study showed that mutations of even two glycosylation sites in GP4 are not critically important for infectious virus recovery, triple and quadruple mutations are lethal (Wei et al., 2012).

The **GP3 glycoprotein**, the focus of my work, is encoded by ORF3. It has 251-265 amino acids and is the most heavily glycosylated envelope protein of PRRSV. GP3 contains two predicted hydrophobic domains: a N-terminal signal peptide (SP) and a C-terminal hydrophobic region (HR). The overall sequence identity between the PRRSV-1 and PRRSV-2 is 58%, but the highest divergence is in the C-terminus of the

protein, where the PRRSV-1 GP3 normally contains additional 3-11 amino acids. GP3 contains 6-8 potential glycosylation sites in the putative ectodomain, depending on strains. One at N195 is not used, which presumably is due to its location in the predicted hydrophobic region at the C-terminus (Das et al., 2011). Interestingly, none of the individual glycosylation site in GP3 has a vital effect on the production of infectious virus and nor does it affect the susceptibility of virus to neutralizing antibody (Wei et al., 2012). However, another study showed that glycan addition at N42, N50 and N131 of GP3 is necessary for infectious virus production (Das et al., 2011), additionally N131 of GP3 is involved in glycan shielding (Vu et al., 2011).

All of the mentioned membrane proteins of PRRSV are essential to generate recombinant infectious PRRSV and EAV virus by reverse genetics. However, in the absence of either GP2, GP3, GP4 or E (but not GP5 or M) virus like-particles are formed. This implicates that the GP5/M complex is required for virus budding. The deletion of any of the four genes encoding GP2, GP3 GP4 or E led to a failure to incorporate the other three (Das et al., 2010; Wissink et al., 2005). GP2, GP3 and GP4 are incorporated as disulfide-linked heterotrimeric complex into the envelope of EAV (Wieringa et al., 2003), but complex formation of the respective proteins of PRRSV has not been investigated. In PRRSV GP2 and GP4 or GP4 alone were shown to associate with CD163, which is the most important receptor of PRRSV (Das et al., 2010). Another studies showed that the minor envelope proteins indeed are responsible for cell tropism of *Arteriviruses* (Tian et al., 2012).

1.6 Translocation and Processing of glycoproteins

As the first step in their biosynthesis membrane and soluble (secretory) proteins must be targeted to the ER using an N-terminal signal sequence/signal peptide (SP) as targeting sequence. At the ER the nascent polypeptide chain is handed over to the translocon, a hetero-oligomeric protein complex serving as a channel for translocation of proteins into the lumen of the rough endoplasmic reticulum (ER) (Rapoport et al., 2017). The translocon has a so-called lateral gate that opens if a hydrophobic polypeptide sequence appears (Fig. 1.4 (Zhang and Miller, 2010)). This allows this part of the translocating polypeptide chain to move from the hydrophilic pore of the translocon into the hydrophobic lipids of the ER membrane. If the amino acid sequence is hydrophilic the lateral gate remains closed and translocation of the protein into the ER lumen continues. This mechanism determines the membrane spanning parts of a protein and also the parts which are in the lumen of the ER (ectodomain) and in the cytosol (endodomain).

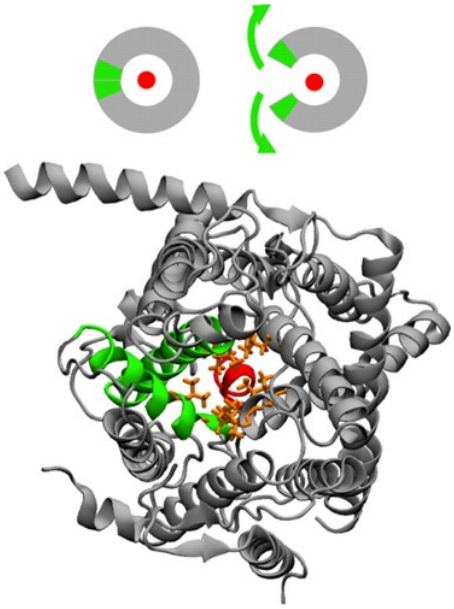

Figure 1.4 The structure of the ER-located translocon. The figure shows the crystal structure of the translocon (viewed from outside the membrane on the cytosolic side) with the translocating polypeptide chain in red, the pore ring in orange, that seals resting translocons from ion influx and the two transmembrane regions in green that form the lateral gate. The two transmembrane regions open if the translocating part of the freshly synthesized protein is hydrophobic, which allows their release into the hydrophobic membrane lipids surrounding the translocon. Used from Zhang et al., PNAS (2010), 107 (12) 5399-5404.

The ectodomain of viral envelope glycoproteins are subjected to co-translational modifications, such as signal peptide cleavage, N-glycosylation and disulphide bond formation. All these events are essential for correct folding, oligomerization and trafficking of PRRSV proteins to the site of viral budding (Braakman and Bulleid, 2011), which is the “late” ER or an early Golgi region. There is no consensus sequence for the SP, but they have a similar overall structure. The SP is approximately 20-30 amino acids long (but might be up to 80), with a typical tripartite structure: an n-region consisting of basic amino acids, a hydrophobic h-region in the middle, and a slightly polar c region (Auclair et al., 2012) (Fig. 1.5).

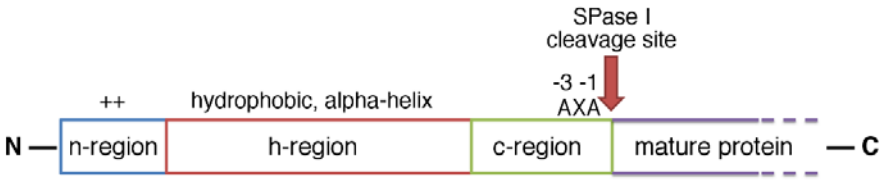

Figure 1.5 The structure of signal peptide. N-region possess positively charged residues (blue), the h- region contain hydrophobic residues (red), and c- region is typically neutral and contains cleavage site. N- amino terminus; C- carboxyl terminus. Modified from Auclair *et al*., Protein Science 21 (2012) 13-25.

In most membrane proteins the SP is not present in the mature protein, as it is cleaved by the signal peptidase (SPase), which is a serine protease and forms a large complex with other proteins, called the signal peptidase complex (SPC) which is located on the luminal side of the ER membrane (Nyathi et al., 2013; Shelness et al., 1993) (Fig. 1.6). The most studied bacterial homologue signal peptidase I recognizes and cleaves signal peptides with the consensus motif alanine-X-alanine in the c-region of the SP. In eukaryotes it is more complex, the -1 and -3 amino acid residues upstream from the cleavage site are typically small neutral residues, such as alanine, glycine, cysteine and serine (Auclair et al., 2012).

N-linked glycosylation is another important co-translational modification, in which a high-mannose core is attached to the amide nitrogen of asparagine in the context of the conserved motif N-X-T/S (asparagine-any amino acid except proline-threonine or serine). It is the most abundant protein modification found in eukaryotes. In the case of viral glycoproteins, N-linked glycosylation is important not only for correct protein folding and subsequent intracellular trafficking, but also for a variety of functions, such as receptor binding, viral membrane fusion, virulence and immune evasion (Vigerust and Shepherd, 2007). The enzyme complex for N-glycosylation (oligosaccharyl transferase, OST) are also associated with the translocon (Shrimal et al., 2015), where the active center is oriented towards the lumen of the ER (Fig. 1.6). Thus, in principle both OST and SPase can perform their activity once a nascent polypeptide chain becomes accessible.

Another important protein co-translational modification occurring in the ER is **disulphide bond formation**. Disulphide bonds are formed by oxidation of sulfhydryl (-SH) groups of two cysteine residues in close spatial proximity. This occurs spontaneously, favored by the reducing milieu in the ER, but unfavorable bonds that

do not allow stable protein folding, are reduced by the protein disulphide isomerases (PDI), a soluble enzyme in the ER lumen (Oka and Bulleid, 2013). This modification possibly competes with signal peptide cleavage and N-glycosylation. Disulphide bond formation generally reduces glycosylation at adjacent sites, probably because a folded and fixed protein conformation cannot be processed by OST(Allen et al., 1995) (Fig. 1.6).

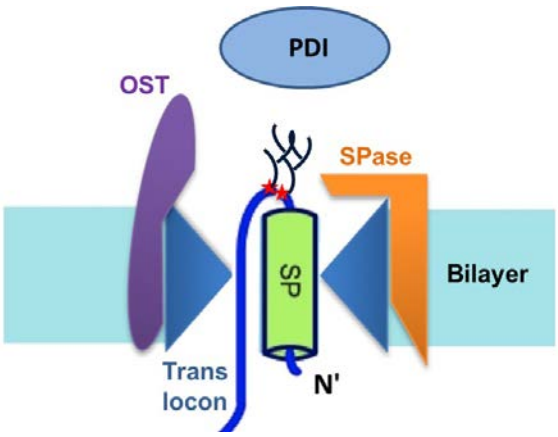

Figure 1.6 Co-translational protein modification. The glycoprotein is subjected to co-translational modifications, such as signal peptide cleavage, N-glycosylation and disulphide bond formation, which are performed by signalpeptidase (SPase), oligosacharytransferase (OST) and protein disulphide isomerases (PDI). The first two are associated with the translocon. Red asterisks: NXT/S, N-glycosylation consensus motif; Branches: N-glycan: SP: signal peptide.

1.7 Aims of study

Previous work from our lab revealed a unique feature for processing of GP3 of EAV: the SP of GP3 from EAV is predicted to be cleaved at two sites G26/S27 or S27/X28 via prediction software SignalP 4.1 (Fig. 1.7).

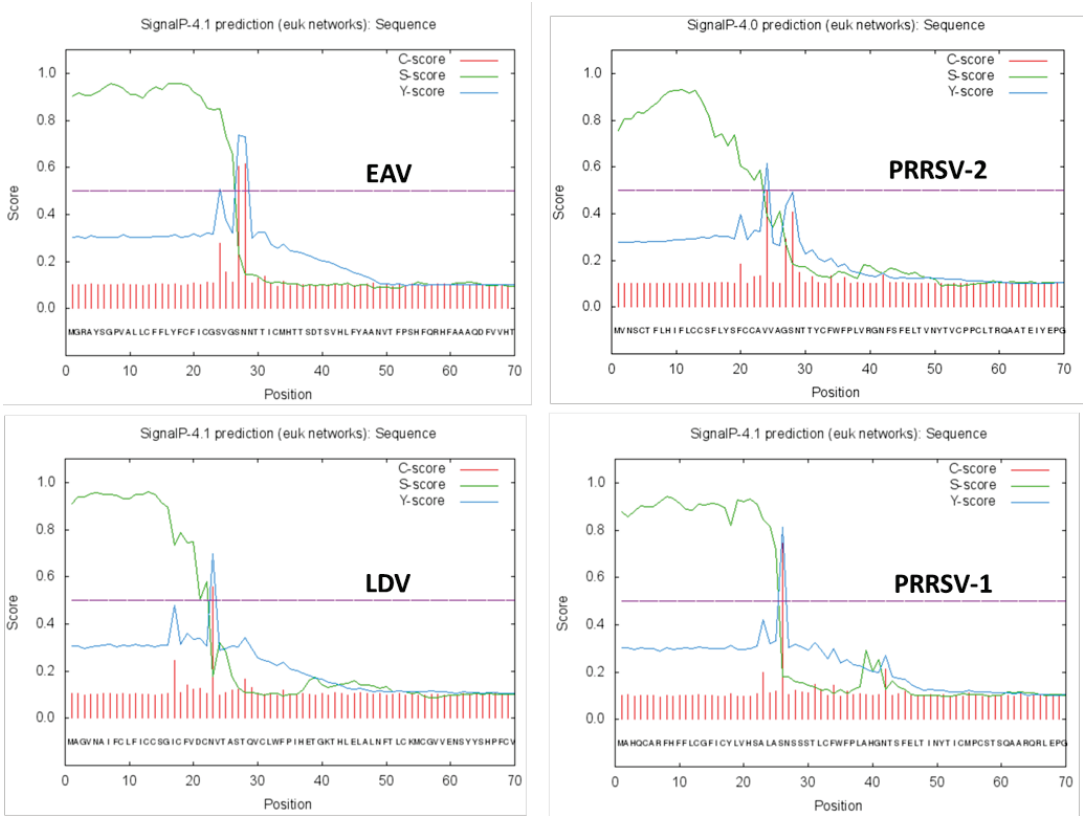

Figure 1.7 Signal peptide cleavage site of GP3 from *Arterivirus* is predicted by SignalP 4.1 software (http://www.cbs.dtu.dk/services/SignalP/). EAV: equine arteritis virus (Bucyrus strain); LDV: lactate dehydrogenase-elevating virus (Plagemann strain); PRRSV-1: porcine reproductive and respiratory syndrome virus of type 1 (Lelystad strain); RRRSV-2: porcine reproductive and respiratory syndrome virus of type 2 (VR-2332 strain). X-axis: amino acid position; C-score (red vertical lines): signal peptide cleavage score calculated for each amino acid; S-score (green line): signal peptide score (high scores indicate that the corresponding amino acid is part of a signal peptide, and low scores indicate that the amino acid is part of a mature protein); y-score (blue line): combined C and S scores, shows the most probable SP cleavage site.

SignalP calculates the probability of a particular sequence to be a signal peptide and gives the output of the most probable SP cleavage site (Petersen et al., 2011). However, the GP3 of EAV has an uncleaved SP. SP cleavage is prevented via two carbohydrates attached near the SP sequence (Matczuk et al., 2013). Deletion of both (but not only one) glycosylation sites of the overlapping sequon NNTT allows signal peptide cleavage (Fig. 1.8), indicating that co-translational attachment of one carbohydrate chain inhibits access of the signal peptidase to the cleavage site (Matczuk et al., 2013). This unique phenomenon was also confirmed in recombinant viruses with disabled glycosylation sites. Furthermore, the infectivity of recombinant

EAV containing GP3 with cleaved signal peptide was not impaired and GP3 without signal peptide associates with GP2/4 in virus particles The functional relevance why evolution created such a complex procedure to prevent signal peptide cleavage from GP3 of EAV is not known, but it was speculated that loss of glycosylation sites (that was described for EAV present in persistently infected stallions) creates a protein without signal peptide which also might have lost epitopes for neutralizing antibodies.

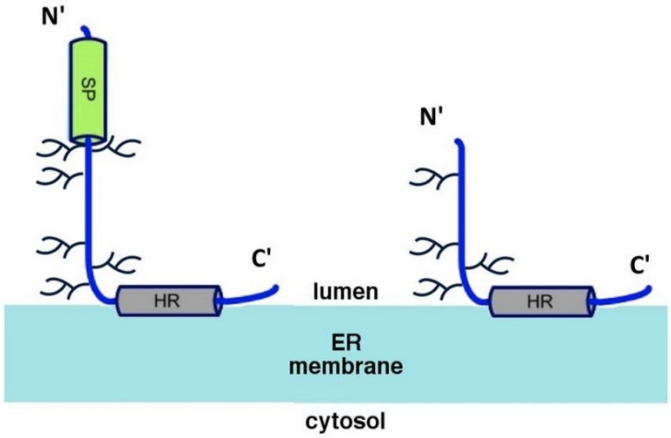

Figure 1.8 Model for signal peptide cleavage and the membrane topology of GP3 of EAV. The uncleaved signal peptide is completely translocated into the lumen of the ER. Signal peptide cleavage is inhibited by two carbohydrates attached to an overlapping sequon in the vicinity of the signal peptide. The protein is anchored by a hydrophobic region to the ER membrane. SP: signal peptide; HR: hydrophobic region; Branches: carbohydrates.

The **first purpose** of this study was to explore whether GP3 proteins from PRRSV (and from LDV) follow the same unusual processing scheme as GP3 from EAV. All GP3 proteins contain a conserved glycosylation site near the signal peptide (Table 2.1). Although more than 1000 nucleotide sequences encoding GP3 proteins especially from PRRSV strains are deposited in databases, biochemical studies on processing of the protein are rare. The signal peptide of GP3 from PRRSV-1 and PRRSV-2 is predicted to be cleaved at site A23/V24 and A25/S26 respectively via prediction software SignalP 4.1 (Fig. 1.7). However, no biochemical processing study existed for GP3 of any of the many PRRSV strains, but it was claimed that the signal peptide serves to retain GP3 in the ER (Kim et al., 2013).On the other hand, the signal peptide of GP3 from LDV is predicted to be cleaved at site C22/N23 (Fig. 1.7), whereas it has been shown that GP3 of LDV retains the signal peptide and it was speculated that this might be due to a carbohydrate attached close to the signal peptide cleavage site (Faaberg and Plagemann, 1997).

The **second purpose** of this study was to apply biochemical methods to experimentally analyze the membrane topology of GP3 of PRRSV. For viral structural proteins, the membrane topology will define which parts of the protein will be hidden inside the virion, and which parts will protrude from its surface. For GP3 of EAV, the protein is tethered peripherally to the ER membrane by the hydrophobic region at its C-terminus, not by the uncleaved signal peptide ((Matczuk et al., 2013), Fig. 1.8). GP3 proteins from Arteriviruses vary in length and hydrophobicity. The GP3 of LDV and especially of both PRRSV-1 and PRRSV-2 strains have longer amino acid sequences and their C-terminal part is much less hydrophobic than the one from EAV. Reduced hydrophobicity in the C-terminus is especially visible in GP3 from PRRSV-1 strains (Fig. 3.9A). Therefore, it is interesting to address whether the membrane anchoring of GP3 of PRRSV is similar to that of EAV?

The **third purpose** of this study was to systematically investigate whether GP3 proteins from PRRSV-1 and PRRSV-2 strains are secreted from transfected cells. Inconsistent results have been published about whether GP3s are structural proteins of *Arterivirus* particles. Only for EAV it has been clearly shown, that GP3 is present in a disulfide-linked trimer with GP2/4 in virus particles (Wieringa et al., 2003; Wieringa et al., 2004). In contrast, GP3 from LDV is only weakly associated with membranes and at least a fraction of the protein is secreted (Faaberg and Plagemann, 1997). Initial studies with GP3 from the PRRSV-2 strain IAF-Klop showed that it is secreted both from transfected and infected cells in a membrane-free form as a disulphide-linked dimer carrying complex type carbohydrates (Gonin et al., 1998; Mardassi et al., 1998; Wieringa et al., 2002). In contrast, a more recent study convincingly demonstrated that GP3 from the PRRSV-2 FL-12 strain is incorporated into virus particles (de Lima et al., 2009) as is GP3 from PRRSV-1 strain Lelystad (Wissink et al., 2005). However, in the latter two studies it was not investigated whether a fraction of GP3 is secreted from cells. These partly conflicting results might be due to the difficulty in purifying PRRSV viruses, the low abundance of GP2/3/4 in particles, the large amino acid variability between GP3 proteins of PRRSV strains and (as a consequence) the lack of cross-reactive antibodies against the minor glycoproteins. Therefore, I analyzed systematically whether GP3 proteins from several PRRSV-1 and PRRSV-2 strains are secreted from transfected cells.

1.8 Reference

Allen, S., Naim, H.Y., Bulleid, N.J., 1995. Intracellular folding of tissue-type plasminogen activator. Effects of disulfide bond formation on N-linked glycosylation and secretion. J Biol Chem 270, 4797-4804.

Auclair, S.M., Bhanu, M.K., Kendall, D.A., 2012. Signal peptidase I: cleaving the way to mature proteins. Protein Sci 21, 13-25.

Beerens, N., Snijder, E.J., 2007. An RNA pseudoknot in the 3 ' end of the arterivirus genome has a critical role in regulating viral RNA synthesis. Journal of Virology 81, 9426-9436.

Benfield, D.A., Nelson, E., Collins, J.E., Harris, L., Goyal, S.M., Robison, D., Christianson, W.T., Morrison, R.B., Gorcyca, D., Chladek, D., 1992. Characterization of swine infertility and respiratory syndrome (SIRS) virus (isolate ATCC VR-2332). J Vet Diagn Invest 4, 127-133.

Braakman, I., Bulleid, N.J., 2011. Protein Folding and Modification in the Mammalian Endoplasmic Reticulum, in: Kornberg, R.D., Raetz, C.R.H., Rothman, J.E., Thorner, J.W. (Eds.), Annual Review of Biochemistry, Vol 80. Annual Reviews, Palo Alto, pp. 71-99.

Brinton, M.A., Di, H., Vatter, H.A., 2015. Simian hemorrhagic fever virus: Recent advances. Virus Res 202, 112-119.

Bryans, J.T., Crowe, M.E., Doll, E.R., McCollum, W.H., 1957. Isolation of a filterable agent causing arteritis of horses and abortion by mares; its differentiation from the equine abortion (influenza) virus. Cornell Vet 47, 3-41.

Burkard, C., Lillico, S.G., Reid, E., Jackson, B., Mileham, A.J., Ait-Ali, T., Whitelaw, C.B.A., Archibald, A.L., 2017. Precision engineering for PRRSV resistance in pigs: Macrophages from genome edited pigs lacking CD163 SRCR5 domain are fully resistant to both PRRSV genotypes while maintaining biological function. Plos Pathogens 13.

Calvert, J.G., Slade, D.E., Shields, S.L., Jolie, R., Mannan, R.M., Ankenbauer, R.G., Welch, S.K.W., 2007. CD163 expression confers susceptibility to porcine reproductive and respiratory syndrome viruses. Journal of Virology 81, 7371-7379.

Choi, Y.J., Yun, S.I., Kang, S.Y., Lee, Y.M., 2006. Identification of 5 ' and 3 ' cis-acting, elements of the porcine reproductive and respiratory syndrome virus: Acquisition of novel 5 ' AU-rich sequences restored replication of a 5 '-proximal 7-nucleotide deletion mutant. Journal of Virology 80, 723-736.

Das, P.B., Dinh, P.X., Ansari, I.H., de Lima, M., Osorio, F.A., Pattnaik, A.K., 2010. The minor envelope glycoproteins GP2a and GP4 of porcine reproductive and respiratory syndrome virus interact with the receptor CD163. J Virol 84, 1731-1740.

Das, P.B., Vu, H.L., Dinh, P.X., Cooney, J.L., Kwon, B., Osorio, F.A., Pattnaik, A.K., 2011. Glycosylation of minor envelope glycoproteins of porcine reproductive and respiratory syndrome virus in infectious virus recovery, receptor interaction, and immune response. Virology 410, 385-394.

de Lima, M., Ansari, I.H., Das, P.B., Ku, B.J., Martinez-Lobo, F.J., Pattnaik, A.K., Osorio, F.A., 2009. GP3 is a structural component of the PRRSV type II (US) virion. Virology 390, 31-36.

Dea, S., Gagnon, C.A., Mardassi, H., Pirzadeh, B., Rogan, D., 2000. Current knowledge on the structural proteins of porcine reproductive and respiratory syndrome (PRRS) virus: comparison of the North American and European isolates. Arch Virol 145, 659-688.

Delputte, P.L., Vanderheijden, N., Nauwynck, H.J., Pensaert, M.B., 2002. Involvement of the matrix protein in attachment of porcine reproductive and respiratory syndrome virus to a heparinlike receptor on porcine alveolar macrophages. Journal of Virology 76, 4312-4320.

Duan, X., Nauwynck, H.J., Pensaert, M.B., 1997. Virus quantification and identification of cellular targets in the lungs and lymphoid tissues of pigs at different time intervals after inoculation with porcine reproductive and respiratory syndrome virus (PRRSV). Vet Microbiol 56, 9-19.

Dunowska, M., Biggs, P.J., Zheng, T., Perrott, M.R., 2012. Identification of a novel nidovirus associated with a neurological disease of the Australian brushtail possum (Trichosurus vulpecula). Vet Microbiol 156, 418-424.

Faaberg, K.S., Plagemann, P.G., 1997. ORF 3 of lactate dehydrogenase-elevating virus encodes a soluble, nonstructural, highly glycosylated, and antigenic protein. Virology 227, 245-251.

Firth, A.E., Zevenhoven-Dobbe, J.C., Wills, N.M., Go, Y.Y., Balasuriya, U.B., Atkins, J.F., Snijder, E.J., Posthuma, C.C., 2011. Discovery of a small arterivirus gene that overlaps the GP5 coding sequence and is important for virus production. J Gen Virol 92, 1097-1106.

Gao, J.M., Xiao, S.Q., Xiao, Y.H., Wang, X.P., Zhang, C., Zhao, Q., Nan, Y.C., Huang, B.C., Liu, H.L., Liu, N.N., Lv, J.H., Du, T.F., Sun, Y.N., Mu, Y., Wang, G., Syed, S.F., Zhang, G.P., Hiscox, J.A., Goodfellow, I., Zhou, E.M., 2016. MYH9 is an Essential Factor for Porcine Reproductive and Respiratory Syndrome Virus Infection. Sci Rep-Uk 6.

Gonin, P., Mardassi, H., Gagnon, C.A., Massie, B., Dea, S., 1998. A nonstructural and antigenic glycoprotein is encoded by ORF3 of the IAF-Klop strain of porcine reproductive and respiratory syndrome virus. Arch Virol 143, 1927-1940.

Harrison, S.C., 2008. Viral membrane fusion. Nat Struct Mol Biol 15, 690-698.

Huang, Y.W., Dryman, B.A., Li, W., Meng, X.J., 2009. Porcine DC-SIGN: Molecular cloning, gene structure, tissue distribution and binding characteristics. Dev Comp Immunol 33, 464-480.

Johnson, C.R., Griggs, T.F., Gnanandarajah, J., Murtaugh, M.P., 2011. Novel structural protein in porcine reproductive and respiratory syndrome virus encoded by an alternative ORF5 present in all arteriviruses. J Gen Virol 92, 1107-1116.

Kappes, M.A., Miller, C.L., Faaberg, K.S., 2013. Highly divergent strains of porcine reproductive and respiratory syndrome virus incorporate multiple isoforms of nonstructural protein 2 into virions. J Virol 87, 13456-13465.

Kim, D.G., Song, C.S., Choi, I.S., Park, S.Y., Lee, J.B., Lee, S.S., 2013. The signal sequence of type II porcine reproductive and respiratory syndrome virus glycoprotein 3 is sufficient for endoplasmic reticulum retention. J Vet Sci 14, 307-313.

Kim, H.S., Kwang, J., Yoon, I.J., Joo, H.S., Frey, M.L., 1993. Enhanced replication of porcine reproductive and respiratory syndrome (PRRS) virus in a homogeneous subpopulation of MA-104 cell line. Arch Virol 133, 477-483.

Kim, J.K., Fahad, A.M., Shanmukhappa, K., Kapil, S., 2006. Defining the cellular target(s) of porcine reproductive and respiratory syndrome virus blocking monoclonal antibody 7G10. Journal of Virology 80, 689-696.

Kreutz, L.C., 1998. Cellular membrane factors are the major determinants of porcine reproductive and respiratory syndrome virus tropism. Virus Res 53, 121-128.

Kuhn, J.H., Lauck, M., Bailey, A.L., Shchetinin, A.M., Vishnevskaya, T.V., Bao, Y., Ng, T.F., LeBreton, M., Schneider, B.S., Gillis, A., Tamoufe, U., Diffo Jle, D., Takuo, J.M., Kondov, N.O., Coffey, L.L., Wolfe, N.D., Delwart, E., Clawson, A.N., Postnikova, E., Bollinger, L., Lackemeyer, M.G., Radoshitzky, S.R., Palacios, G., Wada, J., Shevtsova, Z.V., Jahrling, P.B., Lapin, B.A., Deriabin, P.G., Dunowska, M., Alkhovsky, S.V., Rogers, J., Friedrich, T.C., O'Connor, D.H., Goldberg, T.L., 2016. Reorganization and expansion of the nidoviral family Arteriviridae. Arch Virol 161, 755-768.

Lauck, M., Hyeroba, D., Tumukunde, A., Weny, G., Lank, S.M., Chapman, C.A., O'Connor, D.H., Friedrich, T.C., Goldberg, T.L., 2011. Novel, divergent simian hemorrhagic fever viruses in a wild Ugandan red colobus monkey discovered using direct pyrosequencing. PLoS One 6, e19056.

Lee, C., Yoo, D., 2006. The small envelope protein of porcine reproductive and respiratory syndrome virus possesses ion channel protein-like properties. Virology 355, 30-43.

Mardassi, H., Gonin, P., Gagnon, C.A., Massie, B., Dea, S., 1998. A subset of porcine reproductive and respiratory syndrome virus GP3 glycoprotein is released into the culture medium of cells as a non-virion-associated and membrane-free (soluble) form. J Virol 72, 6298-6306.

Mardassi, H., Massie, B., Dea, S., 1996. Intracellular synthesis, processing, and transport of proteins encoded by ORFs 5 to 7 of porcine reproductive and respiratory syndrome virus. Virology 221, 98-112.

Matczuk, A.K., Kunec, D., Veit, M., 2013. Co-translational processing of glycoprotein 3 from equine arteritis virus: N-glycosylation adjacent to the signal peptide prevents cleavage. J Biol Chem 288, 35396-35405.

Meulenberg, J.J., Petersen-den Besten, A., De Kluyver, E.P., Moormann, R.J., Schaaper, W.M., Wensvoort, G., 1995. Characterization of proteins encoded by ORFs 2 to 7 of Lelystad virus. Virology 206, 155-163.

Nauwynck, H.J., Duan, X., Favoreel, H.W., Van Oostveldt, P., Pensaert, M.B., 1999. Entry of porcine reproductive and respiratory syndrome virus into porcine alveolar macrophages via receptor-mediated endocytosis. J Gen Virol 80 (Pt 2), 297-305.

Notkins, A.L., Scheele, C., 1963. An Infectious Nucleic Acid from the Lactic Dehydrogenase Agent. Virology 20, 640-642.

Nyathi, Y., Wilkinson, B.M., Pool, M.R., 2013. Co-translational targeting and translocation of proteins to the endoplasmic reticulum. Bba-Mol Cell Res 1833, 2392-2402.

Oka, O.B., Bulleid, N.J., 2013. Forming disulfides in the endoplasmic reticulum. Biochim Biophys Acta 1833, 2425-2429.

Petersen, T.N., Brunak, S., von Heijne, G., Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8, 785-786.
Rapoport, T.A., Li, L., Park, E., 2017. Structural and Mechanistic Insights into Protein Translocation. Annu Rev Cell Dev Bi 33, 369-390.

Rascon-Castelo, E., Burgara-Estrella, A., Mateu, E., Hernandez, J., 2015. Immunological features of the non-structural proteins of porcine reproductive and respiratory syndrome virus. Viruses 7, 873-886.

Shanmukhappa, K., Kim, J.K., Kapil, S., 2007. Role of CD151, A tetraspanin, in porcine reproductive and respiratory syndrome virus infection. Virol J 4.

Shelness, G.S., Lin, L.J., Nicchitta, C.V., 1993. Membrane Topology and Biogenesis of Eukaryotic Signal Peptidase. Journal of Biological Chemistry 268, 5201-5208.

Shrimal, S., Cherepanova, N.A., Gilmore, R., 2015. Cotranslational and posttranslocational N-glycosylation of proteins in the endoplasmic reticulum. Semin Cell Dev Biol 41, 71-78.

Snijder, E.J., Kikkert, M., Fang, Y., 2013. Arterivirus molecular biology and pathogenesis. J Gen Virol 94, 2141-2163.

Snijder, E.J., van Tol, H., Roos, N., Pedersen, K.W., 2001. Non-structural proteins 2 and 3 interact to modify host cell membranes during the formation of the arterivirus replication complex. Journal of General Virology 82, 985-994.

Sun, L., Li, Y., Liu, R., Wang, X., Gao, F., Lin, T., Huang, T., Yao, H., Tong, G., Fan, H., Wei, Z., Yuan, S., 2013. Porcine reproductive and respiratory syndrome virus ORF5a protein is essential for virus viability. Virus Res 171, 178-185.

Sun, L., Zhou, Y., Liu, R., Li, Y., Gao, F., Wang, X., Fan, H., Yuan, S., Wei, Z., Tong, G., 2015. Cysteine residues of the porcine reproductive and respiratory syndrome virus ORF5a protein are not essential for virus viability. Virus Res 197, 17-25.

Tauraso, N.M., Shelokov, A., Palmer, A.E., Allen, A.M., 1968. Simian hemorrhagic fever. 3. Isolation and characterization of a viral agent. Am J Trop Med Hyg 17, 422-431.

Tian, D., Wei, Z., Zevenhoven-Dobbe, J.C., Liu, R., Tong, G., Snijder, E.J., Yuan, S., 2012. Arterivirus Minor Envelope Proteins Are a Major Determinant of Viral Tropism in Cell Culture. Journal of virology 86, 3701-3712.

Van Breedam, W., Delputte, P.L., Van Gorp, H., Misinzo, G., Vanderheijden, N., Duan, X., Nauwynck, H.J., 2010. Porcine reproductive and respiratory syndrome virus entry into the porcine macrophage. Journal of General Virology 91, 1659-1667.

Van Gorp, H., Van Breedam, W., Delputte, P.L., Nauwynck, H.J., 2009. The porcine reproductive and respiratory syndrome virus requires trafficking through CD163-positive early endosomes, but not late endosomes, for productive infection. Archives of Virology 154, 1939-1943.

Van Gorp, H., Van Breedam, W., Van Doorsselaere, J., Delputte, P.L., Nauwynck, H.J., 2010. Identification of the CD163 Protein Domains Involved in Infection of the Porcine Reproductive and Respiratory Syndrome Virus. Journal of Virology 84, 3101-3105.

van Nieuwstadt, A.P., Meulenberg, J.J., van Essen-Zanbergen, A., Petersen-den Besten, A., Bende, R.J., Moormann, R.J., Wensvoort, G., 1996. Proteins encoded by open reading frames 3 and 4 of the genome of Lelystad virus (Arteriviridae) are structural proteins of the virion. J Virol 70, 4767-4772.

Veit, M., Matczuk, A.K., Sinhadri, B.C., Krause, E., Thaa, B., 2014. Membrane proteins of arterivirus particles: structure, topology, processing and function. Virus Res 194, 16-36.

Vigerust, D.J., Shepherd, V.L., 2007. Virus glycosylation: role in virulence and immune interactions. Trends Microbiol 15, 211-218.

Vu, H.L., Kwon, B., Yoon, K.J., Laegreid, W.W., Pattnaik, A.K., Osorio, F.A., 2011. Immune evasion of porcine reproductive and respiratory syndrome virus through glycan shielding involves both glycoprotein 5 as well as glycoprotein 3. J Virol 85, 5555-5564.

Wei, Z., Tian, D., Sun, L., Lin, T., Gao, F., Liu, R., Tong, G., Yuan, S., 2012. Influence of N-linked glycosylation of minor proteins of porcine reproductive and respiratory syndrome virus on infectious virus recovery and receptor interaction. Virology 429, 1-11.

Wensvoort, G., Terpstra, C., Pol, J.M., ter Laak, E.A., Bloemraad, M., de Kluyver, E.P., Kragten, C., van Buiten, L., den Besten, A., Wagenaar, F., et al., 1991. Mystery swine disease in The Netherlands: the isolation of Lelystad virus. Vet Q 13, 121-130.

Whitworth, K.M., Rowland, R.R.R., Ewen, C.L., Trible, B.R., Kerrigan, M.A., Cino-Ozuna, A.G., Samuel, M.S., Lightner, J.E., McLaren, D.G., Mileham, A.J., Wells, K.D., Prather, R.S., 2016. Gene-edited pigs are protected from porcine reproductive and respiratory syndrome virus. Nat Biotechnol 34, 20-22.

Wieringa, R., de Vries, A.A., Raamsman, M.J., Rottier, P.J., 2002. Characterization of two new structural glycoproteins, GP(3) and GP(4), of equine arteritis virus. J Virol 76, 10829-10840.

Wieringa, R., de Vries, A.A., Rottier, P.J., 2003. Formation of disulfide-linked complexes between the three minor envelope glycoproteins (GP2b, GP3, and GP4) of equine arteritis virus. J Virol 77, 6216-6226.

Wieringa, R., de Vries, A.A.F., van der Meulen, J., Godeke, G.J., Onderwater, J.J.M., van Tol, H., Koerten, H.K., Mommaas, A.M., Snijder, E.J., Rottier, P.J.M., 2004. Structural protein requirements in equine arteritis virus assembly. Journal of Virology 78, 13019-13027.

Wissink, E.H., Kroese, M.V., Maneschijn-Bonsing, J.G., Meulenberg, J.J., van Rijn, P.A., Rijsewijk, F.A., Rottier, P.J., 2004. Significance of the oligosaccharides of the porcine reproductive and respiratory syndrome virus glycoproteins GP2a and GP5 for infectious virus production. J Gen Virol 85, 3715-3723.

Wissink, E.H., Kroese, M.V., van Wijk, H.A., Rijsewijk, F.A., Meulenberg, J.J., Rottier, P.J., 2005. Envelope protein requirements for the assembly of infectious virions of porcine reproductive and respiratory syndrome virus. J Virol 79, 12495-12506.

Wu, W.H., Fang, Y., Farwell, R., Steffen-Bien, M., Rowland, R.R., Christopher-Hennings, J., Nelson, E.A., 2001. A 10-kDa structural protein of porcine reproductive and respiratory syndrome virus encoded by ORF2b. Virology 287, 183-191.

Wu, W.H., Fang, Y., Rowland, R.R.R., Lawson, S.R., Christopher-Hennings, J., Yoon, K.J., Nelson, E.A., 2005. The 2b protein as a minor structural component of PRRSV. Virus Research 114, 177-181.

Xie, J., Christiaens, I., Yang, B., Breedam, W.V., Cui, T., Nauwynck, H.J., 2017. Molecular cloning of porcine Siglec-3, Siglec-5 and Siglec-10, and identification of Siglec-10 as an alternative receptor for porcine reproductive and respiratory syndrome virus (PRRSV). J Gen Virol 98, 2030-2042.

Yin, Y., Liu, C.L., Liu, P., Yao, H.C., Wei, Z.Z., Lu, J.Q., Tong, G.Z., Gao, F., Yuan, S.S., 2013. Conserved nucleotides in the terminus of the 3' UTR region are important for the replication and infectivity of porcine reproductive and respiratory syndrome virus. Archives of Virology 158, 1719-1732.

Yun, S.I., Lee, Y.M., 2013. Overview: Replication of porcine reproductive and respiratory syndrome virus. J Microbiol 51, 711-723.

Zhang, B., Miller, T.F., 3rd, 2010. Hydrophobically stabilized open state for the lateral gate of the Sec translocon. Proc Natl Acad Sci U S A 107, 5399-5404.

Chapter 2

Differences in signal peptide processing between GP3 glycoproteins of Arteriviridae

Minze Zhang, Michael Veit*

Institute of Virology, Free University Berlin, Robert von Ostertag-Straße 7-13, 14163 Berlin, Germany.

[Virology. 2018 Apr; 517:69-76. doi: 10.1016/j.virol.2017.11.026. Epub 2017 Dec 8]

2.1 Abstract

We reported previously that carbohydrate attachment to an overlapping glycosylation site adjacent to the signal peptide of GP3 from equine arteritis virus (EAV) prevents cleavage. Here we investigated whether this unusual processing scheme is a feature of GP3s of other Arteriviridae, which all contain a glycosylation site at a similar position. Expression of GP3 from type-1 and type-2 porcine reproductive and respiratory syndrome virus (PRRSV) and from lactate dehydrogenase-elevating virus (LDV) revealed that the first glycosylation site is used, but has no effect on signal peptide cleavage. Comparison of the SDS-PAGE mobility of deglycosylated GP3 from PRRSV and LDV with mutants having or not having a signal peptide showed that GP3´s signal peptide is cleaved. Swapping the signal peptides between GP3 of EAV and PRRSV revealed that the information for co-translational processing is not encoded in the signal peptide, but in the remaining part of GP3.

2.2 Introduction

Arteriviridae are a family of enveloped RNA viruses comprising the prototype member equine arteritis virus (EAV), the porcine reproductive and respiratory syndrome virus (PRRSV), and lactate dehydrogenase- elevating virus (LDV) of mice. Arteriviridae possess the ability to establish a persistent infection in their host. LDV causes a lifelong asymptomatic viremia; PRRSV persists in lymphoid tissues for months, whereas EAV persists only in the reproductive tract of infected stallions (An et al., 2011; Balasuriya et al., 2013; Chand et al., 2012; Meulenberg, 2000; Snijder et al., 2013). PRRSV cannot be eliminated from pig farms by vaccination due to the large variability between the existing strains. Two distinct genotypes, recently considered as two species, termed “European” (PRRSV-1) and “North American” (PRRSV-2) circulate worldwide

(Kuhn et al., 2016). Especially the glycoproteins show strong antigenic drift and exhibit large variation (up to 50%) in their amino acid sequence (Murtaugh et al., 2010; Shi et al., 2010).

Despite the enormous amount of sequence information on PRRSV genomes only limited information is available on the structure and function of the membrane proteins of Arteriviridae. The disulphide-linked GP5/M dimer is required for virus budding whereas a complex composed of the glycoproteins GP2, GP3 and GP4 is essential for cell entry (for recent reviews see (Snijder et al., 2013; Van Breedam et al., 2010; Veit et al., 2014; Zhang and Yoo, 2015).

GP3, the focus of this study, consists of an N-terminal signal peptide, a domain containing six cysteine residues and (depending on the virus) six to eight potential N-glycosylation sites, a hydrophobic region and a C-terminal hydrophilic part. From sequencing of the viral genomes it can be deduced that the open reading frame 3 (ORF3) of Arteriviridae encodes GP3 proteins of different sizes. The smallest protein (163 amino acids) is present in EAV (Wieringa et al., 2002); LDV encodes a protein with 191 amino acids (Faaberg and Plagemann, 1997), whereas the longest proteins with a size from 249 to 265 amino acids are encoded by PRRSV strains (Meulenberg et al., 1997). Aligning the GP3 sequences indicates that the heterogeneity in size is due to a variable length of the hydrophilic C-terminus (see also supplementary Table 1). GP3 is essential for replication of both EAV and PRRSV. Deleting its gene from the viral genome does not prevent budding of virus-like particles from transfected cells, but the particles are not infectious (Wieringa et al., 2004; Wissink et al., 2005). In the case of EAV they do not contain the GP2/3/4 complex and reduced amounts of the small and hydrophobic E protein (Wieringa et al., 2004).

Inconsistent results have been published about whether GP3s are structural proteins of Arterivirus particles. GP3 of LDV might be a nonstructural, secreted protein (Faaberg and Plagemann, 1997). Initial studies with GP3 from type-2 PRRSV also suggested that it is secreted (Mardassi et al., 1998), but more recent studies revealed that GP3 is incorporated into virus particles (de Lima et al., 2009) as is GP3 from type-1 PRRSV strain Lelystad (van Nieuwstadt et al., 1996). For EAV it has been clearly shown, that GP3 is present in a disulphide-linked trimer with GP2/4 in virus particles. Uniquely, the disulphide linkages between GP3 and GP2/4 are not formed in the ER but only after release of virus particles from infected cells (Wieringa et al., 2003).

Our recent work revealed another unique feature for processing of GP3 of EAV: We demonstrated that the overlapping sequon NNTT located just downstream of the signal peptide is efficiently N-glycosylated at both asparagines. Deletion of both (but not only one) glycosylation sites of the overlapping sequon allows signal peptide cleavage, indicating that co-translational attachment of one carbohydrate chain inhibits access of the signal peptidase to the cleavage site (Matczuk et al., 2013). This is consistent with predictions of the reliable bioinformatics tool SignalP (Petersen et al., 2011) and the presence of small and neutral amino acids (Ala, Val, Gly, Ser, Thr, Cys) at the −1 and −3 position with respect to the cleavage site, that serve as recognition feature for the signal peptidase (von Heijne, 1983). We also confirmed this unique phenomenon in recombinant viruses with disabled glycosylation sites. Furthermore, the infectivity of recombinant EAV containing GP3 with cleaved signal peptide was not impaired and GP3 without signal peptide associates with GP2/4 in virus particles (Matczuk and Veit, 2014).

The functional relevance why evolution created such a complex procedure to prevent signal peptide cleavage from GP3 of EAV is not known. One might speculate that covering a potentially cleavable signal peptide by a carbohydrate allows regulation of cleavage. During virus infection of horses, glycosylation of the first site might not occur in certain cell types resulting in a GP3 protein without signal peptide. Furthermore, during viral evolution GP3 might lose the first glycosylation site by mutation. From the 253 EAV GP3 sequences present in the database only two sequences do not contain a glycosylation site adjacent to the signal peptide (Matczuk and Veit, 2014). Interestingly, these two GP3 sequences were derived from a persistently infected stallion. Virus originally isolated from that stallion contains the overlapping glycosylation site GSNNTT in GP3, but 24 month later it was exchanged to GSDNPA and subsequently mutated to GRDNPA (Balasuriya et al., 1999). It is therefore tempting to speculate that loss of both glycosylation sites creates a virus having GP3 with cleaved signal peptide. Such a virus might have been evolutionarily selected since it might escape from the immune system and/or might be able to infect other cell types to sustain the persistent infection. In EAV the known neutralization determinants are located in GP5, but at least GP3 of PRRSV might also be a neutralizing target based on sera collected from Lelystad-infected pigs (Vanhee et al., 2011).

The occurrence of glycosylation sites near the signal peptide is a rather rare feature in glycoproteins. Our bioinformatic analysis revealed that only 1.7% of all human and mouse genes containing a predicted signal peptide also contain an N-glycosylation

site (NXS/T, X≠P) within five amino acids behind the signal peptide (Matczuk et al., 2013). It is thus remarkable that two glycoproteins of Arteriviruses, GP3 and GP5, possess carbohydrates near the signal peptide cleavage site. For GP5 of several PRRSV-2 strains we showed that the presence of one (and even the insertion of a second) efficiently used glycosylation site located just three amino acids downstream of the signal peptide cleavage site does not inhibit processing (Thaa et al., 2013). This was also shown for GP5 of a PRRSV-1 strain, but a particular GP5 protein having an unused glycosylation site immediately subsequent to the signal peptide is only inefficiently processed in transfected cells (Thaa et al., 2017). The two cellular glycoproteins investigated in this regard require signal peptide cleavage for efficient glycosylation at sites located four amino acids downstream of the signal peptide (Chen et al., 2001). Thus, GP3 of EAV is hitherto the only glycoprotein where a carbohydrate attached to such a site inhibits signal peptide processing and we investigated here whether this feature is conserved in GP3 of other Arteriviridae.

The enzyme complexes for N-glycosylation (oligosaccharyl transferase, OST, (Shrimal et al., 2015)) and signal peptide cleavage (signal peptidase, SPase, (Auclair et al., 2012)) are associated with the translocon, a hetero-oligomeric protein complex serving as a channel for translocation of proteins into the lumen of the rough endoplasmic reticulum (ER) (Rapoport et al., 2017). Thus, in principle both OST and SPase can perform their activity once a nascent polypeptide chain becomes accessible. In GP3 from EAV (as in most other glycoproteins) both modifications occur rapidly; the mature protein (fully glycosylated with cleaved signal peptide) was already detected after one minute of metabolic labeling (Matczuk et al., 2013). If OST and signal peptidase compete for neighboring sites it must be a subject of regulation whether signal peptidase or OST has privileged access to the growing polypeptide chain. In the absence of regulation, i.e. if both enzymes would have random access to neighboring sites, a mixed protein population would be produced. In the case of GP3 of EAV a fraction of proteins would and the other would not contain the signal peptide, but this has never been observed in transfected cells (Matczuk et al., 2013). It is unexplored, how the access of OST and signal peptidase to a nascent protein chain in the ER lumen is regulated, but one might assume that the signal peptide (the first part of a growing polypeptide chain that contacts the translocon) selects whether OST or SPase is recruited at first to the translocon. It has been reported that different signal peptides interact with different binding sites within the translocon and that these differences can substantially affect protein biogenesis (Hegde and Bernstein, 2006).

In addition, another co-translational modification (possibly) competing with signal peptide cleavage and N-glycosylation is disulphide bond formation. Disulphide bonds are formed by oxidation of sulfhydryl (-SH) groups of two cysteine residues in close spatial proximity. This occurs spontaneously, favored by the reducing milieu in the ER, but unfavorable bonds that do not allow stable protein folding, are reduced by the protein disulphide isomerases (PDI), a soluble enzyme in the ER lumen (Oka and Bulleid, 2013). Disulphide bond formation generally reduces glycosylation at adjacent sites, probably because a folded and fixed protein conformation cannot be processed by OST(Allen et al., 1995).

The purpose of this study was to explore whether GP3 proteins from PRRSV and LDV follow the same unusual processing scheme as GP3 from EAV, i. e. whether a glycosylation site near the signal peptide affects cleavage. Although more than 600 nucleotide sequences encoding GP3 proteins especially from PRRSV strains are deposited in databases, biochemical studies on processing of the protein are rare. No processing study exists for GP3 of any of the many PRRSV strains, but GP3 of LDV retains the signal peptide and it was speculated that this might be due to a carbohydrate attached close to the signal peptide cleavage site (Faaberg and Plagemann, 1997).

2.3 Material and methods

2.3.1 Plasmids and transfection of cells

The nucleotide sequences of open reading frame 3 of the following five strains of PRRSV was synthesized by Bio Basic Inc. (Markham Ontario, Canada): Lelystad virus: low pathogenic PRRSV-I prototype strain (Meulenberg et al., 1993), GenBank accession number: M96262.2; VR-2332: low pathogenic PRRSV-2 prototype strain (Benfield et al., 1992), accession number: AY150564.1; IAF-Klop: low pathogenic PRRSV-2 Québec reference strain (Mardassi et al., 1994), accession number: AF003344; XH-GD: Chinese highly pathogenic PRRSV-II strain, accession number EU624117.1 and Lena: highly pathogenic PRRSV-I strain (Karniychuk et al., 2010), accession number: JF802085.1. All GP3 genes were equipped during synthesis at the 3'end with a sequence encoding the HA-tag (amino acids YPYDVPDYA) plus a small linker (PV).

The nucleotide sequence encoding GP3 of LDV (Plagemann strain, accession number: U15146, (Faaberg and Plagemann, 1997)) with C terminal HA-tag was synthesized by Integrated DNA technologies (IDT, Leuven, Belgium). The chimeras between PRRSV (Lelystadt strain) and EAV (Bucyrus strain, accession number: DQ846750.1) in which

the signal peptides were swapped, were synthesized by Bio Basic Inc. (Markham Ontario, Canada). EAV-PRRSV contains the amino acids 1–26 of GP3 from EAV fused to the amino acids 26–265 of GP3 from PRRSV. PRRSV-EAV contains the amino acids 1–25 of GP3 from PRRSV fused to amino acids 27–163 of GP3 from EAV.

The GP3 genes with HA-tag were subcloned into the plasmid pCMVTNT (Promega, Mannheim, Germany, containing T7 and CMV promotors) using KpnI and NotI restriction sites to create pCMV-TNT-GP3-wt. Using these plasmids as a template, the GP3 mutants described in the text were generated by overlap extension polymerase chain reaction (PCR). The authenticity of all GP3 genes and mutants were verified by sequencing (LGC Genomics, Berlin, Germany) before use in experiment.

BHK-21 (baby hamster kidney cells, ATCC C13) cells were maintained as adherent culture in Dulbecco's Modified Eagle's Medium (DMEM, PAN, Aidenbach, Germany) supplemented with 10% fetal calf serum (FCS) (Perbio, Bonn, Germany) at 37 ℃ in an atmosphere with 5% CO2 and 95% humidity. BHK-21 cells in 6-well plates were transfected with 2.5 µg plasmid DNA using Lipofectamine® 3000 (Thermo Fisher Scientific, Carlsbad, United States) as described by the manufacturer.

2.3.2 In vitro transcription/translation

Unprocessed GP3-wt and GP3-ΔSP with HA-tag were generated by in vitro transcription/translation using the TNT Quick Coupled Transcription/Translation System (Promega, Mannheim, Germany) and the plasmids described above. A reaction (25 µL) was composed of 20 µL rabbit reticulocyte lysate (TNT master mix, also including T7 RNA polymerase), 1 µL methionine (1 mM), and 1 µg of pCMV-TNTplasmid DNA (containing a T7 promotor), complemented by Millipore water. Reactions were incubated for 90 min at 30 ℃. 0.5–1 µL of the reaction was supplemented with reducing SDS-PAGE loading buffer and assessed by SDS-PAGE and Western blot.

2.3.3 Glycosidase treatment

Cells were washed with PBS 24 h after transfection, detached from the dish with trypsin-EDTA (PAN Biotech, Aidenbach, Germany), pelleted, washed twice with PBS, resuspended in 60 µL 1X glycoprotein denaturing buffer (0.5% SDS, 40 mM DTT) and boiled for 10 min at 100 ℃. To test for intramolecular disulphide bonds DTT was omitted from the buffer. An aliquot of the sample (usually 3–17 µL, depending on the expression efficiency) was digested with Peptide-N-Glycosidase (PNGase F, 2.5-5units/µL, 4 h at 37 ℃) or endo-beta-N-acetyl-glucosaminidase (Endo H, 2.5-5units/µL,

1 h at 37 ℃) according to the manufacturer's instruction (New England Biolabs, Frankfurt am Main, Germany). After the deglycosylation reaction, samples were supplemented with reducing SDS-PAGE loading buffer and subjected to SDSPAGE and Western blot.

2.3.4 SDS-PAGE and Western Blot

After sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), typically using 15% polyacrylamide gels, gels were blotted onto polyvinylenedi-fluoride (PVDF) membrane (GE Healthcare, Freiburg im Breisgau, Germany) using standard methodology. After blocking of membranes (blocking solution: 5% skim milk powder in PBS with 0.1% Tween-20, PBST) for 1 h at room temperature (RM), the primary antibody was applied overnight at 4 ℃ in blocking solution: rabbit-anti-HA tag antibody (ab9110, Abcam, Cambridge, UK, 1:10,000) was used to detect translated product with HA tag. After washing (3×10 min with PBST), a suitable horseradish peroxidasecoupled (HRP) secondary antibody (anti-rabbit, Sigma-Aldrich, Taufkirchen, Germany, 1:5000) was applied for 1 h at RM. After washing, signals were detected by chemiluminescence using the ECLplus reagent (Pierce/Thermo, Bonn, Germany) and the Fusion SL camera system (Peqlab, Erlangen, Germany).

2.3.5 Immunofluorescence Assay

For immunofluorescence, transfected BHK-21 cells grown in 6-well plates (Sarstedt, Nümbrecht, Germany), were fixed with 4% paraformaldehyde in phosphate-buffered saline (PBS) for 20 min at 4 °C, washed twice with PBS, permeabilized with 0.5% Triton in PBS for 5 min at room temperature, and washed again twice with PBS. 30 min after blocking with bovine serum albumin (3% in PBS), cells were then incubated with the primary rabbit-anti-HA antibody (ab9110, Abcam, Cambridge, UK, 1:500) at room temperature for 1 h, then washed three times with PBS, incubated with secondary antibody (Alex Fluor 568 goat anti-rabbit IgG(H+L), Invitrogen, Darmstadt, Germany, 1:500) at room temperature for 1 h. Fluorescence analysis was done with a ZEISS Axio Vert. A1 inverted epifluorescence microscope (Zeiss, Göttingen, Germany).

2.4 Results and discussion

2.4.1 Glycosylation at the first site does not affect signal peptide processing of GP3 from PRRSV

Table 2.1 shows the 45 N-terminal amino acids of GP3 proteins from equine arteritis virus (EAV), from lactate dehydrogenase elevating virus (LDV) and from the following five PRRSV strains: a low-pathogenic (Lelystad) and a highly pathogenic (Lena)

"European" PRRSV-1 strain and two low-pathogenic (VR-2332, IAF-Klop) and one highly pathogenic (XH-GD) "North American" PRRSV-2 strains. The bioinformatics tool SignalP (Petersen et al., 2011) predicts signal peptide cleavage of each GP3 protein at the sites indicated by a gap in the sequence.

Equine Arteritis Virus	**D-score**
MGRAYSGPVALLCFFLYFCFICG**S**V**G** S**NNTT**ICMHTTSDTSVHLF	0.82
PRRSV-1 (Lelystad)	
MAHQCARFHFFLCGFICYLVHS**A**L**A** S**NSS**STLCFWFPLAHGNTSF	0.85
PRRSV-1 (Lena)	
MACQCSCFYFLFCGFVGYLICG**A**L**A** SNS**NST**LCFWFPLAHGNASF	0.55
PRRSV-2 (VR-2332)	
MVNSCTFLHIFLCCSFLYSF**C**C**A** VVAGS**NTT**YCFWFPLVRGNFSF	0.68
PRRSV-2 (IAF-Klop)	
MANSCTLLHIFLCCSFLHPFCC**V**V**A** ASA**NAS**YCFWFPLVRGNFSF	0.70
PRRSV-2 (XH-GD)	
MANSCTFLHIFLRCSFLYSF**C**C**A** VVANS**NAT**FCFWFPLVRGNFSF	0.56
LDV	
MASFNAIICVFICFSCCCI**V**D**C** **NVT**APTKTCLWFPIHETGKTHLE	0.79

Table 2.1 N-terminal amino acids of Gp3 proteins from Arteriviruses. The table shows the first 45 amino acids of GP3 from EAV prototype strain Bucyrus, type 1 prototype PRRSV strain Lelystad, type 1 highly pathogenic strain Lena, type 2 prototype PRRSV strain VR-2332, type 2 strain IAF-Klop, the Chinese type 2 highly pathogenic strain XH-GD and for lactate dehydrogenase-elevating virus (LDV). The signal peptide cleavage site predicted by SignalP with the highest probability is indicated as a gap in the amino acid sequence. The small and uncharged amino acids at −1 and −3 position serving as recognition sequence by the signal peptidase are marked in bold. D-score (cut-off value: 0.45) denotes the probability for cleavage at the indicated site as determined by SignalP. For all GP3 proteins, especially those having a rather low D-score further cleavage sites were predicted, albeit with lower probability. The first glycosylation site in each sequence is marked in bold and underlined. A second glycosylation site at position 42 in GP3 of the PRRSV strains and a conserved cysteine residue (in GP3 from PRRSV at position 33) are also underlined.

However, this usually reliable tool does not take into account protein modifications, such as glycosylation. Each GP3 protein contains a putative N-glycosylation site NXS/T (underlined) downstream of the predicted cleavage site. In GP3 from LDV it is located immediately adjacent to the cleavage site, in GP3s from EAV and PRRSV

Lelystad one residue and in GP3s from the other four PRRSV strains three or five residues further downstream. However, for GP3s from the PRRSV strains VR-2332, Lena, IAF-Klop and XH-GD an alternative cleavage site (small und uncharged amino acids at the −1 and −3 positions) is predicted with lower probability that would move it closer to the glycosylation site.

To analyze GP3 processing we fused the gene from the PRRSV-1 prototype strain Lelystad at its C-terminus to an HA-tag for subsequent antibody detection. Immunofluorescence of transfected BHK-cells shows that GP3-wt localizes to intracellular reticular structures, probably the ER, like authentic GP3 (Kim et al., 2013; Mardassi et al., 1998). To determine whether the first glycosylation site is used and inhibits signal peptide processing, asparagine 27 was exchanged by glutamine and both the resulting mutant GP3-N27Q and GP3-wt were expressed in BHK cells. SDS-PAGE and western-blotting (Fig. 2.1b) shows a distinct reduction in the molecular weight of Gp3-N27Q relative to Gp3-wt. Since a signal peptide and one carbohydrate chain have the same molecular weight (~3 kDa), the observed reduction might be either due to signal peptide cleavage, loss of the carbohydrate or caused by a combination of both. To distinguish between these possibilities all carbohydrates were removed by digestion with PNGase F prior to SDS-PAGE and western blotting now reveals that the mobility of GP3-wt and GP3-N27Q are identical. Thus, we conclude that the glycosylation site N27 is used, but the attached carbohydrate has no effect on signal peptide processing.

Since GP3 proteins from PRRSV strains exhibit large amino acid variability, especially in the C-terminal part of the signal peptide and the subsequent amino acids (see Table 2.1) we wondered whether there is also variability in the usage of the first glycosylation site or in its effect on signal peptide processing. We therefore performed the same experiments with GP3 proteins from the other four PRRSV-strains shown in Table 2.1, all of them fused by the same linker to an HA-tag. However, an identical result was obtained: the first glycosylation site is efficiently glycosylated, (as previously described for GP3 of the PRRSV-2 FL-12 strain (Das et al., 2011), but its removal has no effect on the SDS-PAGE mobility of the deglycosylated protein and hence on signal peptide cleavage (Fig. 2.2).

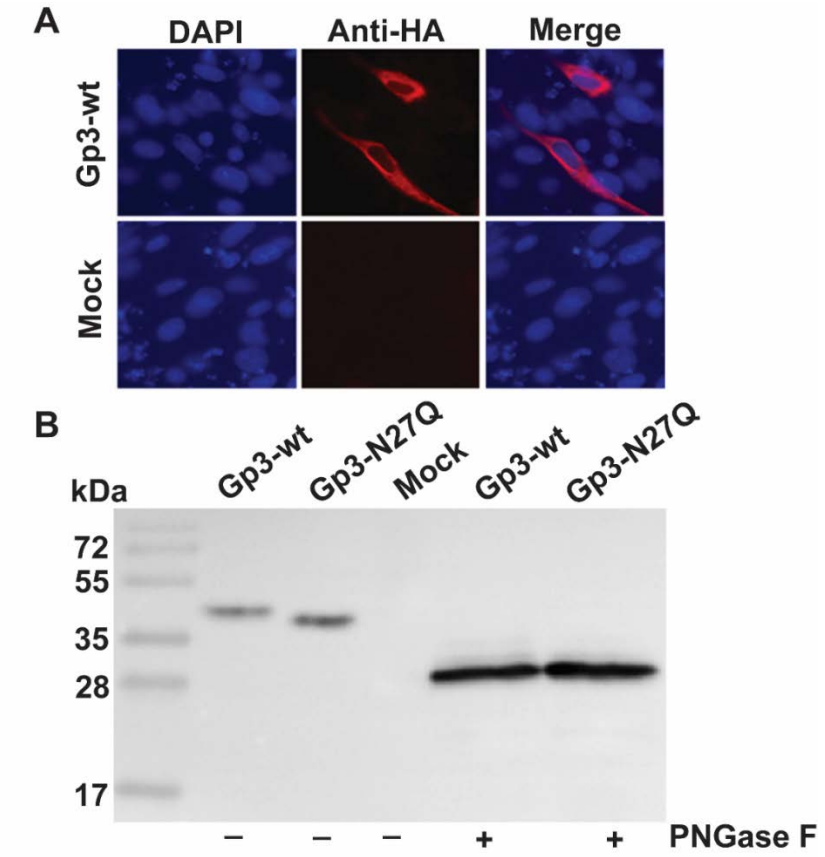

Fig. 2.1. A carbohydrate attached to the first site of Gp3 from PRRSV Lelystad has no effect on signal peptide cleavage. A. GP3-wt containing a C-terminal HA-tag was expressed in BHK-21 cell and visualized by immunofluorescence with anti-HA antibody and goat anti rabbit antibody coupled to Alex-568. Nuclei of cells were stained with DAPI. **B**. GP3-wt and a mutant with a N to Q exchange of the first glycosylation site (GP3-N27Q) were expressed in BHK-21 cells, cell lysates were not digested (-) or digested (+) with PNGase F prior to Western blotting with anti-HA antibodies. The left lane shows molecular weight (MW) markers with the indicated sizes in kDa. Mock are untransfected cells. The predicted MW of unglycosylaed GP3-HA containing the signal peptide is 31.8 kDa, without signal peptide 29 kDa.

GP3 proteins from PRRSV contain another conserved glycosylation site in close distance to the signal peptide, namely at position 42 (Table 2.1). Therefore, we investigated whether attachment of a carbohydrate to this site, possibly in combination with a carbohydrate present at the first site, might prevent access of the signal peptidase to its cleavage site. Expression of the mutant GP3-N42Q from the VR-2332 strain revealed that this site is also glycosylated, but its exchange does not affect signal peptide cleavage, even when it is removed simultaneously with the first glycosylation site (mutant GP3- N29QN42Q, left part of Fig. 2.2). In sum, GP3 proteins from PRRSV apparently follow a different processing scheme as GP3 from EAV; a carbohydrate

attached to the first (or second) glycosylation site does not affect signal peptide cleavage.

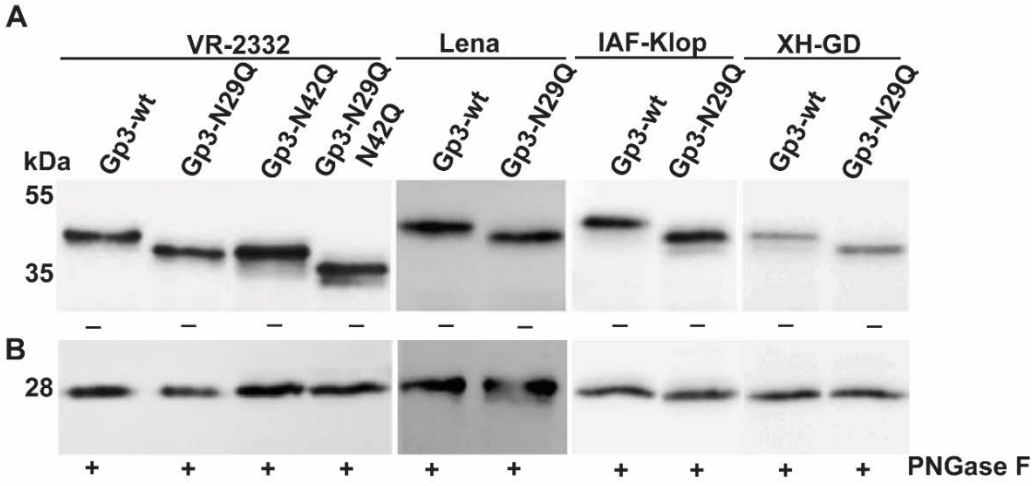

Fig. 2.2. A carbohydrate attached to the first site of Gp3 from other PRRSV-1 and −2 strains has no effect on signal peptide cleavage. GP3-wt and mutants with a N to Q substitution at the first glycosylation site (N29Q) were expressed in BHK-21 cells and detected by Western blotting with anti-HA antibody. In GP3 from VR-2332 the second glycosylation site was also exchanged (N42Q), in the mutant N29QN42Q simultaneously with the first site. Cell lysates were not digested (**A**) or (**B**) digested with PNGase F prior to Western blotting. VR-2332: PRRSV-2 prototype strain; IAF-Klop: PRRSV-2 strain; XH-GD: highly pathogenic Chinese PRRSV-2 strain and Lena: highly pathogenic PRRSV-1 strain. The predicted molecular weights for unglycosylated GP3-HA with and without the signal peptide are 30.2 kDa and 27.6 kDa for VR-2332, 29.9 kDa and 27.1 kDa for Lena, 30.1 kDa and 27.4 kDa for IAF-Klop and 30 kDa and 27.3 kDa for XH-GD.

Another co-translational protein modification occurring in the lumen of the ER is the formation of disulphide bonds. GP3 proteins of PRRSV contain several cysteines within the signal peptide and a highly conserved cysteine at position 33, that is also conserved in GP3 from other Arteriviruses (see Table 2.1 and supplementary Table 1). These cysteines may form a (possibly transient) disulphide bond that might affect signal peptide cleavage. However, comparing the SDS-PAGE mobility of GP3-wt under reducing and non-reducing conditions did not reveal a size difference (shown here in Fig. 2.3a for GP3 from Lelystad, but similar results were observed for the other GP3s) suggesting that GP3 does not contain a widely-spaced intramolecular disulphide bond.

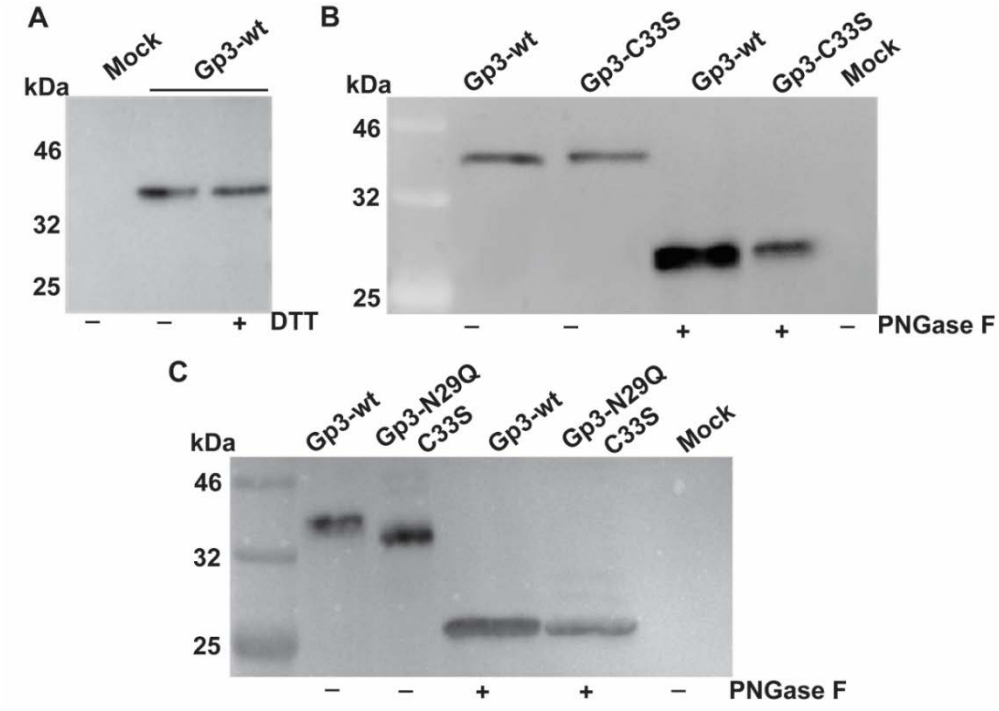

Fig. 2.3. The conserved cysteine C33 has no effect on signal peptide cleavage. A. GP3-wt from the Lelystad strain of PRRSV was expressed in BHK-21 cell, cell lysates were separated by SDS-PAGE under reducing (+DTT) or non-reducing (- DTT) conditions and Gp3 was detected by Western blotting with anti-HA antibody. B. GP3-wt and mutant GP3-C33S from PRRSV Lelystad were expressed in BHK-21 cells, cell lysates were not digested (-) or digested (+) with PNGase F prior to Western blotting. C. GP3-wt and the double mutant GP3-N29QC33S from PRRSV VR-2332 were expressed in BHK-21 cells, cell lysateswere not digested (-) or digested (+) with PNGase F prior to Western blotting.

Furthermore, exchange of the conserved cysteine 33 by serine in GP3 from Lelystad has no effect on glycosylation and signal peptide cleavage (Fig. 2.3b). Moreover, exchanging cysteine 33 simultaneously with asparagine 29 in GP3 from VR-2332 did not strongly alter the electrophoretic mobility of the deglycosylated protein (Fig. 2.3c) excluding the possibility that glycosylation at N29 and a disulphide-linkage involving cysteine 33 are both required to affect signal peptide processing. Thus, we did not obtain evidence that cysteine 33 substantially affects co-translational processing of GP3. However, note that deglycosylated GP3-wt in both Fig. 2.3b and c appears as a double band, whereas deglycosylated GP3-C33S and GP3-N29QC33S forms (mainly) a single band.

Thus, one might guess that cysteine 33 (or a short-distance disulphide linkage it might form) might slightly affect the position of signal peptide cleavage. We therefore attempted to analyse signal peptide cleavage by mass spectrometry (Thaa et al.,

2013). However, this was unsuccessful since we could not immunoisolate sufficient amounts of protein from transfected cells and hence we investigated signal peptide cleavage biochemically.

2.4.2 The signal peptides of GP3 from PRRSV (Lelystad) and LDV are cleaved

To analyse signal peptide cleavage biochemically we did not utilize the classical approach using in vitro translation in the presence of microsomes (Blobel and Dobberstein, 1975) since (due to inefficient translocation and glycosylation) it usually produces multiple protein bands that complicate interpretation of the data (Thaa et al., 2013). Instead, we compared the SDS-PAGE mobility of deglycosylated wildtype GP3 from the PRRSV-1 prototype strain Lelystad expressed in cells with mutants for which we known whether they contain or do not contain the signal peptide. In one mutant (GP3-ΔSP) the whole signal peptide was deleted and thus the SDS-PAGE mobility of the protein indicates the size of unglycosylated GP3 with cleaved signal peptide (Fig. 2.4a). In the other mutant (GP3-USP) the small amino acid alanine located at −1 and −3 positions of the predicted cleavage site was exchanged by the large amino acids phenylalanine (−1) and tyrosine (−3), respectively. Small amino acids are a prerequisite for recognition of a signal peptide by the signal peptidase (Hegde and Bernstein, 2006; von Heijne, 1983) and indeed SignalP predicts an uncleaved signal peptide for this mutant. Expression of these constructs and Western blotting (Fig. 2.4b) shows that GP3-wt and GP3-USP have a similar molecular weight (~40 kDa), but are much larger than GP3-ΔSP (~28 kDa). This proves that the N-terminal amino acids of GP3 indeed function as signal peptide. In their absence GP3 is not translocated into the lumen of the ER and therefore none of its seven N-glycosylation sites is used.

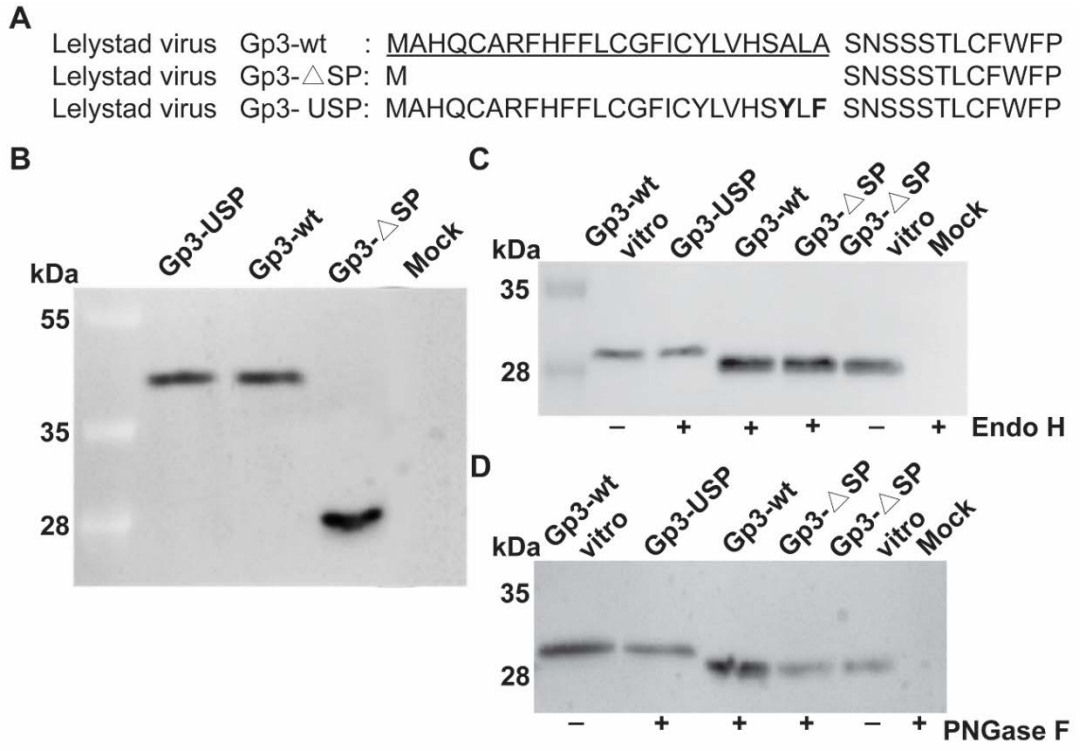

Fig. 2.4. The signal peptide of GP3 from PRRSV Lelystad is cleaved. A. N-terminal sequences (residues 1–37) of GP3-wt and mutants from PRRSV Lelystad. The underlined sequence (residues 1–25) is the predicted signal peptide with the cleavage site shown as a gap. GP3-ΔSP: mutant with deleted signal peptide, Gp3-USP: Small residues A23 and A25 were exchanged to Y and F, respectively (highlighted in bold). The signal peptide is predicted to be uncleaved. **B**. Gp3-wt, Gp3-ΔSP and Gp3-USP were expressed in BHK-21 cells and detected by Western blotting with anti-HA antibody. **C**+**D**. Samples were digested with Endo H (C) or PNGase F (D) prior to Western blotting. Gp3-wt and Gp3-ΔSP were also translated in vitro in the absence of microsomes. The resulting product is intrinsically unprocessed, i.e. not glycosylated and (in the case of Gp3-wt) still containing the signal peptide.

To determine whether the signal peptide is cleaved we used a further size marker, namely GP3 synthesized in vitro in the absence of membranes. The resulting product remains unglycosylated, but retains the signal peptide. GP3 synthesized in vitro exhibits the same molecular weight as GP3-USP deglycosylated with Endo H indicating that the introduced mutations indeed inhibited signal peptide cleavage (Fig. 2.4c). Likewise, GP3-ΔSP expressed in cells has the same SDS-PAGE mobility as GP3-ΔSP synthesized in vitro. Importantly, GP3-wt deglycosylated with Endo H has the same molecular weight as GP3-ΔSP, but is smaller than GP3-USP and GP3-wt synthesized in vitro suggesting that the signal peptide is cleaved inside cells.

Endo-H cleaves between the two innermost N-acetylglucosamine (GlcNAc) residues of the carbohydrate side chain, resulting in the retention of one GlcNAc residue per glycosylation site and we wondered whether this might affect the SDS-PAGE mobility of deglycosylated samples. However, when the same samples were digested with PNGase F, which removes the complete carbohydrate chain from GP3, an identical result was obtained (Fig. 2.4d). Furthermore, signal peptide cleavage was also observed if GP3 was expressed in other cell types (CHO, 293 T) including Marc-145 cells, which support replication of PRRSV (data not shown).

Cleavage of the signal peptide is not consistent with the claim that it serves as an ER-retention signal for GP3 (Kim et al., 2013). This conclusion was solely based on the observation that an artificial construct containing GP3´s signal peptide fused to the N-terminus of GFP exhibits an ER-like localization in confocal microscopy. Since no biochemical studies to test translocation of the construct into the ER lumen were performed, the hydrophobic signal peptide might have attached GFP to the cytosolic site of the ER.

Thus, GP3 from EAV and PRRSV clearly differ in signal peptide processing. We therefore asked how GP3 from another member of the Arteriviridae, the lactate dehydrogenase-elevating virus (LDV) is processed. Gp3 from LDV contains the shortest signal peptide (22 amino acids) and just one cleavage site is predicted which is immediately followed by a glycosylation site (Table 2.1). We again created a GP3 mutant where the first glycosylation site is exchanged (GP3-N23Q) and one with a deletion of the predicted signal peptide (GP3-ΔSP, Fig. 2.5a). Western-blotting showed that GP3-wt has a higher molecular weight as GP3-N23Q indicating carbohydrate attachment to the first glycosylation site (Fig. 2.5b). Both proteins are much larger than GP3-ΔSP; the latter is not glycosylated since it is not targeted to the ER. After deglycosylation with PNGase F, GP3-wt, GP3-N23Q, GP3-ΔSP showed the same SDS-PAGE mobility, but clearly ran faster than GP3-wt synthesized in vitro, the size marker for an unglycosylated GP3 containing the signal peptide (Fig. 2.5c). Thus, GP3 from LDV and PRRSV are processed in an identical manner; the signal peptide is cleaved regardless of the presence of a carbohydrate in its vicinity.

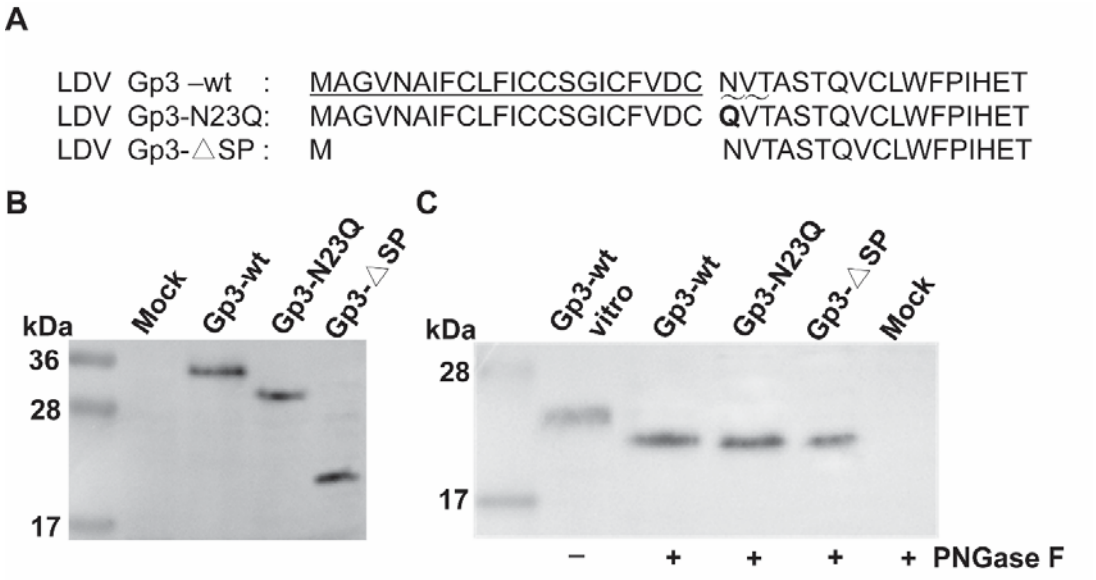

Fig. 2.5. The signal peptide of Gp3 from LDV is cleaved. A. N-terminal sequences (residues 1–37) of Gp3-wt and mutants from lactate dehydrogenaseelevating virus (LDV). The underlined sequence (residues 1–22) is the predicted signal peptide with the cleavage site shown as a gap. The sequon NVT (underlined with a wavy line) is the first glycosylation site. Gp3-N23Q: a mutant with an exchange of the glycosylation site NVT to QVT; Gp3-ΔSP: mutant with deleted signal peptide, **B**. Gp3-wt, Gp3-ΔSP and Gp3-USP were expressed in BHK-21 cells and detected by Western blotting with anti-HA antibody. **C**. Samples were digested with PNGase F prior to Western blotting. Gp3-wt was also translated in vitro in the absence of microsomes. The resulting product is intrinsically unprocessed, i.e. not glycosylated and still containing the signal peptide. The predicted MW of unglycosylated Gp3-HA containing the signal peptide is 22.7 kDa, without signal peptide 20.4 kDa.

Our result with GP3 from LDV conflicts with published data using in vitro translation in the presence of microsomes (Faaberg and Plagemann, 1997). In our hands, in vitro translation with microsomes not always produces reproducible results since the membranes are fragile, sensitive to heat and to cycles of freezing and thawing. Even with fresh membrane the three co-translational activities (translocation into the ER lumen, N-glycosylation and signal peptide cleavage) are rarely performed with high efficiency and in that experiment the signal peptidase might not have worked properly. Alternatively, the interplay between the co-translational modifications N-glycosylation and signal peptide cleavage, discussed in the next paragraph, might be differentially regulated in microsomes compared to living cells.

2.4.3 Swapping signal peptides between EAV and PRRSV does not affect processing

Finally, we asked which domain in GP3, the signal peptide or the remaining part of the protein contains the information whether a carbohydrate attached to a site adjacent to the signal peptide is capable to inhibit its processing. We hypothesized that the temporal order of both co-translational modifications determines the processing scheme. For GP3 from EAV the oligosaccharyl transferase (OST) must have first access to the translocating polypeptide chain to attach a carbohydrate that inhibits signal peptide processing. For GP3 from PRRSV and LDV the sequence of events might be reversed, i.e. the signal peptidase removes the signal peptide and OST then adds carbohydrates. How the sequence of both co-translational modifications is regulated is unknown, but we hypothesize further that the signal peptide itself may play a role. One can imagine that the signal peptide of GP3 from EAV recruits OST to the translocon to perform rapid glycosylation that prevents signal peptide cleavage whereas the signal peptide of GP3 from PRRSV recruits the signal peptidase to the translocon to allow signal peptide cleavage prior to glycosylation.

To test this hypothesis, we created chimeras between GP3 of PRRSV (Lelystad) and EAV in which the signal peptides were swapped. The first set of mutants contain the signal peptide of GP3 from PRRSV fused to the ectodomain of GP3 from EAV, either from wild type GP3 (PRRSVEAV) or from a mutant with the overlapping sequon NNTT changed to QQTT (PRRSV-EAV NNQQ, see Fig. 2.6a). Upon expression in BHK-cells and after removal of carbohydrates by PNGase F digestion PRRSV-EAV exhibits a higher molecular weight as PRRSV-EAV NNQQ indicating that PRRSV-EAV retains the signal peptide, whereas it is cleaved in the mutant (Fig. 2.6c). Comparing the SDS-PAGE mobility of both proteins without PNGase F digestions revealed that PRRSV-EAV has a much higher molecular weight (36 kDa) than the mutant PRRSV-EAV NNQQ (28KDa) suggesting that both sites of the overlapping sequon NNTT are used. Thus, GP3 of EAV follows the same unusual processing scheme when it is equipped with the signal peptide of GP3 from PRRSV.

The reverse set of mutants (named EAV-PRRSV) contains the signal peptide of EAV and the ectodomain of PRRSV, the later again with or without the first glycosylation site (Fig. 2.6b). The biochemical processing analysis revealed that the first glycosylation site is used and that the deglycosylated proteins have an identical size. Furthermore, their SDS-PAGE mobility is identical to the mobility of the authentic GP3 from PRRSV indicating that the signal peptide is cleaved (Fig. 2.6d). Thus, exchange of signal peptides between GP3 of EAV and PRRSV does not affect processing. The

chimeras behaved like the authentic GP3 proteins, i. e. the signal peptides do not influence the processing. Similar results were obtained if the signal peptides of GP3 of EAV and GP5 from PRRSV (where the signal peptide is cleaved despite the presence of a carbohydrate three amino acids downstream of the signal peptide (Thaa et al., 2013) were swapped (data not shown).

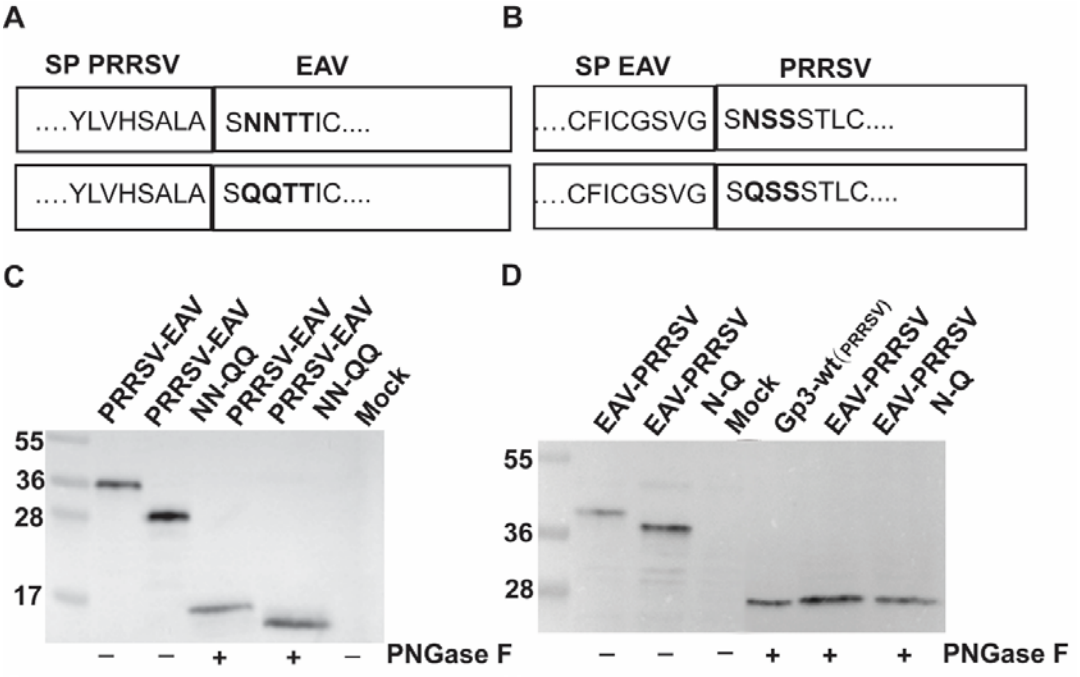

Fig. 2.6. Exchange of signal peptides between Gp3 of EAV and PRRSV does not affect processing. A. The signal peptide (SP) of Gp3 from PRRSV (Lelystad) was fused to the ectodomain of EAV, either wild-type or the mutant NN-QQ with an exchange of the overlapping glycosylation site NNTT to QQTT. **B**. The predicted signal peptide of Gp3 from EAV was fused to the predicted ectodomain of PRRSV (Lelystad), either wild-type or the mutant N-Q with an exchange of the glycosylation site NSS to QSS. Both domains were fused exactly as predicted by SignalP (Table 2.1). The resulting amino acid sequence at the connection is indicated inside the boxes. The chimeras contain an HA-tag fused to their C-termini. **C**. Chimeras shown in A were expressed in BHK-21 cells, and cell lysates were digested with PNGase F (+) or not digested (-) prior to Western blotting with HA antibodies. **D**. Chimeras shown in B were expressed in BHK-21 cells, and cell lysates were digested with PNGase F prior to Western blotting with HA antibodies. Gp3-wt (PRRSV) is Gp3-wt from PRRSV (Lelystad virus).

Thus, our hypothesis that the signal peptide of EAV recruits the oligosaccharyl transferase to the translocon to allow rapid glycosylation to prevent signal peptide cleavage must be rejected. The information for the unusual processing of the GP3 from EAV is rather encoded in the remaining part of the protein. Are there certain features in their amino acid sequences that distinguish GP3 of EAV from GP3s of PRRSV and LDV? Multiple sequence alignment of GP3s revealed that the proteins from PRRSV

and LDV are more similar to each other than to GP3 from EAV (Table 2.2). Whereas the amino acid identity score between GP3s of PRRSV and LDV is 35%, GP3 from EAV has amino acid identity with GP3 from LDV and PRRSV of only 21% and 24%, respectively. The alignment also shows that the six cysteine residues of GP3 from PRRSV and LDV are at identical or very similar positions suggesting that both proteins exhibit a similar folding. Likewise, the five glycosylation sites in GP3 from LDV align also in a similar position as five (out of seven) sites in GP3 from PRRSV (Lelystad). In contrast, in GP3 from EAV only the first cysteine (outside the signal peptide) matches perfectly with the first cysteine in GP3 from PRRSV and LDV and the third cysteine of EAV´s GP3 aligns roughly with the sixth cysteine in GP3 from PRRSV and LDV. Thus, GP3 from EAV might have a peculiar structural element that is not present in GP3 from PRRSV and LDV and one might speculate that this might determine the unusual processing scheme. The closer relationship between PRRSV and LDV is also reflected in the recently improved taxonomy of Arteriviridae, where both viruses are grouped as one genus Rodartevirus, whereas EAV is the single member of the genus Equartevirus (Kuhn et al., 2016).

Another member of the Arteriviruses is the simian haemorrhagic fever virus (SHFV), which is grouped into the genus Simarterevirus. Interestingly, the genes encoding the minor glycoproteins GP2, GP3 and GP4 are duplicated in their genomes (Brinton et al., 2015). The GP3 proteins of SHFV (gene bank number AF180391.2), which are functionally non-redundant (Vatter et al., 2014) and exhibit only 18% amino acid identity, are predicted to have a cleaved signal peptide, but they contain two (ORF 3-1, VYC TLNFTNSS) or one glycosylation site (ORF3-2, CSA NQT) adjacent to the predicted cleavage site. Since the amino acid identity between the GP3 proteins of SHFV and EAV (GP3-1= 19.5%; GP3-2 = 15.6%) and SHFV and PRRSV (GP3-1 = 22.2%; GP3-2=17.4%) is very similar no educated guess can be made how the GP3s of SFHV are processed.

In sum, we showed that a carbohydrate attached to a glycosylation site adjacent to the signal peptide of GP3 from PRRSV and LDV has no effect on cleavage. Likewise, cleavage was not affected if the conserved cysteine 33 was exchanged. By using GP3 mutants with and without signal peptide we showed that the signal peptide is cleaved from GP3 of LDV and PRRSV which is in accordance with bioinformatic predictions. Thus, GP3 proteins from PPRSV and LDV are clearly differentially processed compared to GP3 of EAV. Thus, the presence of a carbohydrate adjacent to the cleavage site alone is not sufficient to inhibit signal peptide cleavage. We hypothesized that the signal peptide itself determines the processing scheme, but found that GP3 from EAV with the signal peptide of PRRSV and GP3 from PRRSV with the signal

peptide of EAV behave like the corresponding wild type proteins. Therefore it is rather an unknown element in the remaining part of the protein that determines whether the signal peptide is cleaved.

2.5 Acknowledgements

This work was supported by the German Research Foundation (DFG) [grant numbers VE 141-13-1]; Minze Zhang is recipient of a PhD fellowship from the China Scholarship Council (CSC).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.virol.2017.11.026.

Supplementary data

Multiple sequence alignment of GP3 proteins

```
GP3EAVBU MGRAYSGPVALLCFFLYFCFICGSVGS-----NNTTICMHT----TSDTSVHLFYAANVT   51
GP3PRRSL ----MAHQCARFHFFL-CGFICYLVHSALASNSSSTLCFWFPLAH-GNTSFELTINYTIC   54
GP3LDVPL ----MAGVNAIFCLFICCSGICFVDC---NVTASTQVCLWFPIHETGKTHLELALNFTLC   53
              :   * : :*:    **           .: :*:       ..* ..*    .:

GP3EAVBU FPS-----HFQR-----------------HFAAAQDFVVHTGYEY-------AGVTMLV   81
GP3PRRSL MPCSTSQAARQRLEPGRNMWCKIGHDRCEERDHDELLMSIPSGYDNLKLEGYYAWLAFLS   114
GP3LDVPL KMCGVVENSYYS-----HPFC-VTDDTCEESSSQYLDFDFPDDVQ--LLSSIYGTIAINA   105
           .                                  : .  . :        . :::

GP3EAVBU HLFANLVLTFPSLVNCSRPVNVFANA--SCVQVVCSHTNS--TTGLGQ—L--SFSFVDE    133
GP3PRRSL FSY--AAQFHPELFGIGNVSRVFVDKRHQFICAEHDGHNSTVSTGHNISALYAAYYHHQ    172
GP3LDVPL LVF--AIHQVPSMF--GNVSAVFFSHNSLCYNSTPVEHNQNVTT------S---LAIFD    152
           :       *.:.  ..   ** .            *.  :*             :

GP3EAVBU D-----LRLHIRPTLICWFALLLVHFLPMPRCRGS-------------------------    163
GP3PRRSL DGGNWFHLEWLRPLFSSWLVLNISWFLRRSPVSPVSRRIYQILRPTRPRLPVSWSFRTSI   232
GP3LDVPL DVEKWYVLEYIRPLFSSWLVLNVSYFLRRSTAGRATRRL---------------------   191
         *         :** : .*:.* :  **

GP3EAVBU ---------------------------------    163
GP3PRRSL VSDLTGSQQRKRKFPSESRPNVVKPSVLPSTSR   265
GP3LDVPL ---------------------------------    191
```

Supplementary table 1: Sequence alignment of Gp3 from EAV Bucyrus, PRRSV Lelystad and LDV Plageman.

Alignment was done with Clustal Omega http://www.ebi.ac.uk/Tools/msa/clustalo/ using default settings. An * (asterisk) indicates positions which have a single, fully conserved residue, a : (colon) indicates conservation between amino acids with strongly similar properties and a . (period) with weakly similar properties. Cysteines within the signal peptide are highlighted in grey and cysteines in the remaining part of the protein in yellow. N-glycosylation sites are in b.

2.6 References

Allen, S., Naim, H.Y., Bulleid, N.J., 1995. Intracellular folding of tissue-type plasminogen activator. Effects of disulfide bond formation on N-linked glycosylation and secretion. J Biol Chem 270, 4797-4804.

An, T.Q., Tian, Z.J., Leng, C.L., Peng, J.M., Tong, G.Z., 2011. Highly pathogenic porcine reproductive and respiratory syndrome virus, Asia. Emerg Infect Dis 17, 1782-1784.

Auclair, S.M., Bhanu, M.K., Kendall, D.A., 2012. Signal peptidase I: cleaving the way to mature proteins. Protein Sci 21, 13-25.

Balasuriya, U.B., Go, Y.Y., MacLachlan, N.J., 2013. Equine arteritis virus. Vet Microbiol 167, 93-122.

Balasuriya, U.B., Hedges, J.F., Nadler, S.A., McCollum, W.H., Timoney, P.J., MacLachlan, N.J., 1999. Genetic stability of equine arteritis virus during horizontal and vertical transmission in an outbreak of equine viral arteritis. J Gen Virol 80 (Pt 8), 1949-1958.

Benfield, D.A., Nelson, E., Collins, J.E., Harris, L., Goyal, S.M., Robison, D., Christianson, W.T., Morrison, R.B., Gorcyca, D., Chladek, D., 1992. Characterization of swine infertility and respiratory syndrome (SIRS) virus (isolate ATCC VR-2332). J Vet Diagn Invest 4, 127-133.

Blobel, G., Dobberstein, B., 1975. Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. J Cell Biol 67, 835-851.

Brinton, M.A., Di, H., Vatter, H.A., 2015. Simian hemorrhagic fever virus: Recent advances. Virus Res 202, 112-119.

Chand, R.J., Trible, B.R., Rowland, R.R., 2012. Pathogenesis of porcine reproductive and respiratory syndrome virus. Curr Opin Virol 2, 256-263.

Chen, X., VanValkenburgh, C., Liang, H., Fang, H., Green, N., 2001. Signal peptidase and oligosaccharyltransferase interact in a sequential and dependent manner within the endoplasmic reticulum. J Biol Chem 276, 2411-2416.

Das, P.B., Vu, H.L., Dinh, P.X., Cooney, J.L., Kwon, B., Osorio, F.A., Pattnaik, A.K., 2011. Glycosylation of minor envelope glycoproteins of porcine reproductive and respiratory syndrome virus in infectious virus recovery, receptor interaction, and immune response. Virology 410, 385-394.

de Lima, M., Ansari, I.H., Das, P.B., Ku, B.J., Martinez-Lobo, F.J., Pattnaik, A.K., Osorio, F.A., 2009. GP3 is a structural component of the PRRSV type II (US) virion. Virology 390, 31-36.

Faaberg, K.S., Plagemann, P.G., 1997. ORF 3 of lactate dehydrogenase-elevating virus encodes a soluble, nonstructural, highly glycosylated, and antigenic protein. Virology 227, 245-251.

Hegde, R.S., Bernstein, H.D., 2006. The surprising complexity of signal sequences. Trends Biochem Sci 31, 563-571.

Karniychuk, U.U., Geldhof, M., Vanhee, M., Van Doorsselaere, J., Saveleva, T.A., Nauwynck, H.J., 2010. Pathogenesis and antigenic characterization of a new East European subtype 3 porcine reproductive and respiratory syndrome virus isolate. BMC Vet Res 6, 30.

Kim, D.G., Song, C.S., Choi, I.S., Park, S.Y., Lee, J.B., Lee, S.S., 2013. The signal sequence of type II porcine reproductive and respiratory syndrome virus glycoprotein 3 is sufficient for endoplasmic reticulum retention. J Vet Sci 14, 307-313.

Kuhn, J.H., Lauck, M., Bailey, A.L., Shchetinin, A.M., Vishnevskaya, T.V., Bao, Y., Ng, T.F., LeBreton, M., Schneider, B.S., Gillis, A., Tamoufe, U., Diffo Jle, D., Takuo, J.M., Kondov, N.O., Coffey, L.L., Wolfe, N.D., Delwart, E., Clawson, A.N., Postnikova, E., Bollinger, L., Lackemeyer, M.G., Radoshitzky, S.R., Palacios, G., Wada, J., Shevtsova, Z.V., Jahrling, P.B., Lapin, B.A., Deriabin, P.G., Dunowska, M., Alkhovsky, S.V., Rogers, J., Friedrich, T.C., O'Connor, D.H., Goldberg, T.L., 2016. Reorganization and expansion of the nidoviral family Arteriviridae. Arch Virol 161, 755-768.

Mardassi, H., Athanassious, R., Mounir, S., Dea, S., 1994. Porcine reproductive and respiratory syndrome virus: morphological, biochemical and serological characteristics of Quebec isolates associated with acute and chronic outbreaks of porcine reproductive and respiratory syndrome. Can J Vet Res 58, 55-64.

Mardassi, H., Gonin, P., Gagnon, C.A., Massie, B., Dea, S., 1998. A subset of porcine reproductive and respiratory syndrome virus GP3 glycoprotein is released into the culture medium of cells as a non-virion-associated and membrane-free (soluble) form. J Virol 72, 6298-6306.

Matczuk, A.K., Kunec, D., Veit, M., 2013. Co-translational processing of glycoprotein 3 from equine arteritis virus: N-glycosylation adjacent to the signal peptide prevents cleavage. J Biol Chem 288, 35396-35405.

Matczuk, A.K., Veit, M., 2014. Signal peptide cleavage from GP3 enabled by removal of adjacent glycosylation sites does not impair replication of equine arteritis virus in cell culture, but the hydrophobic C-terminus is essential. Virus Res 183, 107-111.

Meulenberg, J.J., 2000. PRRSV, the virus. Vet Res 31, 11-21.

Meulenberg, J.J., Hulst, M.M., de Meijer, E.J., Moonen, P.L., den Besten, A., de Kluyver, E.P., Wensvoort, G., Moormann, R.J., 1993. Lelystad virus, the causative agent of porcine epidemic abortion and respiratory syndrome (PEARS), is related to LDV and EAV. Virology 192, 62-72.

Meulenberg, J.J., Petersen den Besten, A., de Kluyver, E., van Nieuwstadt, A., Wensvoort, G., Moormann, R.J., 1997. Molecular characterization of Lelystad virus. Vet Microbiol 55, 197-202.

Murtaugh, M.P., Stadejek, T., Abrahante, J.E., Lam, T.T., Leung, F.C., 2010. The ever-expanding diversity of porcine reproductive and respiratory syndrome virus. Virus Res 154, 18-30.

Oka, O.B., Bulleid, N.J., 2013. Forming disulfides in the endoplasmic reticulum. Biochim Biophys Acta 1833, 2425-2429.

Petersen, T.N., Brunak, S., von Heijne, G., Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8, 785-786. Rapoport, T.A., Li, L., Park, E., 2017. Structural and Mechanistic Insights into Protein Translocation. Annu Rev Cell Dev Bi 33, 369-390.

Shi, M., Lam, T.T., Hon, C.C., Hui, R.K., Faaberg, K.S., Wennblom, T., Murtaugh, M.P., Stadejek, T., Leung, F.C., 2010. Molecular epidemiology of PRRSV: a phylogenetic perspective. Virus Res 154, 7-17.

Shrimal, S., Cherepanova, N.A., Gilmore, R., 2015. Cotranslational and posttranslocational N-glycosylation of proteins in the endoplasmic reticulum. Semin Cell Dev Biol 41, 71-78.

Snijder, E.J., Kikkert, M., Fang, Y., 2013. Arterivirus molecular biology and pathogenesis. J Gen Virol 94, 2141-2163.

Thaa, B., Kaufer, S., Neumann, S.A., Peibst, B., Nauwynck, H., Krause, E., Veit, M., 2017. The complex co-translational processing of glycoprotein GP5 of type 1 porcine reproductive and respiratory syndrome virus. Virus Res 240, 112-120.

Thaa, B., Sinhadri, B.C., Tielesch, C., Krause, E., Veit, M., 2013. Signal peptide cleavage from GP5 of PRRSV: a minor fraction of molecules retains the decoy epitope, a presumed molecular cause for viral persistence. PLoS One 8, e65548.

Van Breedam, W., Delputte, P.L., Van Gorp, H., Misinzo, G., Vanderheijden, N., Duan, X.B., Nauwynck, H.J., 2010. Porcine reproductive and respiratory syndrome virus entry into the porcine macrophage. Journal of General Virology 91, 1659-1667.

van Nieuwstadt, A.P., Meulenberg, J.J., van Essen-Zanbergen, A., Petersen-den Besten, A., Bende, R.J., Moormann, R.J., Wensvoort, G., 1996. Proteins encoded by open reading frames 3 and 4 of the genome of Lelystad virus (Arteriviridae) are structural proteins of the virion. J Virol 70, 4767-4772.

Vanhee, M., Van Breedam, W., Costers, S., Geldhof, M., Noppe, Y., Nauwynck, H., 2011. Characterization of antigenic regions in the porcine reproductive and respiratory syndrome virus by the use of peptide-specific serum antibodies. Vaccine 29, 4794-4804.

Vatter, H.A., Di, H., Donaldson, E.F., Baric, R.S., Brinton, M.A., 2014. Each of the eight simian hemorrhagic fever virus minor structural proteins is functionally important (vol 462, pg 351, 2014). Virology 464, 461-461.

Veit, M., Matczuk, A.K., Sinhadri, B.C., Krause, E., Thaa, B., 2014. Membrane proteins of arterivirus particles: structure, topology, processing and function. Virus Res 194, 16-36.

von Heijne, G., 1983. Patterns of amino acids near signal-sequence cleavage sites. Eur J Biochem 133, 17-21.

Wieringa, R., de Vries, A.A., Raamsman, M.J., Rottier, P.J., 2002. Characterization of two new structural glycoproteins, GP(3) and GP(4), of equine arteritis virus. J Virol 76, 10829-10840.

Wieringa, R., de Vries, A.A., Rottier, P.J., 2003. Formation of disulfide-linked complexes between the three minor envelope glycoproteins (GP2b, GP3, and GP4) of equine arteritis virus. J Virol 77, 6216-6226.

Wieringa, R., de Vries, A.A.F., van der Meulen, J., Godeke, G.J., Onderwater, J.J.M., van Tol, H., Koerten, H.K., Mommaas, A.M., Snijder, E.J., Rottier, P.J.M., 2004. Structural protein requirements in equine arteritis virus assembly. Journal of Virology 78, 13019-13027.

Wissink, E.H., Kroese, M.V., van Wijk, H.A., Rijsewijk, F.A., Meulenberg, J.J., Rottier, P.J., 2005. Envelope protein requirements for the assembly of infectious virions of porcine reproductive and respiratory syndrome virus. J Virol 79, 12495-12506.

Zhang, Q.Z., Yoo, D.W., 2015. PRRS virus receptors and their role for pathogenesis. Veterinary Microbiology 177, 229-241.

Chapter 3

Glycoprotein 3 of porcine reproductive and respiratory syndrome virus exhibits an unusual hairpin-like membrane topology

Minze Zhang[a], Ludwig Krabben[a], Fangkun Wang[ab], Michael Veit[a]#

[a] Institut für Virologie, Freie Universität Berlin, Berlin, Germany.

[b] Shandong Provincial Key Laboratory of Animal Biotechnology and Disease Control and Prevention, Shandong Agricultfural University, Taian City, China.

3.1 ABSTRACT

The glycoprotein GP3 of the Arterivirus porcine reproductive and respiratory syndrome virus (PRRSV) consists of a cleaved signal peptide, a highly glycosylated domain, a short hydrophobic region and an unglycosylated C-terminal domain. GP3 is supposed to form a complex with GP2 and GP4 in virus particles, but secretion of GP3 from cells has also been reported.

We analyzed the membrane topology of GP3 from various PRRSV strains. A fraction of the protein is secreted from transfected cells; GP3 from PRRSV-1 strains to a greater extent than GP3 from PRRSV-2 strains. This secretion behavior is reversed after exchange of the variable C-terminal domain. A fluorescence protease protection assay shows that the C-terminus of GP3, fused to GFP, is resistant against proteolytic digestion in permeabilized cells. Furthermore, glycosylation sites inserted into the C-terminal part of GP3 are used. Both experiments indicate that the C-terminus of GP3 is translocated into the lumen of the endoplasmic reticulum. Deletion of the conserved hydrophobic region greatly enhances secretion of GP3 and fusion of this domain to GFP promotes membrane anchorage. Bioinformatics suggests that the hydrophobic region might form an amphipathic helix. Accordingly, exchanging only a few amino

acids in its hydrophilic face prevents and in its hydrophobic face enhances secretion of GP3. Exchanging the latter amino acids in the context of the viral genome did not affect release of virions, but released particles were not infectious. In sum, GP3 exhibits an unusual hairpin-like membrane topology that might explain why a fraction of the protein is secreted.

IMPORTANCE

The porcine reproductive and respiratory syndrome virus (PRRSV) is the most important pathogen in the pork industry. It causes persistent infections that lead to reduced weight gain of piglets; highly pathogenic strains even kill 90% of an infected pig population. PRRSV cannot be eliminated from pig farms by vaccination due to the large amino acid variability between the existing strains, especially in the glycoproteins. Here we analyzed basic structural features of glycoprotein 3 (GP3) from various PRRSV strains. We show that the protein exhibits an unusual hairpin-like membrane topology; membrane anchoring might occur via an amphipathic helix. This rather weak membrane anchor explains why a fraction of the protein is secreted from cells. Interestingly, PRRSV-1 strains secrete more GP3 than PRRSV-2. We speculate that secreted GP3 might play a role during PRRSV infection of pigs; it might serve as a decoy to distract antibodies away from virus particles.

3.2 INTRODUCTION

Arteriviruses are a family of enveloped positive stranded RNA viruses comprising the prototype member equine arteritis virus (EAV), lactate dehydrogenase-elevating virus (LDV) of mice and the porcine reproductive and respiratory syndrome virus (PRRSV), currently a major pathogen in the pork industry (1, 2). PRRSV was previously divided into two distinct genotypes termed “European” and “North American”, but because of the low nucleotide identity they are now classified as two genera, PRRSV-1 and PRRSV-2, respectively (3). The early isolates Lelystad virus (LV, PRRSV-1, (4, 5)) and VR-2332 (PRRSV-2, (6)) serve as respective prototype strains. Since their discovery, both genotypes have spread worldwide and diversified rapidly by mutation and recombination, including the occurrence of highly pathogenic variants in China ((7), related to PRRSV-2) and Eastern Europe ((8)), strain Lena, related to PRRSV-1). PRRSV cannot be eliminated from pig farms by vaccination due to the large variability between the existing strains. Especially the glycoproteins show strong antigenic drift and exhibit large variation (up to 50%) in their amino acid sequence (9, 10).

Arterivirus particles contain a multitude of membrane proteins, the disulphide-linked GP5/M dimer and the GP2/3/4 complex, the small and hydrophobic E-protein and the ORF5a protein (11). The glycoproteins GP2, GP3 and GP4 as well as the E-protein are essential for replication of both EAV and PRRSV. Deleting their genes individually from the viral genome does not prevent budding of virus-like particles from transfected cells, but the particles are not infectious (12, 13). More detailed experiments with EAV showed that if just one of the genes encoding GP2, GP3 or Gp4 were deleted the resulting particles do not contain the GP2/3/4 complex and reduced amounts of the E protein (12). It was therefore concluded that the GP2/3/4 complex is essential for virus entry, in the case of PRRSV probably due to its interaction with the key cellular receptor CD 163 (14-17).

Despite the large amount of sequence information on PRRSV genomes only limited information is available on the structure of their membrane proteins. GP2 and GP4 are typical type I transmembrane proteins that form an intermolecular disulfide linkage in the ER. GP3, the focus of this study, consists of an N-terminal signal peptide, a 180 amino acids long and highly glycosylated domain, a hydrophobic region (20 aa) and a variable unglycosylated C-terminal domain (50-60 aa). Some, but not all glycosylation sites are required for virus replication and loss of certain sites affect the antigenicity of GP3 (14, 18, 19). A multitude of linear antigenic regions were identified in GP3 proteins from both PRRSV-1 and PRSSV-2 strains, but antibodies directed against most epitopes do not neutralize viral infectivity (20-26). However, it was recently reported that PRRSV viruses which escaped from the antibody response in experimentally infected pigs contain mutations not only in its main glycoprotein GP5, but also in GP3 (27, 28). In transfected and infected cells GP3 is retained in the endoplasmic reticulum (ER), the viral budding site.

Inconsistent results have been published about whether GP3s are structural proteins of Arterivirus particles. Only for EAV it has been clearly shown, that GP3 is present in a disulfide-linked trimer with GP2/4 in virus particles. Uniquely, the disulfide linkages between GP3 and GP2/4 are not formed in the ER, but only after release of virus particles from infected cells (12, 29). In contrast, GP3 from LDV is only weakly associated with membranes and at least a fraction of the protein is secreted (30). Initial studies with GP3 from the PRRSV-2 strain IAF-Klop showed that it is secreted both from transfected and infected cells in a membrane-free form as disulphide-linked dimer carrying complex type carbohydrates (31). In contrast, a more recent study convincingly demonstrated that GP3 from the PRRSV-II FL-12 strain is incorporated into virus particles (32) as is GP3 from PRRSV-1 strain Lelystad (33). Since in the

latter two studies it was not investigated whether a fraction of GP3 is secreted from transfected cells, we systematically examined whether this is the case for GP3 proteins from various PRRSV-1 and PRRSV-2 strains.

The GP3 protein seems to take a special position among the arterivirus glycoproteins; especially its membrane topology has not been resolved and remains speculative. For GP3 from EAV we showed that carbohydrates attached to overlapping glycosylation sites (NNTT) adjacent to the signal peptide prevents signal peptide cleavage. However, the uncleaved signal peptide does not act as a membrane anchor. Membrane attachment is caused by the hydrophobic C- terminus of GP3, which does not span the membrane, but rather attaches the protein peripherally to ER membranes (34, 35). In contrast, the signal peptide of GP3 from PRRSV (and of LDV) is cleaved despite the presence of a carbohydrate in its vicinity (36) indicating that there are differences in protein processing between the Arterivirus genera Rodartevirus (LDV and PRRSV) and Equartevirus (EAV).

Here we systematically investigated whether GP3 proteins from PRRSV-1 and PRRSV-2 strains are secreted from transfected cells. We also show that the C-terminal part of GP3 is translocated into the lumen of the ER. Membrane anchoring is achieved by a short hydrophobic region that might form an amphipathic helix.

3.3 MATERIAL AND METHODS

3.3.1 Cells

Cell lines CHO-K1 (Chinese hamster ovary cells, ATCC CCL-61), BHK-21 (baby hamster kidney cells, ATCC C13), 293T (Human embryonic kidney cells, ATCC CRL-3216), and MARC-145 (simian kidney epithelial cells derived from MA-104, ATCC CRL-6489) were maintained as adherent culture in Dulbecco's Modified Eagle's Medium (DMEM, PAN, Aidenbach, Germany) supplemented with 10% fetal calf serum (FCS) (Perbio, Bonn, Germany), 100U of penicillin per ml, and 100 mg of streptomycin per ml at 37°C in an atmosphere with 5% CO2 and 95% humidity.

3.3.2 Plasmids and mutagenesis

The nucleotide sequences of open reading frame 3 of the following five strains of PRRSV was synthesized by Bio Basic Inc. (Markham Ontario, Canada): Lelystad virus (LV): low pathogenic PRRSV-1 prototype strain (4), GenBank accession number: M96262.2; VR-2332: low pathogenic PRRSV-2 prototype strain (37), accession

number: U87392.3; IAF-Klop: PRRSV-2 Québec reference strain (38), accession number: AF003344; XH-GD: Chinese highly pathogenic PRRSV-2 strain, accession number EU624117.1 and Lena: highly pathogenic PRRSV-1 strain (8), accession number: JF802085.1. All GP3 genes were equipped during synthesis at the 3'end with a sequence encoding the HA-tag (amino acids YPYDVPDYA) plus a small linker (PV). The GP3 genes with HA-tag were subcloned into the plasmid pCMV-TNT (Promega, (Mannheim, Germany), containing T7 and CMV promotors) using KpnI and NotI restriction sites.

Using these plasmids as templates, site-directed mutagenesis was performed by overlap extension polymerase chain reaction (PCR) using standard molecular biology techniques to generate the following GP3-HA mutants: VR-2332 GP3 with point mutations L227N, A246N and Lelystad (LV) GP3 with point mutations F228N and V255S to introduce N-glycosylation sites, VR-2332 GP3 with point mutations N195Q and Lelystad virus GP3 with point mutations N194Q to delete a N-glycosylation site, VR-2332 GP3 with mutations in the predicted amphipathic helix: GP3-2A (L194A W198A), GP3-3A (L184A F187A W191A) and GP3-5A (L184A F187A W191A L194A W198A),. XH-GD GP3 with mutations in the predicted amphipathic helix: GP3-3H (N195S S197L W198L), GP3-4H (R185L P186L S189F S190F) and GP3-7H (R185L P186L S189F S190F N195S S197L W198L). These mutations in the predicted amphipathic helix were also performed in the context of mutant GP3A246N of VR-2332: GP3-3H (N195S S197L W198L), GP3-4H (R185L P186L S189F S190F) and GP3-7H (R185L P186L S189F S190F N195S S197L W198L).

To delete parts from the C-terminus of GP3 of VR-2332, nucleotides encoding the amino acids sequence 1-229 (GP3-$\triangle$C2) or 1-177 (GP3-$\triangle$C1) of GP3 were amplified by PCR using forward primers equipped with KpnI site or reversed primers encoding the HA-tag (amino acids YPYDVPDYA) including the small linker (PV) and equipped with a NotI site. Mutant GP3-$\triangle$C2 lacks the between strains variable region of GP3 and GP3-$\triangle$C1 also the hydrophobic region. The resulting sequences were subcloned into the plasmid pCMV-TNT using the same enzymes. This strategy was also used for construction of mutant GP3-$\triangle$C1 of Lelystad virus (LV) that contains the amino acids sequence 1-171 and mutants GP3-$\triangle$C that contains the amino acids residues 1-203 of GP3 of Lelystad virus (LV) or 1-204 of GP3 of XH-GD. GP5-HA has been described (39).

The two chimeras LV-XH-GD and XH-GD-LV were constructed using the GP3-wt plasmids of the PRRSV strains Lelystad virus (LV) and XH-GD as a template for

overlap extension polymerase chain reaction (PCR) using standard molecular biology techniques. The chimera LV-XH-GD contains the amino acids 1-203 of GP3 of LV (including hydrophobic region) plus the amino acids 205-254 of GP3 of XH-GD (complete C-terminus excluding hydrophobic region). The chimera XH-GD-LV contains the amino acids 1-204 of GP3 of XH-GD (including hydrophobic region) plus the amino acids 204-265 of GP3 of LV (complete C-terminus excluding hydrophobic region). Both chimeras contain the HA-tag (amino acids YPYDVPDYA) and the small linker (PV) as described above.

GP3-YFPs were constructed using plasmids GP3-HA (wild type) of the five PRRSV strains described above as a template. The full length GP3 gene was amplified by PCR using primers equipped with BamHI or KpnI sites and subcloned into pEYFP-N1 (Clontech, Saint-Germain-en-Laye, France) using the same enzymes. This procedure inserts the linker sequence RPVATM between GP3 and YFP, which is derived from the cloning site of the vector. GP4-YFP was constructed as described (34) using the same strategy.

To fuse different regions from the C-terminus of GP3 of VR-2332 to the C-terminus of GFP, nucleotides encoding the amino acids sequences 178-254 (encoding the complete C-terminus including the hydrophobic region), 178-229 (the between strains variable region is not present) or 181-200 (hydrophobic region only) were amplified by PCR (primers equipped with BglII / Hind III or BamHI sites) and cloned to the corresponding sites in in pEGFP-C1 (Clontech, Saint-Germain-en-Laye, France). The resulting constructs are called GFP-HR1 (containing amino acids 181-200), GFP-HR2 (178-229) and GFP-HR3 (178-254), respectively. The cloning procedure adds the following amino acids between GFP and the GP3 sequence: GFP-HR1: RPDSDLEL; GFP-HR2 and GFP-HR3: SGL. The GP3 gene of all plasmids was sequenced (LGC Genomics, Berlin, Germany) before use in experiments.

3.3.3 Transfection and analyzing intracellular and secreted GP3

Cells in 6-well plates were transfected with 2.5 µg plasmid DNA using Lipofectamine® 3000 (Thermo Fisher Scientific, Carlsbad, United States) as described by the manufacturer. Twenty-four hours after transfection of cells, cellular supernatants were removed and cleared from cell debris by low speed centrifugation (3000xg, 10 min, 4°C). Proteins in the supernatant were precipitated with trichloroacetic acid (TCA, 10% final concentration) and pelleted (15000g, 30min, 4°C). Pellets were washed three times with 100% ethanol, dried and resuspended in reducing SDS-PAGE buffer.

To prepare intracellular GP3, cells were washed with PBS, detached from the dish with trypsin-EDTA (PAN Biotech, Aidenbach, Germany), pelleted, washed twice with phosphate-buffered saline (PBS, 137mM NaCl, 2.7mM KCl, 10mM Na2HPO4, 1.8mM KH2PO4, pH 7.4), resuspended in glycoprotein denaturing buffer (0.5% SDS, 40 mM DTT) and boiled for 10 min at 100°C.

To analyze secretion of GP3 in virus-infected cells, MARC-145 cells were infected with XH-GD or Lelystad virus (LV) at a moi of ~ 1 four hours after transfection and samples (cellular lysates and supernatants) were processed 24 hours after transfection as described above.

To analyze glycosylation samples were digested with Peptide-N-Glycosidase (PNGase F, 2.5-5units/µL, 4 h at 37°C) or endo-beta-N-acetyl-glucosaminidase (Endo H, 2.5-5units/µL, 1 h at 37°C) according to the manufacturer's instruction (New England Biolabs, Frankfurt am Main, Germany). For limited PNGase F digestion, a serial 10-fold dilution of PNGase F (starting with 1 unit/µL) was prepared for incubation with the substrate. Deglycosylated or untreated samples were supplemented with reducing (in one experiment non-reducing) SDS-PAGE loading buffer and subjected to SDS-PAGE and Western blot.

3.3.4 Membrane fractionation of transfected cells

Two different membrane fractionation assays were performed. The first method separates microsomes and cytosol (Fig. 3.5), the second firstly prepares microsomes and then separates their luminal (soluble) and membranous content (Fig. 3.4). Twenty-four hours after transfection of 293T cells cellular supernatants were discarded. Cells were washed twice with ice-cold PBS and treated with digitonin (30µM) for permeabilization for 25 minutes on ice. Nuclei and cell debris were removed by low speed centrifugation (700xg, 3min, Eppendorf centrifuge). Microsomes (ER and Golgi) were recovered from the supernatant by high speed centrifugation (100000g, 1h, 4°C, 60 min., Beckman TL100 centrifuge, rotor TLA100.3). The supernatant of this centrifugation contains the cytosolic proteins, which were precipitated with TCA as described above.

Pellets of microsomes were either resuspended in reducing SDS-PAGE buffer (method 1) or were opened (method 2) using ultrasonication (Misonix XL-2000). Membranous (membrane-associated protein) and soluble fractions (soluble protein in the lumen of ER) were separated by centrifugation (100000g, 1h, 4°C). The membranous pellet was washed once in PBS and resuspended in reducing SDS-PAGE buffer. Soluble proteins were precipitated with TCA as described above.

3.3.5 SDS-PAGE and Western Blot

After sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) using 15% or 12% polyacrylamide gels were blotted onto polyvinylidene difluoride (PVDF) membrane (GE Healthcare, Freiburg im Breisgau, Germany). After blocking of membranes (blocking solution: 5% skim milk powder in PBS with 0.1% Tween-20 (PBST)) for 1 h at room temperature, antibodies in blocking solution were applied for overnight at 4°C: rabbit-anti-HA tag antibody (ab9110, Abcam, Cambridge, UK, 1:10.000) was used to detect GP3 with HA tag and GP5 with HA tag, Mouse-anti-N antibody (13E2) (reactive against PRRSV-1 strains, 1:1000), Mouse monoclonal anti-GP5 antibody (reactive against PRRSV-2 strains, 1: 1000). After washing (3x10 min with PBST), suitable horseradish peroxidase-coupled secondary antibody (anti-rabbit or anti-mouse, Sigma-Aldrich, Taufkirchen, Germany, 1:5.000) was applied for 1 hour at room temperature. After washing, signals were detected by chemiluminescence using the ECLplus reagent (Pierce/Thermo, Bonn, Germany) and a Fusion SL camera system (Peqlab, Erlangen, Germany).

3.3.6 Immunofluorescence Assay

Transfected CHO-K1 or infected MARC-145 cells grown in 6-well plates (Sarstedt, Nümbrecht, Germany) were fixed with paraformaldehyde (4% in phosphate-buffered saline (PBS)) for 20 min at 4°C, washed twice with PBS, permeabilized with 0.5% Triton in PBS for 10 min at room temperature, and washed again twice with PBS. After blocking (blocking solution: 3% bovine serum albumin (BSA) in PBS with 0.1% Tween-20 (PBST)) for 30 min at room temperature the cells were incubated with a mouse monoclonal anti-GP5 antibody diluted in blocking solution (reactive against PRRSV-2 strains, 1: 100) at room temperature for 1 hour, then washed three times with PBS, incubated with secondary antibody (Alex Fluor 568 goat anti-mouse IgG(H+L), Invitrogen, Darmstadt, Germany, 1:1000). Pictures were recorded using a ZEISS Axio Vert. A1 inverse epifluorescence microscope.

3.3.7 Fluorescence protease protection assay

The fluorescence protease protection assay was performed as described (34). Briefly, CHO-K1 cells seeded on 6-well plates were transfected with plasmids GP3-YFP of all PRRSV strains or GP4-YFP of EAV as described above. Twenty-four hours after transfection cells were treated with digitonin (30µM) for 1min, and then proteinase K (50ng/ml, Sigma-Aldrich) was added. Pictures were recorded using a ZEISS Axio Vert. A1 inverse epifluorescence microscope.

3.3.8 Mutagenesis and reverse genetics with full-length PRRSV cDNA clone

The full-length infectious PRRSV cDNA clone pOKXH-GD (a kind gift of Prof. Guihong Zhang, South China Agricultural University) was constructed from the XH-GD PRRSV strain (GenBank accession no. EU624117) and vector pOKq (40). The entire genome of PRRSV is present in four plasmids: the plasmid pOK-A contains a CMV promotor plus the PRRSV nucleotide sequences 1 to 6465 (fragment A), pOK-B nucleotide sequences 6466 to 9301 (fragment B), pOK-C 9302 to 12586 (fragment C), and pOK-D nucleotide sequences 12587 to 15345 plus a terminal BGH RNA transcription terminator sequence (fragment D).

Site-directed mutagenesis using overlap extension polymerase chain reaction (PCR) was performed with the pOK-D plasmid to introduce the same mutations in the hydrophobic region as in GP3 of VR-2332. GP3-2A contains the mutations L194A plus W198A, pGP3-3A the mutations L184A + F187A + W191A and pGP3-5A all five mutations, L184A + F187A + W191A + L194A + W198A. It was not possible to create the mentioned mutations in GP3 without affecting the coding sequence of the overlapping GP4 gene: the mutation L184A in GP3 changed A2G in GP4, mutation F187A changed F5C; mutation W191A changed L9C; mutation L194A changed F12C and mutation W198A changed V16G. The corresponding amino acids in GP4 are all located in the middle of the signal peptide, which is cleaved in EAV after amino acid 21 (11) and predicted by SignalP (http://www.cbs.dtu.dk/services/SignalP/) to be cleaved in PRRSV XH-GD after amino acid 22. Since this part of the signal peptide tolerates amino acid substitutions (Fig. 3.3A) and since SignalP also predicts in all mutants a cleaved signal peptide, the mutations are unlikely to affect the functionality of GP4.

The GP3 genes of the pOK-D plasmids containing the desired mutations were completely sequenced to exclude unspecific second site mutations introduced by PCR. The C fragment was cut from pOK-C plasmid by digestion with BamHI and HindIII and ligated to the pOK-D plasmid digested with the same enzymes. The B fragment was cut from pOK-B plasmid by digestion with BamHI and AflII and ligated into the resulting pOK-CD plasmid digested with the same enzymes. Finally, the A fragment was cut from the pOK-A plasmid by digestion with AflII and NotI and ligated with the resulting pOK-BCD plasmid digested with the same enzymes. The complete plasmid was also constructed from wild type pOK-D in parallel. After each ligation the resulting plasmids were digested with four different restriction enzymes. The resulting band pattern was

identical between the three mutants and wild-type indicating that the procedure did not introduce deletions or insertions into the PRRSV cDNA.

The plasmids (2.5 µg) containing the complete PRRSV genome were transfected into 80% confluent CHO-K1 or BHK-21 cells grown in 6-well plates using Lipofectamine® 3000 (Thermo Fisher Scientific, Carlsbad, United States) as described by the manufacturer. After 48 hours the supernatant was removed, and cells were subjected to immunofluorescence assay using anti-GP5 specific antibody. The supernatant was cleared by low speed centrifugation and 500µl was used to infect (PBS washed) MARC-145 cells grown to 70-80% confluence on 6-well plates. After incubation for 1 h at 37 °C, the inoculums were removed, cells were washed twice with PBS, and cells were incubated in culture medium (DMEM with 2% FCS) for 48 h. Cells were then subjected to immunofluorescence assay using anti-GP5 specific monoclonal antibody.

3.3.9 RNA isolation and RNA quantification by real time PCR

To calculate the amount of viral genomic RNA (as a measure of total viral particles) released into the supernatant of transfected or infected cells, quantitative real time PCR was performed. Supernatants (1-2ml) were cleared by low speed centrifugation (10min, 5000g) and 30µl RNA is extracted from 400µl supernatant from either transfected or infected cells using PureLink viral RNA/DNA Mini Kit (Invitrogen, Carlsbad, United States). 8µl RNA is then supplemented with 1µl 10xbuffer and 1µl RNase-free DNase and incubated at 37° for 45min to remove genomic DNA. The reaction is stopped by adding 1µl DNase stop solution. Samples are then reverse transcribed into cDNA using High Capacity cDNA reverse transcription kit (Applied Biosystems, Carlsbad, United States). The cDNA yield was 800-900ng/µl in a volume of 21µl. Samples are diluted to 300ng/µl for qRT-PCR.

A standard curve was generated by serial dilution (108-101 copies/µl) of the full-length cDNA clone pGP3-wt plasmid. The diluted plasmids and the cDNA from samples were used for qRT-PCR, which was performed using the StepOnePlus real-time PCR system (Applied Biosystems) with SYBR green as fluorophore. The primers amplified a 180 bp fragment (427bp - 606bp) from the GP3 gene. Gene copy numbers were calculated with StepOne Software v2.3 (Applied Biosystems). Results from two independent transfections or infections are shown as number of viral RNA (mean including standard deviation) released into the supernatant of transfected or infected 6-well cell culture plates (~5x105 cells).

3.4 RESULTS

3.4.1 A fraction of GP3 is secreted from transfected cells, but the amount varies between GP3s from PRRSV-1 and PRRSV-2 strains

To systematically investigate secretion of GP3 we expressed the ORF3 gene from the following five PRRSV strains in BHK-21 cells: from a low-pathogenic (Lelystad) and a highly pathogenic (Lena) “European” PRRSV-1 strain and two low-pathogenic (VR-2332, IAF-Klop) and one highly pathogenic (XH-GD) “North American” PRRSV-2 strain. All constructs were fused at their C-termini to an HA-tag for subsequent antibody detection. Western-blotting of cellular lysates and supernatants showed that GP3 from all five PRRSV strains is secreted, but to a different extent (Fig. 3.1A). Whereas the band intensity of intracellular GP3 is similar for all five GP3 proteins, the extracellular GP3 bands from the PRRSV-1 strains Lelystad and Lena are much more prominent than the corresponding band from the three PRRSV-2 strains. Similar results were obtained using transfection of CHO-K1 cells (data not shown). Quantification of band intensities (lysate and supernatant) revealed that approximately 20-30% of total GP3 from PRRSV-1 strains were released, whereas ~5-10% of GP3 from PRRSV-2 strains was secreted.

Whereas the SDS-PAGE mobilities of the intracellular forms of GP3 from the strains Lelystad, Lena and VR-2332 are identical (~42 kDa), GP3 from XH-GD exhibits a slightly lower (~39 kDa) and GP3 from the IAF-Klop strain a clearly higher molecular weight (~45 kDa). Although the number of amino acids differ in GP3s of the strains investigated here (254 (including signal peptide) in PRRSV-2 strains, 249 in Lena and 265 in Lelystad), the calculated molecular weights are almost identical (30 – 32 kDa). However, GP3s vary in the number of predicted N-glycosylation sites; seven in Lelystad, Lena and VR-2332, six in XH-GD, but eight in IAF-Klop (see Fig. 3.3A for localization of glycosylation sites). To investigate how many glycosylation sites are used we carried out a limited digestion of cell lysates with PNGase F prior to western blotting. Counting of bands revealed that GP3 from Lelystad and VR-2332 contain six, GP3 from XH-GD five and GP3 from IAF-Klop seven oligosaccharides (Fig. 3.1B-E). Thus, in each of the GP3s all (except the one in the hydrophobic region, see Fig. 3.6) predicted glycosylation sites are used which explains the different SDS-PAGE mobilities of the glycosylated proteins.

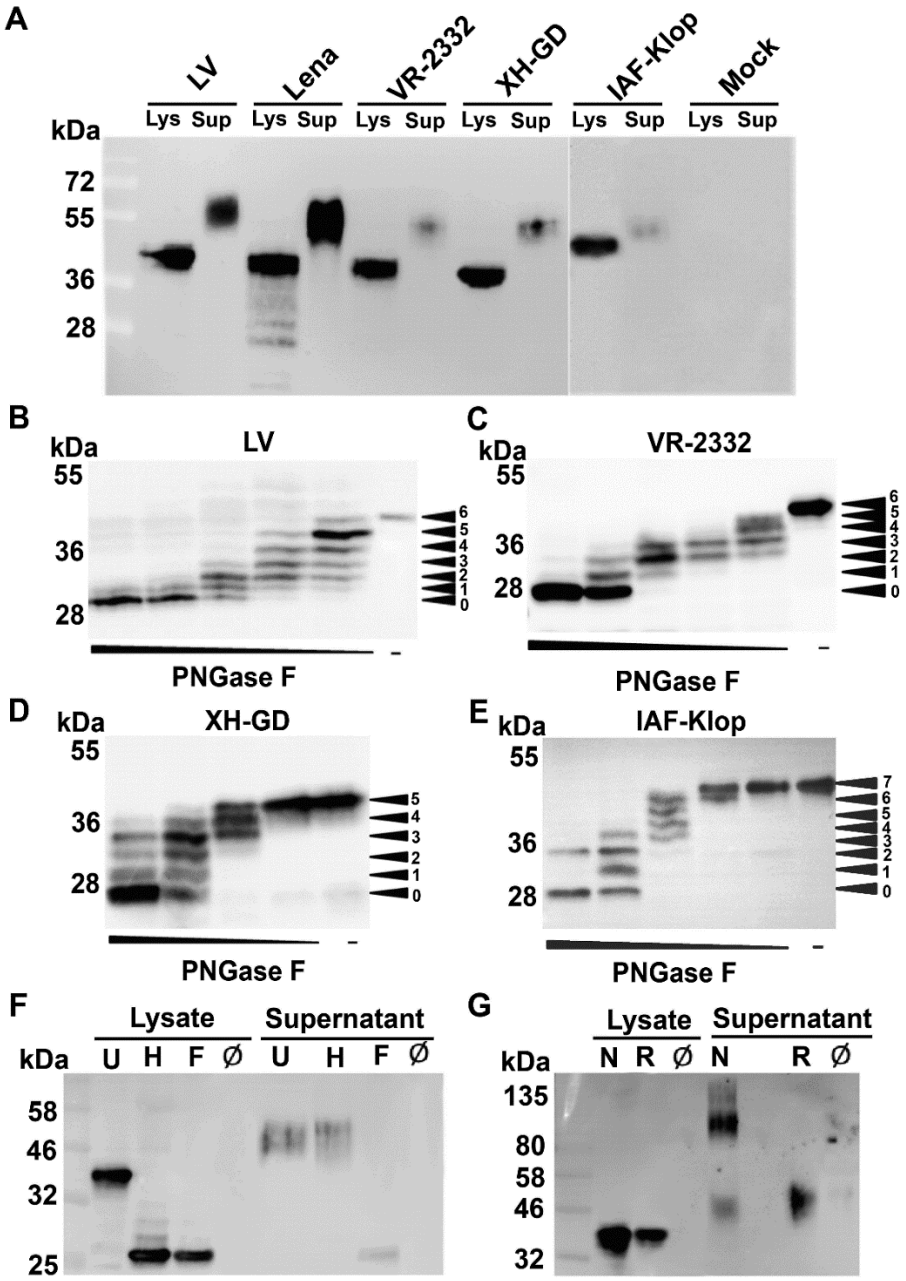

FIG.3.1. A fraction of GP3 is secreted from transfected cells

(A) GP3 (equipped with a C-terminal HA-tag) from five PRRSV-1 and PRRSV-2 strains was expressed in BHK-21 cell. 24 hours after transfection the supernatant was removed, cleared and proteins were precipitated with TCA. Cells were washed, trypsinized, pelleted and lysed in denaturing buffer. Samples were subjected to reducing SDS-PAGE and Western blotting with anti-HA antibodies.

To compare the amount of secreted GP3 between strains, different aliquots of the cell lysate were loaded onto the gel to achieve the same band intensity for intracellular GP3. The aliquot of the supernatant was adjusted accordingly for each GP3. Since ~ 5% (1-12.5%) of the total lysate, but much more (17.5-100%) of the supernatant was loaded onto the gel, the amount of secreted GP3 is overrepresented in this blot. Lelystad virus (LV): PRRSV-1 prototype strain,

VR-2332: PRRSV-2 prototype strain; IAF-Klop: PRRSV-2 strain; XH-GD: highly pathogenic Chinese PRRSV-2 strain and Lena: highly pathogenic PRRSV-1 strain. Mock are untransfected cells.

(B-E) GP3-wt from Lelystad (B), VR-2332 (C), XH-GD (D) and from IAF-Klop virus (E) were expressed in BHK-21 cells, cell lysates were not digested (-) or digested with serial 10-fold dilutions of PNGase F prior to Western blotting with anti-HA antibodies. This causes a ladder-like appearance of bands that allows counting of the total number of carbohydrates linked to GP3 (arrowheads).

(F) GP3-wt from VR-2332 was expressed in BHK-21 cells, cell lysates and supernatants were not digested (U) or digested with PNGase F (F) or Endo H (H) prior to Western blotting with anti-HA antibodies.

(G) GP3-wt from VR-2332 was expressed in BHK-21 cells, cell lysates and supernatants were subjected to reducing (R) or non-reducing (N) SDS-PAGE and Western blotting. The left lanes show molecular weight markers with the indicated sizes in kDa. Ø indicates untransfected cells.

GP3 present in the supernatant displayed a clearly higher molecular weight than intracellular GP3 and a smear band pattern often observed for glycoproteins subjected to terminal and heterogeneous glycosylation in the medial and trans-Golgi. To prove that GP3 was secreted via the exocytic pathway, intracellular and secreted GP3 from VR-2332 were digested prior to SDS-PAGE with PNGase F (that cleaves all types of N-linked carbohydrates) or Endo-H (which cleaves only high-mannose type carbohydrates typical for ER-localized proteins). The carbohydrates from intracellular GP3 are completely sensitive against both enzymes confirming that GP3 is retained in the ER or a cis-Golgi compartment (31, 36). In contrast, GP3 present in the supernatant is completely resistant against Endo-H digestion indicating that all its carbohydrates are terminally glycosylated (Fig. 3.1F). The much higher molecular weight of secreted versus intracellular GP3 is thus due to addition of several N-acetyl-glucosamine, galactose and sialic acid moieties to each of the glycosylation sites in the medial and trans Golgi. Note also that intracellular and secreted GP3 have an identical molecular weight after deglycoslyation with PNGase F indicating that the secreted form is not a result of proteolytic digestion, as has been reported for the secreted fraction of other viral spike proteins (41). Cleavage of GP3 just upstream of the hydrophobic region would remove 80 amino acids thereby yielding a product with a much higher SDS-PAGE mobility.

To assess the oligomeric nature of GP3 we subjected intracellular and secreted forms of GP3 from PRRSV-2 strain VR-2332 to non-reducing SDS-PAGE. Whereas intracellular GP3 revealed no size difference in the presence or absence of DTT, the majority of secreted GP3 showed a much higher molecular weight under non-reducing conditions, most likely representing a disulfide-linked dimer (Fig. 3.1G). Identical results were obtained if GP3 from the PRRSV-1 strain Lelystad was digested with glycosidases or subjected to non-reducing SDS-PAGE (data not shown). Likewise, GP3 of IAF-Klop was also reported to be secreted as an Endo-H resistant, disulfide-linked dimer (31).

3.4.2 GP3 is also secreted in the context of a virus infection

One might argue that secretion of GP3 is an artifact of the expression system and the protein is not released from virus-infected cells, for example because it binds to the transmembrane proteins GP2 and/or GP4. To investigate this we transfected MARC-145 cells with GP3 from either XH-GD or from Lelystad virus, and infected four hours later with the corresponding virus. Cell lysates and supernatants (prepared 24 hours after transfection) were then tested with anti HA-antibodies to detect GP3 and with anti-GP5 and anti-N antibodies to monitor infection with XH-GD and Lelystad virus, respectively. The western blot shows that virus infection does neither change expression levels of GP3 inside cells nor its molecular weight (Fig. 3.2). In addition, GP3 was also detected in (almost) similar amounts in the supernatant of virus-infected cells. However, we cannot exclude that fusion of the HA-tag to GP3 prevents binding to GP2 and/or GP4, but for the IAF-Klop strain it was clearly shown that native GP3 is also secreted in a soluble, membrane-free form from virus-infected cells (31).

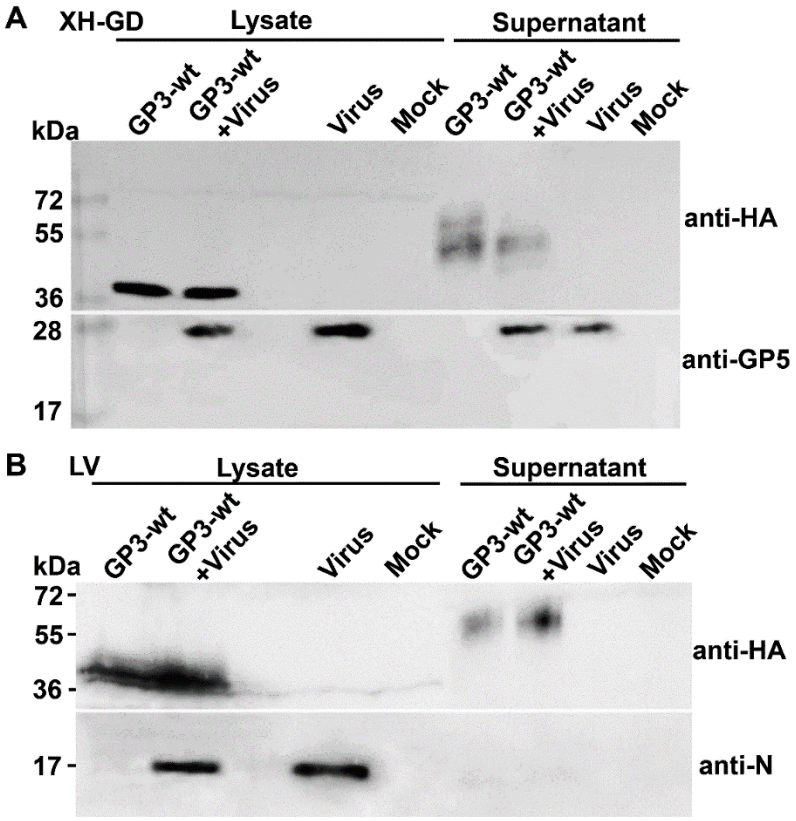

FIG. 3.2. GP3 is also secreted from virus infected cells

(A) and (B) MARC-145 cells were transfected with GP3-wt of XH-GD (A) or Lelystad virus (LV) (B) and 4 hours later infected or not infected with the corresponding virus at an moi of ~1. 24 hours after transfection, when some cytopathic effects appeared, supernatants and cells were processed as described in figure legend 3.1. Samples were subjected to reducing SDS-PAGE and Western blotting with anti-HA (to detect GP3-HA), anti-GP5 (for XH-GD) and anti-N (for LV) antibodies. The left lanes show molecular weight markers with the indicated sizes in kDa. “Virus” indicates infected, but untransfected cells; “Mock” untransfected, non-infected cells. Note that the glycosylation pattern of GP3 slightly differs between transfected and infected cells.

3.4.3 The C-terminus of GP3 is translocated into the lumen of the ER

A hydrophobic plot of GP3 revealed two hydrophobic regions; one is the N-terminal signal peptide, which we have shown to be cleaved in transfected cells (36), the second is an internal ~20 amino acid long hydrophobic region (HR). Since GP3 is a highly variable protein we also defined its conserved and variable parts. 619 ORF3 nucleotide sequences from both PRRSV-1 (162) and PRRSV-2 (457) strains present in the NCBI-database were translated into the corresponding amino acid sequences. From these data a consensus sequence for GP3 of both PRRSV-1 and PRRSV-2 containing the most abundant amino acid at each position was compiled and the

percent conservation at each position was plotted against the amino acid number (Fig. 3.3A). The plot shows that the signal peptide (especially in its C-terminal part which contains the cleavage site) and the C-terminal part downstream of the hydrophobic region are the most variable regions of GP3. In contrast, the hydrophobic region is highly conserved. The domain between the signal peptide and the hydrophobic region is fairly conserved; completely non-variable parts are interrupted by one or a few variable positions. Features which mainly determine the structure of GP3, such as six (possibly disulfide-bond forming) cysteine residues and also seven sites for N-linked glycosylation (at position 42, 50, 131, 152, 160 and 195 in VR-2332) are completely conserved in this region. Another highly conserved glycosylation site is present adjacent to the signal peptide, but its position varies between PRRSV-1 (amino acid 27) and PRSSV-2 strains (amino acid 29). No glycosylation sites or cysteine residues are present in the variable C-terminal part of GP3, suggesting that this region might exhibit a flexible structure.

Considering the basic features of GP3 (Fig. 3.3A) two membrane topologies are possible: It might be a classical type I membrane protein with an N-terminal luminal glycosylated domain, the hydrophobic region is spanning the membrane and the C-terminus is a cytoplasmic tail. Alternatively, GP3 might form a loop-like structure; both the N- and C-terminus of the protein are present in the lumen of the ER and the hydrophobic region peripherally anchors GP3 to the membrane.

We performed two different experiments to determine the membrane topology of GP3. First we used a fluorescence protease protection assay to investigate whether the C-terminus of GP3 is translocated into the ER lumen. The complete GP3 sequence of five PRRSV strains was fused at its C-terminus to YFP. The fluorophore will be protected from proteases added to perforated cells when it protrudes into the ER lumen, but not when exposed to the cytosol. As a control for a typical type I membrane protein with a C-terminus exposed to the cytoplasm, we used GP4 from equine arteritis virus (EAV) fused in a similar manner to YFP (34).

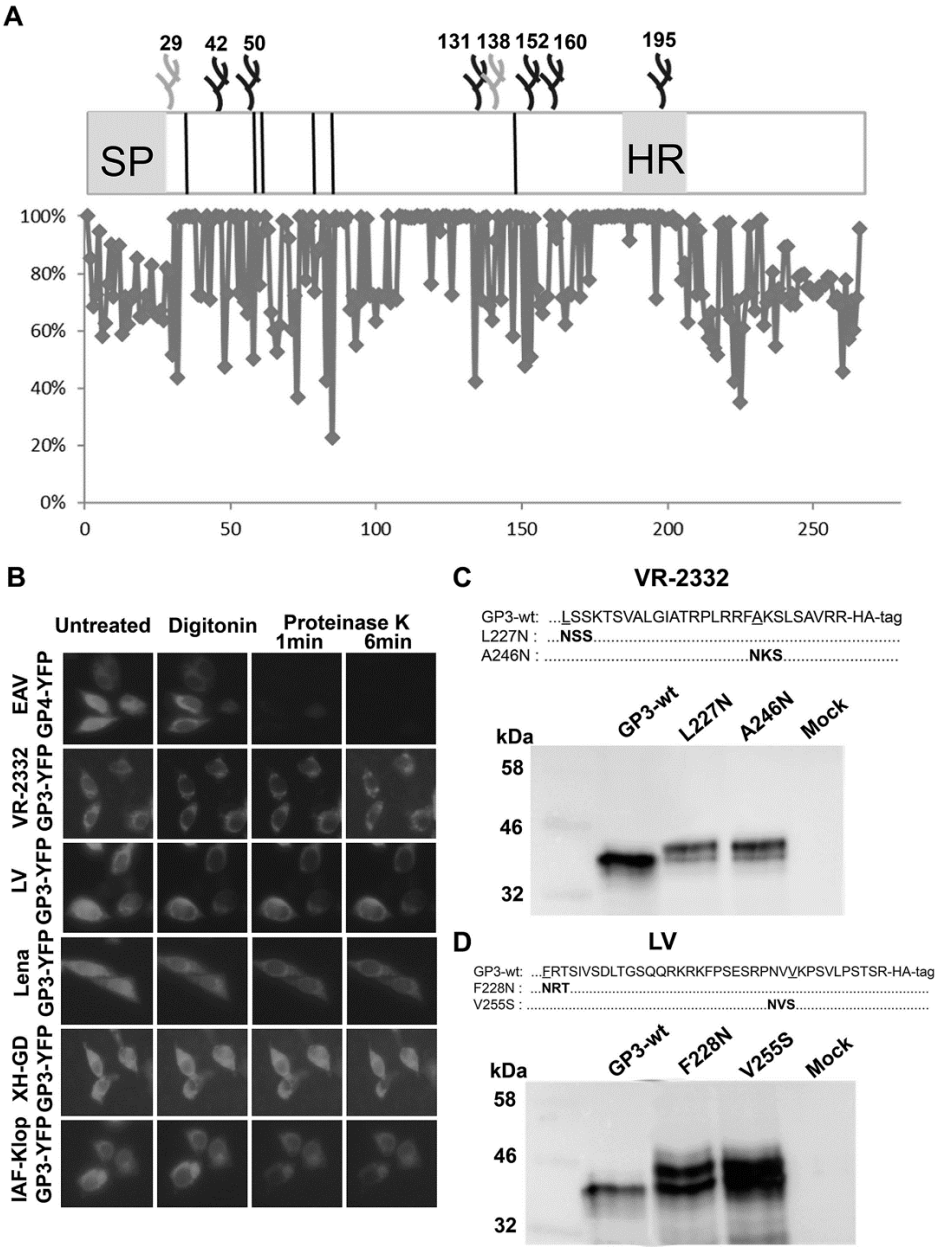

FIG. 3.3. The C-terminus of GP3 is translocated into the lumen of the ER

(A) Primary structure of GP3 with signal peptide (SP), hydrophobic region (HR), six conserved cysteines (lines) and glycosylation sites (branches, numbering of sites is for PRRSV-2 strains), The position of the first site differs between PRRSV-1 and PRRSV-2 strains, glycosylation site 138 is an additional site present in GP3 from the IAF-Klop strain, GP3 from XH-GD lacks the site at position 152. The graph shows the percent conservation (Y-axis) of amino acids at each position (X-axis) of a consensus sequence compiled from all PRRSV-1 and PRRSV-2 GP3 sequences present in the database.

(B) CHO-K1 cells expressing GP3-YFP from 5 different PRRSV strains and (as a control) a type I transmembrane protein (GP4-YFP from equine arteritis virus (EAV)) were treated with

digitonin for 1 min and with proteinase K for 1 and 6 min. After each time point, the same microscopic field was recorded with an epifluorescence microscope.

(C) and (D) Two additional glycosylation sites were inserted into the C-terminal part of GP3 from VR-2332 (C) or Lelystad virus (LV) (D) and constructs were expressed in BHK-cells. 24 hours after transfection cells were lysed and samples subjected to SDS-PAGE and Western blotting with anti-HA antibody. The C-terminal sequences of GP3 (amino acids 227-254 in VR-2332, 228-265 in LV) are shown above the blot. Bold letters indicate the N-glycosylation site inserted by exchange of one amino acid (underlined). The SDS-PAGE mobility of molecular weight makers are indicated at the left side of the blot. Mock: untransfected cells.

Fluorescence microscopy of transfected CHO-K1 cells showed comparable reticular ER staining for GP4-YFP and GP3-YFP of the five different PRRSV strains. Upon addition of the mild detergent digitonin, which solubilizes the plasma membrane, but not internal membranes, the fluorescence emitted from GP3-YFP and GP4-YFP was not diminished. However, only one minute after addition of Proteinase K, the same microscopic image section from cells expressing GP4-YFP did not show any fluorescence, indicating that YFP was digested. In contrast, even after six minutes of protease treatment, the fluorescence from GP3-YFP expressing cells was not even reduced (Fig. 3.3B). We conclude that the C-terminus of GP3 from PRRSV-1 and PRRSV-2 strains is exposed to the lumen of the ER.

To confirm this we inserted sites for N-linked glycosylation into the C-terminal part of GP3 from PRRSV-1 (Lelystad) and PRRSV-2 (VR-2332) strains. Since the active center of the oligosaccharyltransferase is exposed to the lumen of the ER, it can attach carbohydrates only to sites present in the same compartment (42). To minimize possible effects of the mutation on the structure of GP3 we exchange only one amino acid to create an N-glycosylation site at position 227 (mutation L to N) and 246 (A to N) of GP3 from VR-2332 and at position 228 (F to N) and 255 (V to S) of GP3 from Lelystad.

Western-blotting of cellular lysates showed one band for wildtype GP3, but two bands for all four mutants (Fig. 3.3C +D). The lower band ran at the same height as the corresponding wildtype GP3, whereas the other band showed a size increase consistent with the attachment of an additional glycan. In both mutants of GP3 from VR-2332 the upper band account for ~80% of total GP3, whereas in GP3 of Lelystad both fractions constitute ~ 50%. In principle, there are two possible explanations why the introduced sites are not stoichiometrically glycosylated: (i) The C-terminus might

not be translocated into the ER lumen in every synthesized GP3 molecule or (ii) C-terminal glycosylation sites might be occasionally skipped because ongoing protein folding hides them from the oligosaccharyltransferase as a quantitative study of the mouse N-glycoproteome suggested (43). The latter possibility is more likely since YFP attached to the C-terminus of GP3 was completely resistant against proteolytic digestion in the fluorescence protection assay (Fig. 3.3B).

3.4.4 A hydrophobic region in the C-terminal part of GP3 is the membrane anchor

Having established that the C-terminus of GP3 is translocated into the lumen of the ER, we next asked whether the hydrophobic region serves as a membrane anchor. We deleted two regions from the C-terminus of GP3 from VR-2332. The mutant GP3-ΔC2 contains amino acids 1-229; the variable C-terminal region is thereby removed, but the hydrophobic region is still present. In GP3-ΔC1 (containing amino acids 1-177) the hydrophobic region was also deleted (Fig. 3.4A). Western-blotting of cell lysates and supernatants showed that a larger fraction of GP3-ΔC1 is now secreted from transfected BHK-cells, whereas the extracellular amount of GP3-ΔC2 did only slightly increase relative to GP3-wt (Fig. 3.4B). Likewise, deletion of the whole C-terminus including the hydrophobic region of GP3 from the PRRSV-1 reference strain Lelystad increased secretion several fold (Fig. 3.4C).

From this experiment we conclude that deletion of the hydrophobic region enhances secretion of GP3, but it does not prove that GP3 is a membrane protein. One could imagine that the hydrophobic region functions as an ER retention signal and then its removal would allow secretion of a completely soluble protein. We therefore investigated whether intracellular GP3 is membrane-bound and whether the deletion of the hydrophobic region converted intracellular GP3 into a soluble form. Microsomes were prepared from transfected cells, opened by ultrasonication, membranous and luminal fractions were separated by ultracentrifugation and analyzed by Western blotting. The transmembrane protein GP5-HA from PRRSV and calreticulin, a marker for the lumen of the ER, were employed as controls to check the accuracy of the fractionation procedure. GP5-HA, as expected, and also wildtype GP3-HA are only detected in the membrane fraction. GP3-ΔC1 is also mainly present in the membrane fraction, but a clear signal was also detected in the luminal fraction. The amount of soluble GP3-ΔC1 is probably underestimated, since calreticulin, a soluble protein, also fractionates partly with the membrane, most likely due to the non-quantitative removal of luminal content from the membrane (Fig. 3.4D). Note also that the soluble fraction of GP3-ΔC1 has the same molecular weight as membrane-bound GP3-ΔC1. A band

with terminal glycosylated carbohydrates that increases the molecular weight is not visible. Thus, once GP3 is released from the membrane it is rapidly secreted from cells as it is more thoroughly been described in the discussion.

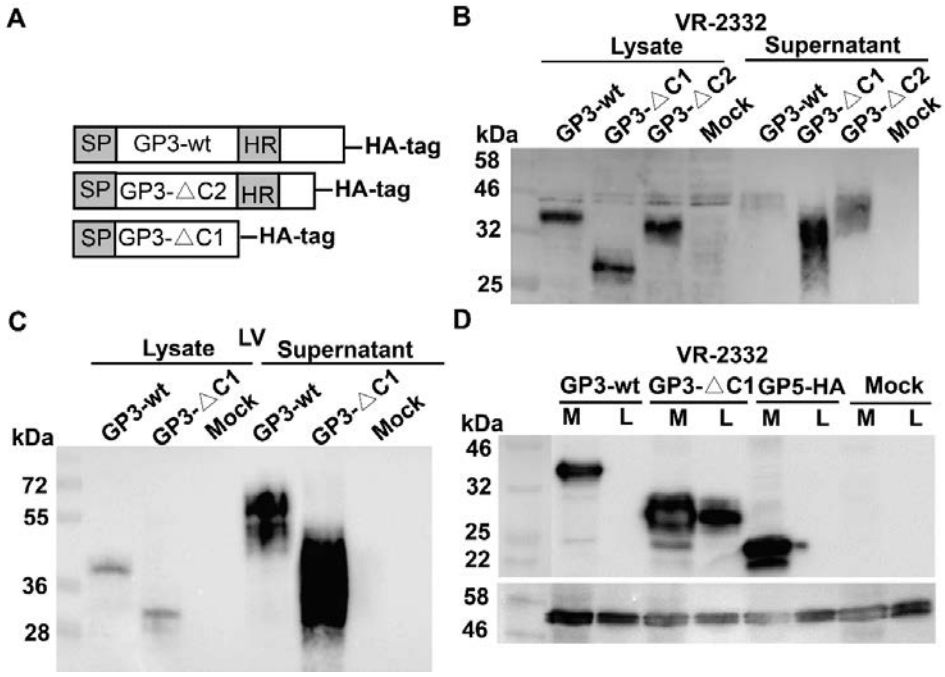

FIG. 3.4. The hydrophobic region at C-terminus of GP3 is a membrane anchor

(A) Scheme of constructs used in these experiments. In GP3-ΔC2 the (between PRRSV strains) variable C-terminus (amino acids 230 to 254) was deleted. GP3-ΔC1 contains a deletion of The C-terminal part including the hydrophobic region (amino acids 178 to 229 in VR-2332, 172 to 265 in Lelystad virus). SP: signal peptide; HR: hydrophobic region.

(B) Release of C-terminal deleted GP3 from VR-2332 into the supernatant of transfected cells.

(C) Release of C-terminal deleted GP3 from Lelystad virus into the supernatant of transfected cells. GP3-HA wild type and mutants GP3-△C1, GP3-△C2 were expressed in BHK-21 cell. 24 hours after transfection the supernatant was removed, cleared and proteins were precipitated with TCA. Cells were washed, trypsinized, pelleted and lysed in denaturing buffer. Samples were subjected to reducing SDS-PAGE and Western blotting with anti-HA antibodies.

(D) Release of C-terminal deleted GP3 into the lumen of the ER. 293T cells were transfected with GP3 wild type, mutant GP3-△C1 or (as control) transmembrane protein GP5-HA. 24 hours after transfection cells were gently opened by digitonin, microsomes were prepared by ultracentrifugation, resuspended in buffer and opened by ultrasonication. Membranous (M) and luminal (L) fractions were separated by ultracentrifugation and subjected to Western blotting with antibodies against HA-tag (upper panel) or against calreticulin, a luminal protein of the ER (lower panel).

Having shown that the hydrophobic region serves as a membrane-anchor of GP3, we asked whether this function can be transferred to another, otherwise soluble protein. We made three constructs in which either the complete C-terminus of GP3, the C-terminus excluding the highly variable region or only the 20 amino acids of the hydrophobic region were fused to the C-terminus of GFP. Fluorescence microcopy of transfected CHO-K1 cells showed that wild-type GFP is present throughout the cell (including the nucleus), whereas all three chimeras localize to an extranuclear, reticular organelle, presumably the ER (Fig. 3.5A).

Fractionation of whole cells into membranes and cytosol (100.000xg pellet and supernatant, respectively) and Western-blotting clearly shows that wild-type GFP is present in the soluble fraction, whereas all three GFP chimeras containing C-terminal GP3 sequences are exclusively membrane-bound (Fig. 3.5B). Thus, the short hydrophobic region in the C-terminal part of GP3 is sufficient to cause quantitative membrane anchoring of an otherwise soluble protein.

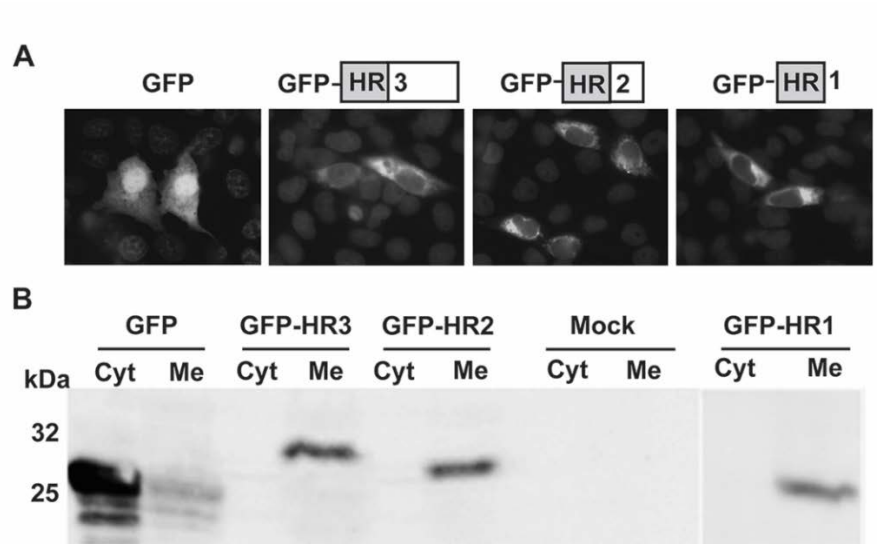

FIG. 3.5. The hydrophobic region of GP3 attaches GFP to membranes

(A) Scheme of GFP-constructs used in these experiments. GFP-HR3 contains the whole C-terminal part of GP3 from VR-2332 (amino acids 178 to 254). GFP-HR2 contains the C-terminal part of GP3 except amino acids 230 to 254, GFP-HR1 contains only the hydrophobic region (amino acids 181 to 200). Localization of GFP and GFP-chimeras in transfected CHO cells visualized with an epifluorescence microscope.

(B) CHO cells were transfected with GFP wild type and GFP chimeras, 24 hours after transfection, membrane (Me) and cytosolic (Cyt) fractions were separated by ultracentrifugation and subjected to Western blotting with antibodies against GFP. The SDS-PAGE mobility of molecular weight makers are indicated at the left side of the blot. Mock: untransfected cells.

3.4.5 Exchange of hydrophobic amino acids in the hydrophobic region by alanine increases secretion of GP3

The hydrophobic region is responsible for membrane anchoring of GP3, but the mechanism by which this association is accomplished remains unclear. The bioinformatics program Jpred (http://www.compbio.dundee.ac.uk/jpred/) predicts that the region might form an α-helix (Fig. 3.6A). Furthermore, this peptide is highly conserved (99 -100%) through GP3 proteins of all PRRSV-1 and PRRSV-2 strains, except at two positions. Position 187 is phenylalanine (98% conservation) in PRRSV-2 strains, but the equivalent position 186 in PRRSV-1 strains (including Lelystad, but not Lena) is leucine (72% conservation). Position 196 is valine in 95% of PRRSV-2 strains, but position 195 is isoleucine in 95% of PRRSV-1 strains (Fig. 3.6A). The second, slightly variable position is in the middle of an otherwise completely conserved N-glycosylation site NXS. If this site is used the bulky and hydrophilic carbohydrate would certainly not allow binding of GP3 to the ER membrane by hydrophobic forces. Instead, one could imagine that a cellular carbohydrate-binding transmembrane protein, such as Calnexin (44), might anchor GP3. To determine whether the glycosylation site is used, we exchanged the asparagine of NXS to a glutamine generating the mutants N194Q of Lelystad virus and N195Q of VR-2332. Transfection of BHK cells and western blot revealed that wildtype GP3 and the mutants have exactly the same size providing clear evidence that the glycosylation site is not used (Fig. 3.6B+C), as reported previously for GP3 of the PRRSV-2 FL12 strain (45).

Membrane binding by an amphipathic helix is another mechanism to mediate membrane binding of proteins (46). We therefore employed the bioinformatic tool heliquest (http://heliquest.ipmc.cnrs.fr/) to investigate whether the hydrophobic region might form an aliphatic helix. The distribution of polar versus hydrophobic residues in this putative helix shows that it indeed displays amphipathic characteristics. One surface of the helix contains hydrophobic amino acids, such as leucine, phenylalanine and tryptophan, which favor a lipid environment. The opposite surface also contains hydrophobic amino acids, but they are interrupted by hydrophilic (serine) or even charged ones (arginine), which prefer an aqueous environment (Fig. 3.7A).

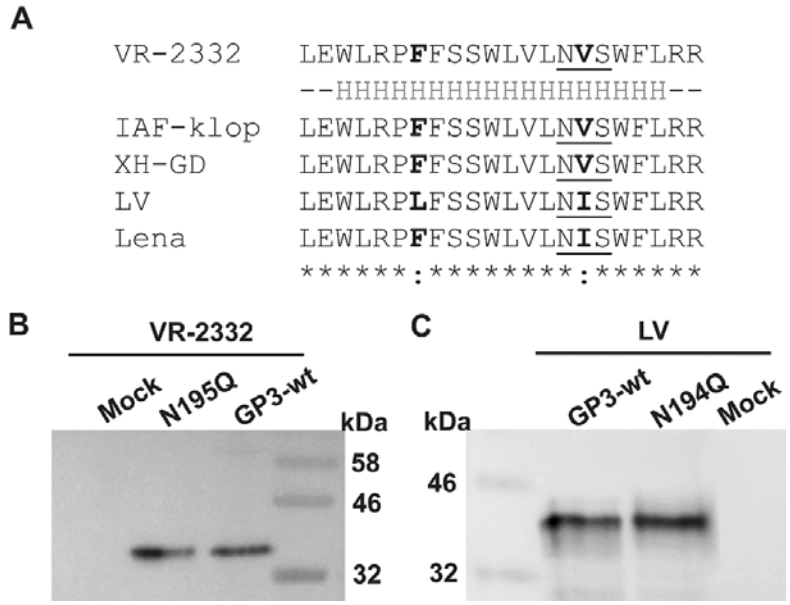

FIG. 3.6. The hydrophobic region is highly conserved and contains an unused glycosylation site.

(A) Sequence comparison of the hydrophobic region of GP3. The bioinformatics tool Jpred (http://www.compbio.dundee.ac.uk/jpred/) predicts that amino acids 183-200 of GP3 might from an α-helix (H below the sequence). The region is highly conserved between PRRSV strains. Only two conservative amino acid exchanges (in bold) are present in the used GP3 proteins. The conserved glycosylation site is underlined.

(B) and (C) The glycosylation site is not used. Asparagine (N195 in VR-2332 and N194 in Lelystad) of the glycosylation site NXS was exchanged to glutamine (Q) and mutants were expressed in BHK-cells. 24 hours after transfection cells were lysed and samples subjected to SDS-PAGE and Western blotting with anti-HA antibody.

To test the hypothesis that the (presumed) amphipathic helix anchors GP3 to the membrane, we replaced residues in its hydrophobic face by alanine. In GP3-2A leucine 194 and tryptophan 198 were replaced, the mutant GP3-3A contains exchanges of leucine 184, phenylalanine 187 and tryptophan 189. GP3-5A combines all five exchanges such that the predicted membrane insertion face of the helix is completely replaced by alanine (Fig. 3.7B). GP3-HA wildtype and the mutants were expressed in BHK-21 cells, and the distribution of GP3 protein in cells and supernatants was analyzed. The Western-blot clearly shows that secretion of all mutants is greatly enhanced to a similar extent as removal of the complete C-terminal part of GP3 in the mutant GP3-ΔC1 (Fig. 3.7C).

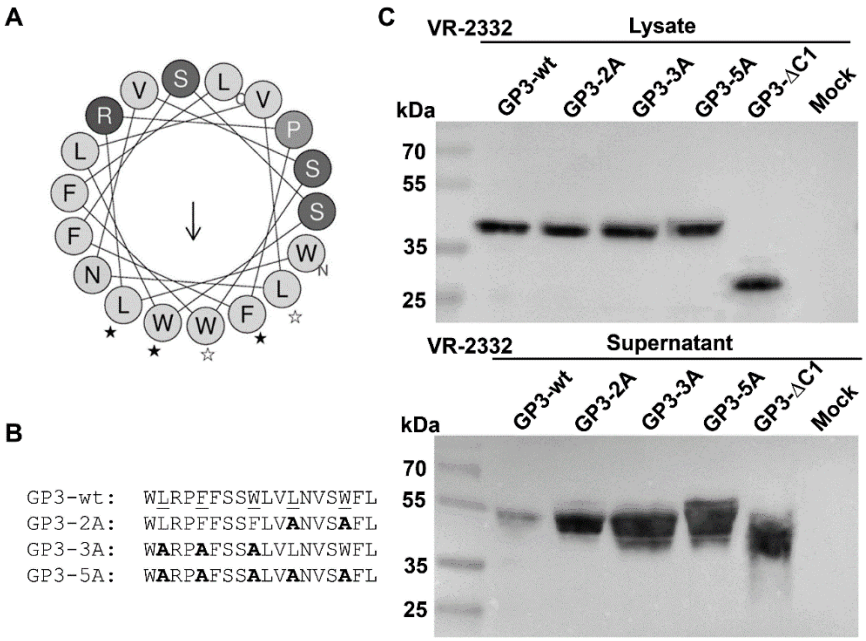

FIG. 3.7. The hydrophobic region might form an amphiphilic helix and exchange of residues in the hydrophobic face enhance secretion of GP3.

(A) The bioinformatics tool HeliQuest (http://heliquest.ipmc.cnrs.fr/) predicts that the hydrophobic region of GP3 (amino acids WLRPFFSSWLVLNVSWFL) might form an aliphatic helix with a mean hydrophobicity <H> of 1.131 and a hydrophobic moment <μH> of 0.302. The location of the mutations is indicated by white (mutant 2A) and black asterisks (mutant 3A).

(B) Exchange of amino acids in the hydrophobic face by alanines. Underlined amino acids correspond to the amino acids in the lower face of the predicted amphiphilic helix, as indicated by asterisks in (A).

(C) BHK-cells cells were transfected with GP3 wild type, and mutant GP3-2A, GP3-3A or GP3-5A. 24 hours after transfection the supernatant was removed, cleared and proteins were precipitated with TCA. Cells were washed, trypsinized, pelleted and lysed in denaturing buffer. Samples were subjected to reducing SDS-PAGE and Western blotting with anti-HA antibodies.

3.4.6 Substitution of hydrophilic by hydrophobic amino acids in the amphiphilic region prevents secretion of GP3

To further substantiate that the hydrophobic region determines the strength of membrane anchoring of GP3 we replaced hydrophilic with more hydrophobic residues. The mutant GP3-4H contains an exchange of arginine 185 and proline 186 by leucine plus a replacement of serine 189 and 190 by phenylalanine (black asterisks in Fig, 3.8A). The exchanged amino acids are solely located in the hydrophilic face of the predicted amphiphilic helix. In the mutant GP3-3H also amino acids in the hydrophobic

face were exchanged; asparagine 195 was replaced by a serine and serine 197 plus tryptophan 198 were replaced by isoleucines (white asterisks in Fig. 3.8A). GP3-7H combines all seven exchanges (Fig. 3.8B). When these three mutations were created in GP3 of the strain XH-GD secretion of the protein into the cellular supernatant was completely prevented indicating that the protein is now firmly membrane-anchored (Fig. 3.8C).

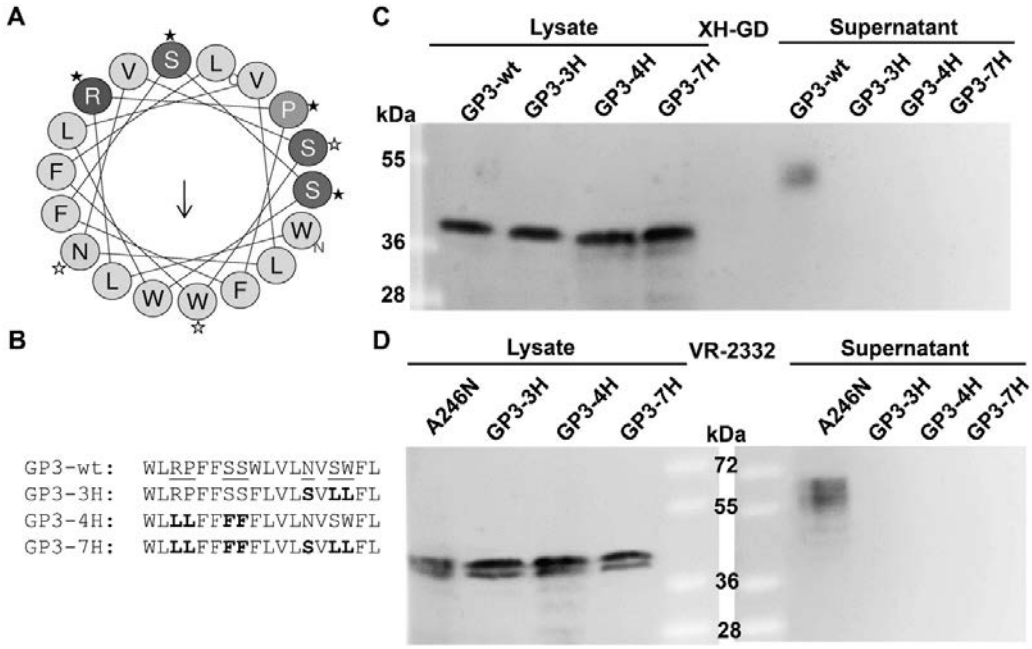

FIG. 3.8. Exchange of residues in the hydrophilic face by hydrophobic amino acids prevents secretion of GP3.

(A) Predicted aliphatic helix of GP3 with location of mutations indicated by white (mutant 3H) and black asterisks (mutant 4H).

(B) Exchange of hydrophilic amino acids by the more hydrophobic amino acids (leucine (L), phenylalanine (F) and serine (S).

(C) BHK-cells cells were transfected with mutant GP3-wt and GP3-3H, GP3-4H or GP3-7H of XH-GD.

(D) BHK-cells cells were transfected with mutant GP3-A246N (which contains an additional glycosylation site the C-terminal region) and mutant GP3-3H, GP3-4H or GP3-7H (D), which also contain this additional glycosylation site. 24 hours after transfection the supernatant was removed, cleared and proteins were precipitated with TCA. Cells were washed, trypsinized, pelleted and lysed in denaturing buffer. Samples were subjected to reducing SDS-PAGE and Western blotting with anti-HA antibodies.

In principle, this could be achieved by two mechanisms. Replacing hydrophilic by hydrophobic amino acids creates an uninterrupted stretch of 18 hydrophobic residues, which is long enough to span the ER membrane (47). This would create a transmembrane protein with the C-terminal region exposed to the cytosol. Alternatively, despite the mutations, the C-terminus (including the hydrophobic region) is completely translocated into the lumen of the ER. To distinguish between both possibilities we also tested the three mutants in the context of the GP3 mutant A264N of VR-2332, which contains the additional (and used) glycosylation site in the C-terminal part (Fig. 3.3C). Similar to GP3 from XH-GD the mutations prevent secretion of GP3 into the cellular supernatant. Comparing the intracellular band pattern revealed that in all three mutants the additional glycosylation site is used to the same extent as in GP3 without mutations in its hydrophobic region (Fig. 3.8D). Thus, the C-terminal part is translocated into the ER lumen; the protein does not adopt a transmembrane topology. In sum, we obtained clear evidence that the hydrophobic region determines membrane anchoring of GP3 probably by forming an amphiphilic helix, but a direct structural analysis is required to prove the later assumption.

3.4.7 The between PRRSV-1 and PRRSV-2 variable C-terminus determines the amount of secreted GP3

Having established that the highly conserved hydrophobic region is the main membrane anchor of GP3 we asked whether other between PRRSV-1 and PRRSV-2 strains variable features modulate how much of the wild type protein is secreted. A Kyte–Doolittle hydropathy plot revealed a remarkable difference in the biophysical properties of the C-terminal part of GP3; it is rather hydrophobic in GP3 from the PRRSV-2 strain XH-GD, but strongly hydrophilic in the PRRSV-1 strain Lelystad (Fig. 3.9A). We hypothesized that the different hydrophobicity in the C-terminus might affect the strength of membrane attachment of GP3s. We therefore deleted the C-terminus from GP3 from both PRRSV strains. This deletion considerably enhanced secretion of GP3 from XH-GD (Fig. 3.9C), but had no obvious effect on secretion of GP3 from Lelystad (Fig. 3.9B) which is already secreted in higher amounts compared to GP3 from XH-GD (Fig. 3.1A). Exchanging the C-terminus between both GP3 proteins completely reversed this secretion behavior. GP3 from Lelystad with the C-terminal domain of XH-GD is secreted in lower amounts than the corresponding wild-type protein (Fig. 3.9D), whereas GP3 from XH-GD with the C-terminal domain of Lelystad is secreted in higher amounts (Fig. 3.9E). In sum, the between PRRSV-1 and PRRSV-2 strains variable C-terminal domain modulates how much of GP3 is secreted.

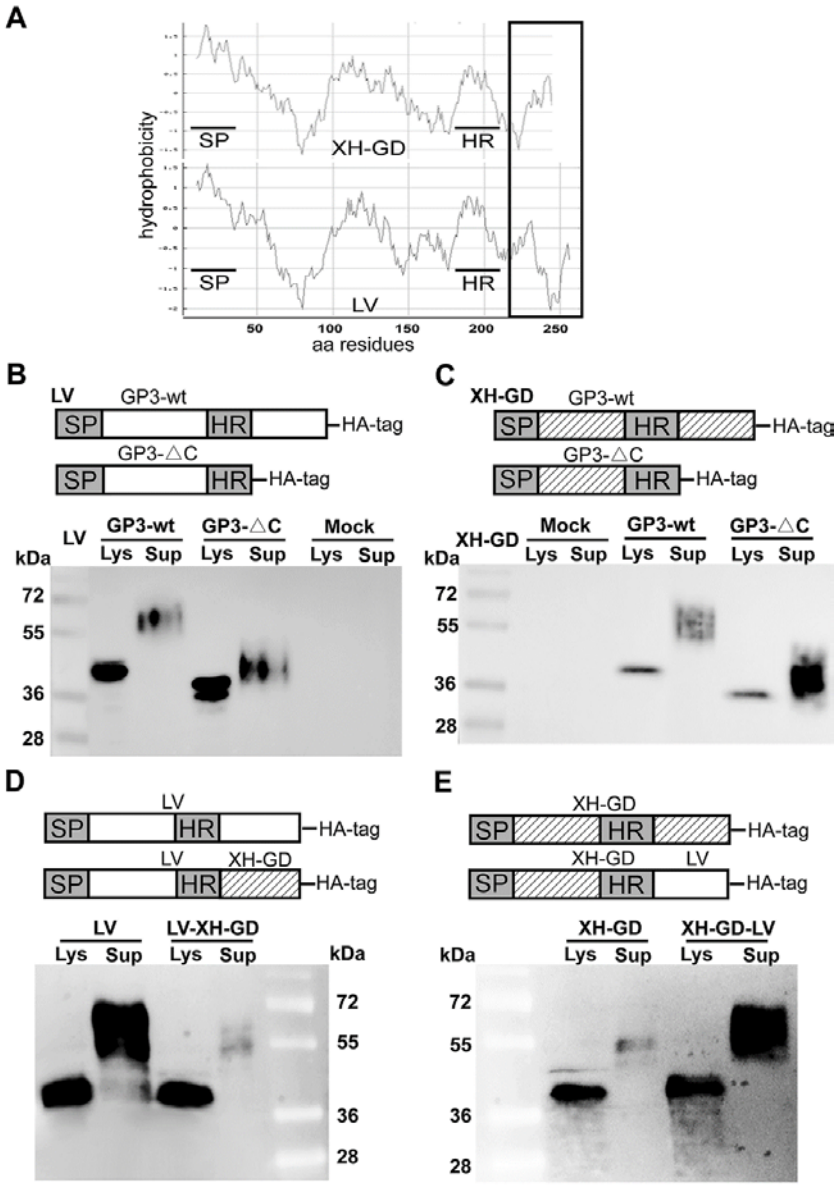

FIG.3.9. The variable C-terminus modulates secretion of GP3.

(A) Kyte-Doolittle hydropathy plot of GP3 from Lelystad (LV) and XH-GD PRRSV strains. y-axis: hydropathy score, x-axis: amino acid position. SP: signal peptide, HR: hydrophobic region. The black rectangle highlights the variable C-terminal region which is hydrophilic in GP3 from Lelystad, but hydrophobic in Gp3 from XH-GD strain.

(B) and (C) Deletion of the between strains variable C-terminal domain from GP3. In GP3-ΔC from Lelystad (LV) amino acids 204-265 and in GP3-ΔC from XH-GD amino acids 205-254 were deleted. Constructs of Lelystad (LV) (B) and XH-GD (C) were expressed together with the corresponding wild type proteins in BHK-21 cell.

(D) and (E) Exchange of the variable C-terminal domain of GP3 between PRRSV-1 and PRRSV-2 strains. (D) Creating the chimera LV-XH-GD: amino acids 1-203 of GP3 from Lelystad were fused to amino acids 205-254 of GP3 from XH-GD. (E) Creating the chimera XH-

GD-LV: amino acids 1-204 of GP3 from XH-GD were fused to amino acids 204-265 of GP3 from LV. The chimeras contain an HA-tag fused to their C-termini. Constructs of Lelystad (LV) (D) and XH-GD (E) were expressed together with the corresponding wild type proteins in BHK-21 cell. 24 hours after transfection the supernatant was removed, cleared and proteins were precipitated with TCA. Cells were washed, trypsinized, pelleted and lysed in denaturing buffer. Samples were subjected to reducing SDS-PAGE and Western blotting with anti-HA antibodies.

3.4.8 Mutation of the hydrophobic region inhibits replication of PRRSV

Finally, we analyze whether mutations in the hydrophobic region of GP3 affect virus replication. We used a reverse genetics system based on a full-length cDNA clone of the Chinese PRRSV-2 strain XH-GD cloned into a plasmid downstream of a CMV promotor. We created the same three mutants that caused enhanced secretion of GP3-HA from VR-2332, but this was not possible without altering also amino acids at the N-terminus of the overlapping GP4 gene. However, the changes are located entirely in the (variable) signal peptide of GP4 and are predicted by SignalP (http://www.cbs.dtu.dk/services/SignalP/) not to change its cleavability and hence the functionality of the mature GP4 protein (see “materials and methods” for details).

Full-length cDNA was transfected into CHO-K1 cells (which lack the receptors for cell entry of PRRSV) and 48 hours later the supernatant was used to infect MARC-145 cells, which support virus replication. Immunofluorescence using antibodies against GP5 revealed that a similar percentage of cells synthesize the viral glycoprotein, i. e. the transfection efficiency was comparable with wildtype or mutant cDNA (Fig. 3.10A). However, infected cells show a fluorescence signal only for wildtype virus, not for any of the three mutants. Not any immunofluorescence signal was detected if the passage one supernatant was subjected to two further passages in MARC-145 cells (not shown). Likewise, two further attempts to rescue infectious mutant virus (including transfection of BHK-21cells) also failed (not shown). We conclude that PRRSV mutants are able to synthesize viral proteins in transfected cells, but cells apparently do not release infectious virus particles.

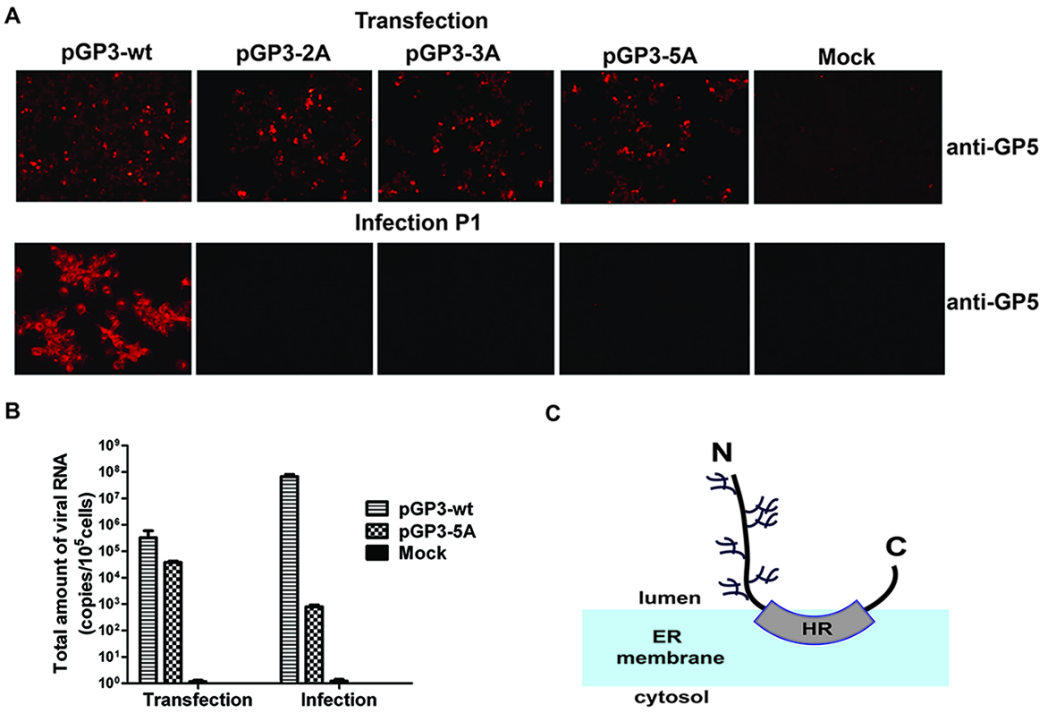

FIG. 3.10. Mutations in the predicted amphipathic helix prevent virus replication

(A) Immunofluorescence of cells transfected (upper panel) or infected (lower panel) with full-length clone of the PRRSV strain XH-GD, either wildtype (pGP3-wt) or the three alanine mutants pGP3-2A, pGP3-3A or pGP3-5A having the same mutations as shown in Fig. 3.7A and 3.7B.

Upper panel: CHO-K1 cells were transfected, ~48 h later the supernatant was removed and permeabilized cells were stained with anti-GP5 monoclonal antibody and Alex-568 anti-mouse secondary antibody.

Lower panel: MARC-145 cells were infected with the supernatant and 24 h post infection stained with anti-GP5 monoclonal antibody and Alex-568 anti-mouse secondary antibody.

(B) Quantification of viral genomes (pGP3-wt and pGP3-5A) released into the supernatant of transfected or infected cells. 24 hours post-transfection or post-infection, viral RNA was isolated from cleared culture supernatants and reverse transcribed into cDNA. Full-length plasmid of infectious clone pGP3-wt was as used as a standard to estimate the number of gene copies. Results from two independent transfections or infections are shown as number of viral RNA (mean including standard deviation) released into the supernatant of 1x105 transfected or infected cells.

(C) Model of GP3. HR: hydrophobic region which attaches GP3 to the membrane. N and C indicate the N – and C-terminus of the protein. The branches are the carbohydrates attached to GP3 of VR-2332. See Figs. 3.1B-E and 3.2A for number of carbohydrates attached to GP3 proteins of other PRRSV strains.

For both PRRSV and EAV it was reported that blockade of GP3 expression in an infectious clone prevents rescue of infectious virus, but virus-like particles (lacking GP2, GP3 and GP4) were released into the supernatant of transfected cells (12, 13). To analyze whether cells transfected with pGP3-5A cDNA having five amino acid substitutions in the helix exhibit the same phenotype we used qRT-PCR to estimate the number of viral genomes released into the supernatant of transfected CHO-K1 and infected MARC-145 cells. Comparable gene copy numbers were detected in the supernatant of cells transfected with wt cDNA (5x105) and pGP3-5A cDNA (8x104) indicating that the mutation does not prevent release of particles. When this supernatant was used to infect MARC-145 cells, 5x108 RNA gene copies of wild type virus were detected 48 hours later (indicating that the input virus multiplied ~1000fold), but very little (103) for the pGP3-5A mutant. Since the number of mutant genomes was greatly reduced (8x104 →103), the viral genomes still present in the supernatant of "infected" cells probably originate from the input virus. In sum, we conclude that the apolar face of the presumed amphipathic helix of GP3 is not essential for budding, but for virus entry.

3.5 DISCUSSION

In this study we investigated secretion, membrane-anchoring and the membrane topology of GP3 from PRRSV-1 and PRRSV-2 strains. We propose that GP3 exhibits an unusual hairpin-like membrane topology (Fig. 3.10C). We show that the C-terminal part of GP3 is exposed to the lumen of the ER since it is protected against proteolytic digestion in perforated cells (Fig. 3.3B) and since inserted N-glycosylation sites are efficiently used (Fig. 3.3C and 3.3D). This excludes that GP3 is a classical type I membrane protein containing a transmembrane region and a cytosolic tail. Nevertheless, the hydrophobic region serves as a membrane anchor. Its deletion enhances secretion of GP3 from transfected cells by several orders of magnitude (Fig. 3.4B+C). The bioinformatics tool heliquest predicts that the short hydrophobic region forms an amphipathic helix (Fig. 3.7A). Amphipathic helices can adopt an orientation parallel to the membrane plane; hydrophobic residues with long side chains insert between fatty acyl-chains of the lipids, whereas polar residues face their polar heads (46). Exchange of only two or three hydrophobic amino acids in the hydrophobic face of the presumed helix by alanine (which are also hydrophobic, but have only a short methylene group as side chain) increases secretion of GP3 to the same extent as removal of the complete C-terminal part of GP3 (Fig. 3.7C). Likewise, replacement of

three or four hydrophilic by hydrophobic residues prevents secretion of GP3 (Fig. 3.8) indicating that the hydrophobic properties of amino acids in this region determine the strength of membrane binding.

Finally, we could not generate infectious PRSSV virus having the same exchanges in the hydrophobic face of the helix of GP3 suggesting that the domain is essential for a complete replication cycle. However, viral genomes of the mutants (and thus probably virus particles) were released into the supernatant of transfected cells (Fig. 3.10). Thus, these mutants show an identical phenotype as mutants of the equine arteritis virus where expression of the GP3 gene was blocked (12, 13), reinforcing the assumption that a functional GP3 is not essential for budding, but for virus entry into target cells.

A multitude of amphipathic helices that interact with membranes have recently been characterized in cellular proteins. Their interactions are mostly reversible and restricted to certain cellular membranes in order to allow the respective protein to fulfil a specific cellular function. The binding specificity is often determined by interactions with certain lipids, for example basic amino acids in the helix interact with negatively charged lipids, such as phosphatidylserine or phosphatidylinositols (46).

Amphipathic helices are characterized by two physico-chemical parameters, the hydrophobic moment (<μH>) and the average hydrophobicity (<H>), which are calculated by heliquest. The hydrophobic moment tells us if a sequence, when considered as helical, exhibits one hydrophobic face and one polar face whereas the hydrophobicity describes the avidity of the helix for lipids (48). Calculating these parameters for the presumed helix of GP3 reveals that its hydrophobic moment is compatible with being amphipathic (0.302) and the average hydrophobicity is high (GP3: 1.131) suggesting a more stable membrane binding than transiently membrane-bound helices. Indeed, the hydrophobic region of GP3 confers complete membrane binding when fused to GFP, an otherwise soluble protein (Fig. 3.5). Nevertheless, its membrane anchoring ability is most likely lower than that of a transmembrane region. Once inserted into the membrane, transmembrane regions are permanently anchored; release (without damaging the bilayer) is not possible. Although shedding of viral type 1 membrane glycoproteins have been described, but they are often generated by proteolytic removal of the transmembrane region (41). However, this is not the case for GP3 since deglycosylated GP3 present inside cells and in the supernatant have identical SDS-PAGE mobility (Fig. 3.1).

Our results are best interpreted if we assume that membrane-bound and soluble GP3 exist in equilibrium. Wildtype GP3 is mostly membrane-bound in the ER, but some GP3 molecules might detach from the membrane and escape from this organelle by inclusion into the lumen of COPII vesicles, the carriers for transport from the ER to the Golgi. In the Golgi, all carbohydrates are processed to an Endo-H resistant, terminally glycosylated form that is secreted from the cells (Fig. 3.1F). Identical results have been reported previously for GP3 of the IAF-Klop strain (31).

Note, however, that GP3 with Endo-H resistant carbohydrates and a higher molecular weight was not observed in cell lysates (Fig. 3.1+3.4) suggesting that transport of GP3 through the Golgi is very fast. This is in line with the general concept that the rate-limiting step for protein transport along the exocytic pathway is exit from the ER; especially small proteins revealed very short half times for Golgi-transit of only 10 minutes (49). Thus, under steady state conditions (transfected cells) most GP3 molecules are either in the ER or already secreted; only a very small fraction of Endo-H resistant GP3 is present in the Golgi and this is therefore not detectable by Western blotting.

This very fast processing also suggests that soluble GP3 once transported to the Golgi does not rebind to its membranes; i. e. the equilibrium is now shifted to the soluble form. The reason might be that the physicochemical nature of the lipid bilayer changes along the exocytic pathway. The cholesterol content increases from ~10% to 45%, which causes tighter packaging of lipids and this might disfavor insertion of hydrophobic amino acid side chains. In addition, the asymmetric lipid distribution characteristic of the plasma membrane begins to build up in the Golgi, e. g. negatively charged lipids are translocated from the luminal to the cytoplasmic side of the membrane (50). These lipids might enhance binding of GP3´s helix to the ER membrane (a positively charged arginine is present in the helix), but their number might be reduced in the luminal part of Golgi membranes. However, detailed experiments with purified GP3 and artificial membranes are required to elucidate the lipid specificity of the helix and the amino acids essential for binding.

Exchange of hydrophobic residues in the apolar face of the helix by alanine shifts the membrane binding equilibrium of GP3 to the soluble version. Heliquest predicts that the average hydrophobicity of the helix is only moderately reduced in the mutants (wt: 1.131; mutant 2A: 0.946; 3A: 0.863; 5A: 0.678), but the hydrophobic moment is severely diminished in each mutant (wt: 0.302; 2A: 0.155, 3A: 0.068, 5A: 0.107) suggesting that this part of the molecule is no longer able to form an amphipathic helix.

As a result mutant GP3s bind to membranes only poorly and a much higher fraction is secreted from cells (Fig. 3.7).

In contrast, exchange of hydrophilic by hydrophobic residues in the polar face of the helix shifts the membrane binding equilibrium of GP3 to the membrane anchored form (Fig. 3.8). Accordingly, the mutations successively increase the hydrophobicity of the hydrophobic region (wt: 1.131, mutant 3H: 1.228; 4H: 1.539; 7H: 1.636) and also decreased the hydrophobic moment (wt: 0.302; 3H: 0.271, 4H: 0.076, 7H: 0.098).

Interestingly, the hydrophobic region also contains a completely conserved glycosylation site (N195 in VR-2332), which is not used (Fig. 3.6), presumably because membrane binding prevents access of the oligosaccharyltransferase to this site. The glycosylation site might become accessible for glycosylation in mutants with diminished membrane binding affinity since both intracellular and secreted forms of GP3-5A exhibit a higher molecular weight compared to wildtype GP3 (Fig. 3.7C). Other Western-blots having better separation revealed an additional band with a slightly higher molecular weight for each mutant consistent with the conclusion that a fraction of ~10% of GP3-2A and GP3-3A and ~50% of GP3-5A are glycosylated at N 195 (not shown).

Exchange of the hydrophobic residues by alanine in the presumed amphipathic helix of GP3 prevented release of infectious virus particles from transfected cells (Fig. 3.10). We assume that a (mostly) soluble version of GP3 does not form a functional GP2/3/4 complex, which is a prerequisite for virus entry (12, 13). We could not test whether GP3 (or GP2 or GP4) are incorporated into virus particles since antibodies against the minor glycoproteins of the XH-GD strain are not available and insertion of an HA-tag into a full-length clone of PRRSV is not possible since the C-terminus of GP3 overlaps with the N-terminus of GP4.

The essential nature of the hydrophobic region is also corroborated by the notion that it is highly conserved between GP3 proteins from all PRRSV-1 and -2 strains (Fig. 3.3A, Fig. 3.6A). A predicted amphipathic helix with a very similar sequence is also present in GP3 of LDV (YIRPLFSSWLVLNVSYFL, three conservative exchanges to PRRSV VR-2332 are underlined, <μH>: 0.215; <H>: 0.988) and it was reported that a fraction of the protein is secreted from cells and that it can be extracted from membranes without detergent suggesting that it is a peripheral membrane protein (30). GP3 from EAV also adopts a hairpin-like topology and is membrane-anchored by a C-terminal hydrophobic domain (34). This region (RPTLICWFALLLVHFLPMPRCRGS) exhibits no sequence homology to the presumed helices of PRRSV and LDV, but

heliquest predicts that it forms an amphipathic helix with similar biophysical properties (<μH>: 0.258; <H>: 1.181 for underlined amino acids) than the helices of PRRSV and LDV. Similar to PRRSV, membrane-anchoring of GP3 of EAV is also essential for virus replication since infectious mutants with a stop codon inserted into this region could not be generated (35). In sum, although experimentally not analyzed for each protein, GP3 proteins from PRRSV, LDV and EAV might have the same hairpin-like membrane topology and the same mechanism of membrane-anchoring, but the presumed amphipathic helix is formed by different amino acids in the genera Rodartevirus (LDV and PRRSV) and Equartevirus (EAV).

Although they contain an (almost) identical hydrophobic region, a larger fraction of GP3s from the PRRSV-1 strains is secreted from transfected cells compared to GP3s from the PRRSV-2 strains (Fig. 3.1). The one (Lena) or two (Lelystad) conservative exchanges in the helix reduced its physico-chemical parameters only slightly (Lena: H: 1.163; μH: 0.273, Lelystad: H: 1.158; μH: 0.269). Instead we have shown that the variable C-terminus modulates how much of GP3 is secreted. The C-terminal part is strongly hydrophilic in the PRRSV-1 strain Lelystad, but rather hydrophobic in the PRRSV-2 strain XH-GD (Fig. 3.9A). Exchanging the C-terminus between both GP3 proteins completely reversed this secretion behavior. GP3 from Lelystad with the C-terminal domain of XH-GD is secreted in lower amounts than the corresponding wild-type protein (Fig. 3.9D), whereas GP3 from XH-GD with the C-terminal domain of Lelystad is secreted in higher amounts (Fig. 3.9E).

Note, however, that the C-terminus is the most variable part of GP3 (Fig. 3.3A, the sequence also reveals insertions or deletions) and is thus remains to be shown whether increased GP3 secretion is a general feature of all PRRSV-1 strains. However, the secreted form of GP3 is likely to be a folded (and hence functional) protein, since it passed the quality control system of the ER (44). Furthermore, since a large fraction of secreted (but not intracellular) GP3 is a disulfide-linked dimer, at least one intermolecular disulfide bond is formed which would be unlikely to happen in a misfolded protein (Fig. 3.1G).

The only other viral structural glycoprotein having a similar membrane anchor is the Erns glycoprotein of Pestiviruses, an unspecific RNAse that suppresses the cellular innate immune response and is involved in the establishment of persistent infection. However, the amphiphilic helix of Erns is longer (60 amino acids) and located at the extreme C-terminus of the protein. Another similarity to GP3 is the dual nature of Erns. Erns is an essential part of the viral envelope, but is also secreted into the extracellular

medium and found in the serum of infected animals. This secreted form is hypothesized to relate to its role as a virulence factor (51-53).

We speculate that secreted GP3 might also play a role during PRRSV infection of pigs, for example as a decoy that serves to distract antibodies away from virus particles. Indeed, antibodies against various regions of the GP3 protein were elicited in the majority of experimentally infected piglets, but most of them had no or only little neutralizing activity (20-26).

3.6 ACKNOWLEDGEMENTS

This work was supported by the German Research Foundation (DFG, grant number VE 141/13-1); Minze Zhang is recipient of a PhD fellowship from the China Scholarship Council (CSC). Fangkun Wang´s sabbatical at the Free University Berlin was supported by the “growth of young teachers program” of the Shandong province. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

3.7 REFERENCES

1. **Chand RJ, Trible BR, Rowland RR. 2012.** Pathogenesis of porcine reproductive and respiratory syndrome virus. Curr Opin Virol 2:256-63.
2. **Snijder EJ, Kikkert M, Fang Y. 2013.** Arterivirus molecular biology and pathogenesis. J Gen Virol 94:2141-63.
3. **Kuhn JH, Lauck M, Bailey AL, Shchetinin AM, Vishnevskaya TV, Bao Y, Ng TF, LeBreton M, Schneider BS, Gillis A, Tamoufe U, Diffo Jle D, Takuo JM, Kondov NO, Coffey LL, Wolfe ND, Delwart E, Clawson AN, Postnikova E, Bollinger L, Lackemeyer MG, Radoshitzky SR, Palacios G, Wada J, Shevtsova ZV, Jahrling PB, Lapin BA, Deriabin PG, Dunowska M, Alkhovsky SV, Rogers J, Friedrich TC, O'Connor DH, Goldberg TL. 2016**. Reorganization and expansion of the nidoviral family Arteriviridae. Arch Virol 161:755-68.
4. **Meulenberg JJ, Hulst MM, de Meijer EJ, Moonen PL, den Besten A, de Kluyver EP, Wensvoort G, Moormann RJ. 1993.** Lelystad virus, the causative agent of porcine epidemic abortion and respiratory syndrome (PEARS), is related to LDV and EAV. Virology 192:62-72.
5. **Wensvoort G, Terpstra C, Pol JM, ter Laak EA, Bloemraad M, de Kluyver EP, Kragten C, van Buiten L, den Besten A, Wagenaar F, et al. 1991**. Mystery swine disease in The Netherlands: the isolation of Lelystad virus. Vet Q 13:121-30.
6. **Collins JE, Benfield DA, Christianson WT, Harris L, Hennings JC, Shaw DP, Goyal SM, McCullough S, Morrison RB, Joo HS, et al. 1992**. Isolation

of swine infertility and respiratory syndrome virus (isolate ATCC VR-2332) in North America and experimental reproduction of the disease in gnotobiotic pigs. J Vet Diagn Invest 4:117-26.

7. **Tian K, Yu X, Zhao T, Feng Y, Cao Z, Wang C, Hu Y, Chen X, Hu D, Tian X, Liu D, Zhang S, Deng X, Ding Y, Yang L, Zhang Y, Xiao H, Qiao M, Wang B, Hou L, Wang X, Yang X, Kang L, Sun M, Jin P, Wang S, Kitamura Y, Yan J, Gao GF. 2007.** Emergence of fatal PRRSV variants: unparalleled outbreaks of atypical PRRS in China and molecular dissection of the unique hallmark. PLoS One 2:e526.

8. **Karniychuk UU, Geldhof M, Vanhee M, Van Doorsselaere J, Saveleva TA, Nauwynck HJ. 2010.** Pathogenesis and antigenic characterization of a new East European subtype 3 porcine reproductive and respiratory syndrome virus isolate. BMC Vet Res 6:30.

9. **Murtaugh MP, Stadejek T, Abrahante JE, Lam TT, Leung FC. 2010.** The ever-expanding diversity of porcine reproductive and respiratory syndrome virus. Virus Res 154:18-30.

10. **Shi M, Lam TT, Hon CC, Hui RK, Faaberg KS, Wennblom T, Murtaugh MP, Stadejek T, Leung FC. 2010.** Molecular epidemiology of PRRSV: a phylogenetic perspective. Virus Res 154:7-17.

11. **Veit M, Matczuk AK, Sinhadri BC, Krause E, Thaa B. 2014**. Membrane proteins of arterivirus particles: structure, topology, processing and function. Virus Res 194:16-36.

12. **Wieringa R, de Vries AA, van der Meulen J, Godeke GJ, Onderwater JJ, van Tol H, Koerten HK, Mommaas AM, Snijder EJ, Rottier PJ. 2004.** Structural protein requirements in equine arteritis virus assembly. J Virol 78:13019-27.

13. **Wissink EH, Kroese MV, van Wijk HA, Rijsewijk FA, Meulenberg JJ, Rottier PJ. 2005.** Envelope protein requirements for the assembly of infectious virions of porcine reproductive and respiratory syndrome virus. J Virol 79:12495-506.

14. **Das PB, Dinh PX, Ansari IH, de Lima M, Osorio FA, Pattnaik AK. 2010.** The minor envelope glycoproteins GP2a and GP4 of porcine reproductive and respiratory syndrome virus interact with the receptor CD163. J Virol 84:1731-40.

15. **Welch SK, Calvert JG. 2010.** A brief review of CD163 and its role in PRRSV infection. Virus Res 154:98-103.

16. **Van Breedam W, Delputte PL, Van Gorp H, Misinzo G, Vanderheijden N, Duan X, Nauwynck HJ. 2010**. Porcine reproductive and respiratory syndrome virus entry into the porcine macrophage. J Gen Virol 91:1659–1667.

17. **Zhang Q, Yoo D. 2015.** PRRS virus receptors and their role for pathogenesis. Vet Microbiol 177:229-41.

18. **Vu HL, Kwon B, Yoon KJ, Laegreid WW, Pattnaik AK, Osorio FA. 2011**. Immune evasion of porcine reproductive and respiratory syndrome virus through glycan shielding involves both glycoprotein 5 as well as glycoprotein 3. J Virol 85:5555-64.

19. **Wei Z, Tian D, Sun L, Lin T, Gao F, Liu R, Tong G, Yuan S. 2012.** Influence of N-linked glycosylation of minor proteins of porcine reproductive and respiratory syndrome virus on infectious virus recovery and receptor interaction. Virology 429:1-11.

20. **Chen JZ, Wang Q, Bai Y, Wang B, Zhao HY, Peng JM, An TQ, Tian ZJ, Tong GZ. 2014**. Identification of two dominant linear epitopes on the GP3 protein of highly pathogenic porcine reproductive and respiratory syndrome virus (HP-PRRSV). Res Vet Sci 97:238-43.

21. **de Lima M, Pattnaik AK, Flores EF, Osorio FA. 2006.** Serologic marker candidates identified among B-cell linear epitopes of Nsp2 and structural proteins of a North American strain of porcine reproductive and respiratory syndrome virus. Virology 353:410-21.

22. **Oleksiewicz MB, Botner A, Normann P. 2002.** Porcine B-cells recognize epitopes that are conserved between the structural proteins of American- and European-type porcine reproductive and respiratory syndrome virus. J Gen Virol 83:1407-18.

23. **Oleksiewicz MB, Botner A, Toft P, Grubbe T, Nielsen J, Kamstrup S, Storgaard T. 2000.** Emergence of porcine reproductive and respiratory syndrome virus deletion mutants: correlation with the porcine antibody response to a hypervariable site in the ORF 3 structural glycoprotein. Virology 267:135-40.

24. **Oleksiewicz MB, Botner A, Toft P, Normann P, Storgaard T. 2001**. Epitope mapping porcine reproductive and respiratory syndrome virus by phage display: the nsp2 fragment of the replicase polyprotein contains a cluster of B-cell epitopes. J Virol 75:3277-90.

25. **Vanhee M, Van Breedam W, Costers S, Geldhof M, Noppe Y, Nauwynck H. 2011**. Characterization of antigenic regions in the porcine reproductive and respiratory syndrome virus by the use of peptide-specific serum antibodies. Vaccine 29:4794-804.

26. **Zhou YJ, An TQ, He YX, Liu JX, Qiu HJ, Wang YF, Tong G. 2006.** Antigenic structure analysis of glycosylated protein 3 of porcine reproductive and respiratory syndrome virus. Virus Res 118:98-104.

27. **Chen N, Trible BR, Kerrigan MA, Tian K, Rowland RR. 2016.** ORF5 of porcine reproductive and respiratory syndrome virus (PRRSV) is a target of diversifying selection as infection progresses from acute infection to virus rebound. Infect Genet Evol 40:167-75.

28. **Evans AB, Loyd H, Dunkelberger JR, van Tol S, Bolton MJ, Dorman KS, Dekkers JCM, Carpenter S. 2017**. Antigenic and Biological Characterization of ORF2-6 Variants at Early Times Following PRRSV Infection. Viruses 9.

29. **Wieringa R, de Vries AA, Rottier PJ. 2003**. Formation of disulfide-linked complexes between the three minor envelope glycoproteins (GP2b, GP3, and GP4) of equine arteritis virus. J Virol 77:6216-26.

30. **Faaberg KS, Plagemann PG. 1997.** ORF 3 of lactate dehydrogenase-elevating virus encodes a soluble, nonstructural, highly glycosylated, and antigenic protein. Virology 227:245-51.

31. **Mardassi H, Gonin P, Gagnon CA, Massie B, Dea S. 1998.** A subset of porcine reproductive and respiratory syndrome virus GP3 glycoprotein is

released into the culture medium of cells as a non-virion-associated and membrane-free (soluble) form. J Virol 72:6298-306.

32. **de Lima M, Ansari IH, Das PB, Ku BJ, Martinez-Lobo FJ, Pattnaik AK, Osorio FA. 2009.** GP3 is a structural component of the PRRSV type II (US) virion. Virology 390:31-6.

33. **van Nieuwstadt AP, Meulenberg JJ, van Essen-Zanbergen A, Petersen-den Besten A, Bende RJ, Moormann RJ, Wensvoort G. 1996.** Proteins encoded by open reading frames 3 and 4 of the genome of Lelystad virus (Arteriviridae) are structural proteins of the virion. J Virol 70:4767-72.

34. **Matczuk AK, Kunec D, Veit M. 2013**. Co-translational processing of glycoprotein 3 from equine arteritis virus: N-glycosylation adjacent to the signal peptide prevents cleavage. J Biol Chem 288:35396-405.

35. **Matczuk AK, Veit M. 2014.** Signal peptide cleavage from GP3 enabled by removal of adjacent glycosylation sites does not impair replication of equine arteritis virus in cell culture, but the hydrophobic C-terminus is essential. Virus Res 183:107-11.

36. **Zhang M, Veit M. 2017.** Differences in signal peptide processing between GP3 glycoproteins of Arteriviridae. Virology.

37. **Benfield DA, Nelson E, Collins JE, Harris L, Goyal SM, Robison D, Christianson WT, Morrison RB, Gorcyca D, Chladek D. 1992.** Characterization of swine infertility and respiratory syndrome (SIRS) virus (isolate ATCC VR-2332). J Vet Diagn Invest 4:127-33.

38. **Mardassi H, Athanassious R, Mounir S, Dea S. 1994.** Porcine reproductive and respiratory syndrome virus: morphological, biochemical and serological characteristics of Quebec isolates associated with acute and chronic outbreaks of porcine reproductive and respiratory syndrome. Can J Vet Res 58:55-64.

39. **Thaa B, Sinhadri BC, Tielesch C, Krause E, Veit M. 2013.** Signal peptide cleavage from GP5 of PRRSV: a minor fraction of molecules retains the decoy epitope, a presumed molecular cause for viral persistence. PLoS One 8:e65548.

40. **Zhang M, Cao Z, Xie J, Zhu W, Zhou P, Gu H, Sun L, Su S, Zhang G. 2013.** Mutagenesis analysis of porcine reproductive and respiratory syndrome virus nonstructural protein 7. Virus Genes 47:467-77.

41. **Dolnik O, Volchkova V, Garten W, Carbonnelle C, Becker S, Kahnt J, Stroher U, Klenk HD, Volchkov V. 2004**. Ectodomain shedding of the glycoprotein GP of Ebola virus. EMBO J 23:2175-84.

42. **Rapoport TA, Li L, Park E. 2017**. Structural and Mechanistic Insights into Protein Translocation. Annu Rev Cell Dev Biol.

43. **Zielinska DF, Gnad F, Wisniewski JR, Mann M. 2010.** Precision mapping of an in vivo N-glycoproteome reveals rigid topological and sequence constraints. Cell 141:897-907.

44. **Ellgaard L, Helenius A. 2003**. Quality control in the endoplasmic reticulum. Nat Rev Mol Cell Biol 4:181-91.

45. **Das PB, Vu HL, Dinh PX, Cooney JL, Kwon B, Osorio FA, Pattnaik AK. 2011.** Glycosylation of minor envelope glycoproteins of porcine reproductive

and respiratory syndrome virus in infectious virus recovery, receptor interaction, and immune response. Virology 410:385-94.

46. **Drin G, Antonny B. 2009**. Amphipathic helices and membrane curvature. FEBS Lett 584:1840-7.

47. **Sharpe HJ, Stevens TJ, Munro S. 2010**. A comprehensive comparison of transmembrane domains reveals organelle-specific properties. Cell 142:158-69.

48. **Eisenberg D, Weiss RM, Terwilliger TC. 1982**. The helical hydrophobic moment: a measure of the amphiphilicity of a helix. Nature 299:371-4.

49. **Rothman JE, Wieland FT. 1996**. Protein sorting by transport vesicles. Science 272:227-34.

50. **van Meer G, Voelker DR, Feigenson GW. 2008.** Membrane lipids: where they are and how they behave. Nat Rev Mol Cell Biol 9:112-24.

51. **Aberle D, Muhle-Goll C, Burck J, Wolf M, Reisser S, Luy B, Wenzel W, Ulrich AS, Meyers G. 2014**. Structure of the membrane anchor of pestivirus glycoprotein E(rns), a long tilted amphipathic helix. PLoS Pathog 10:e1003973.

52. **Fetzer C, Tews BA, Meyers G. 2005.** The carboxy-terminal sequence of the pestivirus glycoprotein E(rns) represents an unusual type of membrane anchor. J Virol 79:11901-13.

53. **Tews BA, Meyers G. 2007.** The pestivirus glycoprotein Erns is anchored in plane in the membrane via an amphipathic helix. J Biol Chem 282:32730-41.

Chapter 4

General discussion

4.1 Signal peptide cleavage of GP3 in different *Arteriviruses (Virology paper)*

The GP3 is one of three minor structural components (GP2-GP3-GP4) of arteriviruses and exists (at least in EAV) in a disulfide-linked trimer with GP2/4 (Wieringa et al., 2003). Our studies previously showed that the signal peptide is not cleaved from GP3 of EAV, although the bioinformatics tool SignalP 4.0 predicts signal peptide cleavage from GP3 of EAV with high probability. Signal peptide cleavage is inhibited by overlapping sequon NNTT (two glycosylation sites) located just downstream of the signal peptide in GP3 of EAV (Matczuk et al., 2013). In this thesis I showed that the signal peptide of GP3 from both PRRSV-1 and PRRSV-2 strains (and from LDV) is cleaved (Fig. 2.4 and Fig. 2.5). The process is not affected if a carbohydrate attached to a glycosylation site adjacent to the signal peptide is removed and the conserved Cys 33 is exchanged, which is in accordance with bioinformatic predictions (Fig. 2.1, Fig. 2.2 and 2.3). Thus, GP3 proteins from PPRSV and LDV are clearly differentially processed compared to GP3 of EAV.

What might be the reason why GP3 proteins from the same virus family are processed differentially? If OST and signal peptidase compete for neighboring sites it must be a subject of regulation whether signal peptidase or OST has privileged access to the growing polypeptide chain (Fig. 1.6). In the absence of regulation, i.e. if both enzymes would have random access to neighboring sites, a mixed protein population would be produced. In the case of GP3 of EAV a fraction of proteins would and the other would not contain the signal peptide, but this has never been observed in transfected cells (Matczuk et al., 2013). It is unexplored, how the access of OST and signal peptidase to a nascent protein chain in the ER lumen is regulated, but one might assume that the signal peptide (the first part of a growing polypeptide chain that contacts the translocon) selects whether OST or SPase is recruited at first to the translocon. It has been reported that different signal peptides interact with different binding sites within the translocon and that these differences can substantially affect protein biogenesis (Hegde and Bernstein, 2006).

Thus, we hypothesized that the signal peptide itself determines the processing scheme, but found that GP3 from EAV with the signal peptide of PRRSV and GP3 from PRRSV with the signal peptide of EAV behave like the corresponding wild type proteins. Therefore, it is rather an unknown element in the remaining part of the protein that determines whether the signal peptide is cleaved or affect the efficiency of signal peptide.

Are there certain features in their amino acid sequences that distinguish GP3 of EAV from GP3s of PRRSV and LDV? Multiple sequence alignment of GP3s revealed that GP3 from PRRSV and LDV are more similar to each other than to GP3 from EAV (**see supplementary Table 1**). Whereas the amino acid identity score between GP3s of PRRSV and LDV is 35%, GP3 from EAV has amino acid identity with GP3 from LDV and PRRSV of only 21% and 24%, respectively. The alignment also shows that the six cysteines residues of GP3 from PRRSV and LDV are at identical or very similar positions suggesting that both proteins exhibit a similar folding. Likewise, the five glycosylation sites in GP3 from LDV align also in a similar position as five (out of seven) sites in GP3 from PRRSV (Lelystad virus). In contrast, in GP3 from EAV only the first cysteine (outside the signal peptide) matches perfectly with the first cysteine in GP3 from PRRSV and LDV and the third cysteine of EAV´s GP3 aligns roughly with the sixth cysteine in GP3 from PRRSV and LDV. Thus, GP3 from EAV might have a peculiar structural element that is not present in GP3 from PRRSV and LDV and one might speculate that this might determine the unusual processing scheme. The closer relationship between PRRSV and LDV is also reflected in the recently improved taxonomy of Arteriviridae, where both viruses are grouped as one genus Rodartevirus, whereas EAV is the single member of the genus Equartevirus (Kuhn et al., 2016).

In addition, a bioinformatics tool (http://dgpred.cbr.su.se/) is frequently used to predict the insertion of protein segments from the hydrophilic pore of the translocon into the lipids of the ER membrane ((Hessa et al., 2007), Fig. 1.4). The apparent free energy difference (ΔGapp) is calculated based on hydrophobicity of individual amino acids, length of the segment and position of individual amino acids. In principle, a negative value of ΔGapp indicates that the sequence is predicted to be recognized as a transmembrane helix by the translocon and integrated into the membrane as an anchor. On the other hand, a positive value of ΔGapp predicts that this part of the polypeptide chain moves through the pore of the translocon into the ER lumen.

If this algorithm is applied for the predicted signal peptides of GP3 from the investigated Arteriviruses the results differ between GP3 proteins. For GP3 of EAV a negative ΔGapp is calculated, -0.441 or -0.152 depending on at which site signal peptide cleavage actually occurs. However, GP3 of EAV containing the signal peptide is secreted from transfected cells if the hydrophobic region is deleted, indicating that the signal peptide is not embedded within the membrane (Matczuk et al., 2013). It might be that glycosylation at the two sites adjacent to the signal peptide with large and hydrophilic carbohydrates pulls the signal peptide out of the membrane. In contrast, for the predicted signal peptides of GP3 from PRRSV-1 and PRRSV-2 a positive ΔGapp value is calculated suggesting that it does not act as a membrane-anchor even

if it would be present in the mature protein. However, note that the signal pepide cleavage sites are just predictions and cleavage might occur at other sites and this might significantly change the ΔGapp value.

Signal peptide sequence	Cleavage site	ΔG_{app}
EAV		
MGRAYSGPVALLCFFLYFCFICGSVG	G26/S27	-0.441
MGRAYSGPVALLCFFLYFCFICGSVGS	S27/N28	-0.152
PRRSV-1 (Lelystad)		
MAHQCARFHFFLCGFICYLVHSALA	A25/S26	+1.908
PRRSV-2 (VR-2332)		
MVNSCTFLHIFLCCSFLYSFCCA	A23/V24	+0.103

Table 4.1. The prediction of insertion of protein segments into the ER membrane. ΔGapp: the apparent free energy difference (http://dgpred.cbr.su.se/). In principle, a negative value of ΔGapp indicates that the sequence is predicted to be recognized as a transsmembrane helix by the translocon and integrated into the membrane. Cleavage site: the predicted signal peptide cleavage site. EAV: equine arteritis virus (Bucyrus strain); PRRSV-1(Lelystad): porcine reproductive and respiratory syndrome virus of type 1; RRRSV-2 (VR-2332): porcine reproductive and respiratory syndrome virus of type 2.

4.2 The membrane topology and secretion of GP3 of *PRRSV (J. Virology paper)*

Inconsistent results have been published about whether GP3s are structural proteins of Arterivirus particles. Here I systematically investigated whether GP3 proteins from PRRSV-1 and PRRSV-2 strains are secreted from transfected cells. I found that a fraction of the GP3 from variable strains is secreted from transfected cells; GP3 from PRRSV-1 strains to a greater extent than GP3 from PRRSV-2 strains (Fig. 3.1A). This secretion behavior is reversed after exchange of the variable C-terminal domain (Fig. 3.9). I also propose that GP3 exhibits an unusual hairpin-like membrane topology (Fig. 3.10C). I show that the C-terminal part of GP3 is exposed to the lumen of the ER since it is protected against proteolytic digestion in perforated cells (Fig. 3.3B) and since introduced N-glycosylation sites are efficiently used (Fig. 3.3C and 3D). Membrane anchoring is achieved by a short hydrophobic region that might form an amphipathic helix (Fig. 3.4B +C, Fig. 3.7A and Fig. 3.8A). Accordingly, exchanging only a few amino acids in its hydrophilic face prevents (Fig. 3.8) and in its hydrophobic face enhances secretion of GP3 (Fig. 3.7).

The essential nature of the hydrophobic region is also corroborated by the notion that it is highly conserved between GP3 proteins from all PRRSV-1 and -2 strains (Fig. 3.3A, Fig. 3.6A). A predicted amphipathic helix with a very similar sequence is also present in GP3 of LDV (YIRPLFSSWLVLNVSYFL) and it was reported that a fraction of the protein is secreted from cells and that it can be extracted from membranes without detergent suggesting that it is a peripheral membrane protein (Faaberg and Plagemann, 1997). GP3 from EAV also adopts a hairpin-like topology and is membrane-anchored by a C-terminal hydrophobic domain (Matczuk et al., 2013). This region (RPTLICWFALLLVHFLPMPRCRGS) exhibits no sequence homology to the presumed helices of PRRSV and LDV, but heliquest predicts that it forms an amphipathic helix with similar biophysical properties to the helices of PRRSV and LDV. Similar to PRRSV, membrane-anchoring of GP3 of EAV is also essential for virus replication since infectious mutants with a stop codon inserted into this region could not be generated (Matczuk and Veit, 2014). In summary, although experimentally not analyzed for each protein, GP3 proteins from PRRSV, LDV and EAV might have the same hairpin-like membrane topology and the same mechanism of membrane-anchoring, but the presumed amphipathic helix is formed by different amino acids in the genera Rodartevirus (LDV and PRRSV) and Equartevirus (EAV).

Although they contain an almost identical hydrophobic region, a larger fraction of GP3 from the PRRSV-1 strains is secreted from transfected cells compared to GP3 from the PRRSV-2 strains (Fig. 3.1A). Instead we have shown that the variable C-terminus modulates how much of GP3 is secreted. The C-terminal part is strongly hydrophilic in the PRRSV-1 strain Lelystad, but rather hydrophobic in the PRRSV-2 strain XH-GD (Fig. 3.9A). Exchanging the C-terminus between both GP3 proteins completely reversed this secretion behavior. GP3 from Lelystad with the C-terminal domain of XH-GD is secreted in lower amounts than the corresponding wild-type protein (Fig. 3.9D), whereas GP3 from XH-GD with the C-terminal domain of Lelystad is secreted in higher amounts (Fig. 3.9E). Note, however, that the C-terminus is the most variable part of GP3 (Fig. 3.3A, the sequence also reveals insertions or deletions) and is thus remains to be shown whether increased GP3 secretion is a general feature of all PRRSV-1 strains. However, the secreted form of GP3 is likely to be a folded (and hence functional) protein, since it passed the quality control system of the ER (Ellgaard and Helenius, 2003). Furthermore, since a large fraction of secreted (but not intracellular) GP3 is a disulfide-linked dimer, at least one intermolecular disulfide bond is formed which would be unlikely to happen in a misfolded protein (Fig. 3.1G).

Hydrophobic region	ΔG_{app}
EAV	
RPTLICWFALLLVHFLPMPRCRGS	+0.751
LDV	
YIRPLFSSWLVLNVSYFL	+2.090
PRRSV-1 (Lelystad)	
WLRPLFSSWLVLNISWFL	+1.291
PRRSV-2 (VR-2332)	
WLRPFFSSWLVLNVSWFL	+1.719
PRRSV-2 2A	
WLRPFFSSWLV**A**NVS**A**FL	+2.699
PRRSV-2 3A	
W**A**RP**A**FSS**A**LVLNVSWFL	+2.229
PRRSV-2 5A	
W**A**RP**A**FSS**A**LV**A**NVS**A**FL	+3.301
PRRSV-2 3H	
WLRPFFSSWLVL**S**V**LL**FL	+0.225
PRRSV-2 4H	
WL**LL**FF**FF**WLVLNVSWFL	-1.998
PRRSV-2 7H	
WL**LL**FF**FF**WLVL**S**V**LL**FL	-3.353

Table 4.2. The prediction of insertion of hydrophobic region into the ER membrane. ΔGapp: the apparent free energy difference (http://dgpred.cbr.su.se/). In principle, a negative value of ΔGapp indicates that the sequence is predicted to be recognized as a transmembrane helix by the translocon and integrated into the membrane. EAV: equine arteritis virus (Bucyrus strain); LDV: lactate dehydrogenase-elevating virus (Plagemann strain); PRRSV-1(Lelystad): porcine reproductive and respiratory syndrome virus of type 1; RRRSV-2 (VR-2332): porcine reproductive and respiratory syndrome virus of type 2; RRRSV-2 2A, 3A and 5A: mutated GP3s with increased secretion; PRRSV-2 3H, 4H and 7H: mutated GP3 that is not secreted.

Moreover, the bioinformatics tool (http://dgpred.cbr.su.se/) was also employed to predict whether the hydrophobic region of GP3 from *Arterivirus* is laterally inserted into the ER membrane or whether it is first translocated through the hydrophilic pore of the translocon and attaches with the ER membrane after translocation (Table 4.2). The tool predicts a pronounced positive value for the free energy ΔGapp of the hydrophobic

region of GP3 from EAV, PRRSV-1 and PRRSV-2 and LDV suggesting that it is not able to open the lateral gate of the translocon and thus translocated into the lumen of the ER. As expected ΔGapp becomes even more positive if hydrophobic amino acids are exchanged by the less hydrophobic residue alanine (mutants 2A, 3A and 5A of GP3 from PRRSV-2, that showed enhanced secretion in transfected cells). If the hydrophobicity is decreased by exchange of hydrophobic with hydrophilic residues in the mutants PRRSV-2 3H, 4H and 7H it causes a substantial decrease in ΔGapp. For 4H and 7H it becomes negative which suggests that in those mutants the hydrophobic region is released from the translocon into the lipids of the ER membrane and thus should in principle act as a transmembrane anchor, i.e. the C-terminus of the protein remains in the cytosol. However, this was not the case; A glycosylation site inserted into the C-terminal part of GP3 downstream of the hydrophobic region was efficiently used (Fig 3.8D) providing evidence that the C-terminus was translocated into ER lumen and thus the hydrophobic region of mutant GP3 is not a transmembrane anchor. The only other viral structural glycoprotein having a similar membrane anchor is the Erns glycoprotein of Pestiviruses, an unspecific RNAse that suppresses the cellular innate immune response and is involved in the establishment of persistent infection. However, the amphiphilic helix of Erns is longer (60 amino acids) and located at the extreme C-terminus of the protein. Another similarity to GP3 is the dual nature of Erns. Erns is an essential part of the viral envelope, but is also secreted into the extracellular medium and found in the serum of infected animals. This secreted form is hypothesized to relate to its role as a virulence factor (Aberle et al., 2014; Fetzer et al., 2005; Tews and Meyers, 2007).

We speculate that secreted GP3 might also play a role during PRRSV infection of pigs, for example as a decoy that serves to distract antibodies away from virus particles. Indeed, antibodies against various regions of the GP3 protein were elicited in the majority of experimentally infected piglets, but most of them had no or only little neutralizing activity (Chen et al., 2014; de Lima et al., 2006; Oleksiewicz et al., 2002; Oleksiewicz et al., 2000; Oleksiewicz et al., 2001; Vanhee et al., 2011; Zhou et al., 2006).

4.3 The distribution of antibody epitopes in GP3 of PRRSV

Virus-specific antibodies are produced after one or two weeks upon PRRSV infection. However, these antibodies are not able to decrease virus replication *in vitro* in primary porcine alveolar macrophages (PAM). Neutralizing antibodies appear only later (three

to four weeks post infection), and titers are usually low, which is probably too late and too little to influence the acute phase of viremia. Despite this weak virus neutralizing antibody response, it is known that the presence of sufficient amounts of virus-neutralizing antibodies at the onset of infection can offer certain protection against virus replication in the lungs, viremia and transplacental spread of the virus, indicating that PRRSV-specific antibodies can contribute to protective immunity.

The antigenic regions (AR) are potential inducers of virus-neutralizing antibodies, therefore exploring its location on the membrane proteins is one interesting topic in PRRSV research. Antigenic characterization of PRRSV by the use of mouse monoclonal antibodies has led to the discovery of neutralizing epitopes in GP4 of PRRSV-1 and GP5 of both PRRSV-1 and PRRSV-2 strains, and it has been suggested that also M and GP3 can act as targets for neutralizing antibodies (Cancel-Tirado et al., 2004; Costers et al., 2010; Meulenberg et al., 1997; Ostrowski et al., 2002; Wissink et al., 2003). Although GP3 antibodies are present at very low levels in pigs, they have been reported to play a role in clearing the viral infection (Plana Duran et al., 1997) and may be involved in viral neutralization (Cancel-Tirado et al., 2004). Using a phage display library or pepscan technology several epitopes containing amino acid regions 53-116, 141-152, 181-192 and 243-260 were identified in GP3 of PRRSV-1 (Oleksiewicz et al., 2002; Vanhee et al., 2011). Among them, the region 57-68 is especially described as an neutralizing epitope (Vanhee et al., 2011). Four B-cell linear epitopes in GP3 of PRRSV-2 virus were also identified, which are located at amino acids between 61-105 (de Lima et al., 2006), some epitopes among them or partly overlapping regions were also found in other PRRSV-2 strain (Chen et al., 2014; Zhou et al., 2006).

Our studies showed that the hydrophobic region at the C-terminus of GP3 does not span the ER-membrane, suggesting that the C-terminus of the protein is located in the lumen of the ER. Thus, all identified epitopes are located in the ectodomain of GP3 (and thus at the surface of virus particles), with the exception of region 181-192, as it is predicted to form an amphipathic helix anchored in the membrane. Among these epitopes, the region 57-68 is described in both types PRRSV, but the amino acid conservation of this region is quite low, at an average 79.55%.

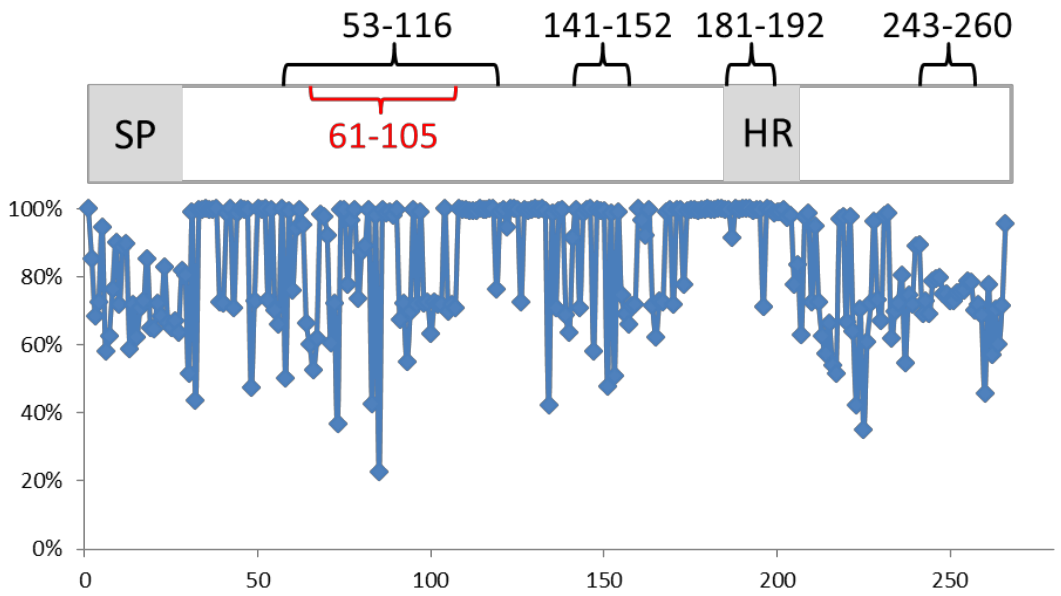

Fig 4.1 Primary structure and sequence comparison of GP3 from PRRSV. The antigenic regions (black brackets) were reported in GP3 from PRRSV-1 strains, including region 53-116, 141-152, 181-192 and 243-260, antigenic regions 61-105 (red bracket) in GP3 from PRRSV-2 strains. The graph shows the percentage conservation (Y-axis) of amino acids at each position (X-axis) of a consensus sequence compiled from all PRRSV-1 and PRRSV-2 GP3 sequences present in the database.

4.4 Final remarks and outlook

We speculate that secreted GP3 might play a role during PRRSV infection of pigs, for example as a decoy that serves to distract antibodies away from virus particles. To prove this idea, I am planning to generate mutant virus containing GP3 with either increased or decreased secretion. The aim is to investigate whether these mutations will influence the biological characteristics and pathogenicity of the virus and/or the immune response of the host. Since the C-terminus of the GP3 gene (including the hydrophobic region) is overlapping with the N-terminus of the GP4 gene the number of mutations that can be performed without altering amino acids of GP4 are limited. Exchange of hydrophobic by hydrophilic amino acids in the hydrophobic region enhanced secretion of GP3 in transfected cells (Fig. 3.7), but the resulting virus-like particles were not infectious (Fig. 3.10). The reverse type of experiment, i.e. exchanging hydrophilic amino acids by hydrophobic ones prevented secretion of GP3 from transfected cells (Fig. 3.8) and currently I am analyzing whether infectious virus particles with these mutations can be generated. To overcome the limitations associated with the overlapping genes of GP3 and GP4 I want to untangle the respective ORFs in the PRRSV genome. If this is feasible much more mutations in the

hydrophobic region are possible since they do not affect the functionality of GP4. For EAV it has been shown that such constructs are viable, and viral titers only slightly reduced compared to wild-type EAV (de Vries et al., 2000).

4.5 Reference

Aberle, D., Muhle-Goll, C., Burck, J., Wolf, M., Reisser, S., Luy, B., Wenzel, W., Ulrich, A.S., Meyers, G., 2014. Structure of the membrane anchor of pestivirus glycoprotein E(rns), a long tilted amphipathic helix. PLoS Pathog 10, e1003973.

Cancel-Tirado, S.M., Evans, R.B., Yoon, K.J., 2004. Monoclonal antibody analysis of porcine reproductive and respiratory syndrome virus epitopes associated with antibody-dependent enhancement and neutralization of virus infection. Vet Immunol Immunopathol 102, 249-262.

Chen, J.Z., Wang, Q., Bai, Y., Wang, B., Zhao, H.Y., Peng, J.M., An, T.Q., Tian, Z.J., Tong, G.Z., 2014. Identification of two dominant linear epitopes on the GP3 protein of highly pathogenic porcine reproductive and respiratory syndrome virus (HP-PRRSV). Res Vet Sci 97, 238-243.

Costers, S., Lefebvre, D.J., Van Doorsselaere, J., Vanhee, M., Delputte, P.L., Nauwynck, H.J., 2010. GP4 of porcine reproductive and respiratory syndrome virus contains a neutralizing epitope that is susceptible to immunoselection in vitro. Arch Virol 155, 371-378.

de Lima, M., Pattnaik, A.K., Flores, E.F., Osorio, F.A., 2006. Serologic marker candidates identified among B-cell linear epitopes of Nsp2 and structural proteins of a North American strain of porcine reproductive and respiratory syndrome virus. Virology 353, 410-421.

de Vries, A.A., Glaser, A.L., Raamsman, M.J., de Haan, C.A., Sarnataro, S., Godeke, G.J., Rottier, P.J., 2000. Genetic manipulation of equine arteritis virus using full-length cDNA clones: separation of overlapping genes and expression of a foreign epitope. Virology 270, 84-97.

Ellgaard, L., Helenius, A., 2003. Quality control in the endoplasmic reticulum. Nat Rev Mol Cell Biol 4, 181-191.

Faaberg, K.S., Plagemann, P.G., 1997. ORF 3 of lactate dehydrogenase-elevating virus encodes a soluble, nonstructural, highly glycosylated, and antigenic protein. Virology 227, 245-251.

Fetzer, C., Tews, B.A., Meyers, G., 2005. The carboxy-terminal sequence of the pestivirus glycoprotein E(rns) represents an unusual type of membrane anchor. J Virol 79, 11901-11913.

Hessa, T., Meindl-Beinker, N.M., Bernsel, A., Kim, H., Sato, Y., Lerch-Bader, M., Nilsson, I., White, S.H., von Heijne, G., 2007. Molecular code for transmembrane-helix recognition by the Sec61 translocon. Nature 450, 1026-U1022.

Kuhn, J.H., Lauck, M., Bailey, A.L., Shchetinin, A.M., Vishnevskaya, T.V., Bao, Y., Ng, T.F., LeBreton, M., Schneider, B.S., Gillis, A., Tamoufe, U., Diffo Jle, D., Takuo, J.M., Kondov, N.O., Coffey, L.L., Wolfe, N.D., Delwart, E., Clawson, A.N., Postnikova, E., Bollinger, L., Lackemeyer, M.G., Radoshitzky, S.R., Palacios, G., Wada, J., Shevtsova, Z.V., Jahrling, P.B., Lapin, B.A., Deriabin, P.G., Dunowska, M., Alkhovsky, S.V., Rogers, J., Friedrich, T.C., O'Connor, D.H., Goldberg, T.L., 2016. Reorganization and expansion of the nidoviral family Arteriviridae. Arch Virol 161, 755-768.

Matczuk, A.K., Kunec, D., Veit, M., 2013. Co-translational processing of glycoprotein 3 from equine arteritis virus: N-glycosylation adjacent to the signal peptide prevents cleavage. J Biol Chem 288, 35396-35405.

Matczuk, A.K., Veit, M., 2014. Signal peptide cleavage from GP3 enabled by removal of adjacent glycosylation sites does not impair replication of equine arteritis virus in cell culture, but the hydrophobic C-terminus is essential. Virus Res 183, 107-111.

Meulenberg, J.J., van Nieuwstadt, A.P., van Essen-Zandbergen, A., Langeveld, J.P., 1997. Posttranslational processing and identification of a neutralization domain of the GP4 protein encoded by ORF4 of Lelystad virus. J Virol 71, 6061-6067.

Oleksiewicz, M.B., Botner, A., Normann, P., 2002. Porcine B-cells recognize epitopes that are conserved between the structural proteins of American- and European-type porcine reproductive and respiratory syndrome virus. J Gen Virol 83, 1407-1418.

Oleksiewicz, M.B., Botner, A., Toft, P., Grubbe, T., Nielsen, J., Kamstrup, S., Storgaard, T., 2000. Emergence of porcine reproductive and respiratory syndrome virus deletion mutants: correlation with the porcine antibody response to a hypervariable site in the ORF 3 structural glycoprotein. Virology 267, 135-140.

Oleksiewicz, M.B., Botner, A., Toft, P., Normann, P., Storgaard, T., 2001. Epitope mapping porcine reproductive and respiratory syndrome virus by phage display: the nsp2 fragment of the replicase polyprotein contains a cluster of B-cell epitopes. J Virol 75, 3277-3290.

Ostrowski, M., Galeota, J.A., Jar, A.M., Platt, K.B., Osorio, F.A., Lopez, O.J., 2002. Identification of neutralizing and nonneutralizing epitopes in the porcine reproductive and respiratory syndrome virus GP5 ectodomain. J Virol 76, 4241-4250.

Plana Duran, J., Climent, I., Sarraseca, J., Urniza, A., Cortes, E., Vela, C., Casal, J.I., 1997. Baculovirus expression of proteins of porcine reproductive and respiratory syndrome virus strain Olot/91. Involvement of ORF3 and ORF5 proteins in protection. Virus Genes 14, 19-29.

Tews, B.A., Meyers, G., 2007. The pestivirus glycoprotein Erns is anchored in plane in the membrane via an amphipathic helix. J Biol Chem 282, 32730-32741.

Vanhee, M., Van Breedam, W., Costers, S., Geldhof, M., Noppe, Y., Nauwynck, H., 2011. Characterization of antigenic regions in the porcine reproductive and respiratory syndrome virus by the use of peptide-specific serum antibodies. Vaccine 29, 4794-4804.

Wieringa, R., de Vries, A.A., Rottier, P.J., 2003. Formation of disulfide-linked complexes between the three minor envelope glycoproteins (GP2b, GP3, and GP4) of equine arteritis virus. J Virol 77, 6216-6226.

Wissink, E.H., van Wijk, H.A., Kroese, M.V., Weiland, E., Meulenberg, J.J., Rottier, P.J., van Rijn, P.A., 2003. The major envelope protein, GP5, of a European porcine reproductive and respiratory syndrome virus contains a neutralization epitope in its N-terminal ectodomain. J Gen Virol 84, 1535-1543.

Zhou, Y.J., An, T.Q., He, Y.X., Liu, J.X., Qiu, H.J., Wang, Y.F., Tong, G., 2006. Antigenic structure analysis of glycosylated protein 3 of porcine reproductive and respiratory syndrome virus. Virus Res 118, 98-104.

Zusammenfassung

Membrantopology und Prozessierung des Glykoproteins 3 des Porzinen Reproduktiven und Respiratorischem Syndrom-Virus

Das Porzine Reproduktive und Respiratorische Syndrom-Virus (PRRSV) ist einer der bedeutensten Infektionskrankheiten bei Schweinen. PRRSV infiziert Schweine jeden Alters, verursacht Fehl- und Totgeburten bei Säuen und Atemwegserkrankungen bei Ferkeln. In den meisten Fällen sind die Symptome mild, allerdings bei deutlich reduzierter Gewichtszunahme. Dies führt weltweit zu erheblichen finanziellen Einbußen in der Schweinemast. In China sind sogar hoch-pathogene Stämme aufgetaucht, die über 90% der infizierten Schweine töteten. Bisher sind Impfstoffe daran gescheitert, das Virus zu eliminieren. Eine Ursache hierfür sind die starken Variationen zwischen den Stämmen und die Fähigkeit dem Immunsystem des Wirtes zu entkommen.

Das Glykoprotein GP3 besteht aus einem N-terminalen Signalpeptid, einer 180 aminosäuren-langen stark glykosylierten Domäne, einem konservierten hydrophoben Bereich (20 Aminosäuren) und einer variablen nicht-glykosylierten C-terminalen Domäne (50-60 Aminosäuren). GP3 formt vermutlich im Viruspartikel einen Komplex mit zwei weiteren Glykoproteinen (GP2 und GP4); es wurde jedoch auch berichtet, dass infizierte Zellen GP3 sekretieren.

In dieser Arbeit analysierte ich die Membrantopologie von GP3 aus PRRSV Typ-1 und Typ-2 Stämmen. Als Erstes habe ich entdeckt, dass das N-terminale Signalpeptid von PRRSV-GP3 (ebenso für GP3 vom laktatdehydrogenase-erhöhenden Virus) abgeschnitten wird, unabhängig davon, ob in unmittelbarer Nähe eine Kohlenhydratkette an das Protein gehängt wurde. Dieses Ergebnis unterscheidet sich vom GP3 des Equinen Arteritis-Virus, bei dem eine Kohlenhydratkette an ähnlicher Position die Prozessierung des Signalpeptids verhindert.

Zum Zweiten, habe ich bestätigt, dass ein Anteil des GP3 (Wildtyp) von transfizierten Zellen sekretiert wird; GP3 von PRRSV-1 Stämmen (Lelystad, Lena) wird hierbei deutlich stärker sekretiert als GP3 von PRRSV-2 Stämmen (VR2332, IAF-Klop, XH-GD). Der Effekt dreht sich um, wenn die variable C-terminalen Domäne entsprechend ausgetauscht wird. Im Unterschied zu intrazellulärem GP3 weist sekretiertes komplex-prozessierte Kohlenhydratketten auf. Dies spricht dafür, dass diese Moleküle den

sekretorischen Pfad in der Zelle komplett durchlaufen haben. Da intrazelluläres und sekretiertes GP3 nach der de-Glykosylierung die gleiche Wanderungsgeschwindigkeit in denaturierenden Proteingelen zeigten, wurde die sekretierte Form nicht proteolytisch gespalten.

Als Nächstes verwendete ich einen über Fluoreszenz nachgewiesenen Protease-Zugangs-Versuch, um zu zeigen, dass der C-Terminus von GP3 verknüpft mit GFP in permeabilisierten Zellen gegen proteolytischen Verdau geschützt ist. Weiterhin habe ich nachgewiesen, dass in den C-Terminus eingefügte Glykosylierungsstellen genutzt werden. Beide Experimente deuten darauf hin, dass der C-terminale Teil des GP3's ins Lumen des Endoplasmatischen Retikulums transloziert wird.

Die Deletion der hydrophoben Region verstärkte die Sekretion von GP3; die Deletion des variablen C-Terminus zeigte diesen Effekt nicht. Weiterhin führte die Verknüpfung der hydrophoben Region von GP3 mit GFP zur kompletten Membranbindung dieses üblicherweise löslichen Proteins. Bioinformatische Analysen sagen voraus, dass die hydrophobe Region eine amphipatische Helix bilden könnte. Dem entsprechend verhinderte der Austausch von einigen Aminosäuren im hydrophilen Bereich die Sekretion von GP3. Der Austausch dieser Aminosäuren im viralen Genom beeinflusste nicht die Freisetzung von Viruspartikeln, allerdings waren diese nicht infektiös. Dieser Befund stimmt mit der GP3 zugewiesenen Rolle beim Virus-Eintritt in die Zelle überein.

Zusammengefasst: Gp3 weist eine ungewöhnliche Haarnadel-artige Membrantopology auf. Das Signalpeptid wird abgespalten und der C-Terminus befindet sich im Lumen des Endoplasmatischen Retikulums. Die Membranbindung wird durch eine kurze hydrophobe Region gewährleistet, die wahrscheinlich eine amphipatische Helix bildet. Diese eher schwache Membranverankerung könnte erklären, warum ein Teil dieses Proteins sekretiert wird. Wir spekulieren, dass sekretiertes GP3 als "Köder" benutzt wird, der Antikörper von den Viruspartikeln ablenken könnte.

Summary

Membrane Topology and Processing of the Glycoprotein 3 of Porcine Reproductive and Respiratory Syndrome Virus

The porcine reproductive and respiratory syndrome virus (PRRSV) causes one of the most important infectious disease of pigs. PRRSV infects pigs of all ages, where it causes reproductive failure in sows and respiratory problems in piglets. Usually, symptoms are mild, but lead to reduced weight gain, which causes huge financial losses in the pork industry worldwide. In China even highly pathogenic strains emerged that kill 90% of infected pigs. So far, vaccines failed to eliminate the virus, which is due to the large variation between strains and their ability to escape the immunity of the host.

The glycoprotein GP3 consists of an N-terminal signal peptide, a 180 amino acids long and highly glycosylated domain, a hydrophobic conserved region (20 aa) and a variable unglycosylated C-terminal domain (50-60 aa). GP3 is supposed to form a complex with two other glycoproteins (GP2 and GP4) in virus particles, but secretion of the protein from infected cells has also been reported.

Here I analyzed the membrane topology of GP3 from type-1 and -2 PRRSV strains. First, Ifound that the N-terminal signal peptide of GP3 (and also from lactate dehydrogenase-elevating virus) is cleaved despite the presence of a carbohydrate in its vicinity. This is in contrast to GP3 of equine arteritis virus where a carbohydrate attached at a similar position prevents processing.

Second, I confirmed that a fraction of wild-type GP3 is secreted from transfected cells; GP3 from PRRSV-1 strains (Lelystad, Lena) to a greater extent than GP3 from PRRSV-2 strains (VR-2332, IAF-Klop, XH-GD). This secretion behavior is reversed after exchange of the variable C-terminal domain. In contrast to intracellular GP3, secreted GP3 contains complex-type carbohydrates, indicating that it passed through the secretory pathway. Since intracellular and secreted GP3 have identical SDS-PAGE mobility after deglycosylation, the secreted form is not derived from proteolytic cleavage.

Next I used a fluorescence protease protection assay to show that the C terminus of GP3, fused to GFP, is resistant against proteolytic digestion in permeabilized cells. Furthermore, glycosylation sites inserted into the C-terminal part of GP3 are used. Both

experiments indicate that the C-terminal part of GP3 is translocated into the lumen of the endoplasmic reticulum.

Deletion of the conserved hydrophobic region, but not of the variable C-terminus greatly enhances secretion of GP3. In addition, fusion of the hydrophobic region of GP3 to GFP promotes complete membrane anchorage of this (otherwise soluble) protein. Bioinformatics suggests that the hydrophobic region might form an amphipathic helix. Accordingly, exchanging only a few amino acids in its hydrophilic face prevents and in its hydrophobic face enhances secretion of GP3. Exchanging the latter amino acids in the context of the viral genome did not affect release of virions, but released particles were not infectious. This is consistent with the proposed role of GP3 in virus entry.

In sum, GP3 exhibits a very unusual hairpin-like membrane topology. The signal peptide is cleaved and the C-terminus is exposed to the lumen of the ER. Membrane attachment is caused by a short hydrophobic region, which might form an amphiphilic helix. This rather weak membrane anchoring might explain why a fraction of the protein is secreted. We speculate that secreted GP3 might function as a “decoy”, which distracts antibodies away from virus particles.

Publications

Zhang M, Veit M. Differences in signal peptide processing between GP3 glycoproteins of Arteriviridae. **Virology**, 2018 Apr; 517:69-76.

Minze Zhang, Ludwig Krabben, Fangkun Wang, and Michael Veit. Glycoprotein 3 of porcine reproductive and respiratory syndrome virus exhibits an unusual hairpin-like membrane topology, **Journal of Virology**, 2018 May 16. pii: JVI.00660-18.

Pradip Dey, Tobias Bergmann, Jose L. Cuellar-Camacho, Svenja Ehrmann, Mohammad S. Chowdhury, **Minze Zhang**, Rainer Haag, Walid Azab. Multivalent flexible-Nanogels exhibit broad-spectrum antiviral activity by blocking virus entry, resubmitted. **ACS Nano**. 2018 Jun 19. doi: 10.1021/acsnano.8b01616.

MZ Zhang, JX Xie, L Sun, ZP Cao, HL Gu, SC Deng, et al., Phylogenetic analysis and molecular characteristics of porcine reproductive and respiratory syndrome virus isolates in Southern China. **Microbial Pathogenesis** (2013).65:67-72.

MZ Zhang, ZP Cao, JX Xie ,WJ Zhu ,P Zhou, HL Gu,et al. (2013) Mutagenesis analysis of porcine reproductive and respiratory syndrome virus nonstructural protein 7. **Virus Genes**. 47:467–477

Wei C, **Zhang M**, Chen Y, Xie J, Huang Z, Zhu W, Xu T, Cao Z, Zhou P, Su S, Zhang G. Genetic evolution and phylogenetic analysis of porcine circovirus type 2 infections in southern China from 2011 to 2012. **Infect Genet Evol.** 2013. Apr 4.

Acknowledgements

There are always some or a few decision, which are important to one's life, to pursue phD in this lab is one in my life. It grants me not only academic achievement even accompany with difficulties and depression within this journey, but also gives me chance to meet a number of amazing people. Without their help and support I can't get this point.

First, I would like to express deep and sincere appreciation to my supervisor PD. Dr. Michael Veit, for his continuing encouragement, always patience to listen my problems and find a solution, also his enthusiasm through my research program of naughty GP3 of PRRSV. He always gives his maximum support to my academic participation and performance, such as writing an attractive recommendation letter for application of travel funding, giving a constructive comment to my talk. I would say this wonderful experience working with him would deeply forever influence my future work.

I am very grateful to Prof. Dr. Nikolaus Osterrieder and Dr. Karsten Tedin (from Institut für Mikrobiologie und Tierseuchen) for their great support, especially constructive suggestion, comments and guidance. It is really my honor and luck to have them in my mentor committee.

I would also like to thank my colleague Dr. Ludwig Krabben for his interests on my project, often discussion with offering valuable suggestion, also his help for performing one of tricky experiments in my paper and translation of summary of the my thesis into german version.

Moreover, I wish to extend my gratitude to all staff members, Dr. Susanne Kaufer, Dr. Chris Höfer, phD student Bodan Hu, Xuejiao Han, Mohamed Rasheed Gadalla, Atika Hadiati and technician Elke Dyrks, Angelika Thomele from Veit's group, Prof. Benedikt Kaufer, PD Dr. Kerstin Borchers, Dr. Walid Azab, Dr. Dusan Kunec, Dr. Darren J. Wight, Dr. Ahmed Kheimar, Dr. Tobias Bergmann, phD student Luca Danilo Bertzbach, Anirban Sanyal and others from Institut für Virologie. They are so friendly and helpful, i am happy be part of this big family of Virology.

Thanks to China Scholarship Council (CSC) for their support. Without this I will not be possible to afford chance to study here.

Particularly, I am deeply indebted to my parents and my sister for their support from spirit and economics. Last, I deeply thank my beloved wife Fengfeng Tang and my

lovely son Ruiyang Zhang, for their endless love and support, for bringing me strength and belief that everything will be ok.

Selbständigkeitserklärung

Hiermit bestätige ich, dass ich die vorliegende Arbeit selbständig angefertigt habe. Ich versichere, dass ich ausschließlich die angegebenen Quellen und Hilfen Anspruch genommen habe.

Berlin, den 16.08.2018

Minze Zhang

Druck: Mensch und Buch Verlag Berlin