Virology xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/virology

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

ARTICLE INFO

Keywords:
Arteriviridae
EAV
PRRSV
LDV
GP3
Co-translational protein processing
Signal peptide cleavage
N-glycosylation
Disulphide-linkage

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.

1. 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 non-structural, 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

* Correspondence to: Freie Universität Berlin, Fachbereich Veterinärmedizin, Institut für Virologie Robert-von-Ostertag-Straße 7–13, DE-14163 Berlin, Germany.
E-mail address: michael.veit@fu-berlin.de (M. Veit).

https://doi.org/10.1016/j.virol.2017.11.026
Received 6 September 2017; Received in revised form 27 November 2017; Accepted 30 November 2017

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. Material and methods

2.1. 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