Aus dem Institut für Virologie
des Fachbereichs Veterinärmedizin
der Freien Universität Berlin
In Kooperation mit dem Friedrich-Loeffler Institut (FLI)
Bundesforschungsinstitut für Tiergesundheit
Institut für Immunologie (IFI), Greifswald -Insel Riems

Functional and structural analysis of pestivirus E1 glycoprotein

Inaugural-Dissertation
zur Erlangung des Grades eines
Doktors der Veterinärmedizin
an der
Freien Universität Berlin

vorgelegt von

Yu Mu
Tierärzt
aus Heilongjiang, Volksrepublik China

Berlin 2021

Journal-Nr.:4258

Gedruckt mit Genehmigung
des Fachbereichs Veterinärmedizin
der Freien Universität Berlin

Dekan: Univ.-Prof. Dr. Jürgen Zentek

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

Zweiter Gutachter: Prof. Dr. Gregor Meyers

Dritter Gutachter: PD Dr. rer. nat. Michael Veit

Deskriptoren: pestivirus, viral morphology, glycoproteins, genetic analysis, polymerase chain reaction

Tag der Promotion: 19.04.2021

To My Beloved Family

Ode to the Plum Blossom

Tune: 'Song of Divination'

December 1961

Then spring departed in wind and rain;
With flying snow it's back again.
Though the cliff is covered with hundreds of meters high ice,
still there is a flowery twig nice.
Though sweet and fair, with other flowers she won't rival,
But only heralds spring's arrival.
When mountain flowers cover all the hills,
She smiles amongst them still.

Table of contents

Table of contents.....I
List of Figures..... V
List of Tables.....VII
Abbreviations..... VIII
Chapter 1: General introduction..... 1
1.1 Taxonomy, host and disease..... 1
1.1.1 Taxonomy..... 1
1.1.2 Host..... 2
1.1.3 Disease..... 3
1.2 Viral particle and genome..... 3
1.3 The viral life cycle..... 4
1.4 Envelope proteins..... 6
1.5 The conventional protein trafficking pathway..... 8
1.5.1 Translocation at the ER site..... 8
1.5.2 Signal sequence and topogenesis of membrane proteins..... 9
1.5.3 *N*-linked glycosylation..... 10
1.5.4 Disulfide bond formation..... 13
1.5.5 Cargo exit from ER mediated by COPII..... 14
1.5.6 ER retention signal..... 15
Chapter 2: Objectives of the study..... 16
Chapter 3: Materials..... 18
3.1 Cells..... 18
3.2 Viruses..... 18
3.3 Bacterial strains..... 18
3.4 Medium..... 18
3.5 Antibodies..... 19
3.6 Chemicals..... 20
3.7 Commercial Kits..... 22
3.8 Prepared Solutions and Buffer..... 22
3.9 Regent..... 25
3.10 Radio activity chemicals..... 25
3.11 Primers..... 26
3.12 Plasmids..... 35

3.12.1 Commercial Vector Plasmids 35
3.12.2 Prepared plasmids available in the laboratory 36
3.12.3 New plasmid constructs in this study 36
3.13 Equipments 39
3.14 Consumables 40
3.15 Software, application and program 41
Chapter 4: Methods 42
4.1 Molecular cloning methods 42
4.1.1 Normal PCR 42
4.1.2 Overlap fusion PCR 42
4.1.3 QuikChange® PCR 43
4.1.4 Agarose gel electrophoresis 44
4.1.5 Preparative agarose gel electrophoresis 44
4.1.6 Restriction analysis 44
4.1.7 Ligation of DNA fragments 45
4.1.8 Sequencing 45
4.1.9 *In vitro*-transcription 46
4.1.10 Phenol-chloroform extraction 47
4.1.11 Reverse Transcriptase PCR (RT-PCR) 47
4.2 Microbiological methods 48
4.2.1 Heat shock transformation of *E.coli* 48
4.2.2 Isolation of plasmid DNA from bacteria 48
4.3 Cell biological methods 49
4.3.1 Cultivation of adherent cells 49
4.3.2 Infection of cells 49
4.3.3 Transfection of cells 49
4.3.4 Electroporation of cells with RNA 50
4.3.5 Virus titration 50
4.3.6 Growth Kinetics of Viruses 51
4.3.7 Extraction of viral RNA from cells 51
4.4 Protein analytical methods 51
4.4.1 Indirect immunofluorescence analysis 51
4.4.2 Flow Cytometry Analysis (FACS analysis) 53
4.4.3 Immunoprecipitation 54
4.4.4 Western blot 54

Chapter 5: Results 56
5.1 Overview of pestiviral E1 glycoprotein 56
5.1.1 Amino acid composition analysis 56
5.1.2 Alignments of multiple sequences of pestiviral E1 57
5.1.3 The prediction of secondary structure and transmembrane domain of pestiviral E1 60
5.2 Subcellular localization of pestivirus E1 62
5.3 Studies on the localization of the retention signal in E1 63
5.3.1 The transmembrane anchor of E1 is a determinant for ER retention 63
5.3.2 The ER retention signal is within the middle part of the transmembrane domain of E1 66
5.3.3 The polar amino acids in the middle part of E1 TM domain play an essential role in ER localization of E1 70
5.3.4 Effect of selected mutations in E1 on the replication of BVDV strain CP7 74
5.4 Analysis of the membrane topology of pestiviral E1 76
5.4.1 Membrane topology of E1 separately expressed E1 76
5.4.2 The TM domain of pestiviral E1 forms a hairpin structure before signal sequence cleavage 80
5.5 The Prerequisites of E1 oligomerization and E1-E2 heterodimerization 83
5.5.1 Pestiviral glycoprotein E1 can form homo-oligomers independent of its membrane anchor 83
5.5.2 Both Cys123 and Cys171 play an important role in E1 homo-oligomerization 84
5.5.3 Critical sites for E1-E2 heterodimer formation and E2 homodimer formation 84
5.5.4 E1-E2 heterodimer formation is independent of the transmembrane region of E1 . 86
5.5.5 E1 overrules the retention of E2 via intermolecular disulphide bond formation 87
5.5.6 Both Cys123 in E1 and Cys295 in E2 are important for viral infectivity 90
5.6 The middle hydrophobic region (MHR) affects the secretion of E1 93
5.6.1 E1 is retained within the cell in the absence of the carboxyterminal membrane anchor 93
5.6.2 The middle hydrophobic region (MHR) affects the secretion of E1 97
Chapter 6: Discussions 101
6.1. The organization of domains in pestiviral E1 101
6.2 The retention of pestiviral E1 glycoprotein 102
6.3 The membrane topology of the pestiviral E1 glycoprotein 105
6.4 The oligomerization of E1 and the heterodimerization of E1-E2 107
Zusammenfassung 111

Summary 113
References 115
Appendix 124
Publications 129
Acknowledgements 130
Declaration of authorship/Selbstständigkeitserklärung 131

List of Figures

Figure 1.1 Phylogenetic tree of pestiviruses ····· 3
Figure 1.2 Schematic representation of the pestiviral particle (left) and genome (right) ····· 4
Figure 1.3 Schematic representation of the replication cycle of pestiviruses ····· 6
Figure 1.4 Membrane/secretory protein biogenesis in the ER ····· 9
Figure 1.5 Three different types of single-spanning membrane proteins topogenesis ····· 10
Figure 1.6 *N*-glycosylation ····· 12
Figure 1.7 Schematic representation of disulphide bond formation ····· 14
Figure 3.1 Schematic representation of pCI and pCITE 2a (+) empty vector ····· 36
Figure 4.1 Schematic representative of different permeabilization treatments ····· 53
Figure 5.1 Amino acid composition of pestiviral E1 ····· 56
Figure 5.2 Kyte & Doolittle hydropathy plot of pestivirus E1 glycoprotein ····· 57
Figure 5.3 Conservation of amino acid sequences of pestiviral E1 protein ····· 58
Figure 5.4 NetNglyc 1.0: predicted *N*-glycosylation site (s) of pestiviral E1 ····· 59
Figure 5.5 Mutagenesis analysis of *N*-Glycosylation site of pestiviral E1 ····· 60
Figure 5.6 The prediction of secondary structure in pestiviral E1 ····· 61
Figure 5.7 The consensus prediction of pestiviral E1 membrane topology ····· 61
Figure 5.8 Subcellular localization of E1 glycoprotein ····· 63
Figure 5.9 The transmembrane anchor is responsible for the retention of E1 ····· 64
Figure 5.10 Subcellular localization of pYM-56 (E1-VSVg chimera) ····· 65
Figure 5.11 Co-localization analysis of the parental proteins or chimeras ····· 66
Figure 5.12 14aa within the middle part of transmembrane domain of E1 is responsible for the ER retention ····· 69
Figure 5.13 The polar residues of E1 TM domain play essential roles in ER retention of E1 71
Figure 5.14 Six polar residues of E1 TM domain are important for ER retention of E1 ····· 72
Figure 5.15 Subcellular localization of pYM-53 derived E1 mutant ····· 73
Figure 5.16 Indirect immunofluorescence analysis after electroporation ····· 75
Figure 5.17 Expression and recognition of tag-labeled E^{rns} and E1 proteins ····· 77
Figure 5.18 Membrane topology of E1 analysed by Avi-tag biotinylation assay ····· 79
Figure 5.19 The TM region of pestiviral E1 form a hairpin structure before signal sequence cleavage ····· 82
Figure 5.20 E1 can form homo-oligomer independent to its membrane anchor ····· 83

Figure 5.21 Both Cys123 and Cys171 play an important role in E1 homo-oligomerization ·· 84
Figure 5.22 Cys123 in E1 is the critical site for E1-E2 heterodimer formation; Cys295 in E2 is essential for not only E1-E2 heterodimerization but also for E2 homodimer formation ····· 86
Figure 5.23 Heterodimer formation analysis ····· 86
Figure 5.24 E1 overrules the retention signal of E2 via intermolecular disulphide bond between E1 and E2 ····· 90
Figure 5.25 Indirect immunofluorescence analysis after electroporation ····· 91
Figure 5.26 CPE observation after electroporation ····· 91
Figure 5.27 Indirect immunofluorescence analysis reinfection after electroporation ····· 92
Figure 5.28 The sequence and helical wheel modeling of the putative membrane anchor of E1 ····· 93
Figure 5.29 The secretion/retention analysis of E1/truncated E1 variants ····· 94
Figure 5.30 Subcellular localization of *C*-terminally truncated E1 variants ····· 96
Figure 5.31 Secretion/retention analysis for the *C*-terminally truncated E1 MHR 2 deletion variants ····· 98
Figure 5.32 Secretion/retention analysis for the *C*-terminal truncated E1 entire MHR deletion variants ····· 99
Figure 6.1 Schematic representation of pestiviral E1 envelope protein ····· 102
Figure 6.2 Comparison of putative TMD sequences of envelope proteins in the family *Flaviviridae* ····· 103
Figure 6.3 Schematic model of the reorientation behaviour of the TM domains of pestiviral envelope proteins during the early steps of their biogenesis ····· 106
Figure 6.4 Schematic drawing of the cysteine residue distribution in native pestiviral E1 and E2 glycoproteins ····· 109

List of Tables

Table 1.1 Member species of pestivius according to the 10th online report of ICTV ····· 1
Table 1.2 Some ER proteins/complexes associate with the Sec 61 complex ····· 8
Table 1.3 Three different well characterized topogenic types of single-spanning membrane proteins ····· 10
Table 1.4 Characterized ER export signals of membrane protein ····· 14
Table 4.1 Normal PCR reaction approach ····· 42
Table 4.2 Overview of the normal PCR program ····· 42
Table 4.3 Overlap fusion PCR reaction approach ····· 43
Table 4.4 Overview of the QuikChange® PCR program ····· 44
Table 4.5 Restriction analysis reaction approach for mini prep products ····· 45
Table 4.6 Ligation reaction mix ····· 45
Table 4.7 Sequence PCR reaction mix ····· 46
Table 4.8 Overview of the sequencing PCR program ····· 46
Table 4.9 *In vitro*-transcription (T7) reaction approach ····· 47
Table 4.10 Overview of the RT-PCR program ····· 48
Table 4.11 RT-PCR reaction approach ····· 48
Table 4.12 Lipofectamine™ 2000 DNA transfection procedure ····· 50
Table 4.13 Different permeabilization treatments ····· 52
Table 5.1 Amino acid composition of pestiviral E1 ····· 56
Table 5.2 The prediction of *N*-glycosylation sites of pestiviral E1 ····· 59
Table 5.3 Constructs for mutagenesis analysis of fully conserved *N*-glycosylation sites of pestiviral E1 ····· 60
Table 5.4 The consensus prediction of pestiviral E1 transmembrane region(s) ····· 62
Table 5.5 Mutagenesis and insertion analysis for conserved residues in TMD of E1 ····· 72

Abbreviations

AA	Amino acids
Ac	Acetate
Amp	Ampicillin
APS	Ammonium persulfate
ATP	Adenosine triphosphate
BHK	Baby Hamster Kidney
bp	Base pair
BSA	Bovine serum albumin
BVDV	Bovine viral diarrhea virus
C	Carboxyl terminus
°C	Centigrade
cDNA	Complementary DNA
cm	centimeter
cp	cytopathogenic
CPE	cytopathogenic effect
CSFV	Classical swine fever virus
d	day
DAPI	4',6-diamidino-2-phenylindole
DEPC	Diethyl pyrocarbonate
DNA	Deoxyribonucleic acid
ds	double-stranded
EDTA	Ethylenediaminetetraacetic acid
Endo H	Endoglycosidase H
ER	Endoplasmic reticulum
ERGIC	ER Golgi intermediate compartment
FCS	Fetal calf serum
FITC	Fluorescein isothiocyanate
FLI	Friedrich-Loeffler-Institut
g	gram
g	gravitational acceleration
GAPDH	Glucose-aldehyde-phosphate dehydrogenase
GFP	green fluorescent protein
h	hour
HCV	Hepatitis C virus
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IRES	internal ribosome entry site
k	kilo
kb	Kilobase pairs
kD	Kilodalton
l	Liter
LB	Lysogeny broth

m	milli
m	meter
M	molar (mol/l)
Mab	monoclonal antibody
MOI	multiple of infection
MD	Mucosal Disease
MDBK	Madin-Darby bovine kidney
min	minute
mRNA	Messenger RNA
MW	Molecular weight
N	Amino terminus
ncp	non cytopathogenic
NTR	non translated region
OD	optical density
ORF	open reading frame
OST	Oligosaccharyl transferase
PAGE	Polyacrylamide gel electrophoresis
pAb	polyclonal antibody
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDI	Disulphide isomerases
pH	negative logarithm of the hydrogen ion concentration
PI	persistently infected
PNGase F	Peptide-n-glycosidase
RNA	Ribonucleic acid
Rnase	Ribonuclease
rpm	revolutions per minute
RNP	ribonucleoprotein complex
RT	Room temperature
RT-PCR	Reverse transcriptase based PCR
s	second
SDS	Sodium dodecyl sulfate
SP	Structure protein
SP	Signal peptide
Spase	Signal peptidase
SPC	Signal peptidase complex
SPR	Signal peptidase peptidase
SR	Signal recognition particle receptor
SRP	Signal recognition particle
ss	single-stranded
TAE	Tris-acetate-EDTA-Buffer
TE	Tris-EDTA-Buffer
TEMED	Tetramethylethylenediamine
TM	Transmembrane
TMD	Transmembrane domain
Tris	Tris (hydroxymethyl) aminomethane

U	Activity unit (for enzymes)
UV	Ultraviolet
V	Volt
Vol.	Volume fraction
vs.	versus
WB	Western blot
wt	Wild-type
w/v	Weight per volume

Amino acids

Alanine	A	Ala
Cysteine	C	Cys
Aspartic acid / Aspartate	D	Asp
Glutamic acid / Glutamate	E	Glu
Phenylalanine	F	Phe
Glycine	G	Gly
Histidine	H	His
Isoleucine	I	Ile
Lysine	K	Lys
Leucine	L	Leu
Methionine	M	Met
Asparagine	N	Asn
Pyrrolysine	O	Pyl
Proline	P	Pro
Glutamine	Q	Gln
Arginine	R	Arg
Serine	S	Ser
Threonine	T	Thr
Selenocysteine	U	Sec
Valine	V	Val
Tryptophan	W	Trp
Tyrosine	Y	Tyr

Chapter 1: General introduction

Pestiviruses, members of the family *Flaviviridae*, are among the most important pathogens of farm animals worldwide. The genus *Pestivirus* has a broad host range (mainly pigs and ruminants), induces several clinical manifestations in farm or wild animals and causes very severe financial losses in the livestock farming industry (Tautz et al., 2015). Nowadays, farm animal trade is becoming more and more international, so the higher frequency of transport of susceptible animals increases the risks of pestivirus infection. The detailed characterization of pestiviral molecular biology just started in recent decades. Especially in the last 20 years, scientists working on pestiviruses have made striking achievements with regard to many aspects of pestivirus biology, but there are still many unsolved questions about the life cycle of pestiviruses, for instance, the function and structure of their envelope glycoproteins, which urge us to conduct further intensive studies of pestiviruses.

1.1 Taxonomy, host and disease

1.1.1 Taxonomy

The genus *Pestivirus* belongs to the family *Flaviviridae*. Originally, there are 4 recognized species have been classified into the genus *Pestivirus*, they are BVDV-1, BVDV-2, CSFV and BDV. But lately, many new isolates were found (Tautz et al., 2015). In the latest online report (10th) of ICTV, the genus *Pestivirus* has been subdivided into 11 different species correspondingly. Furthermore, the ICTV also modified the naming convention of pestivirus species by adopting the format 'Pestivirus X' instead of their traditional designation in a host-dependent manner. It is worth noting that only the names of virus species have been changed, virus isolates continue using their original names (Simmonds et al., 2017; Smith et al., 2017), as shown in Table 1.1.

Table 1.1 Member species of pestivius according to the 10th online report of ICTV

Species	Virus name	Virus abbreviation	Representative Isolate	Accession number
Pestivirus A	bovine viral diarrhea virus 1	BVDV1	SD-1/NADL	M96751/M31182
Pestivirus B	bovine viral diarrhea virus 2	BVDV2	XJ-04/890	FJ527854/U18059
Pestivirus C	classical swine fever virus	CSFV	Alfort/187	X87939
Pestivirus D	Border disease virus	BDV	X818	AF037405
Pestivirus E	pronghorn antelope pestivirus	PAPeV		AY781152
Pestivirus F	porcine pestivirus	PPeV	Bungowannah	EF100713
Pestivirus G	giraffe pestivirus	GPeV	H138	AF144617
Pestivirus H	HoBi-like pestivirus	HoBiPeV	Th/04_KhonKaen	FJ040215
Pestivirus I	Aydin-like pestivirus	AydinPeV	04-TR	JX428945
Pestivirus J	rat pestivirus	RPeV	NrPV/NYC-D23	KJ950914
Pestivirus K	atypical porcine pestivirus	APPeV	515	KR011347

According to a biological feature, pestiviruses can also be divided into cytopathogenic (CP) and non-cytopathogenic (NCP). Compared with NCP strains, the CP pestiviruses have stronger effect on the infected cells or tissues that always result in the lysis of the host cells. This interesting biological characteristic is supposed to exist among the whole pestivirus species, since this conclusion has been proven in the original classified species (BVDV-1, BVDV-2, CSFV and BDV) and Pestivirus H (HoBi-like pestivirus) (Aoki et al., 2001; Kolykhalov et al., 1997; Kosmidou et al., 1998; Kupfermann et al., 1996; Meyers et al., 1991; Meyers et al., 1996b; Tautz et al., 1998). Normally, CSFV is non-cytopathogenic (NCP) in the infected cells or tissue cultures. Only a few cases of cytopathogenic (CP) CSFV strains were reported (Kümmerer et al., 2000). The cytopathic effect (CPE) induced by CP CSFV can only occur in the presence of so-called defective interfering particles (DIs) associated with the complementing helper virus (Meyers and Thiel, 1995; Meyers et al., 1996b). In contrast to the CP BVDV that are associated with mucosal disease (MD), CP CSFV always induce very slight pathological symptoms. NCP BVDV can result in persistent infection (PI) making infected animals a permanent source of infection for the surrounding herds. It is worth noting that all infected with CP pestivirus isolates always have a higher amount of total viral RNA than the corresponding NCP isolates (Aoki et al., 2003).

1.1.2 Host

For a long time, infections with pestiviruses were believed to be restricted to clovenhoofed animals. However, some recent metagenomics research has indicated that pestiviruses are not completely restricted to Artiodactyla species, since some clear evidence for the existence of pestiviruses in mammals like bats (BatPeV), rats and whales were recently described (Firth et al., 2014; Jo et al., 2019; Wu et al., 2012). Naturally, infections of Pestivirus H (HoBi-like pestivirus) always occur in cattle, however, a few cases of Pestivirus H have been described in goats which are non-bovine ruminants (Shi et al., 2018). There is no pestivirus infection case in human or poultry reported. The approved species of pestiviruses were roughly divided into two major groups before, including group1- pestiviruses in swine and group 2- pestiviruses in ruminants like cattle, sheep, goats, and a large variety of wild ruminants. With more and more novel pestivirus infection in bats, rats and even whale are discovered, there should be one additional group 3 which is pestiviruses in other mammals.

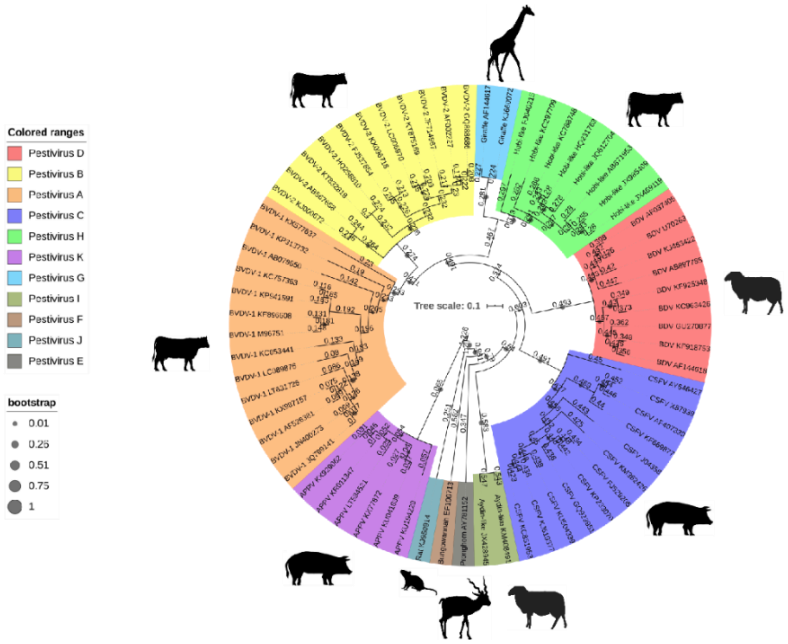


Figure 1.1 Phylogenetic tree of pestiviruses

The phylogenetic tree of pestiviruses was made by iTOL (v5, https://itol.embl.de/). Each member species is in different color. Corresponding sequence alignment and the original phylogenetic tree file were downloaded from the resources page of the ICTV *Flaviviridae* website (https://talk.ictvonline.org/ictv-reports/ictv_online_report/positive-sense-rna-viruses/w/ flaviviridae/371/resources-flaviviridae).

1.1.3 Disease

The most important transboundary viral disease of pestivirus in swine is classical swine fever (CSF). It has enormous impact on animal health and pig industry worldwide. Therefore, outbreaks of CSF have to be reported to the World Organization for Animal Health (OIE) (Edwards et al., 2000). CSF widely spreads in most countries of the world that have a significant pig production. The spreading of CSF is mainly concentrated in the following regions, some countries in South America, several parts of Eastern Europe also including many neighboring countries, as well as Asia, especially China and India. So far, no detailed information about the African situation is available (Blome et al., 2017). The clinical phenotype of classical swine fever can be divided into three types: ①acute type always accompanied by lethally transient course of disease; ②chronic type and ③persistent course. The third type usually occurs in pregnant sows (Moennig et al., 2003).

Bovine viral diarrhea virus (BVDV) is one of the most important infectious agents of cattle herds. BVDV strains of each distinct species (BVDV1 and BVDV2) are further classified as one of two biotypes: cytopathic (CP) and non-cytopathic (NCP) as defined by the lytic activity of the virus in cell culture (Kümmerer et al., 2000; Kupfermann et al., 1996; Meyers and Thiel, 1996; Tautz et al., 2015). CP BVDV strains are not commonly found in acute infections in the field and are mainly involved in outbreaks of mucosal disease whereas NCP BVDV strains are more common in nature and are often associated with the most clinically severe form of acute infection. It causes sustained economic losses to the cattle industry mainly because of reduced reproductive performance and milk production (Brock, 2004). Infection of susceptible pregnant cows with the NCP virus, which has the ability to cross the placenta before the development of fetal immunocompetence, can result in the birth of persistently infected (PI) calves (Bolin et al., 1985; Brownlie, 1990; Brownlie et al., 1984; Moennig et al., 1990).

1.2 Viral particle and genome

Pestivirus virions are enveloped and contain four structural proteins (SPs) including one basic core protein C and three envelope (E) glycoproteins (E^{rns}, E1 and E2) that are present on the viral particles (Thiel et al., 1991; Weiland et al., 1999). The diameter of the pestiviral particle is about 40-60 nm. E1-E2 heterodimers are incorporated into the virions, which were identified to be critical for virus infection (El Omari et al., 2013; Li et al., 2013; Ronecker et al., 2008).Within the particle, there is core protein found to be single-stranded RNA genome. Because of the lipid envelope, pestiviruses are very sensitive to detergent treatment so that they can be inactivated quite easily. The general stability of the virions is low. Under conditions of 37 °C and neutral pH, the half-life of viral particle is about 7h (Depner et al., 1992).But compared to other flaviviruses, pestiviruses are very resistant to low pH.

The pestivirus genome consists of a single stranded RNA with a basic length of about 12.3 kb. It contains only one long open reading frame coding for a polyprotein of ca. 4000 amino acids. Through co- and posttranslational processing by viral and cellular proteases, 12 mature proteins are generated. The arrangement of those 12 recognized pestivirus mature proteins in the

polyprotein is NH_2-N^{pro}/C/E^{rns}/E1/E2/p7/NS2/NS3/NS4A/NS4B/NS5A/NS5B-COOH (Meyers and Thiel, 1996) (see Fig 1.2 right). The first protein encoded by the ORF is a non-structural protein named N^{pro} (Stark et al., 1993). N^{pro} is an autoprotease that cleaves at its own *C*-terminus and thereby releases the *N*-terminus of the core protein. There is one hydrophobic sequence at the *C*-terminus of core protein serving as signal sequence for translocation of the downstream glycoprotein E^{rns}. Cellular signal peptidase (SPase) is responsible for the cleavage at the C/E^{rns} site (Rümenapf et al., 1993), then this *C*-terminal signal sequence will be further processed by signal peptide peptidase (Heimann et al., 2006). Signal peptidase is also responsible for processing at the E^{rns}/E1, E1/E2, E2/p7 and p7/NS2 sites (Bintintan and Meyers, 2010; Harada et al., 2000). The cleavage at the E^{rns}/E1 site is slower than at the E1/E2 site, resulting in the detectable presence of a E^{rns}-E1 precursor in infected and transfected cells.

The next processing step at the NS2/NS3 site is quite special since it is involved in the regulation of pestivirus RNA replication. For a non-cytopathic pestivirus, it is cleaved by the protease activity of NS2 whereas for many CP pestiviruses other proteases are recruited to conduct this cleavage (Lackner et al., 2005; Lackner et al., 2004; Lackner et al., 2006). All sites downstream of the NS2/NS3 site are processed by the NS3 (Lamp et al., 2011; Lamp et al., 2013; Tautz et al., 1997; Wiskerchen and Collett, 1991; Xu et al., 1997). NS4A functions as a cofactor of the NS3 protease and is involved in processing the NS4B/NS5A and the NS5A/NS5B sites.

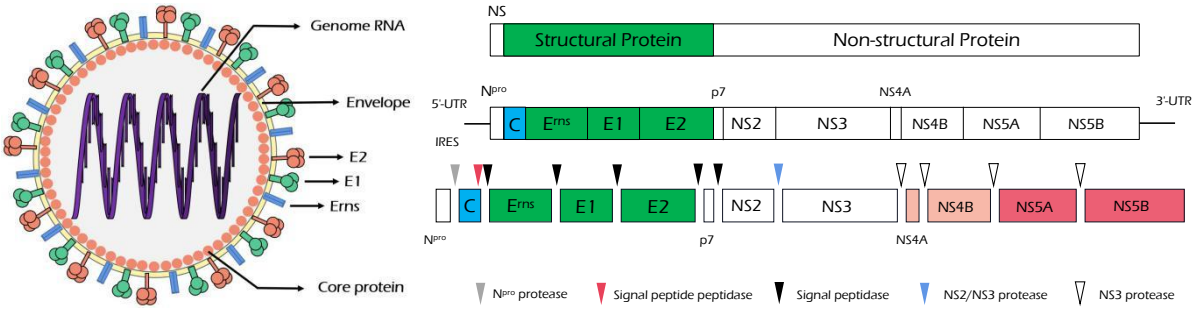


Figure 1.2 Schematic representation of the pestiviral particle (left) and genome (right)

Left, pestiviruses are enveloped viruses with 3 envelope proteins (E^{rns}-blue, E1-green, E2-red) with a single-stranded RNA genome of positive polarity (purple helix inside), which is shown on the right side as a scheme. The genomic RNA contains one long open reading frame (ORF) that is translated into a polyprotein schematically shown below the genome. After co- and post-translational processing by viral and host proteases, 12 mature proteins are generated. Viral and cellular proteases that participate in polyprotein processing are shown at the right bottom. Structural proteins (Core, E^{rns}, E1 and E2) in green/blue, non-structural proteins in white or reddish colour, NS3 proteinase is responsible for the cleavage of reddish proteins.

1.3 The viral life cycle

Pestivirus replication starts with attachment of viral particle to the surface of its host cell. The viral envelope glycoproteins E^{rns} and E2 both play important roles in this process (Hulst and Moormann, 1997). Glycoproteins E^{rns} and E2 have different receptors on the cell surface. Both heparan sulfate (HS) and laminin receptor (LamR) were identified as attachment receptors for E^{rns} (Chen et al., 2015; Munir Iqbal, 2000). E2 is also involved in viral entry since it is the receptor binding protein that is a determinant of pestivirus tropism (Borca et al., 2019; Liang et al., 2003; Reimann et al., 2004). It was shown that E1 and E2 form covalently disulphide linked heterodimers (Thiel et al., 1991), the formation of which is crucial for virus infection. Interestingly, presence of E1-E2 heterodimers in pseudotyped viruses is sufficient to mediate viral entry (Ronecker et al., 2008; Wang et al., 2004), indicating that E^{rns} is dispensable for the

entry step. Pestiviruses enter the host cells via receptor-mediated endocytosis. The main receptor that mediates BVDV entry has been identified so far is CD46 which interacts with BVDV E2 (Maurer et al., 2004). However, it has been shown that CD46 is not sufficient for internalization of virus, indicating that viral entry still needs one or more co-receptors. Recent researches have shown that the host factor MERTK, a member of the receptor protein tyrosine kinases, promotes CSFV entry (Zheng et al., 2020).

After binding to the receptor on the surface of host cells, infectious viruses enter cells by clathrin-dependent endocytosis (Grummer et al., 2004; Lecot et al., 2005). Some small GTPases of the Ras superfamily like Rab5, Rab7, and Rab11, which are required for caveola-dependent endocytosis were shown to be involved in CSFV entry in a recent study, indicating that CSFV might also enter cells via the caveola-mediated pathway (Zhang et al., 2018). After internalization, membrane fusion and uncoating of virions follow and the genomic RNA is released into the cytoplasm. The RNA of pestiviruses is infectious: it functions as both the genomic and messenger RNA. Until now, it is still unclear where exactly in the cell pestivirus viral RNA replication takes place. For the related HCV, the viral replication is believed to occur in a so-called membranous web (Moriishi and Matsuura, 2012; Neufeldt et al., 2016; Wolk et al., 2008). Due to the fact that the pestivirus RNA genome does not contain a cap structure at its 5' end, viruses take advantage of an "internal ribosomal entry site" (IRES) at its 5' NTR to induce efficient translation. The pestiviral genome is initially translated into a polyprotein, which is further processed co- and post-translationally by host and viral proteases. Pestivirus envelope glycoproteins are synthesized at the ER. After translation of a signal sequence located at the end of C protein, the nascent synthesized chain is targeted to the translocon in the ER membrane. After co-translational translocation of the viral envelope proteins, the nucleocapsid complex interacts with the accumulated glycoproteins and buds into the lumen of the ER, resulting in viral particle assembled in the ER lumen. Further modifications of the envelope proteins such as glycosylation and disulfide bond formation also occur in the ER. Further processing of carbohydrate chains occurs in Golgi compartment. Finally, the mature viral particle complex is released from the infected cells via the exocytosis pathway.

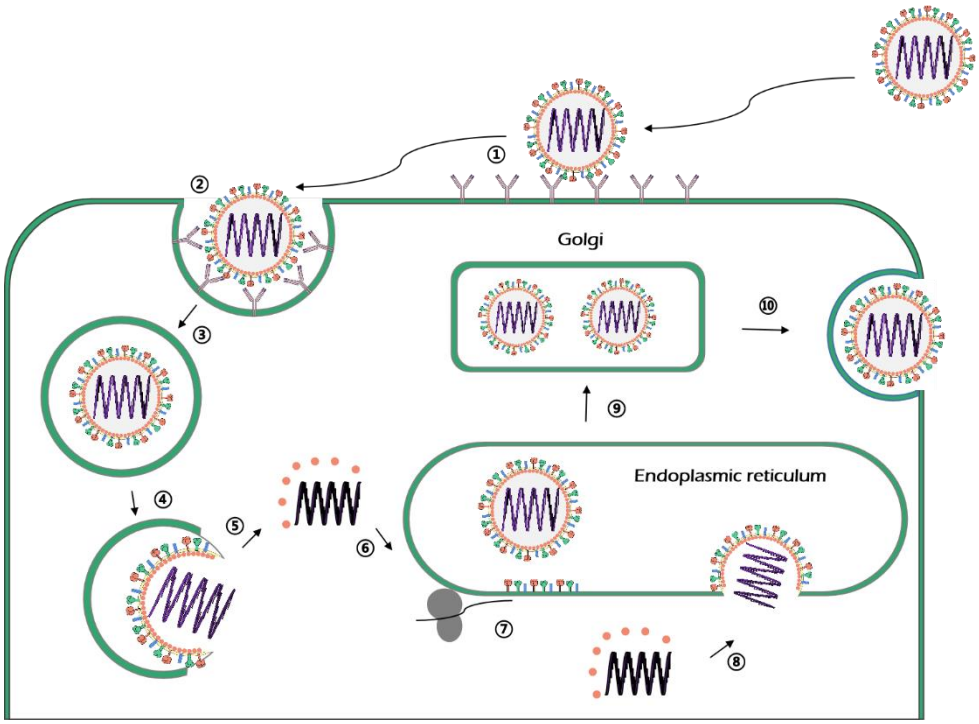


Figure 1.3 Schematic representation of the replication cycle of pestiviruses

①Virions firstly attach to the cell surface; ②Viral envelope protein E2 binds to the main cellular receptor (e.g. CD46 for BVDV); ③then internalization via clathrin-mediated endocytosis; ④Membrane fusion, ⑤the viral genome and capsid protein are released from the early endosome into the cytoplasm of the host cell and uncoating sets the genomic RNA free ⑥The translation and replication of viral genomic RNA are carried out ⑦and the viral proteins are synthesized; ⑧The viral nucleocapsid/RNA complex interacts with glycoproteins and egresses into the lumen side of the ER; ⑨Transport via exocytosis pathway; ⑩The mature pestivirus particles are released from the infected cells.

1.4 Envelope proteins

Pestiviruses contain three envelope proteins E^{rns}, E1 and E2. While E^{rns} represents an attachment protein that is also engaged in pestiviral immune evasion due to its RNase activity, E1 and E2 are most likely the relevant factors for viral entry. E1 is the most mysterious of all viral proteins. The shortage of specific antibodies against E1 making detection and further analysis of E1 difficult, resulted in a lack of knowledge on E1 compared to E^{rns} and E2 with regard to biosynthesis, structure and function. E1 has often been analysed in context with the other two envelope proteins, instead of being used as research object alone. The molecular size of E1 is a mass of 25-33 kDa (depends on the virus species) so only about half of E2. E1E2 can form heterodimers, which is known as the crucial functional complex in viral entry (Ronecker et al., 2008). For HCV, it has been predicted that E2 can serve as fusion protein and possesses a class II fusion fold that harbors a membrane distal fusion loop rich in hydrophobic residues (Garry and Dash, 2003). Unexpectedly, the recently published crystal structure of BVDV1 E2 does not show that it contains a class II fusion protein fold (El Omari et al., 2013; Li et al., 2013), indicating that E2 if being directly involved in fusion at all belongs to a novel structural class of membrane fusion machinery. Otherwise, E1 could represent the fusion protein of pestiviruses as proposed in one publication (Fernandez-Sainz et al., 2014).

E1 has long been believed to be a type I transmembrane protein with a *C*-terminal membrane anchor, but membrane topology of pestivirus E1 has never been analysed. Recently, published data suggested that E1 contains one transmembrane helix with two amphipathic perimembrane helices located upstream of the TM helix (Wang et al., 2014). However, the latter publication is based only on computational modeling tools used to simulate and predict the secondary structure of pestivirus E1 and E2. Crystal structure information about the E1 protein is still not available, and also its membrane topology awaits detailed experimental analysis. In the related hepatitis C virus, which shares many structural and functional properties with pestiviruses, envelope proteins E1 and E2 were identified as type I transmembrane proteins. Surprisingly, there is a dynamic change in the TM regions of both E1 and E2 with changes occurring after signal peptidase cleavage. This unique dynamic behaviour of the TM domain of E1 is supposed to be linked to its multifunctionality, such as membrane anchoring, heterodimerization and retention (Cocquerel et al., 2002).

E1-E2 heterodimers are covalently linked via disulphide bonds (Thiel et al., 1991). One publication suggests that positively charged residues in E1 (lysine and arginine) play an essential role in heterodimer formation. In this paper, the authors declare that cysteine residue at position 668 is not essential for E1-E2 heterodimer formation by using single site mutagenesis (Ronecker et al., 2008). For HCV, similarly, it has been shown that the charged residues within the transmembrane domains of glycoprotein E1 and E2 play an important role in E1-E2 heterodimerization (Cao et al., 2019; Ciczora et al., 2005). However, the interaction between HCV E1 and E2 is non-covalent. Thus, there should be a difference of the interaction

mechanism between envelope proteins in HCV and pestiviruses. HCV E1 was shown form non-covalent trimers on the virions (Falson et al., 2015). Due to the absence of specific antibodies against pestiviral E1, it is still unknown whether E1 of pestiviruses forms oligomers or not.

Some other publications give contradictory conclusions. In their study, the predicted theoretical models suggest that Cys668 in E1 forms a disulphide bond with Cys987 in E2 by using computational secondary structure predictions and E1E2 sequence alignments, along with the geometric constraints imposed by the recently published crystal structure of BVDV E2. (Wang et al., 2014). Additionally, this linkage between E1 and E2 could stabilize the E1-E2 interaction that is required for virus infectivity. In the absence of experimental structural data for E1 or E1-E2 heterodimer, the prerequisites and function of E1 oligomerization and E1-E2 heterodimerization are poorly understood.

The glycoprotein E2 of pestiviruses has a mass of 53-55 kDa. Recently, the membrane topology of mature BVDV E2 has been clarified. E2 is a type I transmembrane protein with an *N*-terminal ectodomain and a hydrophobic membrane anchor at its *C*-terminus (Radtke and Tews, 2017). It represents the target of neutralizing antibodies that can induce protective humoral immunity. Published data also suggested that CSFV E2 can serve as a target for cytotoxic T cells that elicit cellular immunity (Armengol et al., 2002; Ceppi et al., 2005; Franzoni et al., 2013). It is a determinant of cell tropism and host specificity. CD46 is a cellular receptor for BVDV infection that interacts with E2. Since CD46 alone is not sufficient for a successful pestiviral infection, other cellular receptors have to be involved in this step and pestivirus entry is most likely a multi-step process.

CSFV E2 has been shown to form covalently disulphide-linked homodimers (Thiel et al., 1991; Weiland et al., 1990). For BVDV, E2 ectodomain forms covalently disulphide-linked homodimers in the absence of E1. During virus assembly, E2 homodimers are formed much earlier than the formation of E1-E2 heterodimer, this could be due to the slow folding of the E1 protein as well as the slow release of E1 from the endoplasmic reticulum chaperone calnexin (Branza-Nichita et al., 2001). Even though E1-E2 heterodimer is involved in viral attachment and entry step, it is not shown which of the protein is responsible for membrane fusion. Unexpectedly, the recently published crystal structure of BVDV E2 did not reveal the presence of a class II fusion protein fold which was supposed for pestivirus in analogy to hepacivirus. This indicated that fusion machinery of pestiviruses and hepaciviruses could be different from any other previously reported. Until now, the function of the E2 homodimer is still unclear. This point awaits further investigation.

The mature E2 contains 15-17 conserved cysteines depending on the viral species. It was shown that except for C987, all other cysteines in E2 form intramolecular disulphide bonds in the absence of E1 by using computational approaches in the context of geometric constraints deduced from the E2 structures (Li et al., 2013; Wang et al., 2014). In their proposed atomic models, Cys987 in E2 is the critical site not only for E2 homodimer formation but also for E1-E2 heterodimerization. Furthermore, they also suggested that C668 in E1 is most likely the residue engaged in the disulphide bond linkage with E2 to form E1-E2 heterodimers, but those predictions still need detailed experimental verification. A publication by Ronecker and colleagues also suggested that the positively charged amino acid Arg355 in E2 is essential for heterodimerization with E1 (Ronecker et al., 2008). As also shown for E^{rns}, E2 is retained intracellularly showing that the protein contains a retention signal of its own. Recent publication demonstrated that both arginine 355 and glutamine 370 in E2 are important for intracellular retention (Radtke and Tews, 2017). Introduction of mutations of these residues leads to export of E2 to the cells surface.

1.5 The conventional protein trafficking pathway

After uncoating, the pestiviral RNA can be directly used by the system of host cells to synthesize viral proteins. To know how the protein synthesis pathway generally work in host cells will help to further understand some important steps in viral life cycle. The protein trafficking pathway is charged with the synthesis, modification, and delivery of a series of cellular soluble and membrane proteins in both prokaryotes and eukaryotes. In this eukaryotes trafficking route, membrane proteins and soluble cargoes are delivered from the endoplasmic reticulum (ER) to the Golgi apparatus, subsequently transported to the plasma membrane by some specific secretory vesicles. The first biogenesis step in the conventional secretory pathway of soluble and membrane proteins is translocation of the polypeptide mediated by signal sequence to the ER.

1.5.1 Translocation at the ER site

The translocon is assembled by the so-called Sec61 complex which consists of the heterotrimeric Sec61-α, Sec61-β and Sec61-γ in mammalian cells (Gorlich et al., 1992; Gorlich and Rapoport, 1993). Among them, Subunit Sec61-α forms a gated pore of the translocon (Gorlich and Rapoport, 1993). Both Sec61-α and Sec61-γ are highly conserved and critical not only for the function of the translocon channel but also for the cell viability, while the subunit Sec61-β is dispensable.

In mammals, there are plenty of ER proteins and complexes (shown in Table 1.2) that always associate with the Sec 61 complex, present in the ER and exercise translocation behaviour. Both co- and post-translational translocation are observed that depend on Sec61 but in association with different components/complexes. The co-translational mechanism widely exists in nearly all types of cells and is used for soluble and integral membrane proteins of most eukaryotes. Post-translational translocation is more commonly arising in *Escherichia coli* and yeast. The co-translational mechanism is GTP-dependent (Rapoport, 2007), while the post-translational translocation is in a ATP dependent process (Chirico et al., 1988; Deshaies et al., 1991; J A Rothblatt 1986).

Table 1.2 Some ER proteins/complexes associate with the Sec 61 complex

	Protein/complex	Abbreviation	Molecular weight
Chaperones/Targeting components	signal recognition particle	SRP	Six subunits: 9, 14, 19, 54, 68, and 72 kDa
	signal recognition particle receptor	SR	Two subunits: SRα of 72 kDa and SRβ of 30 kDa
Auxiliary components	translocating chain-associating membrane	TRAM	37 kDa
	translocon-associated protein	TRAP	90 kDa
	binding immunoglobulin protein	Bip	78 kDa
	signal peptidase complex	SPC	Five subunits: 12, 18, 21, 22/23, and 25kDa

Modifying enzymes	oligosaccharyltransferase	OST	Core complex: ribophorin I of 66kDa, ribophorin II of 63/64kDa, OST48 of 48kDa and DAD1 of 10kDa

In the co-translational targeting pathway, firstly, the signal sequence/signal peptide (SP) is recognized by the signal recognition particle (SRP). This co-translational targeting stage needs the interaction of SRP with the SP of a nascent polypeptide chain. Subsequently, this SP-SRP complex is directed to the ER membrane by binding to the SRP receptor (SR). The SRP-nascent polypeptide chain complex is transferred by SR to the translocon in a GTP dependent manner (Rapoport, 2007), which mediates the ribosome-channel alignment and initiates translocation of the nascent polypeptide chain, as shown in Fig1.4A.

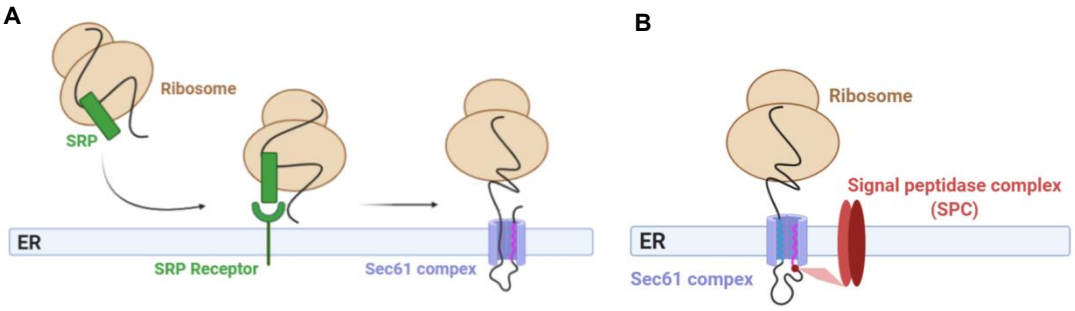


Figure 1.4 Membrane/secretory protein biogenesis in the ER

(A) Co-translational membrane translocation of membrane/secretory proteins. **(B)** Signal peptidase complex cleaves hydrophobic signal sequence/peptide during nascent polypeptide translocation.

1.5.2 Signal sequence and topogenesis of membrane proteins

In mammalian cells, two types of signal sequences are responsible for targeting and integration of polypeptides into the ER. The first one has *C*-terminal translocating activity and consists always of an anchor sequence and a cleavable signal sequence. The second one has *N*-terminal translocating activity and contains reverse signal-anchors (Goder and Spiess, 2001). The characteristic of the signal sequence is an uncharged, hydrophobic stretch of 7-25 amino acids (von Heijne, 1990). It is common to see that secretory and membrane proteins contain the signal sequence at their *N*-terminus. When this *N*-terminal signal enters the Sec61 complex, it induces the translocation of either polypeptide *N*- or their *C*-terminal sequence. Afterwards, the respective hydrophilic portion of the polypeptide is transferred through the channel of the translocon into the ER lumen. According to the so-called 'positive-inside rule' (Beltzer et al., 1991; Hartmann et al., 1989; Heijne, 1986), the orientation of signal sequence is mainly determined by charged residues flanking the apolar sequence, the hydrophobicity of signal sequence itself and folding properties of the *N*-terminal section.

The signal peptidase complex (SPC) cuts off this hydrophobic signal sequence/peptide via endoproteolytic cleavage at a specific cleavage site during translocation through the Sec61 complex (as shown in Fig1.4B). In most cases, the signal peptide is about 20-30 amino acids in length, and composed of three parts: ① n-region: composed by several basic amino acids; ② h-region: a hydrophobic stretch in the middle; ③ c-region: a slightly polar area containing a specific consensus motif. In eukaryotes, according to so called 'von Heijne' rule (von Heijne, 1990), the -1 and -3 position upstream from the cleavage site are always occupied by small,

non-charged residues, such as alanine and glycine. For bacteria, there is a well characterized consensus motif A-X-A (A: alanine, X: any amino acid residue) at the c-region of the signal peptide recognized by homologue signal peptidase I (Auclair et al., 2012). It is worth noting that signal sequence is not only essential for poly-peptide targeting to the ER but also plays an important role in topogenesis of mature protein (Higy et al., 2004).In secretory and single-spanning membrane proteins, topology is highly determined by the orientation of the signal sequence in the membrane. The membrane topology of single-spanning transmembrane proteins can be divided into 3 major types, as is summarized in Table 1.3 and presented in Fig1.5.

Table 1.3 Three different well characterized topogenic types of single-spanning membrane proteins

	Type I	**Type II**	**Type III**
Signal type	Translocating at *C*-terminus of SP		*N*-terminus translocating
Topogenic determinants	Cleavable signal+stop transfer sequence	Signal-anchor	Reverse Signal-anchor
Machinery	SRP/SR/Sec61+ signal peptidase	SRP/SR/Sec61	SRP/SR/Sec61
Orientation	$\mathbf{N}_{exo}/\mathbf{C}_{cyt}$	$\mathbf{N}_{cyt}/\mathbf{C}_{exo}$	$\mathbf{N}_{exo}/\mathbf{C}_{cyt}$

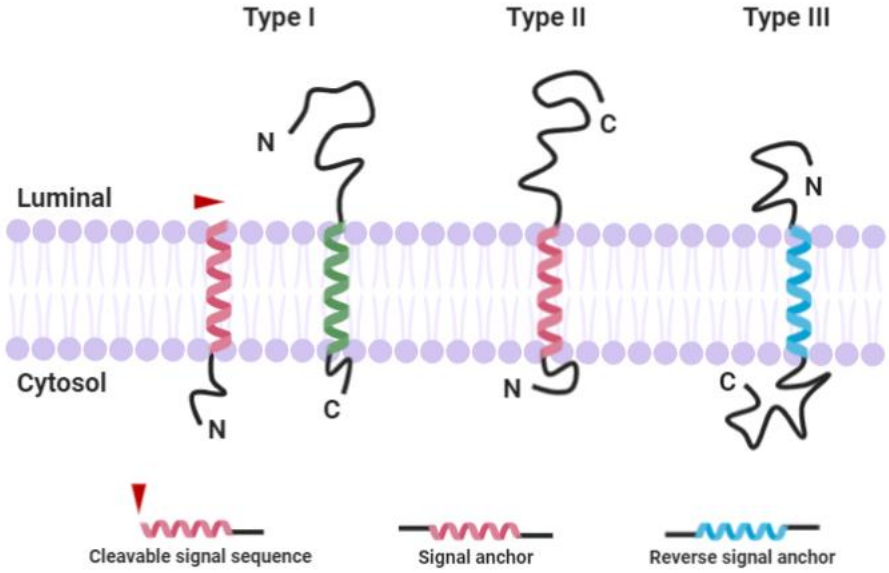


Figure 1.5 Three different types of single-spanning membrane proteins topogenesis

1.5.3 *N*-linked glycosylation

Glycosylation is the most common co-translational protein modification in eukaryotes. It has strong effects on the folding, conformation and stability of protein (Shental-Bechor and Levy, 2009), and also participates in host-pathogen interaction (Blattner et al., 2014; Carbaugh et al., 2019; Falkowska et al., 2014; Wei et al., 2012; Zheng et al., 2018). Some publications suggested that different levels of underprocessed high-mannose-type glycans affect the epitope conformation of viral glycoproteins, and also indicated the importance of *N*-glycans for the structure and function of viral glycoproteins (Tong et al., 2018).

The process of *N*-glycosylation is initially carried out in the ER and the subsequent processing occurs in the Golgi apparatus. It starts with the *en bloc* transfer of $Glc_3Man_9GlcNAc_2$ to the nascent polypeptide by an oligosaccharyltransferase (OST). OST is a multimeric complex that catalyses the *N*-glycosylation in the ER lumen (Bai et al., 2018). *N*-linked glycosylation requires the asparagine mediated tripeptide consensus motif Asn-X-Ser/Thr (X could be any amino acid but not proline). The oligosaccharide chain is attached to the *N*-linked oligosaccharide core - asparagine residue (as shown in Fig 1.6A). Only if the nascent protein folded properly, two glucose residues are removed by glucosidase I and II (as shown in Fig 1.6C). If the final third glucose residue was cut off by ER mannosidase, it signals that this nascent glycoprotein is ready for transport from the ER to the cis-Golgi (Taylor, 2011).This step is considered to act as a quality control step in the ER to monitor protein folding.

However, if this nascent protein is not folded correctly, those three glucose residues are not removed, and in consequence this nascent glycoprotein can't leave the ER. Normally, those unfolded or partially folded proteins require chaperone proteins (like calnexin/calreticulin) to assist their folding.

Glycoproteins without ER retention signal transit from the ER to cis-Golgi when correctly folded. Some glycosyltransferases and glycosidases in the cis-Golgi catalyze further modifications that is addition and removal of sugar residues. Finally, at the medial portion of the Golgi apparatus, some sugar residues are added to the core glycan structure mediated by glycosyltransferases, giving rise to the three main types of *N*-glycans, as shown in Fig 1.6B.

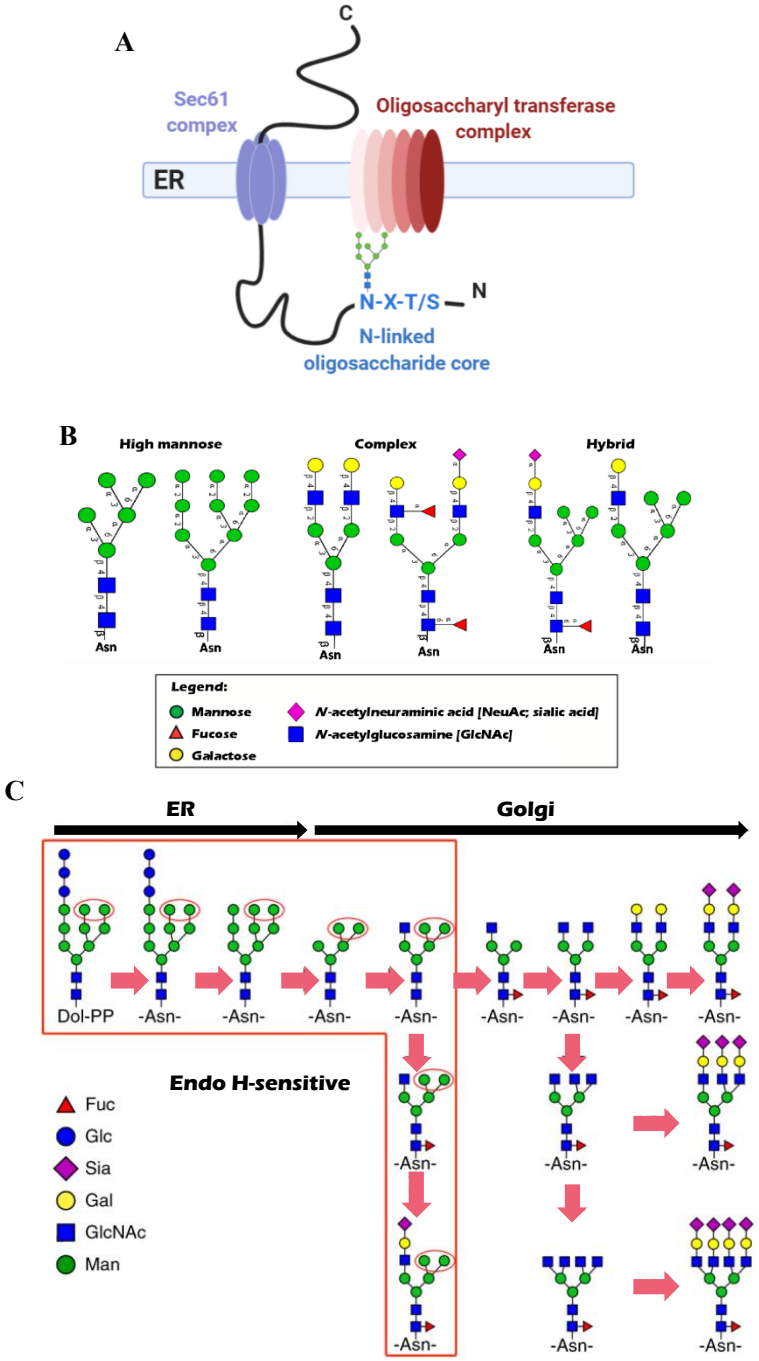


Figure 1.6 *N*-glycosylation

(A) Schematic representation of *N*-glycosylation in the ER; **(B)** Three main types of *N*-glycans, modified from (Sethi and Fanayan, 2015); **(C)** The processing of *N*-linked glycans from the endoplasmic reticulum to Golgi apparatus, modified from (Cao et al., 2018).

For pestiviruses, some publications showed that the glycosylation of envelope proteins can affect virulence. For CSFV, substitution at *N*-linked glycosylation site N116 in E2 induced viral attenuation (Risatti et al., 2007a). Removal of a *N*-linked glycosylation site of E^{rns} reduced viral infectivity (Sainz et al., 2008). Moreover, single mutation of N594A or combined N500A/N513A substitutions in E1 also resulted in CSFV attenuation (Fernandez-Sainz et al., 2009). These data indicated that the *N*-glycosylation of viral envelope protein is important for

folding, conformation and stability of the proteins, which can result in changes in some processes in the viral life cycle like receptor recognition, viral membrane fusion step and immune evasion.

1.5.4 Disulfide bond formation

Disulfide bond formation is another essential co-translational modifications found in proteins that enter the secretory pathway. The essence of disulfide bond formation is building up covalent linkages within and between proteins via the oxidation of sulfhydryl (-SH) groups which have their cysteine residues in close enough proximity. For eukaryotes, the formation of disulfide bonds (also known as oxidative protein folding) mainly occurs in the ER and inter-membrane space of mitochondria (Saaranen and Ruddock, 2019). A family of protein-disulfide isomerases (PDIs) is required for formation of correct disulfide bonds in secretory and cell surface proteins (Farquhar et al., 1991; Laboissiere, 1995). Normally, simple disulfide bond formation can be divided into two steps, dithiol oxidation and disulfide isomerization (as shown in Fig1.7). In eukaryotes, the first step is carried out by sulfhydryl oxidases (like the Ero1, ERV/ALR, and QSOX families in human), whereas the next step is catalyzed by PDI in the ER. Those enzymes are required to not only introduce disulfides between proximal cysteines but also to collapse disulfides which are not present in the final mature proteins to properly ensure the final disulfide formation (Bulleid, 2012; Jansens et al., 2002).

Disulfide bond formation also plays an essential role in the life cycle of viruses. For HCV, some publications demonstrated that cysteine mutations in E2 drastically reduce virus infectivity (McCaffrey et al., 2012). The ‘CxxC’ motif is a key feature for the recognition by protein disulfide isomerase to mediate the isomerization of disulfide bonds in HCV E1 (Castelli et al., 2017; Wahid et al., 2013). The cysteine residue at position 128 of core protein was shown to be a dominant disulfide bond formation site in terms of HCV-like particle production. The studies indicated that this disulfide bond is critical for the HCV virion (Kushima et al., 2010). More recently, a low-molecular-weight PDI inhibitor called origamicin was shown to negatively impact HCV replication by inducing incorrect proteins folding, causing an imbalance in cellular homoeostasis and induction of stress responses (Ozcelik et al., 2018).

The cysteine residue at position 171 of the E^{rns} of CSFV was shown to be critical for the formation of E^{rns} homodimers. Substitution of this cysteine leads to attenuation of the virus (Tews et al., 2009a). Furthermore, restoring the E^{rns} dimerization via cysteine residue downstream of position 171 can also partially recover the virulence of CSFV (Tucakov et al., 2018). This finding further supports the connection between virulence and E^{rns} dimerization.

The relationship between different modifications is possibly competitive. An early study showed that HCV E1 glycosylation can impair the formation of disulfide bond in E1-E2 heterodimer (Tong et al., 2018). The reverse is also true that disulfide bond involving C306 in E1 is most likely be prevented by glycosylation at N305 because of spatial hindrance (Meunier et al., 1999). Interestingly, removal of this *N*-linked glycosylation site increases the immunogenicity of soluble E1 (Fournillier et al., 2001), which also proves that the process of post-translational modification and nascent protein maturation is complex and closely interrelated.

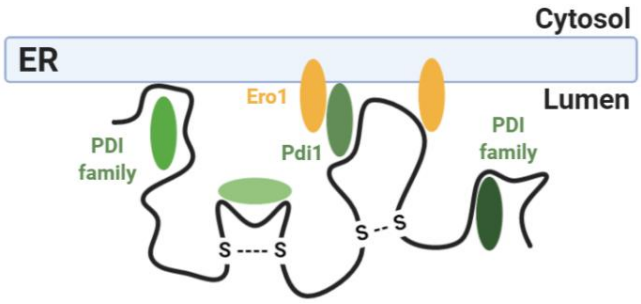


Figure 1.7 Schematic representation of disulphide bond formation

In the ER lumen, disulfide bond formation is reversibly catalyzed by protein disulfide isomerases (such as Pdi1) with Ero1 providing oxidizing equivalents.

1.5.5 Cargo exit from ER mediated by COPII

Secretory and membrane proteins that are folded correctly and not ER-resident proteins, are transported from the ER to the Golgi apparatus. Both the 'quality control' mechanism of ER and the unfolded protein response (UPR) protects the ER folding environment by detecting and responding to the presence of misfolded proteins in its lumen (Ron and Walter, 2007). When misfolded polypeptides accumulate continuously, they trigger the activation of the ER-associated degradation (ERAD) pathway to be degraded. In eukaryotic cells, the transport of newly synthesized proteins out of the ER is carried out via the coat protein complex II (COPII) vesicles (Hughes and Stephens, 2008). The COP II complex consists of five proteins, including a small GTPase Sar1, heteromeric complexes Sec23-Sec24 and Sec13-Sec31 (as shown in Fig 1.8). The components of the COP II complex are recruited and assembled at ER-exit sites (ERES) where COP II complex vesicles bud off. ERES is present in most eukaryotic cells (Anelli and Sitia, 2008).

There are some indications that transmembrane cargo directly binds to COPII subunits Sec24 mediated via some specific signals at it *C*-terminus (Nishimura and Balch, 1997). However, soluble secretory cargo can not be bound directly by coat subunits, since there is no consensus, but evidence suggests that transmembrane receptors might link certain luminal cargo to COPII. Several ER export signals on transmembrane protein have been identified, as shown in Table 1.4.

Table 1.4 Characterized ER export signals of membrane protein

	Export Signal	Representative Protein
Di-acidic motifs	IYTDIEMNRLGK	VSV-G
	ANSFCYENEVAL	Kir2.1
	QSPIQLKDLESQI	Sys1p
	AEKMDIDTGR	Gap1p
Di-hydrophobic motifs	YIMYRSQQEAAAKKFF	ERGIC53
	YYMFRINQDIKKVKLL	Emp46p
	YLRRFFKAKKLIE	p24d1

YQPDDKTKGILDR	Erv41p
KLFYKAQRSIWGKKSQ	Erv46p

Underlined residues are required for export from the endoplasmic reticulum (ER).

1.5.6 ER retention signal

In eukaryotes, there are mechanisms to selectively retain proteins in the ER or Golgi apparatus. Selective export and retrieval of proteins between the endoplasmic reticulum (ER) and Golgi apparatus is vital for eukaryotic cell function. Both the ER and the Golgi apparatus maintain specific resident membrane proteins and lipids to achieve their structural and functional integrities. There is a well-known carboxylterminal retention signal Lys-Asp-Glu-Leu (KDEL) mediating an essential step in the retrieval of ER luminal proteins from the Golgi by the KDEL receptor in a pH-dependent manner (Brauer et al., 2019). Coat protein complex I (COPI) vesicles are known to mediate retrograde protein traffic from Golgi to ER. Some retention signals which interact with the COPI complex, such as the well-known *C*-terminal KKXX or KXKXX motif, function specifically in retrieving ER-resident membrane proteins from the Golgi via interacting with COPI coat proteins (Jackson et al., 1990; Nilsson et al., 1989). For KDEL-bearing proteins, deletion of the retrieval motif leads to their secretion. A growing number of ER retention signals have been identified. However, in some cases, the retention signals do not fully retain the protein in the cell, these proteins are still secreted at different rates and in generally very slow. The retention mechanism is complex and varies, not immobilization. HCV envelope protein E1 and E2 both contain retention signal of their own, located within their transmembrane domains. These signals consist of some charged residues (Ciczora et al., 2005; Cocquerel et al., 1999; Cocquerel et al., 1998; Duvet, 1998). Recently, a publication showed that this situation is also true for BVDV E2 (Radtke and Tews, 2017). Pestiviral glycoprotein E^{rns} was also shown to contain a retention signal located in its unusual *C*-terminal membrane anchor (Burrack et al., 2012; Tews and Meyers, 2007). Besides the classical retrieval mechanisms described for proteins with a KDEL or a KKXX signal, retention of native proteins without retrieval can also occur in the ER.

Chapter 2: Objectives of the study

Pestiviruses contain three envelope proteins: E^{rns}, E1, and E2. Among them, E1 is the worst characterized. There is still no good specific antibody for E1, making the detection and further characterization of this protein difficult. Therefore E1 has mainly been analysed in the context with the other two envelope proteins. My aim for the present study was to take advantage of tagged E1 for detection to be able to address the following questions:

1. Subcellular Localization of E1

Pestiviruseses are known to bud at intracellular membranes, since all the glycoproteins are not accessible on the surface of transfected or infected cells. It indicates that retention signal(s) have to be present that ensure accumulation of the envelope proteins in defined intracellular sites. In this study, I analyzed whether E1 alone stay within the cell, and if so, in which compartment E1 is concentrated. Tagged E1 was analyzed via co-localization with marker proteins for different cellular compartments (e.g. ER and Golgi apparatus) via confocal microscopy.

2. Membrane topology of E1

The *N*-terminus of E1 is generated by signal peptidase cleavage at the unusual E^{rns} membrane anchor/E1 site, so that the *N*-terminus of E1 should be located in the ER lumen. The length of the hydrophobic region at the *C*-terminus of E1 is too long for a normal single span transmembrane domain, so that the membrane topology of the mature E1 protein is difficult to predict and studies on the membrane topology of E1 are still missing. To analyse the membrane topology of E1, we established a plasmid construct for expression of a double tagged E1 protein with one tag located at the *N*-terminus and the other at the *C*-terminus. By using selective permeabilization combined with fluorescence microscopy, we wanted to determine whether the *C*-terminus of E1 is accessible from the cytosolic side. To verify the results we learned from the selective IF, we used a so-called Avitag biotinylation assay as a second technical approach. Furthermore, we also tested the membrane topology of E1 in E1E2 fusion proteins before signal sequence cleavage at the E1/E2 site to see whether the fusion of the two proteins has impact on E1 topology.

3. Retention of E1

Since pestiviruses bud intracellularly, E1 has to stay within the cell either by a retention mechanism or via interaction with other viral proteins like E2. If E1 is identified to contain an intracellular retention signal of its own, we would like to do further characterization of the relevant element to search for the respective sequence motif for this retention behaviour. To prove the crucial sequence that we found is really responsible for the retention of E1, it is fused to a typical surface protein (eg. VSV-g). Analyses were done using FACS analysis and IF to see whether this fusion protein demonstrates intracellular retention or not. To hunt for the critical residue(s) of the retention signal of E1, we used mutagenesis or deletion analysis combined with FACS and IF assay. If the critical sites are identified, we also would like to test those sites in the context of the BVDV CP7 infectious clone, to investigate the influence on the live virus when the retention of the viral envelope protein E1 is impaired.

4. Prerequisites of E1 oligomer formation and E1-E2 heterodimerization

It is well known that E1 forms disulphide linked heterodimers with E2 that are needed for virus infectivity and crucial for viral entry. Due to the absence of specific antibodies for E1, so far it has not been shown whether E1 can also form disulphide linked homodimers or oligomers. Basically, covalent disulphide linkages can be formed by cysteine residues. In the present study, we want to hunt for the critical site(s) for E1 oligomerization (if possible) and E1-E2 heterodimerization. To achieve this purpose, we established several E1, E2 mutants lacking the cysteine residues supposed to be involved in disulphide bond formation. In addition, it was intended to characterize the E1E2 interaction platform upon the co-expression of E1 mutants with E2 mutant. Furthermore, if the sites important for E1 oligomerization (if possible) and E1-E2 heterodimerization are identified, we want to further test those sites in the live virus with the help of BVDV CP7 infectious clone bearing specific mutations, to see the effect of disulfide bond formation within E1/E2 on the viral life cycle.

Taken together, the aim of this study is to analyse the structure and function of the pestivirus E1 protein at the molecular level. These results will help to further understand several important processes in the pestiviral life cycle like intracellular budding, protein retention and oligomerization.

Chapter 3: Materials

3.1 Cells

The following eukaryotic cell lines were used for the transient expression of proteins and for the cultivation of viruses:

Cell lines	Origin
BHK-21(Baby hamster kidney cell)	T. Rümenapf (Universität Gießen)
RK-13(Normal Rabbit Kidney Epithelial Cells)	ATCC#: CCL-37
MDBK-B2(Madin-Darby Bovine Kidney Epithelial Cells)	ATCC#: CCL-22

3.2 Viruses

The following viruses were used for the transient expression of proteins and for the cultivation of viruses:

Name	Origin
BVDV CP7	Received from Cornell University, Ithaca, USA
Vaccinia Virus MVA T7	Received from Dr. Gerd Sutter, LMU München, Germany

3.3 Bacterial strains

Bacteria strains *E.coli* HB101 and Top10F' were used for the amplification of the plasmids in this study.

HB101 genotype: *sup*E44, Δ (*mcr*C-*mrr*), *rec*A13, *ara*-14, *pro*A2, *lac*Y1, *gal*K2, *rps*L20, *xyl*-5, *mtl*-1, *leu*B6, *thi*-1

Top 10 F' genotype: *mcr*A, Δ (*mcr*BC-*hsd*RMS-*mrr*), *end* A1, *rec*A1, *rel*A1, *gyr*A96, F 80*lac*Z Δ M15, deoR, nupG, araD139, F (lacIq, Tn10 (Tetr)), galU, D lacX74, galK, D (araleu) 7697

3.4 Medium

The media as follows were used for cell culture or experiments with mammalian cells.

Name	Component		Manufacturer
ZB5d	MEM Eagle (Hank's salts, Sigma M4642)	5.32 g/L	Cell bank
	MEM (Earles'salts, Gibco/Invitrogen 61100)	4.76 g/L	
	NaHCO3 (Roth 6885.1)	1.25 g/L	
	NEA (Biochrome K 0293, 100x)	10 ml/L	
	Na pyruvate (Merck 1.06619)	120 mg/L	

ZB5	MEM Eagle (Hank's salts, Sigma M4642)	5.32 g/L	Cell bank
	MEM (Earles'salts, Gibco/Invitrogen 61100)	4.76 g/L	
	NaHCO3 (Roth 6885.1)	1.25 g/L	
	NEA (Biochrome K 0293, 100x)	10 ml/L	
	Na pyruvate (Merck 1.06619)	120 mg/L	
	FCS	100ml/L	
Opti-MEM			Gibco, USA

The medium as follows were used for bacterial culture or transformation.

Name	Component	
LB-Medium	Bacto-Tryptone	10g
	Bacto-Yeast-Extract	5g
	NaCl	5g
	Distilled water	1L
LB-Ampicillin	Agar in LB-Ampicilin	
LB-Ager	100 µg/ml Ampicillin in LB-Medium	
LB++	KCl	10mM
	$MgSO_4$	20mM
	in LB-Medium	

3.5 Antibodies

The following antibodies were used for western blot analysis, immunofluorescence and immunoprecipitations. The amounts or dilutions used are given in each case.

	Name	Antigen	Host	Dilution	Origin
Primary antibody	α-HA	HA tag	Mouse	1:1000(IF)	Abcam, UK
	α-HA	HA tag	Rabbit	1:1000(WB)	Abcam, UK
	α-V5	V5 tag	Mouse	1:1000(IF)	Invitrogen, USA
	α-Flag M2	Flag tag	Mouse	1:1000(IF)	Sigma-Aldrich®, USA
	α-AU1	AU1 tag	Mouse	1:1000(IF)	Abcam, UK
	α-BVDV MIX	BVDV E2	Mouse	1:100(IF) 100µl(IP)	FLI-Tübingen
	α-WB 214	BVDV E2	Mouse	1:250(WB)	Weybridge, England

α-f48	CSFV E2	Mouse	1:100(IF)	FLI
α-Code 4	NS3	Mouse	1:10(IF)	FLI
α-VSVg	VSVg [ecto-]	Rabbit	1:100(IF)	PD Dr. Stefan Finke

	Name	Antigen	Host	Dilution	Origin
Secondary antibody	α-Mouse-FITC	Mouse IgG	Mouse	1:100 (IF/FACS)	Dianova, Hamburg
	α-Mouse-Alexa-Fluor-448	Mouse IgG	Mouse	1:250(IF)	Thermo Scientific, USA
	α-IgG1-Cy3	Mouse-IgG1	Mouse	1:100(IF)	Dianova, Hamburg
	α-IgG2a-Alexa-Fluor-488	Mouse-IgG2a	Mouse	1:250(IF)	Thermo Scientific, USA
	α-Rabbit-Alexa-Fluor-555	Rabbit IgG	Mouse	1:250(IF)	Thermo Scientific, USA
	α-Avidin-PO	Biotin		1:20000(WB)	Invitrogen, USA
	α-Mouse-PO	Mouse IgG light chain	Goat	1:10000(WB)	Dianova, Hamburg
	α-Rabbit-PO	Rabbit IgG light chain	mouse	1:10000(WB)	Dianova, Hamburg

3.6 Chemicals

All chemicals used in this study were obtained in analytical quality.

Name	Manufacturer
Acetic acid	Hoechst, Dortmund
Aceton	Roth, Karlsruhe
Acrylamid (40 %) (29:1)	AppliChem, Darmstadt
Agarose	Gibco, Scotland
Ampicillin	Ampicillin
APS	Merck, Darmstadt
Bacto-Agar	Becton Dickinson, USA
Bacto-yeast extract	Becton Dickinson, USA

Bacto-tryptone	Becton Dickinson, USA
BSA	Roche, Mannheim
$CaCl_2$	Roth, Karlsruhe
Coomassie Brilliant Blue G250	Serva, Heidelberg
DAPI	Serva, Heidelberg
DEPC	Sigma, Munich
Digitonin	Serva, Heidelberg
dNTPs	New England BioLabs, Frankfurt
DTT	Roche, Switzerland
EDTA	Roth, Karlsruhe
Ethanol	Roth, Karlsruhe
FCS	Biochrom, Berlin
GelRed	Phenix Research, USA
Glucose	Merck, Darmstadt
Glycine	Roth, Karlsruhe
Glycerin (87 %)	Roth, Karlsruhe
H_3PO_4	Roth, Karlsruhe
HCl	Roth, Karlsruhe
HEPES	Sigma, Munich
KCl	Roth, Karlsruhe
KH_2PO_4	Roth, Karlsruhe
Magermilchpulver	Hobbybäcker Versand, Bellenberg
Methanol	Roth, Karlsruhe
β-Mercaptoethanol	MP Biomedicals, Heidelberg
$MgCl_2$	Merck, Darmstadt
$MgSO_4$	Roth, Karlsruhe
Mowiol	Roth, Karlsruhe
NaCl	Roth, Karlsruhe
Na_2CO_3	Roth, Karlsruhe
Na_2HPO_4	Roth, Karlsruhe
NaOH	Roth, Karlsruhe
Paraformaldehyde	Sigma, Munich

Phenol, tris-saturated	Roth, Karlsruhe
Saponin	Roth, Karlsruhe
SDS	Roth, Karlsruhe
Sucrose	Roth, Karlsruhe
TEMED	Serva, Heidelberg
Tris	Invitrogen, USA
Tricin	Roth, Karlsruhe
Triton-X100	Sigma, Munich
Tween-20	Sigma, Munich

3.7 Commercial Kits

Name	Manufacturer
BigDye® Terminator v.3.1 Cycle Sequencing Kit	Applied Biosystems, Weiterstadt
SuperSignal™ West Pico Chemiluminescent substrate Kit	Thermo Fisher Scientific Inc.
RiboMAX™ Large Scale RNA Production System- T7	Promega, Mannheim
Nucleo-Spin®-Gel and PCR clean- up	Macherey-Nagel, Düren
QIAGEN Plasmid Midi Kit	QIAGEN, Germany
SuperScript™ III One-Step RT-PCR System	Thermo Fisher Scientific Inc.
Trizol® Reagent	ambion RNA, USA

3.8 Prepared Solutions and Buffer

Name	Composition		Purpose
Blocking buffer(TGG 2% BSA)	9% (w/v) 2%	Glucose BSA in PBS	Avitag Biotinylation analysis (WB)
Blocking buffer	5 %(w/v)	Magermilchpulver in PBS-T	Normal Western Blot
Coomassie-Stock	50% (v/v) 10% (v/v) 0.25% (w/v)	Methanol Acetic acid Coomassie blue	RIP gel stainning

Digitonin solution	20 mM	HEPES (pH 6.9)	Selective permeabilization assay
	0.3 M	Sucrose	
	0.1 M	KCl	
	2.5 mM	$MgCl_2$	
	1.0 mM	EDTA	
	5 µl/ml	Digitonin	
Decoloring solution	20% (v/v)	Methanol	RIP gel decoloring
	3%	Glycerol	
Fixing solution	30%	Methanol	RIP gel fixing
	10% (v/v)	Acetic acid	
Jagow anode buffer (10x)	2 M	Tris (pH 8.9)	SDS-PAGE
Jagow cathode buffer (10x)	1 M	Tris	SDS-PAGE
	1 M	Tricine	
	1% (w/v)	SDS	
		pH 8.25	
Lämmli-Puffer	0.25 M	Tris	Preparation for WB transfer buffer
	1.925 M	Glycine	
		pH 8.3	
Solution I	50 mM	Tris	Plasmid mini prep
	10 mM	EDTA	
	50 mM	Glucose	
		pH 8.0	
Solution II	0.2 M	NaOH	
	1% (v/v)	Triton-X100	
Solution III	3 M	NaAcetat (pH 4.8)	
Mowiol DAPI	6 g	mowiol	Confocal sample fixation
	6 ml	aqua bidest	
	12 ml	0.2 M Tris buffer (pH 8)	
	0.1%	DABCO	
	0.2%	DAPI	
PBS	0.5 mM	$MgCl_2$	Cell culture, washing step for IF and FACS
	0.9 mM	$CaCl_2$	

PBS-A	137 mM	NaCl	
	2.7 mM	KCl	
	7.4 mM	Na_2HPO_4	
	1.5 mM	KH_2PO_4	
PBS-Tween	137 mM	NaCl	Washing step for WB
	2.7 mM	KCl	
	7.4 mM	Na_2HPO_4	
	1.5 mM	KH_2PO_4	
	0.2% (v/v)	Tween-20 in PBS-A	
4% PFA	4% (w/v)	paraformaldehyde in PBS-A	Fixation for IF/FACS samples
SDS-Sample-Buffer	62.5 mM	Tris-HCl (pH 6.8)	SDS-PAGE Samples treatment
	2% (w/v)	SDS	
	10% (v/v)	Glycerol	
	6 M	Urea	
	5% (v/v)	Mercaptoethanol	
	0.01% (w/v)	Brpmphenol blue	
	0.01% (w/v)	Phenol red	
SDS-Concentration-Gel	0.075% (w/v)	SDS	SDS-PAGE
	10% (w/v)	Acrylamide (29: 1)	
	0.75 M	Tris-HCl (pH 8.45)	
	0.08% (w/v)	GSP	
	0.09% (w/v)	TEMED	
SDS-Separation-Gel	0.1% (w / v)	SDS	SDS-PAGE
	8%,10%(w/v)	acrylamide (29: 1)	
	1 M	Tris-HCl (pH 8.45)	
	5.5%	Glycerol	
	0.08% (w/v)	GSP	
	0.09% (w/v)	TEMED	
TAE-Buffer	40 mM	Tris	Electrophoresis
	5 mM	Na acetate	

	1 mM	EDTA	
		pH 7.8	
Transfer Buffer (Westernblot)	18% (v/v)	Methanol	WB transfer step
		in Lämmlipuffer	
	1% (v/v)	Triton-X100	
Cell-Lysis-Buffer	2mM	EDTA	Cell-Lysis
		in PBS	

3.9 Regent

Name	Manufacturer
Pfu DNA Polymerase	Promega Corporation, USA
Platinum™ Taq DNA Polymerase	Invitrogen, USA
T4 DNA Ligase	New England BioLabs Inc., Frankfurt
Restriction enzymes	New England BioLabs Inc., Frankfurt
RNase	Serva Electrophoresis, Heidelberg
Lipofectamin™ 2000	Invitrogen, USA
1 Kb plus DNA Ladder	Invitrogen, USA
170 kDa PageRuler Prestained Protein Ladder	Thermo Scientific, USA
250 kDa PageRuler Prestained Protein Ladder	Thermo Scientific, USA
SepharoseA	GE Healthcare, Freiburg
Gel Loading Dye, Purple (6X)	New England BioLabs Inc., Frankfurt

3.10 Radio activity chemicals

Name	Manufacturer
Tran^{35}S-Label (1175 Ci/mmol,10.5 mCi/ml)	MP Biomedicals, USA
^{35}S-Methionin (1175 Ci/mmol,10 mCi/ml)	MP Biomedicals, USA
14C molecular weight standard CFA626	GE Healthcare, Munich
14C molecular weight standard CFA645	GE Healthcare, Munich

3.11 Primers

The primers listed below were used for PCR and sequencing and were ordered from primer synthesis company-Metabion (Munich). Primers provided desalinated and lyophilized. By default, a stock solution of the oligonucleotides was used with a concentration of 100mM (0.1 nmol/µL). Working solutions with a concentration of 1mM (10 pmol/µL) were then prepared from the stock. The sequences of the oligonucleotides are listed in the 5'-3' direction.

Name	Sequence (5'-3')
yM_1for	GCATTCTTGGTTTCTCTGGTGAAGGTAGTG
yM_1rev	CTACCTTCACCAGAGAAACCAAGAATGC
yM_2for	GGAAAAGACGGCATCCACCTTCAACTACAC
yM_2rev	GTAGTTGAAGGTGGATGCCGTCTTTTCC
yM_3for	CTCATAACAGGGAGGCAAGGGTACCCAG
yM_3rev	CTGGGTACCCTTGCCTCCCTGTTATGAG
yM_4for	CCAGTCACAATGGGATCCTCTCCCTATTGTGAG
yM_4rev	CTCACAATAGGGAGAGGATCCCATTGTGACTGG
yM_5for	GATCGTACCCATACGACGTCCCAGACTACGCT
yM_5rev	AGCGTAGTCTGGGACGTCGTATGGGTACGATC
yM_6for	GAGAATTCACGCGTGCTACCTCTAGAATG
yM_6rev	CATTCTAGAGGTAGCACGCGTGAATTCTC
yM_7for	GTACGGCAAACCGATTCCGAACCCGCTGCTGGGCCTGGATAGCACCGG
yM_7rev	CCGGTGCTATCCAGGCCAGCAGCGGGTTCGGAATCGGTTTGCCGTAC
yM_8for	GTCACAATGGGATCCTCTCCCTATTGTG
yM_8rev	CACAATAGGGAGAGGATCCCATTGTGAC
new HA F	GATC C TACCCATACGACGTCCCAGACTACGCT G
new HA R	CAGCGTAGTCTGGGACGTCGTATGGGTAGGATC
yM-HA tag A1	GTGATGCGGTTTTGGCAGTACAC
yM-13HA tag A2	AGCGTAATCTGGAACATCGTATGGGTATCCCATTGTGACTTGAAAGAAAACCAGGG
yM-19HA tag A2	AGCGTAATCTGGAACATCGTATGGGTA-TGCATATGCCCCAAACCATGTCTTACTC
yM-HA tag B1	TACCCATACGATGTTCCAGATTACGCTGCCTCTCCCTATTGTGAGGTAGAACG
yM-HA tag B2	CAAATGCGCGTCAGATCTTTTAACGCC
yM-FLAG tag A1	GGTAGAACGGAAGCTTGGTTACATCTGG

yM-FLAG tag A2	CTTATCGTCGTCATCCTTGTAATC-CCCTTGCGCCCCTGTTATGAG
yM-FLAG tag B1	GATTACAAGGATGACGACGATAAG-TACCCAGACTGCAAACCCGGC
yM-FLAG tag B2	CCACACCACTCACAAGACTCAACGG
yM-QC-9 for	CAGACTACGCTGCCTCTCCCTATTGTG
yM-QC-9 rev	CACAATAGGGAGAGGCAGCGTAGTCTG
yM-QC-10for	CTCATAACAGGGGAGGAAGGGGATTAC
yM-QC-10rev	GTAATCCCCTTCCTCCCCTGTTATGAG
yM-QC-11for	CTGGTGACCGGGGCATGATCTAGAGTC
yM-QC-11rev	GACTCTAGATCATGCCCCGGTCACCAG
yM-QC-12for	CTCATAACAGGGAGGCAAGGGGATTAC
yM-QC-12rev	GTAATCCCCTTGCCTCCCTGTTATGAG
yM-QC-13for	CCTGGTTTGGTAGATATGCCGGCAC
yM-QC-13rev	GTGCCGGCATATCTACCAAACCAGG
E1 182Δ QC for (pCR54)	GGC CAA GTG TTA____GGT ATA CTG TGG
E1 182Δ QC rev (pCR54)	CCA CAG TAT ACC____TAA CAC TTG GCC
E1 177E QC for (pCR60)	GTG AAG GTA GTG GAG GGC CAA GTG TTA
E1 177E QC rev (pCR60)	TAA CAC TTG GCC CTC CAC TAC CTT CAC
yM-Avitag-E1 F	GCA CAG AAA ATT GAA TGG CAT GAG - TGAGCGGCCGCTTCGAGC
yM-Avitag-E1 R	TTCGAAAATATCATTCAACCCGGACCC-CCCTTGCGCCCCTGTTATGAGC
yM-Avitag QC F	GAATGATATTTTCGAAGCACAGAAAATTGAATGG
yM-Avitag QC R	CCATTCAATTTTCTGTGCTTCGAAAATATCATTC
yM-5' Avitag E1 F	TCGAAAATATCATTCAACCCGGACCCTCCCATTGTGACTTGAAAGAAAACCAGGG
yM-5' Avitag E1 R	AGCACAGAAAATTGAATGGCATGAG-GCCTCTCCCTATTGTGAGGTAGAACG
yM-QC-16 F	GATAATACCATGGGGAAGGCCCTGTTG
yM-QC-16 R	CAACAGGGCCTTCCCCATGGTATTATC
yM-QC-17 F	CTCATAACAGGGAGACAAGGGGGG
yM-QC-17 R	CCCCCCTTGTCTCCCTGTTATGAG
yM-BirA F	CCGCTCGAG-ggATGAAGGATAACACCGTGCCACTGAAATTG
yM-BirA R	CGACGCGT-TTATTTTTCTGCACTACGCAGGGATATTTCACC

yM-HEVG XhoI F	CCGCTCGAG-GGATGGAGAAAGCCCTATTGGCCTGG
yM-HEVG NotI R	ATAGTTTAGCGGCCGCTTACTTTCCAAGTCGGTTCATCTCTATGT CTG
yM-E^{rns}(BT) NotI R	ATAGTTTAGCGGCCGCTCATGCATATGCCCCAAACCATGTC
yM-3' Avi E^{rns} F	AGCACAGAAAATTGAATGGCATGAG-TGAGCGGCCGCTTCGAG
yM-3' Avi E^{rns} R	TCGAAAATATCATTCAACCCGGACCCTGCATATGCCCCAAACCAT GTCTTAC
yM-Avi MluI and XbaI F	CGACGCGTGGGTCCGGGTTGAATGATATTTTCGAAGCACAGAAA ATTGAATGGCATGAGTGA
yM-Avi MluI and XbaI R	GCTCTAGATCACTCATGCCATTCAATTTTCTGTGCTTCGAAAATAT CATTCAACCCGGACCC
yM-QC-18F	CTCATAACAGGGAGGCAAGGGTACC
yM-QC-18R	GGTACCCTTGCCTCCCTGTTATGAG
yM-3'R for pB153	CGACGCGTTGCATATGCCCCAAACCATGTCTTACTC
yM-3'R for pB154	CGACGCGTTGCATATCGCCCAAACCATGTCTTACTC
yM-QC-19F	GCTAGCCTCGAGGGATGAAGGATAAC
yM-QC-19R	GTTATCCTTCATCCCTCGAGGCTAGC
yM-HEVG XhoI F	CCGCTCGAG-GGATGGAGAAAGCCCTATTGGCCTGG
yM-HEVG NotI R	ATAGTTTAGCGGCCGCTTACTTTCCAAGTCGGTTCATCTCTATGT CTG
yM-nB F	CCGCTCGAG-ggATGAAGGATAACACCGTGCCACTGAAATTG
yM-TA Hseq F	CTTCTCGCACTACTGGCGGCATTCTTGGTTTGTCTGGTGAAGGTA GTG
yM-TA Hseq R	CGCCAGTAGTGCGAGAAGAGTGGTTGTAGCAGCGGTCCAAAT
yM-QG Hseq F	CTTCTCGCACTACTGGCG-GGTATACTGTGGCTGATGCTCATAACAG
yM-QG Hseq R	CGCCAGTAGTGCGAGAAGTTGTAACACTTGGCCTCTCACTACCT TC
yM-QC-20F	GATGCGGCGCAAATAGTCATTGGAG
yM-QC-20R	CTCCAATGACTATTTGCGCCGCATC
yM-VSVG EcoRI 5'F	GGCAAAGAATTCCACCATGAAGTGCC
yM-VSVG MluI 3'R	CGACGCGTAGAGGCAATAGAGCTTTTCCAACTACTGAAC
yM-E1tmd MluI 5'F	CGacgcgtGGAGGAGGAGGAAGTACCACTGC
yM-E1tmd NotI 3'R	ATAGTTTAGCGGCCGCTCACCCTTGCGCCCCTGTTATG
yM-QC-1st G to L F	GGTAGTGAGACTGCAAGTGTTACAAGG

yM-QC-1st G to L R	CCTTGTAACACTTGCAGTCTCACTACC
yM-QC-1st G to A F	GGTAGTGAGAGCGCAAGTGTTACAAGG
yM-QC-1st G to A R	CCTTGTAACACTTGCGCTCTCACTACC
yM-QC-2nd G to L F	CCAAGTGTTACAACTGATACTGTGGCTG
yM-QC-2nd G to L R	CAGCCACAGTATCAGTTGTAACACTTGG
yM-QC-2nd G to A F	CCAAGTGTTACAAGCGATACTGTGGCTG
yM-QC-2nd G to A R	CAGCCACAGTATCGCTTGTAACACTTGG
yM-QC-2G to L F	GTGAGACTGCAAGTGTTACAACTGATACTG
yM-QC-2G to L R	CAGTATCAGTTGTAACACTTGCAGTCTCAC
yM-QC-2G to A F	GTGAGAGCTCAAGTGTTACAAGCTATACTG
yM-QC-2G to A R	CAGTATAGCTTGTAACACTTGAGCTCTCAC
yM-E1 3'GS R	ACTTCCTCCTCCTCCCCCTTGTCTCCCTGTTATGAGCATCAG
yM-E2 5'GS F	GGAGGAGGAGGAAGTTACCCAGACTGCAAACCCGGCTTTTC
yM-E2 3'NotI R	ATAGTTTAGCGGCCGCTCAGATGTATC
yM-5'E1tmd-6L B2 R	CAGTAGCAGTAGGAGAAGCAGACAAACCAAGAATGCAGTGGTTGTAG
yM-5'E1tmd-6L A1 F	CTTCTCCTACTGCTACTG-CAAGTGTTACAAGGTATACTGTGGCTGATG
yM-3'E1tmd-6L B2 R	CAGTAGCAGTAGGAGAAGGAGCATCAGCCACAGTATACCTTGTAAC
yM-3'E1tmd-6L A1 F	CTTCTCCTACTGCTACTG-TGAGCGGCCGCTTCGAG
yM-QC-21F	GCGTTAAAAGATATCACGCGCATTTGGAC
yM-QC-21R	GTCCAAATGCGCGTGATATCTTTTAACGC
yM-E1 TMD 3'Truncation R	ATAGTTTAGCGGCCGCTCATTGTAACACTTGGCCTCTCACTACCTTC
yM-E1 TMD 5'Truncation F	GCATTTGGACCGCTGCTACAACC
yM-E1 TMD 5'Truncation R	CTGTTATGAGCATCAGCCACAGTATACC
yM-QC-21F (for all leucine)	GCATTCTTGGTTTCTCTGGTGCTGGTAG
yM-QC-21R (for all leucine)	CTACCAGCACCAGAGAAACCAAGAATGC
yM-5T MluI F	CGacgcgtCAAGTGTTACAAGGTATACTGTGGCTGATG
yM-NotI E1ecto 3'R	TAAACTATGCGGCCGCTCAGGTTGTAGCAGCGGTCCAAATGC

E1 TMD middle R	AATTGATATCTACCTTCACCAGACAAACCAAGAATGCAG
E1 TMD middle F	aattGATATCTGGCTGATGCTCATAACAGGGG
yM-E1-TMD-MT R (QC)	GTCTGGTGAAGGTATGGCTGATGCTC
yM-E1-TMD-MT F (QC)	GAGCATCAGCCATACCTTCACCAGAC
AL MluI F (VSV-Gecto)	CGACGCGTACTGCATTCTTGGTTTGTCTG
AL NotI R (VSV-Gecto)	TTTATAGCGGCCGCTCACCCTTG
QC-22F	CATAGGGTTAATCATTCTGCTATTCTTGGTTCTCC
QC-22R	GGAGAACCAAGAATAGCAGAATGATTAACCCTATG
QC-23R	GGACTATTCTTGCTACTTATACTGTGGCTG
QC-23R	CAGCCACAGTATAAGTAGCAAGAATAGTCC
QC-24F	CTGGTAGTGGGACTGCTGGTG
QC-24R	CACCAGCAGTCCCACTACCAG
QC-25F	CTGCTGGTGGGACTGCTGATAC
QC-25R	GTATCAGCAGTCCCACCAGCAG
yM-A2-YTDIE-R	TCGGTTCATCTCTATGTCTGTATACCCTTGCGCCCCTGTTATGAG
yM-B1-YTDIE-F	TATACAGACATAGAGATGAACCGA-TGAGCGGCCGCTTCGAG
CP7 NS4A F NcoI	CATG-CCATGG-TCCTCTGCCGAAAATGCCTTGCT
CP7 NS4A R XbaI	GCTCTAGATAATTCCTTTAGTTCAGTCTCCTTCCCCTCAG
QC-27 for K671A F	GTTTGTCTGGTGGCAGTAGTGAGAGG
QC-27 for K671A R	CCTCTCACTACTGCCACCAGACAAAC
QC-28 for R674A F	GTGGCAGTAGTGGCTGGCCAAGTGTTAC
QC-28 for R674A R	GTAACACTTGGCCAGCCACTACAGCCAC
QC-29 for E2 R tm F	CCTGGGTGGCGCTTACGTGCTTTG
QC-29 for E2 R tm R	CAAAGCACGTAAGCGCCACCCAGG
QC-30 C620S for E1 F	CTTGGGTAAATATGTTTCGGTAAGACCAGATTGG
QC-30 C620S for E1 R	CCAATCTGGTCTTACCGAAACATATTTACCCAAG
QC-31 1st G STO F	GGTAGTGCTGGGCCTGGTGTTAC
QC-31 1st G STO R	GTAACACCAGGCCCAGCACTACC
QC-32 2nd G STO F	CTGGTGTTACTGGGTATACTGTGGCTG
QC-32 2nd G STO R	CAGCCACAGTATACCCAGTAACACCAG
QC 35F (for ALR 1st G)	GTTGCTGTTGGGTCTGCTGCTG
QC 35R (for ALR 1st G)	CAGCAGCAGACCCAACAGCAAC

QC 36F (for ALR 2nd G)	CTGCTGCTGGGCTTACTGCTG
QC 36R (for ALR 2nd G)	CAGCAGTAAGCCCAGCAGCAG
QC 37F (for ALO 1st G)	GTGCTGGTAGGTCTGCTGCTG
QC 37R (for ALO 1st G)	CAGCAGCAGACCTACCAGCAC
OL A2 522 R	CTTATCGTCGTCATCCTTGTAATC-CCCTTGCGCCCCTGTTATGAG
OL B1 522 F	GATTACAAGGATGACGACGATAAG-TGAGCGGCCGCTTCGAG
OL A2 332 R	CTTATCGTCGTCATCCTTGTAATC-AAGATGGATACCAACTCGGAGAACCAAG
QC 38F (Q182A/G183A)	CCAAGTGTTAGCAGCTATACTGTGGC
QC 38R (Q182A/G183A)	GCCACAGTATAGCTGCTAACACTTGG
QC 39F (G178A/Q179A)	GTAGTGGCTGCCGCAGTGTTAGC
QC 39R (G178A/Q179A)	GCTAACACTGCGGCAGCCACTAC
QC 40F (G178A/Q179A)	GTAGTGGCTGCCGCAGTGTTACAAG
QC 40R (G178A/Q179A)	CTTGTAACACTGCGGCAGCCACTAC
QC41 F	CCACCACACATAGTACTTTGCCGAG
QC41 R	CTCGGCAAAGTACTATGTGTGGTGG
QC-42F	GTGAAGGTAGTGAGATGACAAGTGTTACAAG
QC-42R	CTTGTAACACTTGTCATCTCACTACCTTCAC
QC-43F	CGCATTTGGACCTGAGCTACAACCACTG
QC-43R	CAGTGGTTGTAGCTCAGGTCCAAATGCG
QC-44F	GTGATTGAAGAGGTGTGACAAGTAATTAAGG
QC-44R	CCTTAATTACTTGTCACACCTCTTCAATCAC
QC-45F	GAGGATGGTAAAATGATGCATGAGATGGGG
QC-45R	CCCCATCTCATGCATCATTTTACCATCCTC
QC-46F	CGCAGTGTTAGCAGCTATACTGTGGC
QC-46R	GCCACAGTATAGCTGCTAACACTGCG
R-Seq 3' BVDV E1	CTGAGTGGTGAGGCCTGTAGCTC
OL A2 E1 TMD 3'T R	CTTATCGTCGTCATCCTTGTAATCTTGTAACACTTGGCCTCTCACTACCTTC
QC-48 F	CATAACAGGGGCGCAAGGGTACCC
QC-48 R	GGGTACCCTTGCGCCCCTGTTATG
d-pmH1 F	CAGTCCTGGTGATTGAAGAGGTGGGT
d-pmH1 R	CAAACCAAGAATGCAGTGGTTGTAGCAGC

QC-49 F	GGCCAAGTGTTAGCAGCTATACTGTGGCTG
QC-49 R	CAGCCACAGTATAGCTGCTAACACTTGGCC
QC-50 F	CGCTGCTGGTTGACAGCTAGGCC
QC-50 R	GGCCTAGCTGTCAACCAGCAGCG
QC-51 F	CTTGGTTCTCCGATGAGGTATCCATCTTTG
QC-51 R	CAAAGATGGATACCTCATCGGAGAACCAAG
QC-52 F	GGGCGCAAGGGTGATACAAGGATG
QC-52 R	CATCCTTGTATCACCCTTGCGCCC
OL BirA AU1 5' B1 F	ATG GAC ACG TAC CGATACATCATGAAGGATAACACCGTGCCACTGAAATTG
OL BirA AU1 5' A2 R	GATGTATCGGTACGTGTCCATCCCTCGAGGCTAGCCTATAGTGAG
GFP F 5' KpnI F	GGGGTACCATGAGTAAAGGAGAAGAACTTTTCACTGGAGTTG
GFP R 3'NotI R	TTTGCGGCCGCTCACAGAATGTTCATCCATGCCATGTGTAATCC
OL 13-AU1 5B1 F	GACACGTACCGATACATC GCCTCTCCCTATTGTGAGGTAGAACG
OL 13-AU1 5A2 R	GATGTATCGGTACGTGTCTCCCATTGTGACTTGAAAGAAAACCAGGG
OL 13-AU1 3A2 R	GATGTATCGGTACGTGTCCCCTTGCGCCCCTGTTATGAGC
OL 13-AU1 3B1 F	GACACGTACCGATACATCTGAGCGGCCGCTTCGAGCAG
Avi TmH F	AGCACAGAAAATTGAATGGCATGAGGGCCAAGTGTTACAAGGTATACTGTGGC
Avi TmH R	TCGAAAATATCATTCAACCCGGACCCTCTCACTACCTTCACCAGACAAACCAAG
QC-53 F	GTAGTGCAGAAAAAATGACGCGTGGTACCTC
QC-53 R	GAGGTACCACGCGTCATTTTTTCTGCACTAC
Avi pmH2 F	AGCACAGAAAATTGAATGGCATGAGGCTGCTACAACCACTGCATTCTTGG
Avi pmH2 R	TCGAAAATATCATTCAACCCGGACCCGGTCCAAATGCGCGTCAGATCTTTTAAC
Avi pmH12 F	AGCACAGAAAATTGAATGGCATGAGGGTCAAGTAATTAAGGTTGTCTTAAGGGCG
Avi pmH12 R	TCGAAAATATCATTCAACCCGGACCCCACCTCTTCAATCACCAGGACTGTG
Seq m BVDV E1 For	GTCACTTGTCGGAGGTGCTACTACTC
QC-54F	GGTGCCCAGGGATACCTAGAGC
QC-54R	GCTCTAGGTATCCCTGGGCACC

QC-55F	CTGGAAAACAAGTCTAGAACATGGTTTGGG
QC-55R	CCCAAACCATGTTCTAGACTTGTTTTCCAG
HA-E1 5' F	GCTCTAGAACATGGTTTGGGGCATATGCA-TACCCATACGATGTTCCAGATTACGCTGC
HA-E1E2 3' R	GGGGTACCCTTGCGCCCCTGTTATG
VSVg 5' F	GCTCTAGAACATGGTTTGGGGCATATGCA-ATGAAGTGCCTTTTGTACTTAGCCTTTTTATTCATTG
VSVg 3' Rc	GGGGTACCCGCAAAGATGGATACCAACTCGGAGAAC
VSVg 3' Rk	GGGGTACCCCTTTCCAAGTCGGTTCATCTCTATGTCTG
QC-60 F	GTGATGTCTAGGGTGATAGCAGCAC
QC-60 R	GTGCTGCTATCACCCTAGACATCAC
BirA KDEL R	CAGTTCATCTTTGGAGGAGGAACCTTTTTCTGCACTACGCAGGGATATTTCACC
BirA KDEL F	GGTTCCTCCTCCAAAGATGAACTGTAAACGCGTGGTACCTCTAGAGTCG
BirA KEDL SeqS F	TGTTATCACCGCCCAAGCCAACAGGGCTTTCCCTCGAGGCTAGCCTATAGTGAG
BirA KEDL SeqS R	ATCTTGCTGTACCAGCCTGTAGCAGCCAAGGATAACACCGTGCCACTGAAATTGATTG
QC-62 F	GTGATAACAATCTTGCTGTACCAGCCTG
QC-62 R	CAGGCTGGTACAGCAAGATTGTTATCAC
QC-63 F	GTGCTGTCCAATTTCTCTCCAGAGACAG
QC-63 R	CTGTCTCTGGAGAGAAATTGGACAGCAC
QC-64 F	GTGCTGTCCAATTTCACTCCAGAGACAG
QC-64 R	CTGTCTCTGGAGTGAAATTGGACAGCAC
QC-65 F	GATTCCGAACCCGCTGCTGGGCCTG
QC-65 R	CAGGCCCAGCAGCGGGTTCGGAATC
ym-H1D F	GATTTCGCTCCAGAGACAGCCAG
ym-H1D R	GGTAAAATACTGCATGAGATGGGGGGT
ym-H2D F	ATCCCACAAGGACACACTGATATACAAG
ym-H2D R	TGGAGCGAAATCGGACAGCACTAC
QC-66 F	GAGATGGGGGGTGATTTCGCTCC
QC-66 R	GGAGCGAAATCACCCCCCATCTC
QC-67 F	GATTTCGCTCCAATCCCACAAGGAC
QC-67 R	GTCCTTGTGGGATTGGAGCGAAATC

QC-68 F	GATGGGGGGTATCCCACAAGGAC
QC-68 R	GTCCTTGTGGGATACCCCCCATC
Seq BirA Signal S R	CCCAGTCACGCAGTGTCTGAATGTG
ym-Seq pCI F	GATAGGCACCTATTGGTCTTACTGACATC
QC-69 F	CTGGTATACAAAGGCTTGCACTCCAGC
QC-69 R	GCTGGAGTGCAAGCCTTTGTATACCAG
QC-70 F	CAAAAACCAACTAGCCCTCACCGTAGAAC
QC-70 R	GTTCTACGGTGAGGGCTAGTTGGTTTTTG
QC-71 F	CTGATATACAAGATTCCGACAAAAACCAAC
QC-71 R	GTTGGTTTTTGTCGGAATCTTGTATATCAG
QC-72 F	GCACTCCAGCCTCCTTGCCTAGG
QC-72 R	CCTAGGCAAGGAGGCTGGAGTGC
QC-73 F	CCTCTCCCTATTCCGAGGTAGAAC
QC-73 R	GTTCTACCTCGGAATAGGGAGAGG
QC-74 F	GTATACAAAGAATTCCACTCCAGCCTG
QC-74 R	CAGGCTGGAGTGGAATTCTTTGTATAC
ym-344 new F	CCGCTCGAGATGGCATCAACTACTGCGTTTCTCATTTGCTTG
ym-345 new R	CGACGCGTTCAACCAGCAGCGAGCTGCTCTGTTAG
ym-BirA ATG F	CCGCTCGAGGGATGGCCCTGTTGGCTTGGGCGGTG
QC-75 F	CCGATTTCGCTCCAATCCCACAAGG
QC-75 R	CCTTGTGGGATTGGAGCGAAATCGG
N-H1D R	ACCCCCCATCTCATGCAGTATTTTACC
QC-76 F	GTGACCGGGAGACAAGGGCG
QC-76 R	CGCCCTTGTCTCCCGGTCAC
Avi half R	TTCGAAAATATCATTCAACCCTGAACCAGC
QC-77F	GTAGTGCTGTCCATCGCTCCAGAGAC
QC-77R	GTCTCTGGAGCGATGGACAGCACTAC
CR74 E1 171A forward	GCATTCTTGGTTGCACTGGTGAAGGTA
CR75 E1 171A reverse	TACCTTCACCAGTGCAACCAAGAATGC
CR76 E1 174A forward	GTTTGTCTGGTGGCAGTAGTGAGAGGC
CR77 E1 174A reverse	GCCTCTCACTACTGCCACCAGACAAAC
CR78 E1 174E forward	GTTTGTCTGGTGGAGGTAGTGAGAGGC

CR79 E1 174E reverse	GCCTCTCACTACCTCCACCAGACAAAC
CR80 E1 174D forward	GTTTGTCTGGTGGTAGTGAGAGGC
CR81 E1 174D reverse	GCCTCTCACTACCACCAGACAAAC
CR82 E1 177A forward	GTGAAGGTAGTGGCAGGCCAAGTGTTA
CR83 E1 177A reverse	TAACACTTGGCCTGCCACTACCTTCAC
CR84 E1 177K forward	GTGAAGGTAGTGAAGGGCCAAGTGTTA
CR85 E1 177K reverse	TAACACTTGGCCCTTCACTACCTTCAC
CR86 E1 177D forward	GTGAAGGTAGTGGGCCAAGTGTTA
CR87 E1 177D reverse	TAACACTTGGCCCACTACCTTCAC
CR88 E1 179A forward	GTAGTGAGAGGCGCAGTGTTACAAGGT
CR89 E1 179A reverse	ACCTTGTAACACTGCGCCTCTCACTAC
CR90 E1 179N forward	GTAGTGAGAGGCAACGTGTTACAAGGT
CR91 E1 179N reverse	ACCTTGTAACACGTTGCCTCTCACTAC
CR92 E1 179E forward	GTAGTGAGAGGCGAGGTGTTACAAGGT
CR93 E1 179E reverse	ACCTTGTAACACCTCGCCTCTCACTAC
CR94 E1 179D forward	GTAGTGAGAGGCGTGTTACAAGGT
CR95 E1 179D reverse	ACCTTGTAACACGCCTCTCACTAC
CR96 E1 182A forward	GGCCAAGTGTTAGCAGGTATACTGTGG
CR97 E1 182A reverse	CCACAGTATACCTGCTAACACTTGGCC
CR98 E1 182N forward	GGCCAAGTGTTAAACGGTATACTGTGG
CR99 E1 182N reverse	CCACAGTATACCGTTTAACACTTGGCC
CR100 E1 182E forward	GGCCAAGTGTTAGAGGGTATACTGTGG
CR101 E1 182E reverse	CCACAGTATACCCTCTAACACTTGGCC
CR102 E1 182D forward	GGCCAAGTGTTAGGTATACTGTGG
CR103 E1 182D reverse	CCACAGTATACCTAACACTTGGCC

3.12 Plasmids

3.12.1 Commercial Vector Plasmids

Name	Manufacturer	Purpose
pCI	Promega, Mannheim	Expression/Cloning target gene
pCITE 2a (+)	Novagen, Merck, Darmstadt	Expression/Cloning target gene

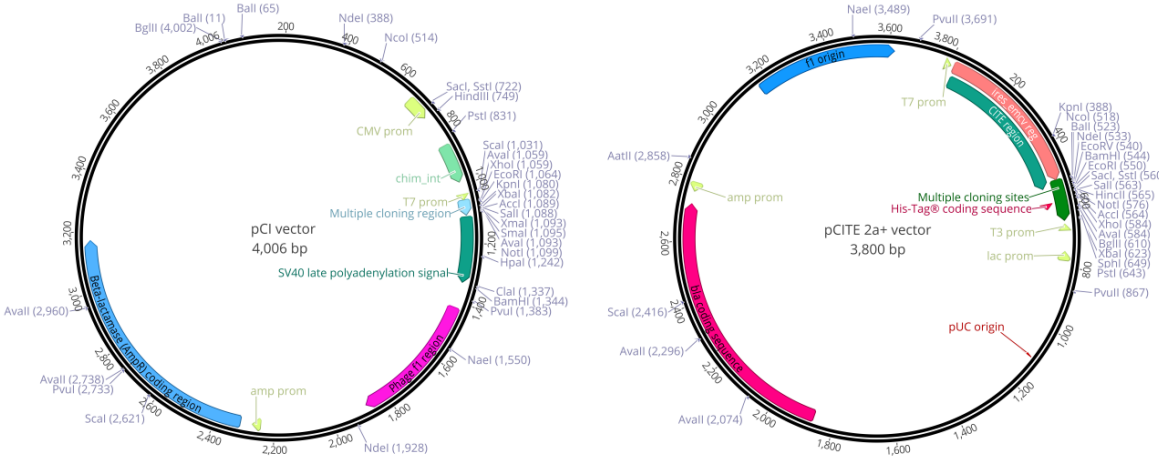


Figure 3.1 Schematic representation of pCI and pCITE 2a (+) empty vector

Maps of the pCI and pCITE 2a (+) vector plasmids. Those two plasmid maps were generated via Geneious Prime® (version 2019.2.3). Main single restriction sites and multiple cloning regions are present in the corresponding regions (depicted by blue and green arrows).

3.12.2 Prepared plasmids available in the laboratory

Plasmids	Source
pB-E^{rns}-V5	Dr. Birke Andrea Tews (Tews and Meyers, 2007)
pB-E^{rns}/TM-V5	Dr. Birke Andrea Tews (Tews and Meyers, 2007)
p798(Full length infectious clone BVDV CP7)	Prof. Dr. Gregor Meyers
pcDNA3-VSVg	Dr. Birke Andrea Tews
pFBD-shortH5-Avi	Prof. Dr. Timm Harder (Postel et al., 2011)
pDsRED-ER	Dr. Birke Andrea Tews
pDsRED-Golgi	Dr. Birke Andrea Tews
pcDNA3-GFP	Dr. Birke Andrea Tews
pCR-13 (BVDV CP7 pCI-Flag-E1)	Dr. Christina Radtke (Radtke and Tews, 2017)
pCR-16 (BVDV CP7 pCI-E2-AU1)	Dr. Christina Radtke (Radtke and Tews, 2017)
pCR-17 (BVDV CP7 pCI-E1-E2)	Dr. Christina Radtke (Radtke and Tews, 2017)

3.12.3 New plasmid constructs in this study

To study the function and structure of glycoprotein E1 of pestiviruses, a number of expression plasmids were prepared. In this study, we used the E1 sequence of the wild-type BVDV strain CP7 as representative for all pestiviruses. The individual constructs were made by using standard molecular biology technology (PCR, restriction digestion and ligation).

Plasmid Number	Description/Sequence
pYM-13	pCI-ss-HA-E1
pYM-14	pCI-ss-HA-E1-V5
pYM-15	pCI-ss-HA-E1(C171S)
pYM-16	pCI-ss-E2-AU1(C295S)
pYM-17	pCI-ss-HA-E1-E2
pYM-18	pCI-ss-2HA-E1-E2
pYM-19	pCI-ss-5HA-E1-E2
pYM-20	pCI-ss-HA-E1-E2-AU1
pYM-21	pCI-ss-HA-E1-Flag-E2-AU1
pYM-22	pCI-ss-HA-E1(C171A)
pYM-23	pCI-ss-HA-E1(K174A)
pYM-24	pCI-ss-HA-E1(K174E)
pYM-25	pCI-ss-HA-E1(K174Δ)
pYM-26	pCI-ss-HA-E1(R177A)
pYM-27	pCI-ss-HA-E1(R177E)
pYM-28	pCI-ss-HA-E1(R177K)
pYM-29	pCI-ss-HA-E1(R177Δ)
pYM-30	pCI-ss-HA-E1(G178L)
pYM-31	pCI-ss-HA-E1(Q179N)
pYM-32	pCI-ss-HA-E1(Q179E)
pYM-33	pCI-ss-HA-E1(Q179A)
pYM-34	pCI-ss-HA-E1(Q179Δ)
pYM-35	pCI-ss-HA-E1(Q182N)
pYM-36	pCI-ss-HA-E1(Q182A)
pYM-37	pCI-ss-HA-E1(Q182Δ)
pYM-38	pCI-ss-HA-E1(G183L)
pYM-39	pCI-ss-HA-E1(K174A,R177A)
pYM-40	pCI-ss-HA-E1(R177E,Q182A)
pYM-41	pCI-ss-HA-E1(G178L,G183L)
pYM-42	pCI-ss-HA-E1(Q182A,G183A)
pYM-43	pCI-ss-Avi-E1

pYM-44	pCI-ss-E1-Avi
pYM-45	pCI-ss-E1-Avi(B)
pYM-46	pCI-ss-E^{rns}-Avi
pYM-47	pCI-ss-E^{rns}-Avi(B)
pYM-48	pCI-ss-E^{rns} TM-Avi
pYM-49	pCI-ss-E^{rns} TM-Avi(B)
pYM-50	pCI-BirA (BirA-cyto)
pYM-51	pCI-ss-BirA (BirA-ER)
pYM-52	pCI-ss-HA-E1(K174A,R177A,G182A,Q183A)
pYM-53	pCI-ss-HA-E1(K174A,R177A,Q178A,G179A,G182A,Q183A)
pYM-54	pCI-ss-HA-E1(C123S)
pYM-55	pCI-ss-HA-E1(C123S,C171S)
pYM-56	pCI-ss-HA-E1ecto-VSVgTMD
pYM-57	pCI-ss-HA-VSVgecto-E1TMD
pYM-58	pCI-ss-HA-E1 (Ct-15aa)
pYM-59	pCI-ss-HA-E1 (Ct-31aa)
pYM-60	pCI-ss-HA-E1 (Ct-52aa)
pYM-61	pCI-ss-HA-E1 (Ct-30aa)
pYM-62	pCI-ss-HA-E1 (Ct-30aa)(C123S)
pYM-63	pCI-ss-HA-E1 (N19A)
pYM-64	pCI-ss-HA-E1 (N100A)
pYM-65	pCI-ss-HA-E1 (D67N,A69S)
pYM-66	pCI-ss-HA-E1 (D67N,A69T)
pYM-68	pCI-ss-HA-E1 MHD(delete H52-S83)
pYM-69	pCI-ss-HA-E1 MHD2(delete E71-S83)
pYM-70	pCI-ss-HA-E1ecto-[8aaVSVgTMD+22aaE1TMD]
pYM-71	pCI-ss-HA-E1ecto-[16aaVSVgTMD+14aaE1TMD]
pYM-72	pCI-ss-HA-E1ecto-[24aaVSVgTMD+6aaE1TMD]
pYM-73	pCI-ss-HA-E1ecto-[22aaE1TMD+8aaVSVgTMD]
pYM-74	pCI-ss-HA-E1ecto-[14aaE1TMD+16aaVSVgTMD]
pYM-75	pCI-ss-HA-E1ecto-[8aaE1TMD+24aaVSVgTMD]
pYM-76	pCI-ss-HA-E1 (C5S)

pYM-77	pCI-ss-HA-E1 (C20S)
pYM-78	pCI-ss-HA-E1 (C24S)
pYM-79	pCI-ss-HA-E1 (C94S)
pYM-80	pCI-ss-HA-E1 MHD(delete H52-S83)(Ct-15aa)
pYM-81	pCI-ss-HA-E1 MHD(delete H52-S83)(Ct-31aa)
pYM-82	pCI-ss-HA-E1 MHD(delete H52-S83)(Ct-52aa)
pYM-83	pCI-ss-HA-E1 MHD2(delete E71-S83)(Ct-15aa)
pYM-84	pCI-ss-HA-E1 MHD2(delete E71-S83)(Ct-31aa)
pYM-85	pCI-ss-HA-E1 MHD2(delete E71-S83)(Ct-52aa)

3.13 Equipments

Device	Supplier
AB Hitachi 3500 Genetic Analyzer	Applied Biosystems, Weiterstadt
Analytical balance Satorius A200S	Satorius, Göttingen
Blotkammer Mini-Trans Blot	Biorad, Munich
Blot chamber Tankblot	Hoefer, USA
CO_2 incubator MCO-19AIC, Sanyo	Ewald Innovationstechnik, Bad Nenndorf
ChemiDoc XRS$^+$ System	Bio-Rad Laboratories, Inc.
Revolver ™ rotary mixer	Labnet International Inc., USA
Flow cytometer MACSQuant Analyzer	Miltenyi Biotec GmbH, Bergisch Gladbach
Ice maker Manitowoc B230	Manitowoc, USA
Developer machine Compact2 ™	PROTEC medical technology, Oberstenfeld
Film cassettes	Agfa, Cologne
Liquid suction system BVC 21NT	Vacuubrand, Wertheim
Fluorescence microscope Axiovert 200M	Zeiss, Göttingen
Fluorescence accessories ApoTome	Zeiss, Göttingen
Gas safety burner flammy S	Schütt, Göttingen
Gel documentation system Quantum	peqlab, Erlangen
Gel electrophoresis chamber horizontal	Feinmechanik, FLI Tübingen
Heating block Thermo-Shaker TS-100	bioSan, Latvia
Heating magnetic stirrer KAMAG RCT	IKA Labortechnik, Staufen im Breisgau
Incubator Kelvitron t Heraeus	Instruments, Hanau

Microwave oven	Panasonic, Japan
PCR device Trio-Thermoblock	Biometra, Göttingen
Calimatic pH meter	Knick, Berlin
Power-Supply PS500XT	Hoefer, USA
Power-Supply 2301 Macrodrive1	LKB Bromma, Sweden
Rotor Ja10	Beckman Coulter, Munich
Rotor 1754, 5960	Hettich centrifuges, Tuttlingen
Multitron Standard shaking incubator	Infors HT, Switzerland
Speed Vac Concentrator Savant	Bachhofer, Reutlingen
Vortex Genie	Scientific Industries, USA
Centrifuge Avanti J-26 XP	Beckman Coulter, Munich
Centrifuge 5415C, 5430R	Eppendorf, Hamburg
Centrifuge Rotina 380R	Hettich centrifuges, Tuttlingen

3.14 Consumables

Name	Manufacturer
Red Caps Tube (15, 50 ml)	Sarstedt AG&Co.KG, Germany
Disposable cannulas	Braun, Melsungen
Disposable syringes	Braun, Melsungen
FACS tubes	BD Bioscience, USA
Whatman 3MM paper	Schleicher & Schuell, Dassel
Nitrocellulose transfer membrane, Protran	Schleicher & Schuell, Dassel
PCR tubes (0.2 ml)	Biozym, Hess. Oldendorf
Pipette tips	Greiner, Frickenhausen
Reaction tubes (0.5 ml, 1.5 ml and 2 ml)	Eppendorf, Hamburg
Cell culture bottles / dishes	Greiner, Frickenhausen
Cell culture plate (6- / 24- / 96-well)	Greiner, Frickenhausen
Centrifuge tubes (5, 15, 30 ml)	Greiner, Frickenhausen

3.15 Software, application and program

	Name	Supplier/Web sites
Desktop application	ImageJ 1.52a	National institutes of Health, USA
	Image Lab (Beta 3)	Bio-Rad Laboratories, Inc.
	MACSQuantify ™ (version 2.10)	Miltenyi Biotec GmbH, Bergisch Gladbach
	Geneious Prime(version 2019.2.1)	Biomatters Ltd.
	DNAMAN (version 9.0.1.116)	Lynnon Corporation.
	Jalview (Version 2.11.1.0)	University of Dundee, Scotland, UK
	GraphPad prism 8	GraphPad Software Inc., USA
	Microsoft Office 365	Microsoft, USA
	Aida Image Analyzer 5.0	Raytest Isotopenmeßgeräte GmbH, Benzstr. 4, D-75334 Straubenhardt, Germany.
	pDRAW32 (Version 1.1.140)	AcaClone Software
Web application	Biorender APP	https://app.biorender.com/
	PredictProtein 2013	https://www.predictprotein.org/
	PRIPRED	http://bioinf.cs.ucl.ac.uk/psipred/psiform.html
	TMHMM Server v. 2.0	http://www.cbs.dtu.dk/services/TMHMM/
	Jpred 4	http://www.compbio.dundee.ac.uk/jpred/
	NetNGlyc 1.0 Server	http://www.cbs.dtu.dk/services/NetNGlyc/
	TOPCONS	http://topcons.cbr.su.se/
	itol	https://itol.embl.de/gallery.cgi
	NetWheels	http://lbqp.unb.br/NetWheels/
	HeliQuest	https://heliquest.ipmc.cnrs.fr/
	Clustal Omega	https://www.ebi.ac.uk/Tools/msa/clustalo/
	WebLogo 3	http://weblogo.threeplusone.com/create.cgi

Chapter 4: Methods

4.1 Molecular cloning methods

The polymerase chain reaction (PCR) was used for the amplification of targeted genes such as cloning, sequencing or mutagenesis. The DNA polymerases (Phusion High-Fidelity DNA Polymerase) used in those experiments have extremely low error rates to ensure high-fidelity PCR. The working concentration of the primers (as shown in **3.11**) that used for the PCR is 10 pmol/µl. The PCR reaction mixture and the cycling programs are listed below.

4.1.1 Normal PCR

The reaction mixture of the standard PCR is shown in **Table 4.1** and the program which was used to amplify target DNA fragments is shown in **Table 4.2**.

Table 4.1 Normal PCR reaction approach

Components	Volume [µl]
DNA-Template	1 (5-10 ng/µl)
Forward Primer	2 (1 mM)
Reverse Primer	2 (1 mM)
dNTPs (10 mM each)	1
5X reaction buffer	5
Pfu DNA polymerase	1 (1.25u/50µl)
H_2O	38
Total	50

Table 4.2 Overview of the normal PCR program

Procedure	Temperature [°C]	Time [s]	
Initial denaturation	95	90	
Denaturation	95	30	x 29 cycles
Annealing (hybridization)	58	30	
Extension	72	17 bp/s	
Final extension	72	300	

4.1.2 Overlap fusion PCR

Overlap PCR is commonly used for cloning large complex fragments, making edits to cloned genes or fusing two or more gene elements together. For this purpose, the primary products

from normal PCR were firstly purified via using agarose gel electrophoresis, then the purified overlapping fragments A and B serve as template for the following overlap PCR.

The key step that decides whether overlap fusion PCR is successful or not is that ensure the overlap part can effectively anneal (two overlap parts can be firmly "sticked" together). It means that the overlap parts should have a certain length (generally, overlap area should be long enough, more than 25bp). Due to their complementary ends, the two purified products from normal PCR are able to hybridize. It is worth noting that the GC content of this overlap part should have a suitable Tm value (such as 65°C). The annealing temperature used in the PCR cycle should be lower than this temperature, otherwise the overlapping parts may not "stick" together. After hybridization, two short fragments form a long mostly single stranded DNA fragment which is made double stranded via DNA synthesis with *Pfu* polymerase in a first step and then can serve as template for further amplification. The forward primer of the first fragment and the reverse primer of the second fragment served as oligonucleotides of the rest 28 cycles. Tables 4.3 shows the reaction mixture approach and the program of overlap fusion PCR is same to the normal PCR.

Table 4.3 Overlap fusion PCR reaction approach

Components	Volume [µl]
Forward Primer	2 (1 mM)
Reverse Primer	2 (1 mM)
dNTPs (10mM each)	1
PCR product 1	1
PCR product 2	1
5X reaction buffer	5
Pfu DNA polymerase	1 (1.25u/50µl)
H_2O	37
Total	50

4.1.3 QuikChange® PCR

The mutagenesis protocol is used to make point mutations, switch amino acids, and delete or insert single or multiple amino acids. The QuikChange site-directed mutagenesis method is performed using a proof-reading DNA polymerase. The basic procedure utilizes a double-stranded DNA (dsDNA) plasmid with an insert of interest and two synthetic oligonucleotide primers containing the desired mutation. The oligonucleotide primers, each complementary to opposite strands of the vector, are extended during temperature cycling by DNA polymerase. Incorporation of the oligonucleotide primers generates a mutated plasmid containing staggered nicks. Following temperature cycling, the product is treated with Dpn I. The Dpn I endonuclease (target sequence: 5′-Gm6ATC-3′) is specific for methylated and hemimethylated DNA and is used to digest the parental DNA template, to select for mutation-containing newly synthesized DNA. DNA isolated from almost all *E.coli* strains is dam methylated and therefore susceptible to Dpn I digestion. The nicked vector DNA containing the desired mutations is then transformed

into TOP10 competent cells. The reaction approach was chosen analogous to the standard PCR (Table 3.1), while the program sequence is shown in **Table 4.4**.

Table 4.4 Overview of the QuikChange® PCR program

Procedure	Temperature [°C]	Time [s]	
Initial denaturation	95	90	
Denaturation	95	30	x 19 cycles
Annealing (hybridization)	58	30	
Extension	72	Variable (17 bp/s)	
Final extension	72	300	

4.1.4 Agarose gel electrophoresis

Agarose gel electrophoresis enabled DNA fragments to be separated according to their size and was used for analytical and preparative purposes. The negatively charged DNA fragments were retarded to different extents depending on their size when they moved through the agarose matrix in an electric field. The rate of migration was dependent on the size of the DNA fragments (the smaller the faster) and the cross-linking of the agarose, which could be controlled by the concentration of the agarose. By default, 1% (in 1x TAE buffer) agarose gels were used. The DNA samples were mixed with Gel Loading Dye, Purple (6X) (see **3.9**) and loaded into agarose gel in TAE buffer at a constant voltage of 100V. The 1 kb Plus DNA marker (see **3.9**) from Invitrogen was used as the size standard. The detection was carried out after 15-30 minutes incubation at room temperature in GelRed (1: 3300), under UV light at a wavelength of 254 nm.

4.1.5 Preparative agarose gel electrophoresis

For the preparative agarose gel electrophoresis, the DNA sample was electrophoresed as described under 4.1.4. The bands were then viewed under UV light at 302 nm and cut out with a scalpel. The cut-out bands were cleaned using the Nucleo-Spin® Extract II kit (see **3.7**) according to the manufacturer's protocol and eluted with 25 µl H_2O.

4.1.6 Restriction analysis

Restriction endonucleases are enzymes that can cut dsDNA (some restriction enzymes are also functional for single stranded DNA) on specific recognition sequences. This method was used to analyze plasmids after mini and midi preparation (**4.2.2**) and to generate specific cleavage fragments in the context of cloning.

The restriction took place according to the manufacturer's instructions in the corresponding buffers with an enzyme concentration of 2-40 U/batch and a total volume of 10-100 µl.

Table 4.5 Restriction analysis reaction approach for mini prep products

Components	Volume [µl]
Mini prep product	1
Restriction enzyme 1	0.2
Restriction enzyme 2	0.2
10X reaction buffer	1
RNase	0.1
H_2O	7.5
Total	10

The reaction time at the temperature recommended for the respective enzyme was 0.5-2 h. The completeness of the reaction and the sizes of the fragments formed were checked with agarose gel electrophoresis (**4.1.4**) and the fragments were purified for use in cloning if necessary (**4.1.5**).

4.1.7 Ligation of DNA fragments

In the ligation of digested and purified DNA fragments with complementary ends, circular DNA was formed by the enzymatic activity of the T4-DNA ligase. This DNA was then used to transform competent *E.coli*. In order to achieve high ligation efficiency, a vector/insert ratio of at least 1:3 was used. The reaction mixture described in **Table 4.6** was incubated at room temperature for 1-2 h or at 15 °C in water bath overnight.

Table 4.6 Ligation reaction mix

Components	Volume [µl]
Vector	Variable (0.1 pmol)
Insert fragment	Variable (0.01 pmol)
10x ligation buffer	2
T4-DNA ligase	1
H_2O	Variable (to 20 µl)
Total	20

4.1.8 Sequencing

Sequence analysis of DNA samples was carried out using the Sanger chain termination method in a PCR based assay. The difference to a normal PCR was the use of only one oligonucleotide primer and the use of ddNTPs in addition to dNTPs. The ddNTPs were coupled with fluorescent dyes, with ddATP, ddCTP, ddTTP and ddGTP carrying dyes with different fluorescence spectra. The ddNTPs do not have a 3'OH group, so their incorporation leads to the termination of the polymerization reaction and the DNA product fragments are labeled with the fluorescent dye of the incorporated ddNTP. The different lengths of ssDNA fragments with specifically labeled 3'ends could then be separated by capillary electrophoresis.

The sequencing reactions were set up with the BigDye® Terminator v.3.1 Cycle Sequencing Kit as shown in **Table 4.7**. The DNA to be sequenced was amplified in the thermal cycler with the following cycle parameters.

Table 4.7 Sequence PCR reaction mix

Components	Volume [µl]
DNA-Template	1 (100-200ng/µl)
Sequencing Primer	1 (0.25 mM)
BigDye® 5X buffer	1.5
BigDye® Sequence Mix	1
H_2O	5.5
Total	10

Table 4.8 Overview of the sequencing PCR program

Procedure	Temperature [°C]	Time [s]	
Initial denaturation	95	60	
Denaturation	95	10	x 25 cycles
Annealing (hybridization)	55	5	
Extension	60	110	

The synthesized sequencing products were then subjected to ethanol precipitation. For this purpose, 1 µl 3M NaAc pH 5 and 25 µl 100% ethanol were incubated together with the reaction mixture for 5 min. After centrifugation at 14,000 rpm for 20 min, the supernatant was removed, the pellet was washed twice with 75% ethanol and dried out in the SpeedVac for 5 min, then dissolved in 20 µl HiDi. The sample was used completely for analysis by the "3130 Genetic Analyzer" from Applied Biosystem (ABI).

The sequence data was evaluated with the Geneious Prime®2019 (**3.15**).

4.1.9 In vitro-transcription

For the transcription, the desired plasmid DNA was first linearized with a restriction enzyme (SmaI, at 25°C for 1-2h) at the 3'end of the sequence to be transcribed. The DNA fragment was purified by preparative agarose gel electrophoresis (shown in **4.1.4** and **4.1.5**), followed by phenol-chloroform extraction (**4.1.10**). The DNA fragment was used together with the components of the RiboMAX™ Large Scale RNA Production System T7 kit (**3.7**) according to the manufacturer's instructions for the in vitro transcription reaction. The reaction mixture was incubated for 4 h at 37°C. Then the obtained RNA was purified by means of phenol-chloroform extraction and ethanol precipitation.

Table 4.9 *In vitro*-transcription (T7) reaction approach

Components	Volume [µl]
dNTP MIX	9
Linearized DNA	12
5X reaction buffer	6
Enzyme MIX, RNA polymerase	3
Total	30

4.1.10 Phenol-chloroform extraction

The phenol-chloroform extraction was used for the purification and precipitation of DNA and RNA. The volume of the sample to be cleaned was first increased to 100 µl, then mixed with 100 µl of tris-saturated phenol and vortexed vigorously. After a centrifugation step of 5 min at 14,000 rpm, the phases separated and the upper aqueous phase could be transferred to a new 1.5 ml reaction tube. The aqueous phase was then mixed with 100µl chloroform, shaken vigorously and centrifuged for 3 min at 14000 rpm to separate the phases. The DNA or RNA was then precipitated from the aqueous phase with 1/8 volume of 2M potassium acetate (pH 5.6) and 2.5 volumes of ethanol. In order to achieve quantitative precipitation, the sample was stored on dry ice for 30 min or at -20°C overnight. The precipitated DNA or RNA was then pelleted at 14,000 rpm for 15 min at 4°C, the pellet washed with 80% ethanol, dried in the SpeedVac and finally taken up in 25µl DEPC water.

4.1.11 Reverse Transcriptase PCR (RT-PCR)

In a reverse transcriptase PCR, RNA was used as a template instead of DNA an appropriate volume of water (DEPC water for RNA). In the first step of the reaction, the RNA used is first reverse transcripted into cDNA before the newly obtained cDNA was amplified using oligonucleotides in the second step. The SuperScript™ III One-Step RT-PCR System Kit was used according to the manufacturer's protocol for the RT-PCR. The program and reaction approach are described in **Table 4.10** and **4.11**. The PCR products obtained in this way were first checked using agarose gel electrophoresis (**4.1.4**), then purified by preparative agarose gel electrophoresis (**4.1.5**) and finally used for sequencing (**4.1.8**). It is worth noting that the concentration of sequencing template (extracted DNA) should be very low (no more than 10 ng).

Table 4.10 Overview of the RT-PCR program

Procedure	Temperature [°C]	Time [s]	
cDNA synthesis and pre-denaturation	45-60	900-1800	
	94	120	
Denature	94	15	x 40 cycles
Anneal	55-66	30	
Extend	68	1kb/60s	
Final extention	68	300	

Table 4.11 RT-PCR reaction approach

Component	Volume[µl]
2X Reaction Mix	25
Template RNA (0.01 pg to 1 µg)	x
Sense primer (10 µM)	1
Anti-sense primer (10 µM)	1
SuperScript™ III RT/Platinum™ *Taq* Mix	2
Autoclaved distilled water	to 50

4.2 Microbiological methods

4.2.1 Heat shock transformation of *E.coli*

For the heat shock transformation, 50 µl of the competent cells were first thawed on ice, mixed with 10 µl ligation mixture or 0.5-2 µg plasmid DNA and incubated on ice for 20 min. The heat shock was carried out for 2 min at 42°C, the competent bacteria then were cooled on ice for 2 min and further 200 µl LB++ medium was added. After an incubation of 30-60 min at 37°C, the bacteria were spread on preheated LB agar plates with ampicillin, then placed in the incubator at 37°C overnight.

4.2.2 Isolation of plasmid DNA from bacteria

4.2.2.1 Mini-preparation

The mini preparation procedure based on the principle of alkaline lysis was used to isolate small amounts of plasmid DNA for analysis. 5 ml LB-Amp medium was inoculated with a single *E.coli* colony from a LB agar plate and incubated overnight at 37 °C in a thermo-shaker.

1.5 ml of the overnight culture was then pelleted by centrifugation (1.5 min, 14,000 rpm). The bacterial pellet was resuspended in 100 µl mini prep solution I. After adding 200 µl cold miniprep solution II, the tube containing the mixture was vortexed and incubated for 5 min on ice until a clear lysate was formed. After adding 150 µl of mini prep solution III, briefly swirling and incubating on ice for 10 minutes, the proteins as well as the cross-linked, long-chain, genomic DNA, were pelleted by centrifugation (10 min, 14,000 rpm). To precipitate the plasmid DNA, the clear supernatant was transferred to a new 1.5 ml reaction tube, 400 µl of isopropanol were added, and the pellet was then pelleted at 14,000 rpm for 10 min. The plasmid DNA was then washed with 200 µl 75% ethanol, dried in the SpeedVac and taken up in 50 µl H_2O.

4.2.2.2 Midi-preparation

The QIAGEN plasmid Midi kit was used according to the manufacturer's instructions to isolate larger amounts of clean DNA. For this purpose, 100 ml of LB-Amp medium were inoculated at 37°C in a thermo-shaker. For "low copy" plasmids, the volume of the LB-Amp medium was increased to 200 ml. At the end, the DNA was taken up in 100µl of demineralized sterile water and the concentration was determined photometrically using NanoDrop.

4.3 Cell biological methods

4.3.1 Cultivation of adherent cells

The cells used in this study were routinely kept in 10 cm cell culture dishes with ZB5 with 10% FCS medium at 37°C and 5% CO_2. The cells were separated and converted every 3-4 days according to their growth rate. For this purpose, they were washed once with trypsin mixture, covered with 5-10 ml of trypsin mixture and incubated for 3 min at 37 °C until they were detached from the bottom of the dish. The cells were then taken up again in medium and seeded to the 24-well plate or 3.5 cm dishes (dependent on the following experiment). For immunofluorescence or protein expression experiments, the cells were seeded in 3.5 cm cell culture dishes one day before transfection. For immunofluorescence, the cells were seeded in 3.5 cm cell culture dishes, 6-well plates or 24-well plates in a suitable dilution one day before transfection. For confocal immunofluorescence experiments, the wells of a 24-well plate were previously fitted with sterile coverslips so that the cell layer could grow on them.

4.3.2 Infection of cells

For infection experiments with BVD viruses, MDBK-B2 cells were seeded the day before in such a way that they were about 80% dense on the day of the infection. The cells were first washed with ZB5d medium and then incubated in ZB5d medium with virus for 4-6 h. After the incubation, the medium was changed and the cells were incubated in ZB5d medium + 10% FCS for 16-24 h before they were used for experiments.

4.3.3 Transfection of cells

In order to transiently express foreign proteins in eukaryotic cells, expression plasmid DNA was transfected into the cells using lipofectamine™ 2000 (see in **3.9**). In lipofection, the

transfection reagent contains positively charged molecules (lipids/polymer), which form complexes with the plasmid DNA and were internalized by the cells via endocytosis.

Lipofectamine™ 2000 was used as a normal reagent for the transfection of BHK21/RK13 cells. For this purpose, cells were seeded to be 70–90% confluent at transfection, then dilute four amounts of Lipofectamine® Reagent in Opti-MEM® Medium and dilute DNA in Opti-MEM® Medium respectively (detailed procedure shown in **Table 4.12**). After adding diluted DNA to diluted Lipofectamine® 2000 Reagent (1:1 ratio), incubate the mixture at room temperature for 5 min. Then drop DNA-lipid complex to cells and slightly shake the plate at the same time. Incubate cells for 24h at 37°C. Then visualize/analyze transfected cells, dependent on the following experiments.

Table 4.12 Lipofectamine™ 2000 DNA transfection procedure

Component	96-well	24-well	6-well
Final DNA per well	100ng	500ng	2500ng
Opti-MEM® Medium	25µL	50µL	150µL
Final lipofectamin™ 2000 regent per well	0.2-0.5µL	1-2.5µL	5-12.5µL

4.3.4 Electroporation of cells with RNA

For RNA electroporation, the MDBK-B2 cells should be used freshly, so the cells were seeded to 10 cm cell culture dishes one day before the day of the electroporation. The cells in 10 cm plate are sufficient for 3 EP samples.

Firstly, cells were detached from the dishes by treatment with trypsin mixture and resuspended in ZB5 with 10% FCS. The cells were then pelleted by centrifugation for 10 min at 1000 rpm, the supernatant was removed and the cells were washed once with ZB5.

The cell pellet was then taken up in 1.3 ml of cold (~4°C) PBS and 0.4 ml used for each electroporation RNA sample. 3-5 µl of RNA were mixed with 0.4 ml of cells and electroporated for 1 sec at 180 V and 980 mF. After a second pulse with the same settings the electroporated cells were immediately rinsed from the cuvette with ZB5 and transferred to two 3.5 cm cell culture dishes. The cells were observed to document the eventual development of a cytopathogenic effect (CPE).

The replication of electroporated RNA was demonstrated by immunofluorescence and the formation of infectious particles by reinfection experiments. For reinfection, the transfected cells were lysed by three cycles of freezing/thawing and then part of the lysate was added to new cells. The successful reinfection was tested by immunofluorescence.

4.3.5 Virus titration

A virus titration was carried out to determine the tissue culture-infectious dose 50 per ml (TCID50 / ml) for a virus passage. For this purpose, dilution series were made in 1:10 steps in a double batch on 96-well plates. 900 µl of ZB5d medium + 10% FCS were placed in 1.5 ml reaction tubes per dilution step. 100 µl of the virus passage to be tested were then added to the first reaction tube, thoroughly mixed and then 100 µl of the dilution (10^{-1}) were pipetted into the wells of the first column of the 96-well plate. A further 100 µl of the first dilution are added

to the next 1.5 ml reaction tube and a series of dilutions with the dilution factor 1:10 is continuously prepared. For cell control, a column of the 96-well plate was loaded with ZB5d medium + 10% FCS.

Then 1.75×10^4 cells in 100 µl ZB5d medium + 10% FCS were added per well and the plate was incubated at 37 °C. After 4-5 days, the titration was evaluated by means of indirect immunofluorescence (antibody: Code4) and the titer was calculated using the following Spearman / Karber formula:

$$\text{TCID}_{50}/\text{ml} = -\left(x_0 - \frac{d}{2} + d \cdot \sum \frac{r_i}{n_i}\right)$$

x_0 : $\log_{10}$ of the reciprocal value of the lowest dilution, where all wells are positive

d: $\log_{10}$ of the dilution factor

n: number of wells per dilution

r: number of positive wells per dilution level

$\sum \frac{r_i}{n_i}$: beginning of the summation at the dilution level x_0

4.3.6 Growth Kinetics of Viruses

To compare the growth kinetics of mutant viruses with that of the wild-type virus over a period of 72 h, growth curves were recorded. For this purpose, $5x10^5$ cells were infected per virus with an MOI of 0.1. For the infection, the cells were incubated for 1 h at 37 °C together with the calculated amount of virus and diluted in ZB5d medium. The cells were then centrifuged at 1000 rpm for 10 min, the supernatant was discarded and the cell pellet was resuspended in 10 ml of ZB5d medium + 10% FCS. 2 ml portions of the cell suspension were then distributed into five 3.5 cm cell culture dishes, incubated at 37 °C and harvested after 8 h, 24 h, 32 h, 48 h and 72 h by freezing at -70 °C. The evaluation was carried out after three freeze/thaw cycles by titration as described in **4.3.5**.

4.3.7 Extraction of viral RNA from cells

The viral RNA was extracted from infected cells in order to be able to analyze its sequence after RT-PCR amplification (**4.1.11**) by means of sequence PCR (**4.1.8**). For this purpose, MDBK-B2 cells were infected 2-4 days before RNA extraction. The viral RNA was extracted with Trizol® Reagent according to the manufacturer's protocol and then stored at -20 °C and used for an RT-PCR.

4.4 Protein analytical methods

4.4.1 Indirect immunofluorescence analysis

Indirect immunofluorescence was used on the one hand to detect BVDV infections and on the other hand to analyze proteins with regard to their location, retention and topology after transient transfection.

4.4.1.1 Staining cells in 24-well plates

To evaluate titrations (**4.3.5**), growth curves (**4.3.6**) and to check success after electroporations (**4.3.4**), the cells were first carefully washed three times with PBS and then fixed with 4% PFA solution (4°C, 30 min). After three more washing steps with PBS, the cells were permeabilized with 0.5% saponin (5 min, 4°C) or 0.05% Triton-X100 (30 min, 4°C) and washed three times with PBS.

Incubation with the primary antibody was carried out in the dilution given in section 3.6 for overnight at 4°C. The excess primary antibody was then removed by three washing steps with PBS and the cells were incubated with the secondary antibody for 1h at 37°C. After final washing step three times with PBS, analysis was done with a fluorescence microscope.

4.4.1.2 Staining cells on coverslips

The required cells were seeded to a 24-well plate with coverslips layered in the wells on day 1 before the following day transfection or infection assay. For the staining of the cells on the following days, the medium was removed as a first step and the cells were carefully washed three times with PBS. The fixation was then carried out using 4% PFA solution for 20 min at 4°C.

After the cells were washed with PBS, depending on the objective, with 0.05% Triton-X100 (in PBS) or 5 µg/ml digitonin (1:1000 dilution from digitonin stock which is in 5 mg/ml) for 30 min at 4°C permeabilized or left in the non-permeabilized state (see **Table 4.13**). After permeabilization of the cells, the analysis was carried out as described above staining. Before embedding in Mowiol DAPI, the cells were washed with PBS three times. After drying, the embedded preparations were sealed with nail polish and analyzed on a fluorescence microscope.

Table 4.13 Different permeabilization treatments

Permeabilization reagent	Aims
Not permeabilized	Only proteins on the cell surface can be detected
5 µg/ml digitonin	Only the plasma membrane is permeabilized; Proteins on the surface in the cytoplasm or on the inner side of intracellular membranes can be detected
0.05% Triton-X100	The plasma membrane and the internal membranes are permeabilized membranes; Proteins can be found on the surface, in the cytoplasm as well as in the cell organelles

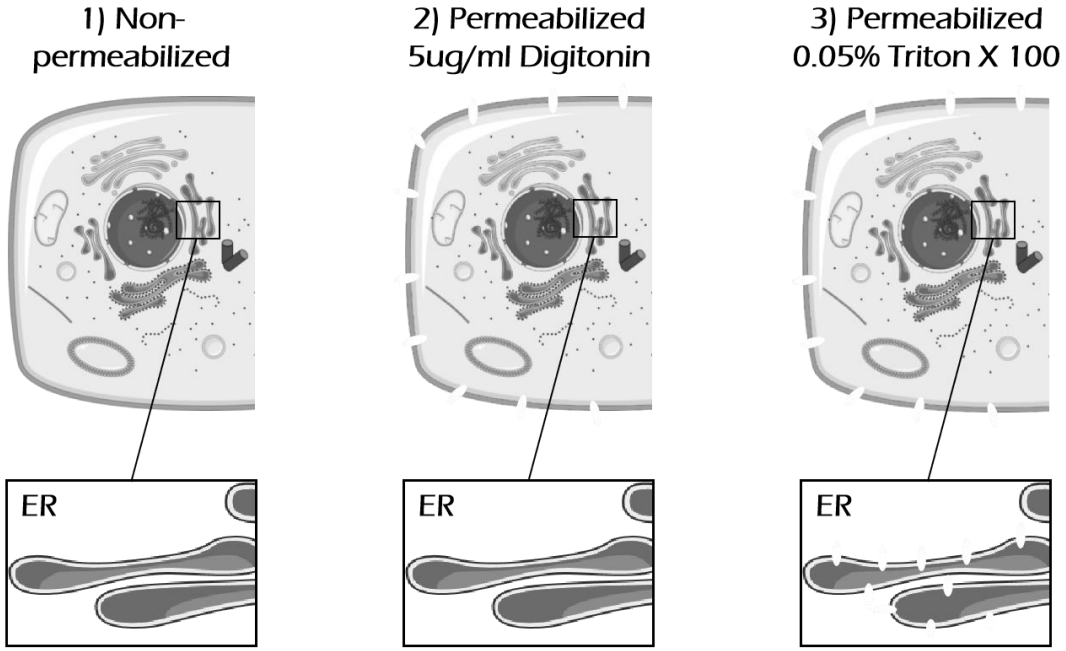


Figure 4.1 Schematic representative of different permeabilization treatments

4.4.2 Flow Cytometry Analysis (FACS analysis)

A flow cytometry approach was used to analyze the retention behavior of fusion proteins and proteins with mutations. The proteins to be examined were transiently expressed in RK13 cells, for which two wells of a 24-well plate were transfected with the same DNA reaction mix (**4.3.3**) as a duplicate. One day before the FACS analysis. First, the RK13 cells were carefully washed with pre-warmed PBS, then added 300 µl PBS with 2mM EDTA for 1h at 4℃ to detach cells from the 24-well plate. One hour later, detach cells gently and transfer cells to FACS tubes. Same transfected cells should be transferred identically into the same tubes. Then add 150 µl 4% PFA to each tube at 4 ℃ for 20-30 min to fix the cells. Next, add 1.25 ml PBS with 2% BSA and 2mM EDTA to bring the samples up to 2 ml total volume, then transfer 1 ml to a new FACS tube. So each sample has two same fixed cells in the FACS tubes, label one tube with ‘P’ for permeabilization and the other with ‘NP’ for non-permeabilization.

To get rid of extra PFA, 1ml 2% BSA and 2mM EDTA in PBS were added to each tube, and the cells were spun down. After removing the supernatant, ‘P’ samples should be permeabilized. To each ‘P’ sample tubes 300 µl 0.5% saponin in PBS with 0.2% BSA and 2mM EDTA were added and incubated for 30 min at 4 ℃. To remove excess saponin, all to ‘P’ sample tubes 2% BSA 2mM EDTA in PBS were added, spun down cells at 1000 rpm for 7 min. Subsequently, the cells were resuspended with 150 µl first antibody (e.g. HA tag antibody), which was diluted with 0.2% BSA and 2mM EDTA in PBS, for 1 h at 4 ℃ on a shaker. After two washing steps with 2% BSA in PBS, cells were incubated for 45-60 min at 37 ℃ with the second antibody (mouse-FITC). Finally, after washing with 2ml PBS with 2% BSA and 2mM EDTA, cells were resuspended in 100 µl 2% BSA in PBS and then measured using the MACSQuant Analyzer.

4.4.3 Immunoprecipitation

For the quantitative detection of proteins after transfection or infection, these expressed proteins were enriched and detected by using specific antibodies. Proteins that were metabolically radioactively labeled during their expression were used for immunoprecipitation. In this experiment, all cellular proteins were either brought into solution and purified or separated according to their properties in relation to a membrane association and then denatured. Antibody and sample were mixed and incubated for 1 h at 37℃ and then at 4℃ also for 1h. After adding Staph.A binds the FC region of antibodies (mainly class IgG) via protein the target protein bound to the antibodies can be precipitated. The samples were incubated for 30 min at room temperature and vortexed every 10 min for three times. The samples were then incubated at 4℃ for at least 16 h before being further purified.

After the overnight incubation, the samples were vortexed again and then underlayered with 500 ml of RIP-sucrose solution. The aggregates of target protein, antibody and staph.A were then pelleted at 1,500 x g for 10 min. The supernatant was discarded by the pump and the samples washed in 700 ml of RIP solution 2 (new prepared 1% Triton X100). Another washing step with 700 ml of RIP solution 3 (new prepared 0.2% Triton X100) followed before the samples were resuspended in 10 ml of RIP solution 4 (new prepared 0.06% Triton X100). The proteins and the aggregates were denatured by adding 40 ml of SDS sample buffer and incubating for 5 min at 95 ℃. Then the mixtures were centrifuged at 6,200 × g for 10 min and the supernatant with the dissolved proteins was transferred to a new reaction tube for further RIP gel analysis.

4.4.4 Western blot

4.4.4.1 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-polyacrylamide gel electrophoresis (SDS-PAGE) with the Jagow buffer was used for the analytical separation of protein mixtures. The proteins previously denatured by SDS were separated electrophoretically according to their molecular mass. The rate of migration of the proteins depends on the retention capacity of the polyacrylamide gel, which was determined by the concentration of the acrylamide.

The polyacrylamide gels used for the discontinuous separation of proteins consisted of a collection gel (4% acrylamide) for focusing the proteins and a separation gel with a uniform concentration of 8, 10 or 12% acrylamide depending on purpose.

First, the glass plates were clamped in the appropriate pouring stand, the separation gel was freshly poured and carefully overlaid with isopropanol to obtain a flat surface. After the polymerization, the isopropanol was removed, the gel washed and then the collection gel poured and the comb sample pockets was inserted into the liquid gels, liquid gel material to generate the loading pockets. Before applying the samples, the pockets were carefully washed with Jagow cathode buffer three times. After the gels were clamped in the electrophoresis apparatus, Jagow cathode buffer and Jagow anode buffer were added. The samples and a size marker were applied to the pockets of the stacking gel. Then a voltage of 80V was applied for 30min, then switch up to 120V for 1h. The electrophoresis was stopped when the running front had reached the bottom of the gels. After successful separation, the proteins were transferred to NC membrane for WB analysis.

4.4.4.2 Transfer of proteins to nitrocellulose membranes

The proteins separated by SDS-PAGE (**4.4.4.1**) were transferred from the SDS-gel to a nitrocellulose membrane with the Bio-rad Mini Trans-Blot® system. A nitrocellulose (NC) membrane with a pore size of 0.2 μm was used. Transfer was done by applying a voltage of 100V for 1 h. After the transfer, the proteins were visualized using an immunodetection system (3.4.7).

4.4.4.3 Immunodetection of the proteins in the Western blot

① Standard procedure

To detect the target proteins on a nitrocellulose membrane (**4.4.4.2**), it was first incubated for one hour with block buffer to saturate non-specific binding sites. After blocking the membrane and a short washing step with PBS-T (**3.8**), the detection antibody was diluted with PBS-T solution (see **3.5**), incubated overnight at 4℃ on the shaker. On the next day, the NC membrane was washed three times for 10 min with PBS-T and then incubated for 1-2 h with a suitable secondary antibody at room temperature. After repeated washing with PBS-T (three times, 10 min), the proteins were detected using the SuperSignal™ West Pico Chemiluminescent substrate kit. Follow the instruction of this kit. Briefly, 5x volume PBS-T plus two components of the kit were mixed 1: 1 and then the membrane was incubated with the solution for 2 min in the dark. Then, ChemiDoc imaging system (ChemiDoc XRS+, as shown in **3.13**) was used for detection of luminescence. By using Image Lab software (3.15) for further analysis.

② Detection of biotinylated proteins in Avitag-biotinylation assay

For the biotinylated proteins detection in Avitag-biotinylation assay, the first step for WB was the same as for the normal procedure. A TGG solution with 2% BSA was used instead of the milk blocking buffer. The membrane was incubated with TGG blocking buffer for 1h at room temperature on the shaker. After washing with PBS-T (three times, for 20 min), the membrane was incubated with diluted avidin-PO (in TGG solution with 2% BSA) for 60 min. To reduce the background, the membrane was washed with PBS-T three times for 20 min after incubation. The luminescence was then detected as usual.

Chapter 5: Results

5.1 Overview of pestiviral E1 glycoprotein

Compared to E^{rns} and E2, glycoprotein E1 is the least characterized among the pestivirus glycoproteins. There is still no crystal structure information for pestiviral E1 protein. In addition, E1 alone has not been analysed in detail. In this section, the E1 of BVDV CP7 strain served as a representative for analysing the biochemical properties of pestiviral E1. Analysis software or prediction application in molecular biology were used in this part to provide an overview of basic features of the glycoprotein E1 of pestiviruses.

5.1.1 Amino acid composition analysis

The length of genomic coding region for the E1 glycoprotein is nucleotides 585-594, which give rise to amino acids for 195-198aa (dependent on the virus species). The molecular size of E1 is just half of E2. Concerning there is still no detailed biochemical properties analysis about E1, we started our investigation with the amino acid composition (AAC) analysis. Exhaustive information about individual types of amino acids and groups of amino acids with similar physicochemical properties of the target protein that we obtain will help to easier understand characteristics and function of this protein. Especially, distinguishing membrane proteins from other types of proteins, because normally the former have a higher content of hydrophobic residues. Furthermore, AAC analysis is always used in combination with sequence alignments to help identify and classify the types of membrane protein and to reveal secondary structure domains and sites of specific function.

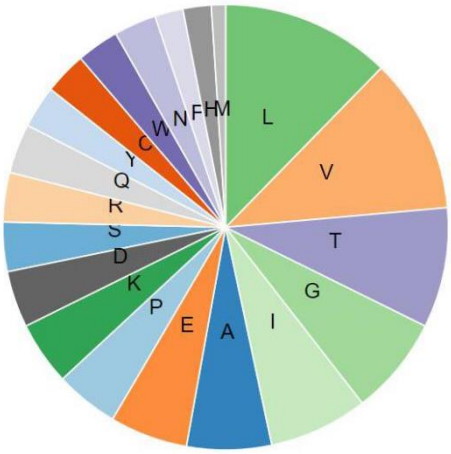


Figure 5.1 Amino acid composition of pestiviral E1

This figure was generated by online program (https://www.predictprotein.org/), E1 sequence of BVDV CP7 as an example.

Table 5.1 Amino acid composition of pestiviral E1

Amino Acid	Proportion	Amino Acid	Proportion
L	**12.31%**	S	3.59%
V	**11.28%**	R	3.59%
T	8.72%	Q	3.59%
G	7.18%	**Y**	**3.08%**
I	**7.18%**	C	3.08%
A	**6.15%**	**W**	**3.08%**
E	5.64%	N	3.08%
P	4.62%	**F**	**2.05%**
K	4.62%	H	2.05%
D	4.10%	M	1.03%

For the AAC analysis in this study, we choose BVDV strain CP7 as a representative strain for all pestiviruses. As shown in **Table 5.1**, the hydrophobic amino acids (red marked residues) account for a large proportion (approximately 40%) of the E1 residues. This data indicated that pestiviral E1 glycoprotein is a hydrophobic envelope protein or has several large hydrophobic areas that may serve as sites of specific function.

The result of the AAC analysis was supported by the 'Kyte & Doolittle' hydropathy plot (**Fig 5.2**) which clearly showed that pestiviral E1 contains four major hydrophobic regions located in the middle (h1 and h2) and at the *C*-terminus (h3 and h4). H4 region (roughly from position 166 to 195) was considered as transmembrane region.

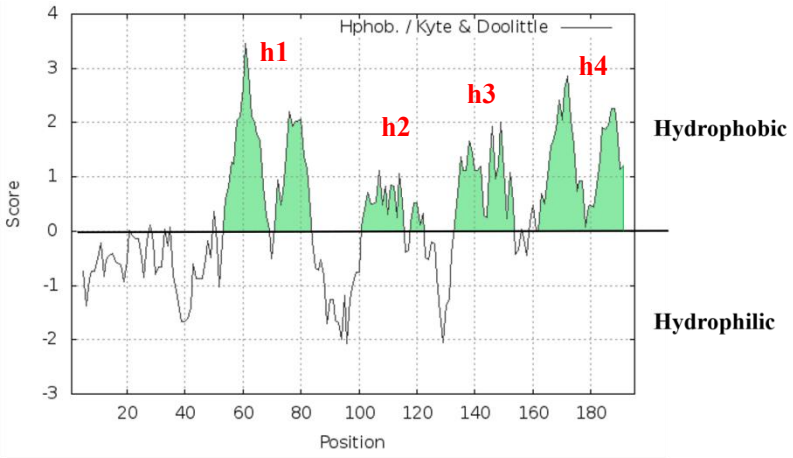


Figure 5.2 Kyte & Doolittle hydropathy plot of pestivirus E1 glycoprotein

BVDV CP7 E1 sequence as a representative for all pestiviruses; Hydropathy plot was drawn with the help of web application-ExPASy (https://web.expasy.org/protscale/). Four major hydrophobic regions are given in green and numbered h1 to h4 from the *N*-terminus to the *C*-terminus.

5.1.2 Alignments of multiple sequences of pestiviral E1

To obtain more detailed information on the conservation of the pestiviruses E1 protein, an alignment of 68 pestivirus E1 sequences throughout 11 ITCV-classifed species (from pestivirus A to K) was carried out via the Clustal Omega online program using default settings (https://www.ebi.ac.uk/Tools/msa/clustalo/, results shown in supplementary 1) and WebLogo3 application (http://weblogo.threeplusone.com /create.cgi, results shown in **Fig 5.3**).

Pestiviral E1 has a mass of about 27-33 kDa (dependent on the species), the result of a sequence alignment of glycoprotein E1 from pestiviruses (**Supplementary Material 1**) demonstrated that most E1 sequences contain 195aa residues apart from two atypical pestivirus isolates (pestivirus-J NrPV/NYC-D23 [KJ950914.1] and porcine Bungowannah [EF100713.2]). The former contains one lysine (K) insertion at position 42/43 and the other 'HR (histidine and arginine)' insertion at position 91/92, making E1 of this pestivirus J isolate 3aa longer than other pestivirus E1 proteins. Porcine Bungowannah E1, also contains an 'NR (aspartic acid and arginine)' insertion at 91/92 site. The molecular weight difference between different pestiviral E1 is partially due to the different number of amino acids, but also sequence variation plays a role.

5.1.2.1 Cysteine residues in E1

The glycoprotein E1 of pestiviruses contains six cysteine residues which can form intra- or inter-molecular disulphide bonds that play essential roles in stabilizing protein structure and can affect the function of this protein. Among them, the first five residues are almost fully conserved throughout all pestivirus species (as shown in **Fig 5.3** and **Supplementary Material 1**). The last cysteine residue at position 171 in species BVDV-2 and pronghorn antelope pestivirus is replaced by F (phenylalanine). Furthermore, the position 171 of E1 of pestivirus J

is also occupied by phenylalanine, but there is an additional cysteine residue at position 170 of this sequence which does not exist in all the other species. It is worth noting that cysteine residues at positions 5, 20 and 123 are 100% conserved throughout all the pestiviral species indicating those cysteine residues must play essential roles in the biological function of E1. The cysteines at positions 24 and 94 are also conserved, except for two atypical pestivirus isolates pestivirus-J NrPV/NYC-D23 (KJ950914.1) and porcine Bungowannah virus (EF100713.2). Cys24 was replaced by G (glycine) in both atypical pestiviruses, a mutation C94G occurs in pestivirus-J NrPV/NYC-D23 while L (leucine) substitution at that site in porcine Bungowannah. The cysteine residues replacement in pairs indicated that Cys24 together with Cys94 most likely form intramolecular disulphide bonds.

Figure 5.3 Conservation of amino acid sequences of pestiviral E1 protein

68 pestivirus E1 sequences throughout 11 ITCV-classified species (from pestivirus A to K) was aligned via WebLogo3 application (http://weblogo.threeplusone.com/create.cgi). The overall height of the stack indicates the sequence conservation at that position, while the height of symbols within the stack indicates the relative frequency of each amino acid at that position (Crooks et al., 2004; Schneider and Stephens, 1990). Fully conserved *N*-glycosylation sites in all genotypes are shown in the box with red solid line, one non-conserved *N*-glycosylation site is marked by a red dash box.

5.1.2.2 *N*-linked glycosylation sites in E1

E1 of pestivirus possesses three potential *N*-linked glycosylation sites (N6, N19 and N100). Among them, N19 and N100 are fully conserved in all major genotypes of pestiviruses, while the first *N*-glycosylation site at N6 is only present in CSFV, BDV, Pestivirus Aydin and Pestivirus Burdur isolates. In this section, first we used the web applicatyion NetNGlyc 1.0 Server (http://www.cbs.dtu.dk/services/NetNGlyc/) to make *N*-glycosylation site prediction for pestivirus E1. Normally, the glycosylation pattern could be interfered by the formation of

disulphide bonds in close distance to the glycosylation motif. As shown in **Fig 5.3**, there are two cysteine residues Cys20 and Cys24 which are located closely to each other, most likely forming of intra- or inter-molecular disulphide bond. Due to the spatial restriction, the *N*-glycosylation at N19 will likely be prevented by this disulphide bond formation. As shown in **Fig 5.4**, it seems that the prediction program has no confidence in the potential *N*-glycosylation site at N19. To determine whether the *N*-glycosylation takes place at N19 and N100, mutagenesis analysis of these conserved glycosylation sites was carried out. Three HA tagged E1 mutants pYM-63 [E1 (N19A)], pYM-64 [E1 (N100A)] and pYM-86 [E1 (N19A and N100A)] were generated via QuikChange® PCR. Wild-type HA-tagged E1 and those glycosylation site(s) defective constructs were expressed by Vaccinia T7 expression system and further analysed by immunoprecipitation with HA-tag antibody under reducing condition.

Table 5.2 The prediction of N-glycosylation sites of pestiviral E1

Position	Potential	Jury agreement	N-Glyc result
19NCTP	0.1414	(9/9)	- - -
100NLTV	0.6729	(9/9)	+ +

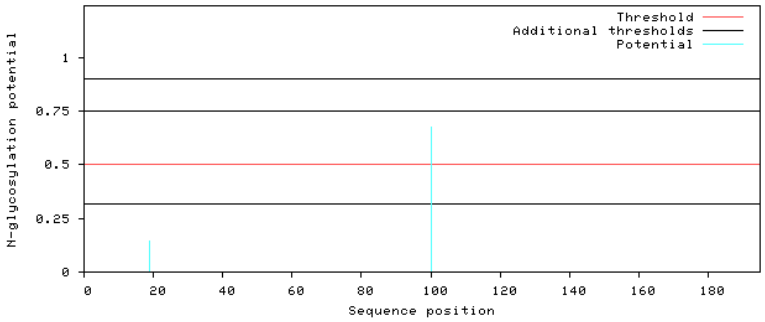


Figure 5.4 NetNglyc 1.0: predicted *N*-glycosylation site (s) of pestiviral E1

Potential *N*-glycosylation sites in pestiviral E1 glycoprotein (BVDV CP7 as an example) predicted using NetNGlyc 1.0 online web server. Amino acids with a score higher than the threshold indicated that this position has more possibilities to be a target for post-translation modification.

Table 5.3 Constructs for mutagenesis analysis of fully conserved *N*-glycosylation sites of pestiviral E1

Lane	1	2	3	4	5	6
Plasmid	pYM-13	pYM-63	pYM-64	pYM-86	pYM-13	Mock
Mutation(s)		E1 N19A	E1 N100A	E1 N19A/N100A		
Treatment					PNGase	

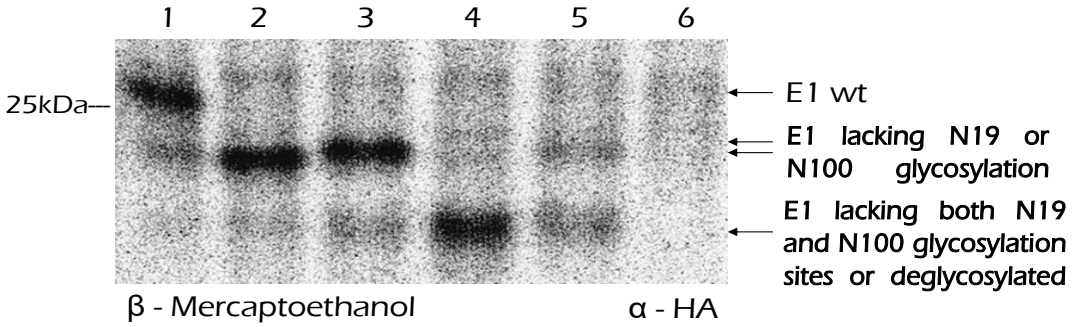


Figure 5.5 Mutagenesis analysis of *N*-Glycosylation site of pestiviral E1

N-glycosylation site(s) defective E1 mutants (in table 5.1) were expressed via Vaccinia virus T7 expression system, the products were precipitated by HA-tag antibody. Cells only infected with vaccinia virus without any other treatment for the mock control. All the samples were processed with the RIP buffer (plus β-mercaptoethanol) under reducing condition. The corresponding samples are shown in **Tab 5.3**.

As shown in **Fig 5.5**, the substitution at both N19 and N100 resulted in molecular weight decrease (lane 2 and 3). When both potential *N*-glycosylation sites were destroyed, an even stronger molecular weight shift was visible (lane 4). After removing all types of *N*-linked glycans of E1 wt by using PNGase treatment, the molecular size of this deglycosylated E1 dropped to the same level as the double *N*-glycosylation site E1 mutant (lane 4 and lane 5). This indicated that both putative *N*-glycosylation sites of E1 are used, so that pestiviral E1 has a mass of 27-32 kDa. It can be inferred that the non-conserved *N*-glycosylation site at N6 which is specific for CSFV, BDV, Pestivirus Aydin and Pestivirus Burdur species was also used. The molecular weight difference in pestiviral E1 is mainly dependent on the question whether this N6 site can be glycosylated. Furthermore, the slight molecular difference between E1 N19A mutant and E1 N100A mutant also indicated that they have different glycan types.

5.1.3 The prediction of secondary structure and transmembrane domain of pestiviral E1

5.1.3.1 The secondary structure prediction of pestiviral E1

The prediction for the secondary structure of E1 in this section was generated with two widely used web applications JPred4 (version 4, http://www.compbio.dundee.ac.uk/jpred4/index.html and PSIPRED Workbench (http://bioinf.cs.ucl.ac.uk/psipred/). The E1 sequence of BVDV CP7 isolate was used as a representative for this study. The predication results are shown in **Fig 5.6** and **Supplementary Material 2**. In combination with the hydropathy plot (shown in **Fig 5.2**), the prediction result from PSIPRED demonstrated that the two hydrophobic regions in the middle part of E1 (h1) most likely form α-helices (L53-D67 and P70-S83) which correspond to the two peaks in the h1 region of the hydropathy plot in **Fig 5.2**. The hydrophobic regions h3

and h4 located at the *C*-terminal of E1 may form three sections of α-helices according to JPred4 (as shown in **Fig 5.6** and **Supplementary Material 2**).

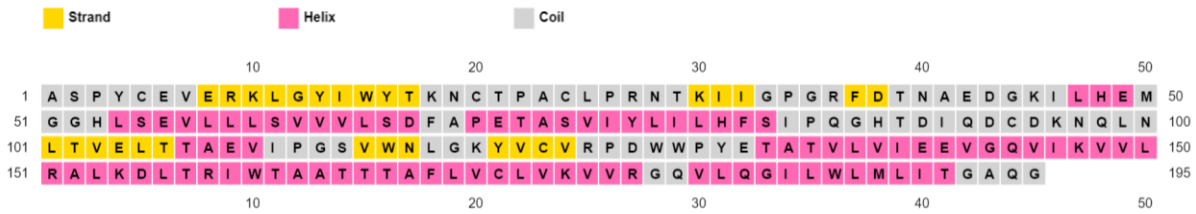


Figure 5.6 The prediction of secondary structure in pestiviral E1

Secondary structure prediction for pestivirus E1 glycoprotein was carried out by PSIPRED 4.0 web-server. The E1 sequence of BVDV CP7 strain was used for prediction in this study. All residues of E1 are shown in small box with different colour. Residues in pink colour indicates that are residues most likely form α-helix at secondary structure level. Similarly, β-strands are highlighted in yellow, coiled regions in grey.

5.1.3.2 The consensus prediction of pestiviral E1 membrane topology

The E1 glycoprotein of pestiviruses is supposed to be a transmembrane protein with *C*-terminal membrane anchor. It is still unclear which type of membrane topology E1 has. For the closely related hepatitis C virus, the last 30aa of E1 were identified to be the membrane anchor, which is also responsible for ER localization of E1 (Cocquerel et al., 1999; Cocquerel et al., 2000b).

To analyse which type of membrane topology pestiviral E1 adopts, the consensus prediction of pestiviral E1 membrane topology in this section was carried out by topology predication algorithm TOPCONS. As shown in **Fig 5.7**, the prediction results from TOPCONS demonstrated that E1 most likely adopt a type I transmembrane topology with its *N*-terminus in the ER lumen while spanning the membrane via the membrane anchor at the *C*-terminus. The transmembrane region was supposed to range from AA 159 to 179. The other five different topology prediction algorithms OCTOPUS, Philius, PolyPhobius, SCAMPI (multiple sequence mode) and SPOCTOPUS were also used in this study. Interestingly, some of them considered that E1 has more than one transmembrane region (**Table 5.4**), since two α-helices located in the middle part were also assumed to adopt the transmembrane configuration. The membrane topology of E1 is therefore questionable and requires more detailed experimental investigation. According to the preliminary data from our lab and confirmed membrane topology of HCV E1, we considered that the putative transmembrane region of pestiviral E1 should be within last 30aa at the *C*-terminal.

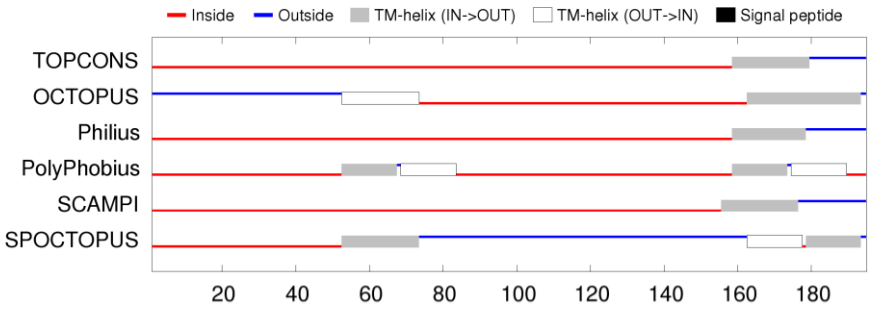


Figure 5.7 The consensus prediction of pestiviral E1 membrane topology

TOPCONS topology prediction for pestiviral E1 glycoprotein (BVDV CP7 as an example). The TOPCONS topology prediction (http://topcons.cbr.su.se) is a consensus predictor that collects data from the other prediction servers listed in the panel. Blue line indicates the part should be out of the cells or specific

organelles, similarly, red line indicates the part should be inside. TM-helices with different orientations were shown as grey or white bar.

Table 5.4 The consensus prediction of pestiviral E1 transmembrane region(s)

Topology prediction algorithms	TM-helix positions (position starting from 1)
TOPCONS	TM1: 159-179
OCTOPUS	TM1: 53-73, TM2: 163-193
Philius	TM1: 159-178
PolyPhobius	TM1: 53-67, TM2: 69-83, TM3: 159-173, TM4: 175-189
SCAMPI (multiple sequence mode)	TM1: 156-176
SPOCTOPUS	TM1: 53-73, TM2: 163-177, TM3: 179-193

5.2 Subcellular localization of pestivirus E1

It is known that E1 forms a covalently linked heterodimeric complex with E2 which is presented on the viral particle surface. In addition, pestiviruses have long been considered to bud intracellularly, as recently confirmed by ultra-structure analysis of the pestivirus Giraffe-1 (Schmeiser et al., 2014). Therefore, the envelope proteins of pestiviruses have to accumulate and bud in some specific intracellular compartments. The glycoprotein E1-E2 complexes of HCV were shown to accumulate in the endoplasmic reticulum (ER) where also virus budding takes place (Cocquerel et al., 2000b). As part of the polyprotein, both E^{rns} and E2 were shown to be concentrated at the ER (Burrack et al., 2012; Radtke and Tews, 2017; Tews and Meyers, 2007). For E1, there are no published data. The formation of E1-E2 heterodimers strongly indicates that E1 should locate in the ER or a compartment close to the ER. Both E^{rns} and E2 have retention mechanisms mediated by retention signals of their own to keep them within the ER. The retention mechanism of pestiviral E1 has never been discussed before. In this section, firstly, we would like to determine the intracellular localization of pestiviral E1. For this purpose, the E1 glycoprotein of BVDV CP7 was analysed together with compartment markers via indirect immunofluorescence.

Since no antibodies directed specifically against pestiviral E1 were available for the detection of transiently expressed E1, the plasmid pCR-13 which was established by Dr. Christina Radtke for BVDV CP7 E1 expression, was used in this study. This plasmid is based on the mammalian expression vector pCI, in which the target sequences are inserted and expressed under control of the CMV promoter. The last 20 amino acids (aa) at the *C*-terminus of the BVDV CP7 core protein (aa sequence is as follow: EKALLAWAIIALVFFQVTMG) served as a signal sequence (SS) for the transient expression of E1. The full-length sequence coding for CP7 E1 (585bp) was inserted right after this SS and these sequences were introduced into the pCI vector. For the detection of transiently expressed E1, a HA (hemagglutinin) tag was inserted at the *N*-terminus of the E1 sequence in pCR-13, right after the signal sequence. This newly made HA tagged E1 construct was named pYM-13. Transient expression of E1 took place in BHK-21 or RK-13 cells (depending on the purpose of the study) via the lipofectamine™ 2000 transfection method or Vaccinia virus T7 expression system. E1 was detected with using HA-specific antibodies.

To investigate the subcellular localization of E1 in different cell organelles, the ER/golgi compartment was visualized in parallel to E1 by the pDsRed-ER/pDsRed-Golgi plasmids co-expressed with HA-tagged E1. The pDsRed-ER/pDsRed-Golgi is designed for fluorescent labeling of the endoplasmic reticulum (ER) or the Golgi apparatus in mammalian cells, respectively. Fluorescence can be observed in living/fixed cells by microscopy or flow cytometry.

Briefly, the ER was labeled by the fluorescent dye pDsRed-ER, which was transiently co-expressed with HA tagged E1 expression plasmid pYM-13. Similarly, the pDsRed-Golgi plasmid was used for Golgi-apparatus labeling. The results were analysed with a confocal fluorescence microscope (as shown in **Fig 5.8**).

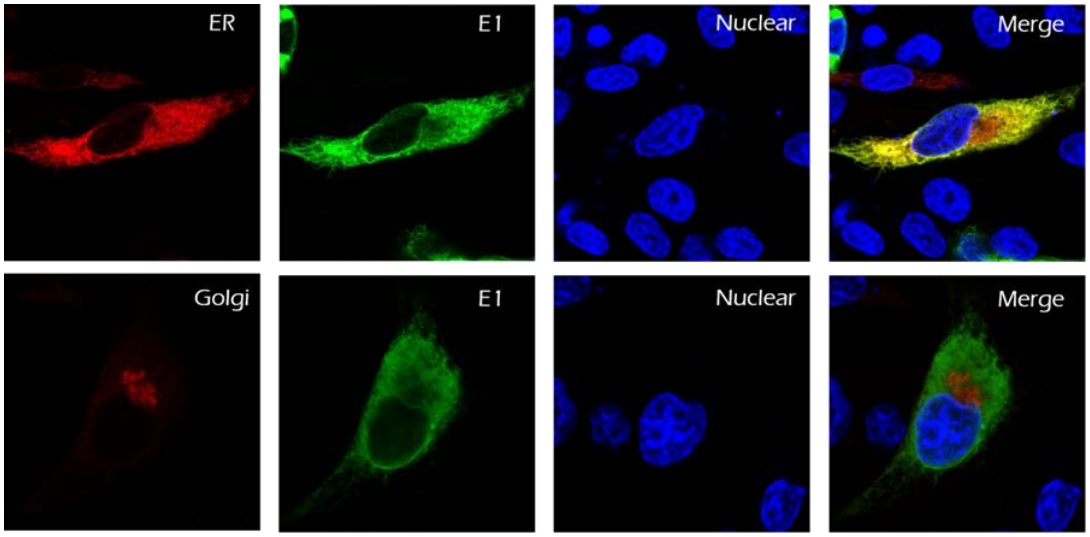


Figure 5.8 Subcellular localization of E1 glycoprotein

HA tagged E1 expression plasmid co-transfected with pDsRed-ER/pDsRed-Golgi respectively for E1 intracellular localization analysis. At 24h post-transfection, cells were fixed by 4% PFA, permeabilized with 0.05% Triton X-100 and stained with specific antibodies against HA (green). Compartments (ER or Golgi) are in red. Nuclei were stained with DAPI.

Colocalization of E1 with the compartment markers suggested that E1 was mainly concentrated in the ER, while showing no detectable localization in the Golgi compartment. This results also demonstrated that E1 has to contain an ER retention signal of its own instead of being localized in the ER via the interaction with other viral proteins (e.g. E2).

5.3 Studies on the localization of the retention signal in E1

The colocalization analysis showed that E1 was mainly distributed in the ER. Pestiviral E1 glycoprotein has never been found on the infected or transfected cell surface. This situation is also true for E^{rns} and E2 which both have been shown to contain ER retention signals. E1 can be retained within the ER by an intrinsic retention signal or by a retrieval signal that makes E1 travel from post-ER compartments back to the ER.

5.3.1 The transmembrane anchor of E1 is a determinant for ER retention

It was shown that the E1 of hepatitis C virus, which is rather closely related to pestiviruses, contains an intrinsic retention signal within its TMD. For both E^{rns} and E2 of pestiviruses the *C*-terminal membrane anchors were found to be responsible for their retention (Burrack et al., 2012; Radtke and Tews, 2017). In order to investigate whether the TMD of E1 plays a similar role, fusion proteins composed of parts of a protein naturally exported to the cell surface and parts of E1 were constructed (as shown in **Fig 5.9A**). A commonly used partner protein for such analyses is vesicular stomatitis virus (VSV) G protein. VSV-g is a typical type III viral fusion protein that is normally expressed on the plasma membrane of the cells. Based on our prediction

data (**see 5.1.3.2**), the putative TMD of E1 should encompass the last 30aa at the *C*-terminus (from position 166 to 195), the rest is supposed to represent the ectodomain of E1. It is known that the TMD of VSV-g is a short peptide that extends from I-465 to C-489 (aa sequence: IASFFFIIGLIIGLFLVLRVGIHLC).

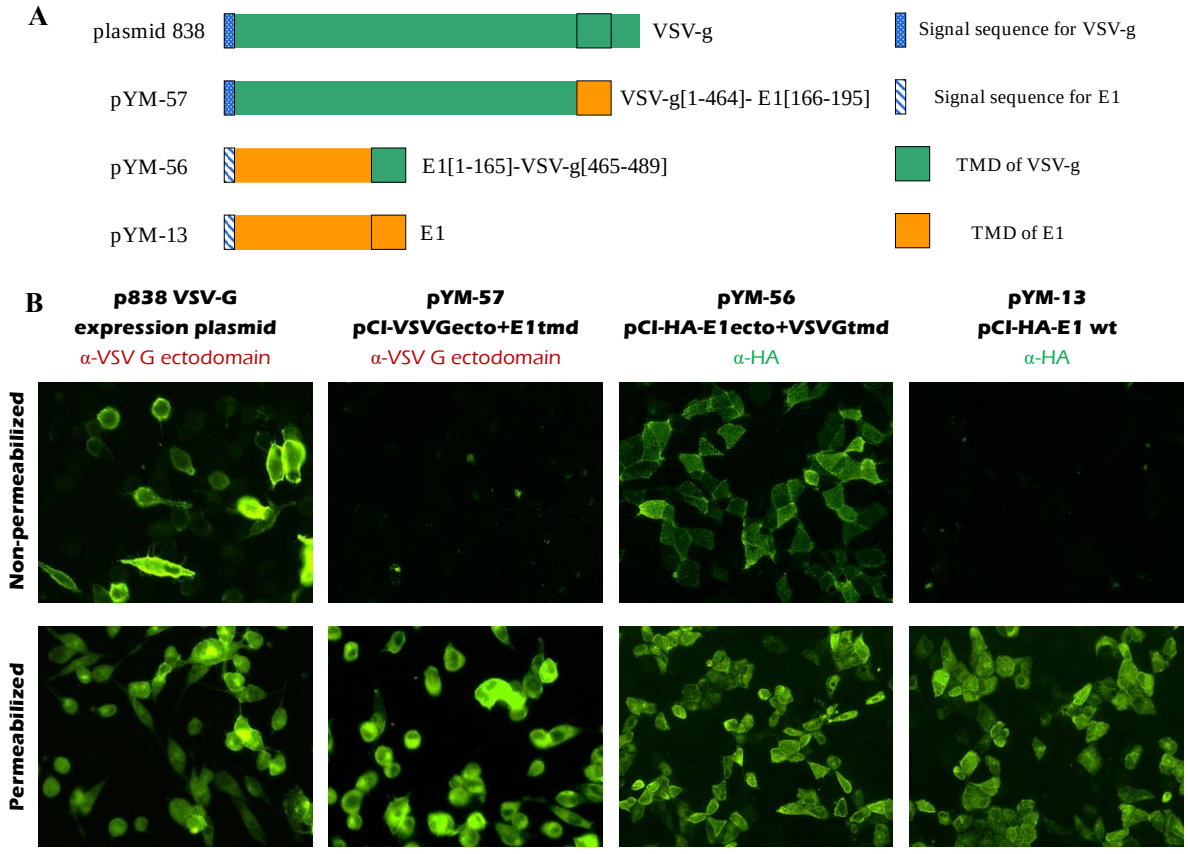


Figure 5.9 The transmembrane anchor is responsible for the retention of E1

(A) Schematic representation of the parental proteins or chimeras used in this study. Two chimeric proteins: pYM-57 (VSVg ectodomain-E1 transmembrane domain): VSV-g[1-464], ectodomain of VSV-g fused to the TMD of E1 [166-195]; pYM-56 (E1 ectodomain-VSVg transmembrane domain): E1[1-165], ectodomain of E1 fused to the TMD of VSV-g [465-489].

(B) Test for cell surface expression of the parental proteins or chimeras. RK-13 cells were transfected with the indicated plasmids. At 24h post-transfection, cells were further analysed by indirect immunofluorescence. The RK-13 cells were fixed with 4% PFA, then permeabilized or not with 0.05% Triton X-100, and immune-stained with anti-HA mAb (secondary FITC-anti-mouse) or anti-VSV-g pAb (secondary FITC-anti-rabbit).

Cells transfected with plasmid 838 expressing full-length VSV-g were used as a control of cell surface expression, since there is no retention signal within VSV-g. The transfected cells were all positive after permeabilization with 0.05% Triton X-100, indicating that all the parental/fusion proteins were expressed (shown in **Fig 5.9B**). For the non-permeabilized cells, VSV-g wt was detected as a rim surrounding the transfected cells. Interestingly, also the chimera composed of E1 ectodomain and VSVg TM was detected on the cell surface. In contrast, the fusion protein VSVg[1-464]-E1[166-195] showed no cell surface expression and was completely retained within the cell, similar to E1 wt (**Fig 5.9B**). This result clearly indicates that the TM anchor of the pestiviral E1 functions as an intracellular localization signal, in the other words, the TM region is responsible and obviously sufficient for the retention of E1.

In this section, the pYM-56 (E1[1-165]-VSVg[465-489]) E1-VSVg chimera was also analysed via confocal microscope to investigate its intracellular localization. E1-VSVg chimera expression plasmid pYM-56 was co-transfected with pDsRed-ER/pDsRed-Golgi into RK-13 cells respectively. As shown in **Fig 5.10**, when the transmembrane region of E1 was replaced by that of VSV-g, this chimera was entirely presented on the cell surface. Neither co-localized with ER nor Golgi apparatus, showing a completely different intracellular localization compared to wild-type E1 (as shown in **Fig 5.8**).

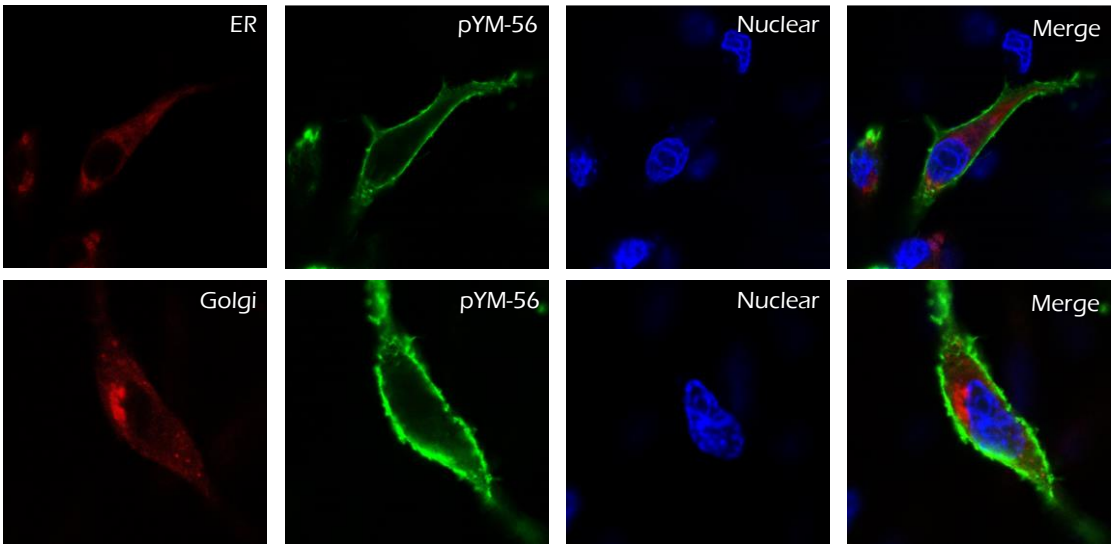


Figure 5.10 Subcellular localization of pYM-56 (E1-VSVg chimera)

HA tagged E1-VSVg chimera (E1[1-165]-VSV-g[465-489]) expression plasmid co-transfected with pDsRed-ER/pDsRed-Golgi, respectively for its intracellular localization analysis. At 24h post-transfection, cells were fixed by 4% PFA, permeabilized with 0.05% Triton X-100 and stained with specific antibodies against HA (green). Compartments (ER or Golgi) are in red. Nuclei were stained with DAPI.

Since VSV-g is a typical plasma membrane protein, it is almost exclusively found on the cell surface. To further prove the cell surface presence of this E1-VSVg chimera (pYM-56). We performed a colocalization assay with wild-type VSV-g protein. Cells were co-transfected with 838 plasmid expressing full-length VSV-g and pYM-13 (E1 wt). After permeabilization with 0.05% Triton X-100, the E1 wt showed totally different cellular localization to VSV-g wt (**Fig 5.11A**). However, it is clear to see a perfect colocalization of this E1-VSVg chimera (pYM-56) with the wild-type VSV-g. As shown in **Fig 5.11B**, under non-permeabilization conditions, both E1-VSVg chimera and VSV-g wt presenting on the cell surface which was detected as impressive rim surrounding the transfected cells. This result further confirmed the cell surface presence of E1-VSVg chimera. Importantly, the VSVg-E1 chimera pYM-57 (VSV-g[1-464]-E1[166-195]) showed absence of cell surface expression and was completely retained within the ER, thus showing a similar cellular localization to E1 wt (**Fig 5.11C**), it means that the TM domain of E1 is sufficient to keep the ectodomain of VSVg within the ER. Taken together, these results indicated that the transmembrane anchor of E1 is a determinant for ER localization.

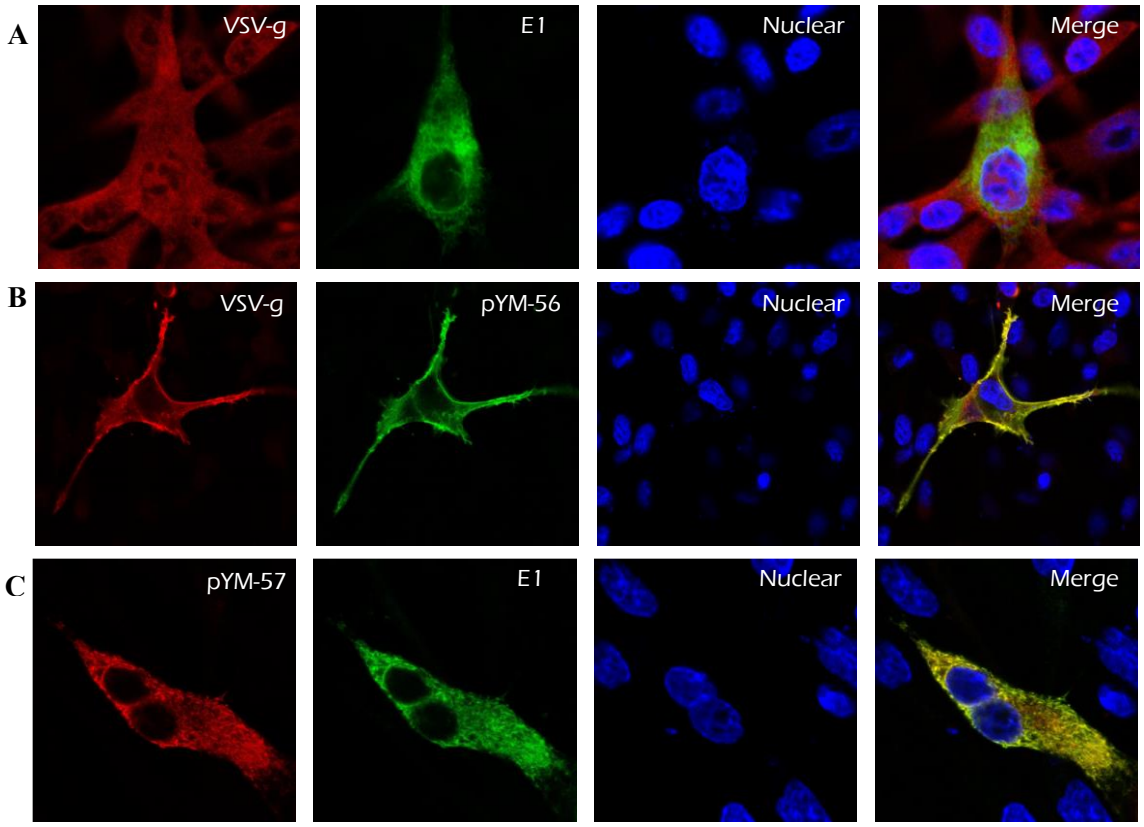


Figure 5.11 Co-localization analysis of the parental proteins or chimeras

(A) Co-localization analysis for the parental protein E1 and VSV-g. RK-13 cells were transfected with the indicated plasmids. At 24h post-transfection, cells were further analysed by indirect immunofluorescence. The RK-13 cells were fixed with 4% PFA, then **permeabilized with 0.05% Triton X-100**, and immune-stained with anti-HA mAb (secondary FITC-anti-mouse) or anti-VSV-g pAb (secondary FITC-anti-rabbit).

(B) Co-localization analysis for the cell surface expression of the parental protein VSV-g and pYM-56 (E1[1-165]-VSV-g[465-489]) chimera. RK-13 cells were transfected with the indicated plasmids. At 24h post-transfection, cells were further analysed by indirect immunofluorescence. The RK-13 cells were fixed with 4% PFA, **without any permeabilization treatments**, and immune-stained with anti-HA mAb (secondary FITC-anti-mouse) or anti-VSV-g pAb (secondary FITC-anti-rabbit).

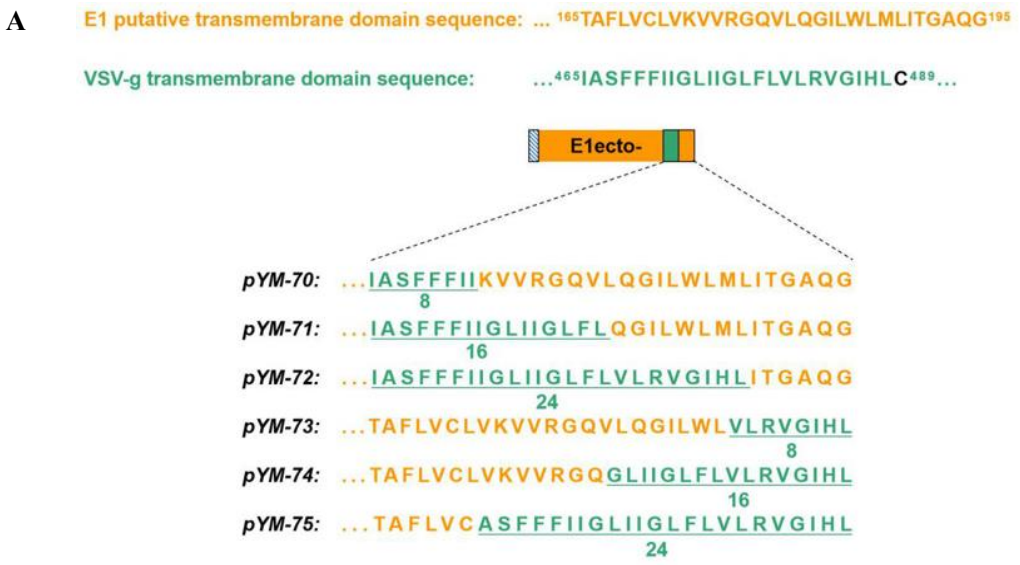
(C) Co-localization analysis for the cell surface expression of the parental protein E1 and pYM-57 (VSV-g[1-464]- E1[166-195]) chimera. RK-13 cells were transfected with the indicated plasmids. At 24h post-transfection, cells were further analysed by indirect immunofluorescence. The RK-13 cells were fixed with 4% PFA, then **permeabilized with 0.05% Triton X-100**, and immune-stained with anti-HA mAb (secondary FITC-anti-mouse) or anti-VSV-g pAb (secondary FITC-anti-rabbit).

5.3.2 The ER retention signal is within the middle part of the transmembrane domain of E1

To further narrow down the area that is responsible for the ER retention of E1, the ectodomain of E1 (1-165) was fused to a series of artificial chimeric transmembrane sequences (as shown in Fig 5.12A). Since the replacement by 100% of VSV-g transmembrane sequence resulted in cell surface expression of E1, the strategy in this study was that the transmembrane sequence from VSV-g partially and progressively substituted for the original sequence of E1 from either *N*-terminal or *C*-terminal end of the putative E1 transmembrane region. It is worth noting that

the total length remained the same as in the original sequence. For instance, the pYM-70 construct contains the E1 ectodomain fused to a chimeric TM domain in which the first 8 residues (TAFLVCLV) were replaced by the first 8 amino acids (IASFFFII) of VSV-g membrane anchor. The rest of the constructs were made in the same manner. Cell surface expression of these chimeras was investigated by IF and flow cytometry (FACS). It is worth noting that the cysteine residue at position 489 in the transmembrane region of VSV-g transmembrane region substitute the original first 8aa (TAFLVCLV) of E1 putative transmembrane anchor sequence at the *N*-terminal. The rest of the constructs were made in the same manner. Cell surface expression of these chimeras was investigated by IF and flow cytometry (FACS). It is worth noting that the cysteine residue at position 489 in the transmembrane region of VSV-g was not used in this study to prevent unpredictable effects from disulphide bond formation.

In a first step, the surface expression of the different chimeras was analysed using FACS analysis. For this purpose, RK-13 cells were transfected with the corresponding expression plasmids. The following day, one sample of each transfected RK-13 dishes was permeabilized with 0.05% Triton-X100 and served as an expression control. A second sample of each transfection reaction group was processed under non-permeabilized condition for cell surface expression analysis. Flow cytometry was used to analyse the presence of cell surface expression of HA tagged chimeric proteins always compared to the pCI empty vector transfected negative control (shown in **Fig 5.12B** and **D left**).



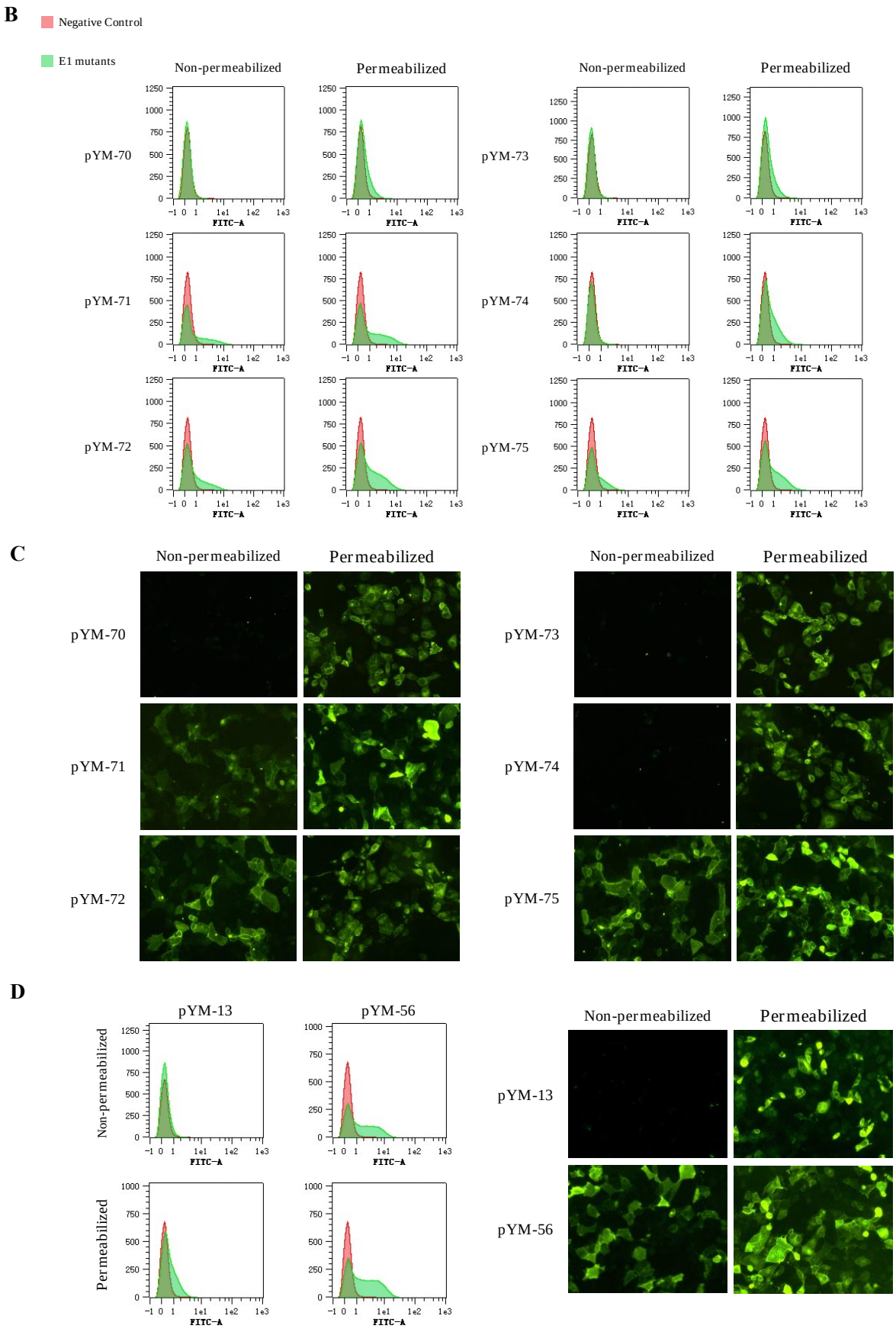

B
Negative Control
E1 mutants
Non-permeabilized
Permeabilized
pYM-70
pYM-71
pYM-72
pYM-73
pYM-74
pYM-75
FITC-A
C
Non-permeabilized
Permeabilized
pYM-70
pYM-71
pYM-72
pYM-73
pYM-74
pYM-75
D
pYM-13
pYM-56
Non-permeabilized
Permeabilized

Figure 5.12 14aa within the middle part of transmembrane domain of E1 is responsible for the ER retention

(A) Schematic representation of the chimeric transmembrane sequences used in this study. The original sequences of TM domains of both E1 and VSV-g are shown above. The sequences of chimeric TM anchors are presented below.

(B) Cell surface expression of HA-tagged chimeric proteins analysed by flow cytometry. The RK-13 cells were transfected with the corresponding expression plasmids. At 24h post-transfection, cells were fixed by 4% PFA and immune-stained with α-HA, α-mouse FITC and then analysed with the MACSQuant. Red peak: The fluorescence signal of RK-13 cells transfected with pCI empty vector under non-permeabilization condition served as a real negative control; Green peak: The fluorescence signal of RK-13 cells transfected with the indicated plasmids under non-permeabilization/permeabilization condition.

(C) Presence of cell surface expression of HA tagged chimeric proteins analysed by IF. The RK-13 cells were transfected with the corresponding expression plasmids. At 24h post-transfection, cells were fixed with 4% PFA, immune-stained with α-HA/α-mouse FITC and then analysed by immuno-fluorescence microscopy.

(D) The control samples for cell surface expression of HA tagged proteins were analysed by both FACS and IF. pYM-13: HA tagged E1 wt, pYM-56: HA tagged E1 ectodomain fused to VSV-g transmembrane sequence. Schematic representation of the constructs (pYM-13 and pYM-56) used in this section are shown in **Fig 5.9A**.

The results of permeabilized samples showed that all fusion proteins were expressed (**Fig 5.12B**). Compared to the pCI control, constructs pYM-71, pYM-72 and pYM-75 transfected cells showed clear cell surface signals. In contrast, the other three fusion proteins showed no cell surface expression and were completely retained in the cell, since there is no significant difference from the pCI negative control.

These FACS results were further confirmed by immunofluorescence analysis (shown in **Fig 5.12C**). RK-13 cells were seeded into a 24-well plate one day before transfection. For each sample two wells were seeded, one for non-permeabilization, the other for permeabilization. On the following day, the RK-13 cells were transfected with the corresponding expression plasmid. At 24h post-transfection, the cells were fixed with 4% PFA and either permeabilized with 0.05% Triton-X100 or processed without any detergent treatment. For the immune-fluorescence staining, α-HA was used as the primary antibody. The cells permeabilized with Triton-X100 served as an expression control, since the antibody was able to penetrate into all compartments. An intact cell membrane, on the other hand, represents an insurmountable barrier for antibodies and, therefore, only surface proteins could be detected without detergent (shown in **Tab 4.13**). The IF results in Fig 5.12C showed that the pYM71, pYM-72 and pYM-75 fusion proteins can be found on the cell surface. Interestingly, the fluorescence signal of pYM-71 transfected cells under non-permeabilized condition was not as strong as that of pYM-72, indicating that there are still some fusion proteins expressed from pYM-71 retained within the cell. For the other samples, signals were not detected on the cell surface. These data fit with the results that we obtained in FACS analysis.

The IF and FACS analysis of the E1/VSV-g fusion proteins indicated that the first 8aa and last 8aa of the E1 transmembrane domain are not important for the retention, since substitution of these residues with VSV-g sequences which do not contain a retention signal has no or nearly no effect on the retention of E1. Fusion proteins start to show significant presence at the cell surface only when the E1 original middle part (K174-L187) was completely replaced by VSV-g sequence. This indicated that the retention signal or at least critical sites for the retention should be present within this middle part.

5.3.3 The polar amino acids in the middle part of E1 TM domain play an essential role in ER localization of E1

To further verify the middle area (from K174 to L187) of the E1 TM domain as critical for the ER retention of E1, two new chimeras were generated (as shown in **Fig 5.13A**). In pYM-76, the middle part sequences of the E1 TM domain were completely replaced by VSV-g sequence (F470-V484), while the rest part still displayed the original sequence. The other construct (pYM-77) still contains the original TM domain middle sequence, changed on both sides by 8aa corresponding to the VSV-g TM sequence. The cell surface expression of these chimeras was also analysed by IF and flow cytometry (FACS). It is worth noting that the cysteine residue at position 489 in the transmembrane region of VSV-g was again not used in this section to prevent unpredictable effects from disulphide bond formation.

As shown in **Fig 5.13B**, both IF and flow cytometry data showed that the cell surface expression of the chimeras only occurs in the absence of the original middle part of the E1 TMD sequence, in other words, the middle 14aa residues located in the E1 TM region are sufficient for the retention of E1.

The conservation of the amino acid sequence of the E1 TM domain throughout all species of pestiviruses was analysed by WebLogo 3 web application (**Fig 5.13C**). Several polar residues are fully conserved including not only four polar, non-charged residues [Glycine (G) and Glutamine (Q)], but also two positively charged residues Arginine (R) and Lysine (K). Those fully conserved residues were hypothesized to be functional for the retention in analogy to other envelope proteins from pestiviral related viruses which are retained because of their polar residues in the TM region (Cocquerel et al., 1999; Cocquerel et al., 1998; Cocquerel et al., 2000b; Radtke and Tews, 2017). To investigate whether these conserved residues in the middle part are essential for the retention of pestivirus E1, mutagenesis analysis was carried out. Firstly, single mutants carrying point mutations at position K174, R177, G178, Q179, Q182 or G183 were investigated via both FACS and IF as described in the previous experiments.

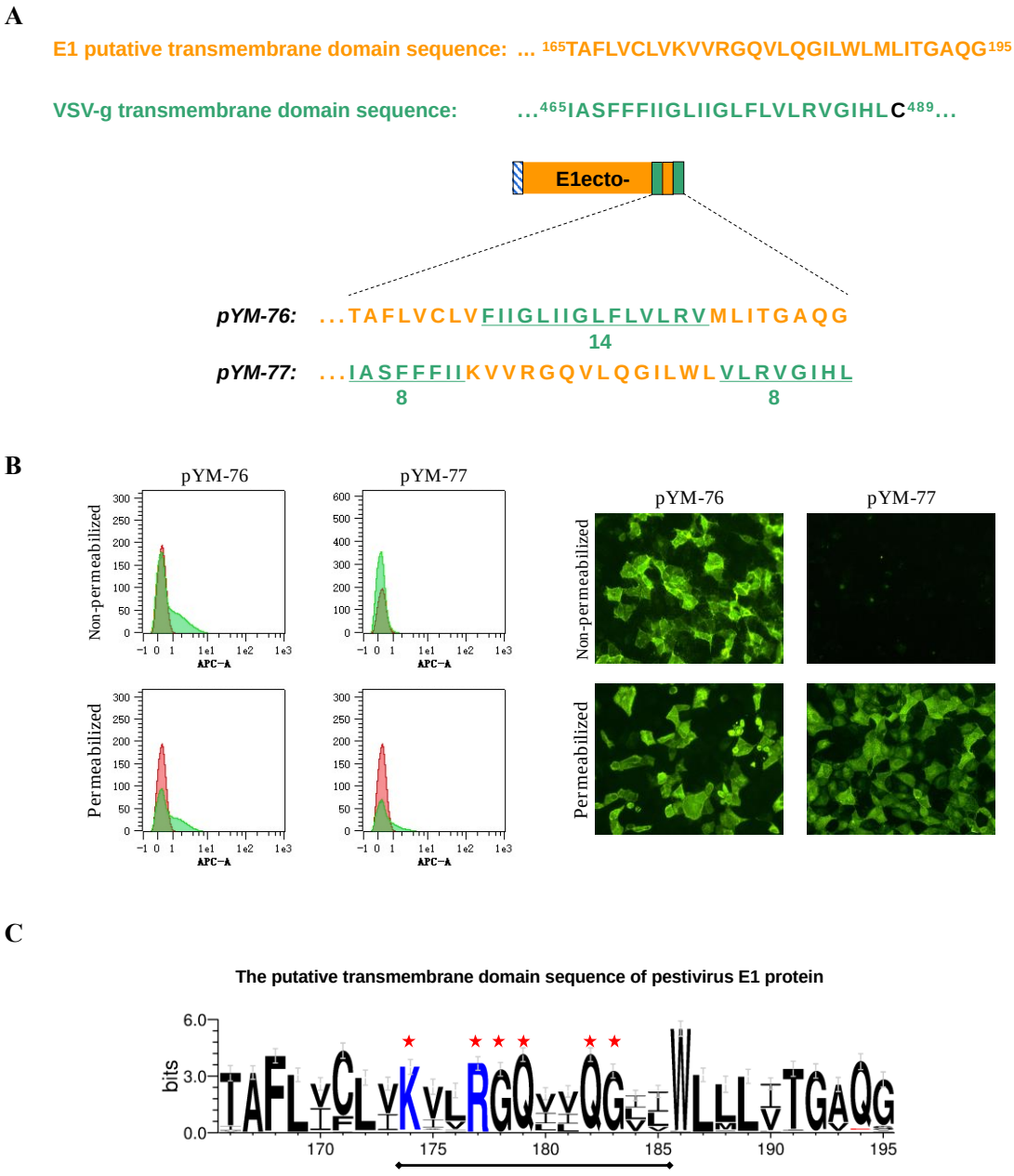


Figure 5.13 The polar residues of E1 TM domain play essential roles in ER retention of E1

(A) Schematic representation of the two new chimeric transmembrane sequences used in this section. The original sequences of the TM domains of both E1 and VSV-g are shown above. The sequences of chimeric TM anchors are presented below.

(B) Cell surface expression of HA tagged chimeric proteins analysed by flow cytometry and IF. The RK-13 cells were transfected with the corresponding expression plasmids. At 24h post-transfection, cells were fixed with 4% PFA and immune-stained with α-HA/α-mouse FITC and then analysed with the MACSQuant and with immuno-fluorescence microscopy. Red peak in FACS: The fluorescence signal of RK-13 cells transfected with pCI empty vector under non-permeabilization condition served as a real negative control; Green peak in FACS: The fluorescence signal of RK-13 cells transfected with the indicated constructs under non-permeabilization/permeabilization condition.

(C) Conservation of amino acid sequences in the putative TM region of pestiviral E1. The sequence logo was generated by WebLogo 3 web application (http://weblogo.threeplusone.com/create.cgi) and demonstrates the alignment of 68 pestivirus E1 sequences throughout all pestiviral species in one letter code. The size of the letters in the sequence logo corresponds to the degree of conservation over the 68 sequences (from pestivirus A to K). The fully conserved residues were highlight with red star, the height of symbols within the stack represents the relative frequency of each amino acid at corresponding position.

Table 5.5 Mutagenesis and insertion analysis for conserved residues in TMD of E1

Single Mutation	Double Mutations	Insertion
K174: A / E / Δ R177: A / E / K / Δ G178: L Q179: N / E / A / Δ Q182: N / A / Δ G183: L	K174A and R177A R177E and Q182A G178L and G183L Q182A and G183A	① Between T166 and A167 'LLALLA' insertion ② Between G178 and Q179 'LLALLA' insertion

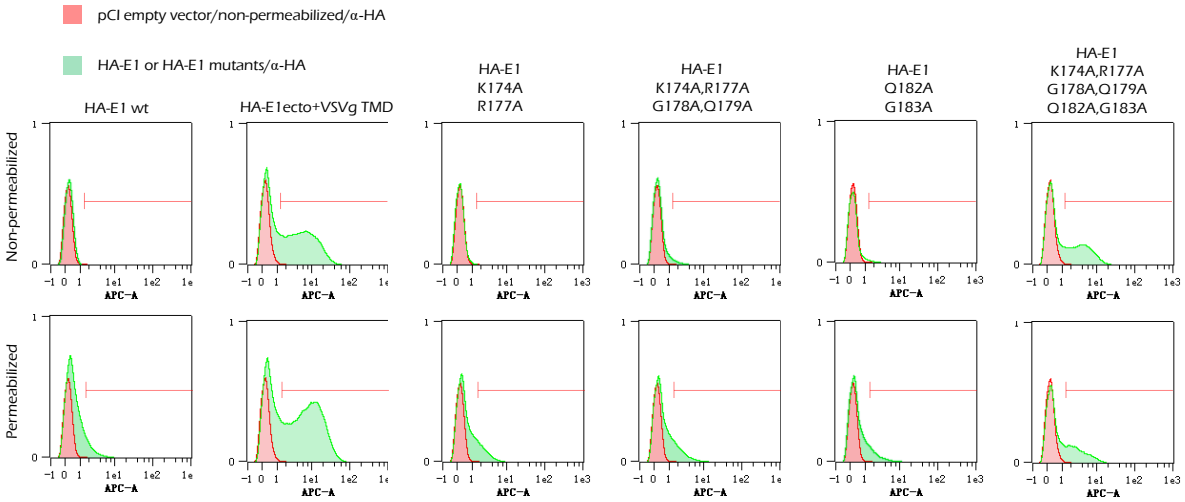


Figure 5.14 Six polar residues of E1 TM domain are important for ER retention of E1

Test of the cell surface expression of HA tagged E1 mutants analysed by FACS. The RK-13 cells were transfected with the corresponding expression plasmids. At 24h post-transfection, cells were fixed with 4% PFA and immune-stained with α-HA/α-mouse FITC and then analysed with the MACSQuant.

Since the FACS and IF data of the pYM-76 and pYM-77 derived fusion proteins further proved that the retention signal was localized within the middle part of E1 transmembrane domain, in addition, the conservation sequence logo showed that 6 polar residues (K174, R177, G178, Q179, Q182 and G183) in this area are fully conserved throughout the whole pestiviral species indicating that they should be important for the retention. To hunt for the critical site which is essential for E1's retention, firstly, a mutagenesis analysis was carried out in which selected polar amino acids were substituted or deleted**.** The corresponding expression plasmids (**Table 5.5 single mutation**) were transfected into RK-13 cells and the surface presence of the proteins was investigated by using both IF and flow cytometry. While all mutations led to slight changes in the surface presence of the HA tagged E1 mutants, none of the tested single substitutions effected the retention of E1 strongly (data not shown). Therefore, the E1 variants with double mutations were generated via QC PCR (**Table 5.5 double mutations** and **Fig 5.14** for two selected double mutants). However, the IF and FACS data still showed that those selected mutations did not increase the surface presence of the mutants.

It is reported that the length of TMDs of membrane proteins can affect the intracellular trafficking and sorting of the proteins. Furthermore, adjusting the length of TMDs of membrane proteins is closely associated with the complexity of communication between subcellular compartments (Bretscher and Munro, 1993). Based on this theory, we made two hydrophobic sequence insertions at T166/A167 or G178/Q179 site to extend the length of the TM region of E1. Surprisingly, no significant subcellular localization change was observed (data not shown). In conclusion, single or double substitutions, or even fragment insertion, in the TM region of E1 could not lead to a significant increase in the surface presence indicating that the retention signal of E1 is apparently not dependent on single amino acids but relies on a stretch of residues so that it can't be destroyed via a replacement of a few residues.

As a next step, two new E1 mutants containing four (pYM-52) and six mutations (pYM-53) were generated for the last attempt. The newly made expression plasmids were transfected into RK-13 cells and the surface presence of the E1 mutants was investigated using flow cytometry. As shown in **Fig 5.14**, four mutations (K174A, R177A, Q182A and G183A) in TMD of E1 did still not lead to surface presence of the E1. However, when all six conserved polar residues were replaced by alanine, E1 presents a plasma membrane localization.

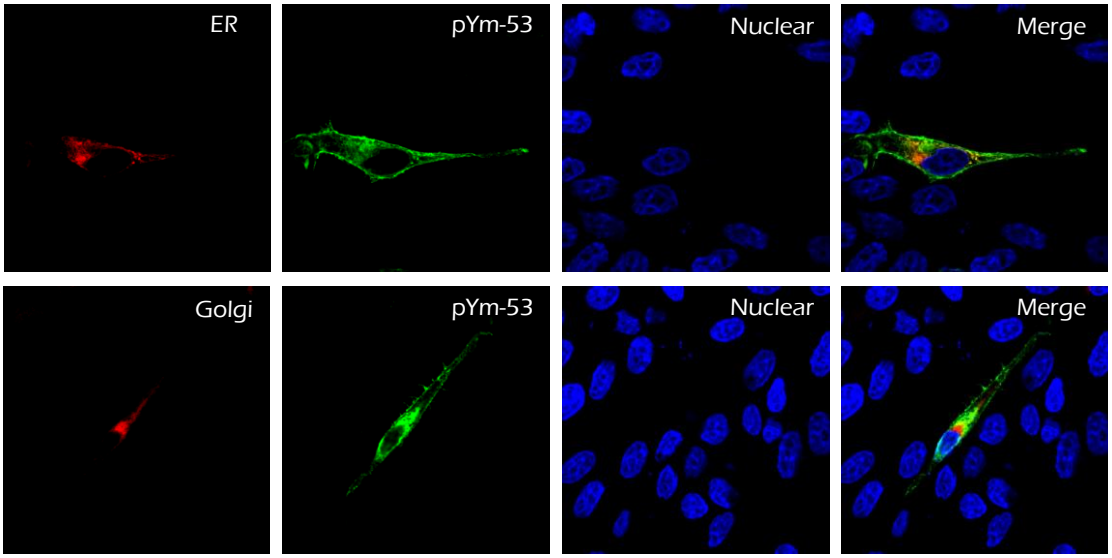


Figure 5.15 Subcellular localization of pYM-53 derived E1 mutant

The pYM-53 construct expressing E1 with six mutations (K174A, R177A, G178A, Q179A, Q182A and G183A) in the TM domain was co-transfected with pDsRed-ER/pDsRed-Golgi, respectively, for the subcellular localization analysis. At 24h post-transfection, cells were fixed by 4% PFA, permeabilized with 0.05% Triton X-100 and stained with specific antibodies against HA (green). Compartments (ER or Golgi) are in red. Nuclei were stained with DAPI.

To investigate the subcellular localization of the pYM-53 derived protein in different cell organelles, the pDsRed-ER or pDsRed-Golgi plasmids were co-expressed with HA tagged E1 mutant. The pDsRed-ER or pDsRed-Golgi is designed for fluorescent labeling of the endoplasmic reticulum (ER) or the Golgi apparatus in mammalian cells. Fluorescence can be observed in living/fixed cells by microscopy or flow cytometry.

Briefly, the ER was fluorescently from labeled by pDsRed-ER, which was transiently co-expressed with HA tagged E1 expression plasmid pYM-53. Similarly, the pDsRed-Golgi plasmid was used for Golgi-apparatus labeling. The results were analysed with a confocal microscope.

As shown in **Fig 5.15**, colocalization analysis of pYM-53 with the ER/Golgi compartment markers demonstrated that the E1 retention defective mutant presented a plasma membrane localization, but was also partially located in the ER and the Golgi compartment. Wild-type E1 was shown to be perfectly colocalized with the ER marker (shown in **Fig 5.8**). The partial ER/ Golgi localization of pYM-53 suggested that the retention signal of E1 has been destroyed, so that this E1 retention defective mutant distributed on the secretion pathway. Compared to the entire cell surface presence of E1-VSVg chimera (pYM-56: E1[1-165]-VSV-g[465-489]), pYM-53 seemed to be different. Since there are some E1 retention defective mutants still retained within the ER also indicated that other factors could have effects on the retention of E1 or maybe the export of the protein goes more slowly. Taken together, these data showed that the fully conserved polar amino acid residues in the middle part of the E1 TM domain play an essential role in ER localization of E1. In addition, those 6 polar residues (Lys174, Arg177, Gly178, Gln179, Gln182 and Gly183) seem to serve as a functional group in retention.

5.3.4 Effect of selected mutations in E1 on the replication of BVDV strain CP7

In the previous section, the mutations which have an influence on the retention of the E1 glycoprotein were identified. In order to find out whether mutations affecting these residues allow the formation of infectious virus particles, selected mutations were inserted into the full-length infectious clone for BVDV CP7 (p798). A publication suggested that the polar charged residues in TM domain of E1 (K174 and R177) play an essential role in BVDV infectivity, and therefore a respective double mutation was also introduced into the p798 as a control.

5.3.4.1 Generation of BVDV CP7 virus mutants with selected E1 mutations

The six mutations K174A, R177A, G178A, Q179A, Q182A and G183A and the double mutation K174A and R177A were first introduced into the infectious clone p798 using common cloning techniques. Plasmid construct p798 contains the cDNA of the entire BVDV CP7 genome (Meyers et al., 1996a). A viral genome-like RNA can be generated via *in vitro* transcription of this cDNA. To check whether this RNA can serve as a replicon that starts the replication of the viral genome, generates infectious virus particles, the obtained RNA was transfected into MDBK-B2 cells via electroporation (EP). For each sample cells after electroporation were separated into two 30mm dishes. One day after electroporation, one dish of each sample was examined for viral protein synthesis using indirect immunofluorescence. The other dish of each sample was further incubated for further subculture and RNA isolation. The viral protein NS3 was detected with the primary monoclonal antibody Code4 (2^{nd} antibody α-mouse FITC).

Immunofluorescence images of each sample after the EP are shown in **Fig 5.16**. The wild-type control and 798-E1 K174A and R177A, 798-E1 with 6 mutation in TMD of E1 all showed a positive signal in the detection of NS3 one day after EP indicating, thus, all the constructs can start the viral RNA replication and gene expression.

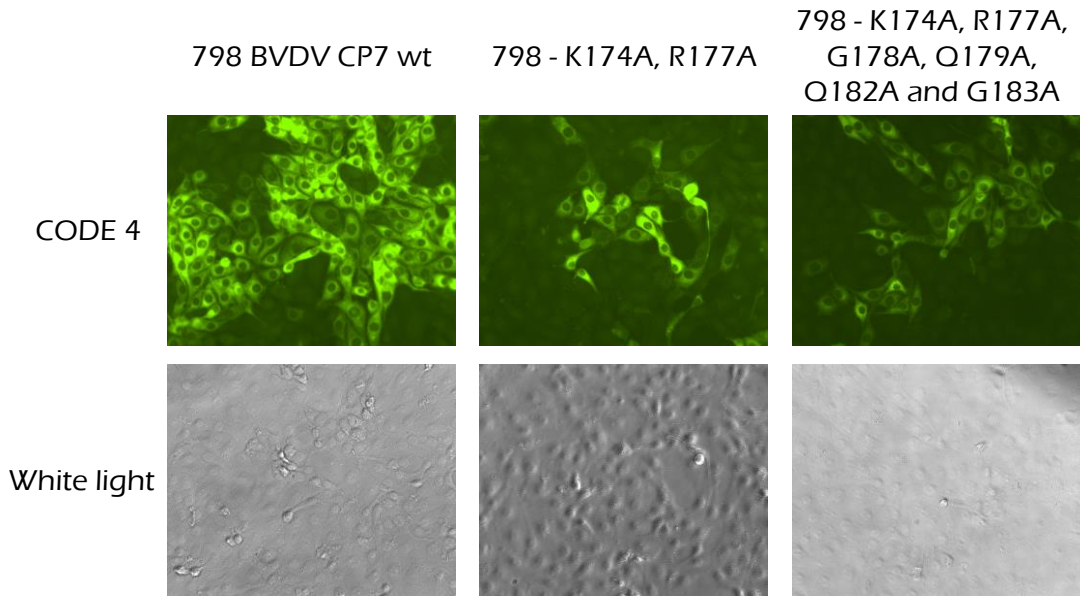


Figure 5.16 Indirect immunofluorescence analysis after electroporation

First, MDBK-B2 cells were transfected by electroporation with the corresponding RNA. One day after EP, the cells were fixed with 4% PFA and permeabilized with 0.05% Triton X-100. The viral protein NS3 was detected using the primary antibody Code4 and α-mouse FITC.

5.3.4.2 Characterization of the CP7 virus mutants

In order to check that the recovered viruses contained the desired mutations and were not reverted, MDBK-B2 cells were infected with the lysates of the transfected cells. This infection was also used to test whether infectious viruses and not just replicons were recovered after EP. 2-4 days after infection, the viral RNAs were extracted using Trizol® Reagent according to the manufacturer's protocol and then used for an RT-PCR with subsequent sequence analysis.

Interestingly, the sequencing results showed that the mutations in the genome derived from the infectious clone p798- K174A and R177A are not reverted, they still existed in the recovered viral genome. However, all substitutions of the infectious clone p798- K174A, R177A, G178A, Q179A, Q182A and G183A, after EP, were fully reverted. This finding further confirmed those six polar residues in the TM domain of E1 served as a functional group in the viral life cycle. This initial result can not exclude the contamination from recovered wild-type BVDV CP7 (p798), it has to be repeated at least three times. I am still working on this experiment, since my contract is going to be finished, hopefully this part of work could be presented before I leave.

5.4 Analysis of the membrane topology of pestiviral E1

Pestiviral proteins are synthesized as a polyprotein cleaved by a signal peptidase and viral proteases. The glycoprotein E1 is translocated as a part of the polyprotein during translation into the ER. For a long time, E1 has been thought to be a membrane protein that is composed of *N*-terminal ectodomain and a *C*-terminal hydrophobic membrane anchor. Often, the *C*-terminal hydrophobic region of viral envelope proteins of positive stranded RNA viruses is composed of two stretches of hydrophobic residues separated by a short segment containing at least one fully conserved positively charged residue (Cocquerel et al., 1999; Cocquerel et al., 1998; Cocquerel et al., 2000b). The first stretch of hydrophobic part is considered to be a membrane anchor of the viral protein, the second part is supposed to act as signal sequences of the following viral protein. The *N*-terminus of E1 is generated by signal peptidase cleavage at the unusual E^{rns} membrane anchor/E1 site. The length of the hydrophobic region at *C*-terminal of E1 making the membrane topology of mature E1 alone difficult to predict.

In the synthesized polyprotein, the internal signal peptides at the *C*-terminus of E1 should be used to target the following E2 glycoprotein to the ER. After signal sequence cleavage, the proposed signal peptide of E2 should remain bound to the *C*-terminus of the E1. It was reported that there is a reorientation of the signal sequences after cleavage at the *C*-terminus of HCV E1 and E2 (Cocquerel et al., 2002) which might in analogy also be true for pestiviral E1. In this section, the membrane topology of E1 should be determined. In addition, we also wanted to clarify whether there are changes in membrane topology of E1 when the signal sequence cleavage occur and the fully processed protein is generated.

5.4.1 Membrane topology of E1 separately expressed E1

5.4.1.1 Membrane topology of E1 alone analysed by selective permeabilization assay

To investigate the membrane topology of separately expressed E1, firstly, the indirect immunofluorescence analysis was carried out with a so-called selective permeabilization assay. The cells were fixed with 4% PFA, then incubated with 0.05% Triton X-100 for 30 min in PBS resulting in the permeabilization of both the plasma membranes and the compartment membranes. Alternatively, cells were treated with a 5 µg/ml digitonin solution for 15 min at 4 °C leading to plasma membrane permeabilization only, while the compartment membranes remain intact and represent an insurmountable barrier for the antibodies used. The 0.05% Triton X-100 permeabilization treatments served as expression control, since the target proteins can be detected in all areas of the cell. In contrast, only proteins on the cell surface and in the cytoplasm can be detected in the digitonin permeabilized preparations.

In this study, two variants of the BVDV glycoprotein E^{rns} with known topology (Tews and Meyers, 2007) were used as controls for the correct selective permeabilization. The plasmid construct pB11 encode the BVDV CP7 wild type E^{rns} with a *C*-terminal V5 tag. All the epitopes of this protein are known to be on the luminal side of the ER and should therefore not be detectable after digitonin permeabilization. pB154 is the other expression plasmid that expresses a variant of E^{rns} with a hydrophobic leucine stretch replacing the original amphipathic helix at the *C*-terminus of E^{rns}. In addition, again a V5 tag was fused to the *C*-terminus of pB154. As a result of the exchange of the amphipathic helix for a hydrophobic region, a transmembrane domain was created, as a result of which the V5 tag is accessible at the cytosolic side of the ER membrane and can therefore also be detected after digitonin permeabilization.

Both E^{rns} variants were expressed in RK-13 cells. The transfected cells were selectively permeabilized and stained with an antibody against the V5 tag (schematic illustration see Fig 5.16A). The staining showed the expected pattern, which confirmed that the selective permeabilization assay was done correctly (**Fig 5.17**).

To determine the membrane topology of E1 alone, in the absence of any other viral protein a double tagged E1 variant which had an *N*-terminal Flag tag and a *C*-terminal V5 tag was used in this study. As shown in **Fig 5.17B**, both tags could be detected in the Triton-X100 permeabilized cells with specific antibodies, which indicated that the proteins were successfully expressed and could be recognized by the corresponding antibodies. After permeabilization with digitonin, only the V5 tag at the *C*-terminus of E1 could be detected. These results showed that the *N*-terminus was on the luminal side of the ER and the *C*-terminus was in the cytosol. Thus, by using the selective permeabilization assay we could preliminarily demonstrate that pestiviral E1 is a type I transmembrane protein that has a *N*-terminal ectodomain in the ER lumen and a *C*-terminal transmembrane anchor.

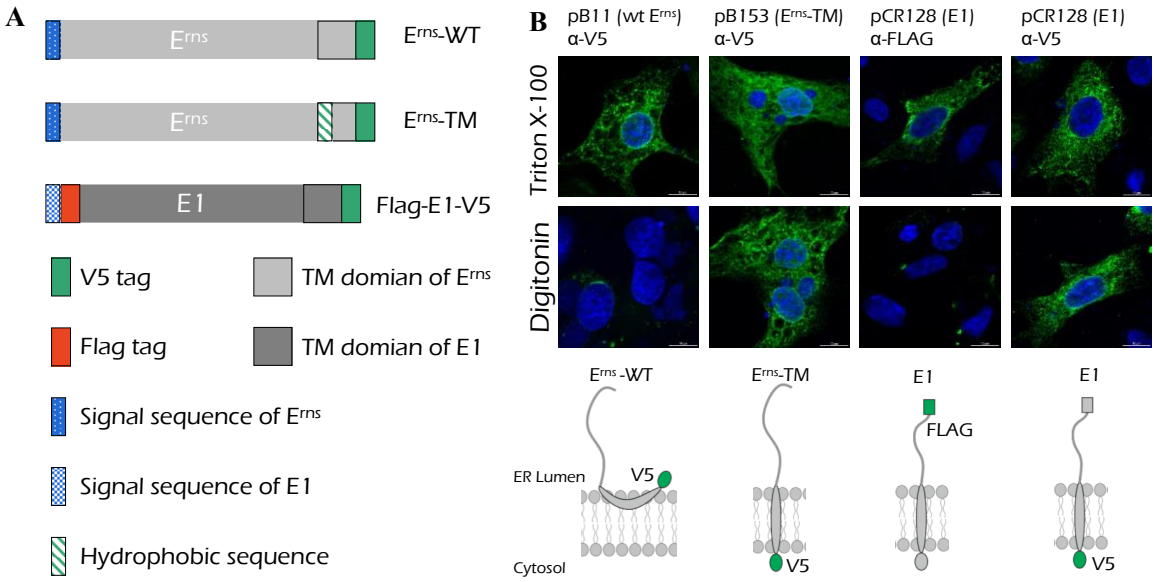


Figure 5.17 Expression and recognition of tag-labeled E^{rns} and E1 proteins

(A) Schematic representation of constructs used in this section: E^{rns} wild type with *N*-terminal V5 tag; E^{rns} TM with leucine stretch instead of the amphipathic helix; double tagged E1 variant which had an *N*-terminal Flag tag and a *C*-terminal V5 tag.

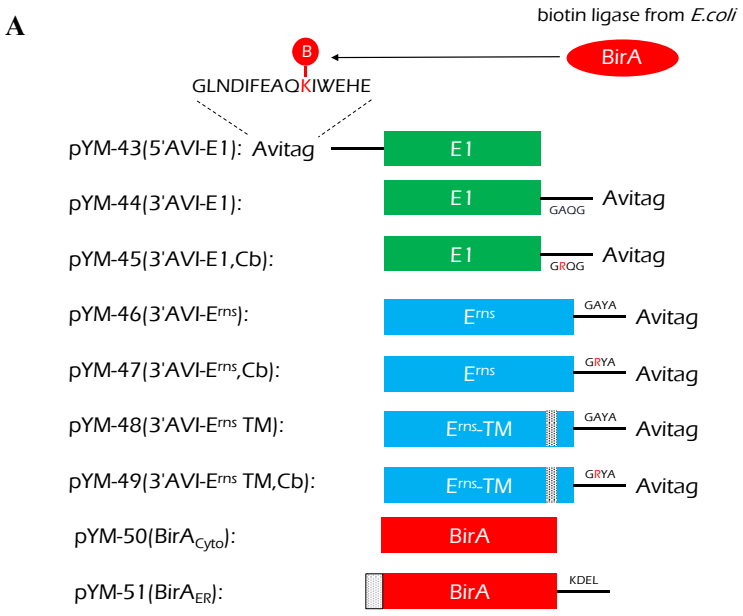
(B) RK-13 cells were transfected with the indicated expression plasmids and fixed with 4% PFA on the following day. Cell membranes were permeabilized with 0.05% Triton-X100, all membranes or the plasma membrane was selectively permeabilized with digitonin. E^{rns}: α-V5, α-mouse Alexa-Fluor-488; E1: α-FLAG, α-mouse Alexa-Fluor-488; α-V5, α-mouse Alexa Fluor 488; Nucleus: DAPI (blue); Below: schematic representation of the membrane topology of the analysed proteins.

5.4.1.2 Membrane topology of E1 analysed by an Avi-tag biotinylation assay

To further verify the conclusion we got from the selective permeabilization assay, a sensitive and selective biotechnology approach called Avi-tag biotinylation assay was used in this study. The Avi-tag is a short peptide of 15 amino acids in length (GLNDIFEAQKIEWHE) which can be covalently attached to biotin in the presence of *E.coli* biotin ligase (BirA). The biotin moiety bound to the Avi-tag can then be detected via (strept)avidin. It is known that the interaction between biotin and streptavidin or avidin is a very strong, sensitive and selective biological

intermolecular force, so the (strept)avidin-biotin binding has been widely used for molecular biology research.

In this section, a short sequence coding for the peptide Avitag as a target for site-specific biotinylation was added genetically to the region coding for *N*-terminus, *C*-terminus of target proteins. By co-expression with the modified BirA biotin ligase, this Avi-Tag can be labeled depending on either the localization of the Avi-tag at the target protein or the distribution of BirA. The advantage of this system over the IF/differential permeabilization approach is that the labelling occurs under native conditions before the cell is destroyed. Specifically, the Avi-tag was added to either the *N*-terminus or *C*-terminus of E1, furthermore, to ensure the Avi-tag at the *C*-terminus can not be cleaved off, we introduced a mutation Alanine to Arginine (A to R) at the -3 position of the SP cleavage site at the *C*-terminus of E1, thereby the von Heigne motif of the signal peptidase cleavage site at the E1 caboxytermius is blocked. Cb version, cleavage blocked 3' Avi-E1 Cb, in **Fig 5.18A**). For the control, E^{rns} wt and E^{rns} TM were also used in this study. It is worth noting that when the Avi-tag was introduced at the *C*-terminus of E^{rns} wt or variant, the cleavage site at *C*-terminal of E^{rns} was also blocked. In this study, we constructed two types of BirA expression plasmids. The plasmid construct pYM-48 ($BirA_{Cyto}$) encoded biotin ligase (BirA) without signal sequence so that all the expressed biotin ligase is located in the cytosol. In contrast, the plasmid pYM-49 based on the pYM-48, additionally contained a signal sequence (signal sequence for BVDV E^{rns}: MALLAWAVITILLYQPVAA) and the well characterized ER retention signal 'KDEL' at the *C*-terminus. Theoretically, this modified BirA (pYM-49) should predominantly be located in the ER.



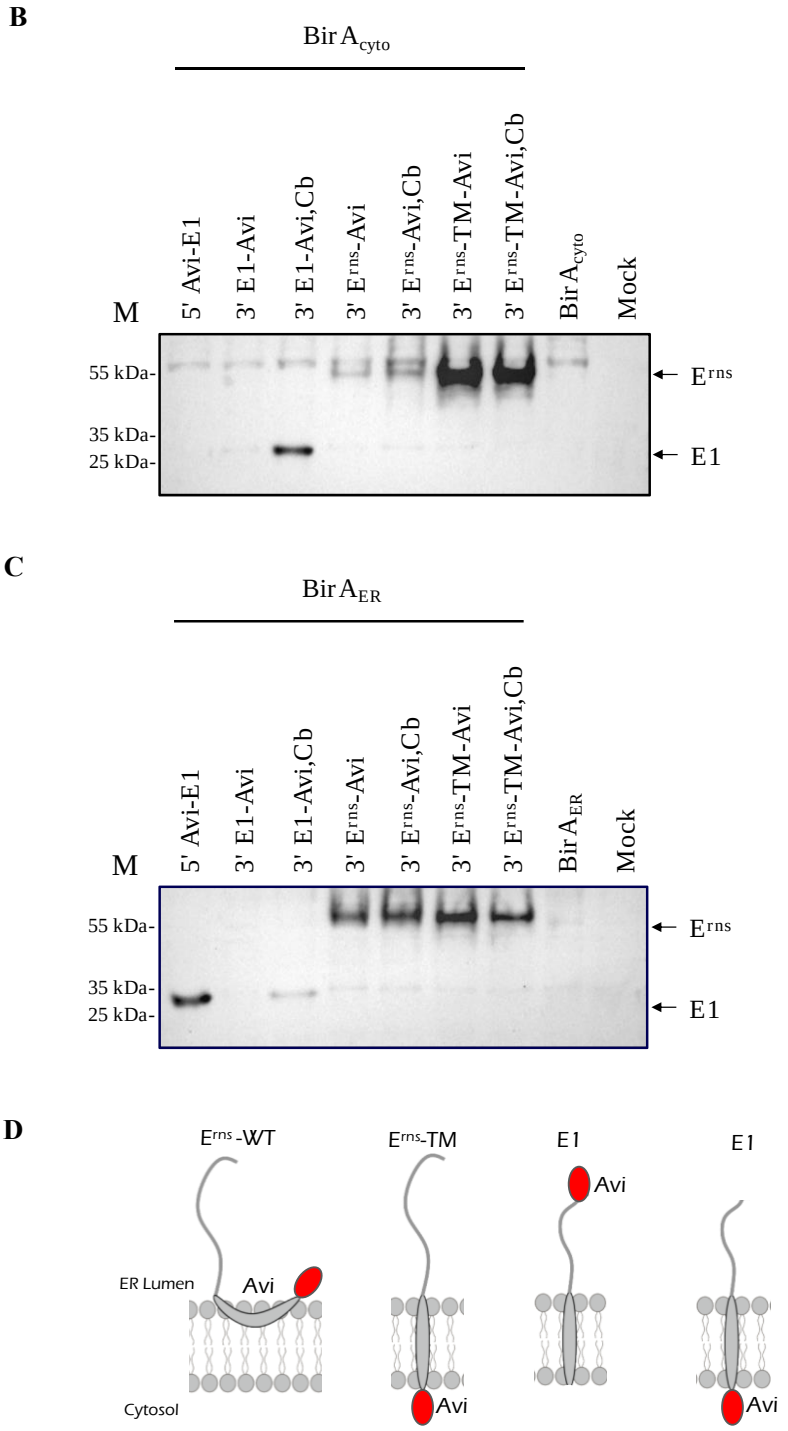


Figure 5.18 Membrane topology of E1 analysed by Avi-tag biotinylation assay

(A) Schematic representation of constructs used in this section.

(B) The appropriate Avi-tagged expression plasmids were co-transfected with pYM-48 (BirA $_{Cytosol}$) into RK-13 cells, and expression products were analysed via Western blot on the following day. The cells were lysed with lysis buffer (+β-ME), then the samples were loaded onto SDS gels and separated electrophoretically. Western blot analysis was carried out using α-streptavidin PO for the biotinylated proteins. Lane numbers refer to construct numbers shown above.

(C) The Avi-tagged expression plasmids were co-transfected with pYM-49 (BirA $_{ER}$) into RK-13 cells, and expression products were analysed via Western blot on the following day. The sample treatment was the same

as described above. Western blot analysis was carried out using α-streptavidin PO for the biotinylated proteins. Lane numbers refer to construct numbers shown above.

(D) Schematic representation of membrane topology of Avi-tagged constructs.

In this BirA mediated Avi-tag biotinylation assay, $BirA_{Cyto}$ or $BirA_{ER}$ were co-transfected with the corresponding Avi-tagged expression plasmids into RK-13 cells. On the following day, the cells were lysed with lysis buffer (+β-ME), then the samples were loaded onto SDS gels and separated electrophoretically. Western blot analysis was performed using peroxidase-coupled avidin (avidin-PO) for detection of biotinylated proteins. As shown in **Fig 5.18B**, when Avi-tagged proteins were co-expressed with $BirA_{Cyto}$, the control plasmid E^{rns}-TM-Avi was detected as the dominant bands (lane 6 and 7) on the blot whereas there was only a very faint band visible for E^{rns} wt-Avi. This indicated that $BirA_{Cyto}$ was at least predominantly expressed in the cytoplasm, since only the tag accessible from the cytosol was biotinylated. For the E1-Avitag fusion protein samples, only the Avi-tag located at the *C*-terminus of E1 with the cleavage site block could be biotinylated, demonstrating that *C*-terminus of E1 is located in the cytosol and the *N*-terminus should be retained in the ER lumen.

When the Avi-tagged constructs were co-expressed with $BirA_{ER}$, both E^{rns} wt-Avi and E^{rns} TM-Avi were biotinylated (**Fig 5.18C**, lane 4, 5, 6 and 7). In the light of published and above presented data, this result suggested that either the *C*-terminus of E^{rns} TM (3'-E^{rns}-Avi) is in part of the cases exposed in the ER lumen, a point that would not have been detected in previous analysis, or the BirA expression construct after introduction of a signal sequence and ER retention signal is not present in the ER alone. In contract, there is still part of $BirA_{ER}$ present in the cytoplasm. That could be the reason why there is still a very slight band showing in the lane 3 in **Fig 5.18C** for E1 with *C*-terminal tag (3'-E1-Avi). The *N*-terminal Avi-E1 (5'-Avi-E1) can be biotinylated in the presence of $BirA_{ER}$, clearly showing that the *N*-terminus of E1 is located in the ER lumen. With these results, the data previously collected from the selective permeabilization assay could be further confirmed.

Taken together, in this section, the membrane topology of pestiviral E1 expressed in the absence of other viral proteins was analysed by using two different independent molecular approaches. We concluded from the results that pestiviral E1 adopts a typical type I transmembrane topology in the absence of other viral proteins.

5.4.2 The TM domain of pestiviral E1 forms a hairpin structure before signal sequence cleavage

Normally, the *C*-terminal part of viral transmembrane proteins in polyproteins can be divided into two parts: the first part is a TM segment that generally consists of mostly 20-25 mostly non-polar residues and the second one is a short hydrophobic sequence (7-12 aa) serving as the signal sequence which directs the translocation of the following precursor. The presence of a signal sequence in the second half of the *C*-terminus of the TM domain of E1 does not fit with a single membrane-spanning topology. Since pestiviral envelope proteins are synthesized as a polyprotein, it is conceivable that the membrane topology of the polyprotein precursor should be different from that found after the signal sequence cleavage occurred since the *C*-terminus of E1 in the polyprotein should be located in the ER-luminal side to allow translocation of the downstream E2. Therefore, the membrane topology of the *C*-terminus of E1 was analysed under the condition that the signal sequence cleavage was hampered.

To achieve this, the pCR-17 (plasmid from Dr. Christina Radtke, shown in **3.12.2**) coding for the BVDV CP7 E1-E2 proteins was used in this study. We introduced a mutation (Ala to Arg) into the *C*-terminus of E1 at position -3 of the cleavage site to block the signal sequence cleavage. For the detection of *N*- and *C*-terminus of E1, a HA epitope (YPYDVPDYA) and a Flag tag sequence (DYKDDDDK) were fused to its *N*- or *C*-terminus, respectively. In addition, E2 was also tagged at the *C*-terminus with an AU1 epitope (DTYRYI) for tracing whether the *C*-terminus of E2 is still accessible from the cytosol when cleavage at the E1/E2 site was blocked (shown in **Fig 5.19A**). To analyse the membrane topology of this un-cleaved E1-E2 precursor, the so-called selective permeabilization assay was carried out as described above. As control, cells permeabilized with 0.05% Triton X-100 were also analysed in parallel. In this experiment, pYM-14 (ss-HA-E1-Flag) and pCR-16 (ss-E2-AU1) were co-transfected together to show the situation after the signal sequence cleavage. The mixture of monoclonal antibodies directed against BVDV CP7 E2 called BVDV Mix was used for the detection of the E2 ectodomain. As shown in **Fig 5.19B**, the *C*-terminus of both E1 and E2 were accessible to the respective antibodies in the digitonin permeabilized cells when cleavage between E1 and E2 occurs. This result fits well to the results and the conclusions described above. Surprisingly, when the cleavage was abolished, all the epitopes were only accessible to their respective antibodies under Triton X-100 permeabilization, but not in digitonin permeabilized cells (**Fig 5.19B**). This data indicates that the *C*-terminal of E1 should orientate toward the luminal side of the ER to adopt a double membrane-spanning structure in the absence of signal sequence cleavage. Moreover, it also suggests that before the cleavage between E1 and E2 has occurred, the *C*-terminus of E2 can't adopt a transmembrane configuration in the ER membrane. The sequence preceding the E2 *C*-terminus is shown as a red line with dotted contour because its behaviour is not clear (**Fig 5.19C**). Since this sequence is highly hydrophobic it is shown as membrane interacting, but this is only speculation at this time being.

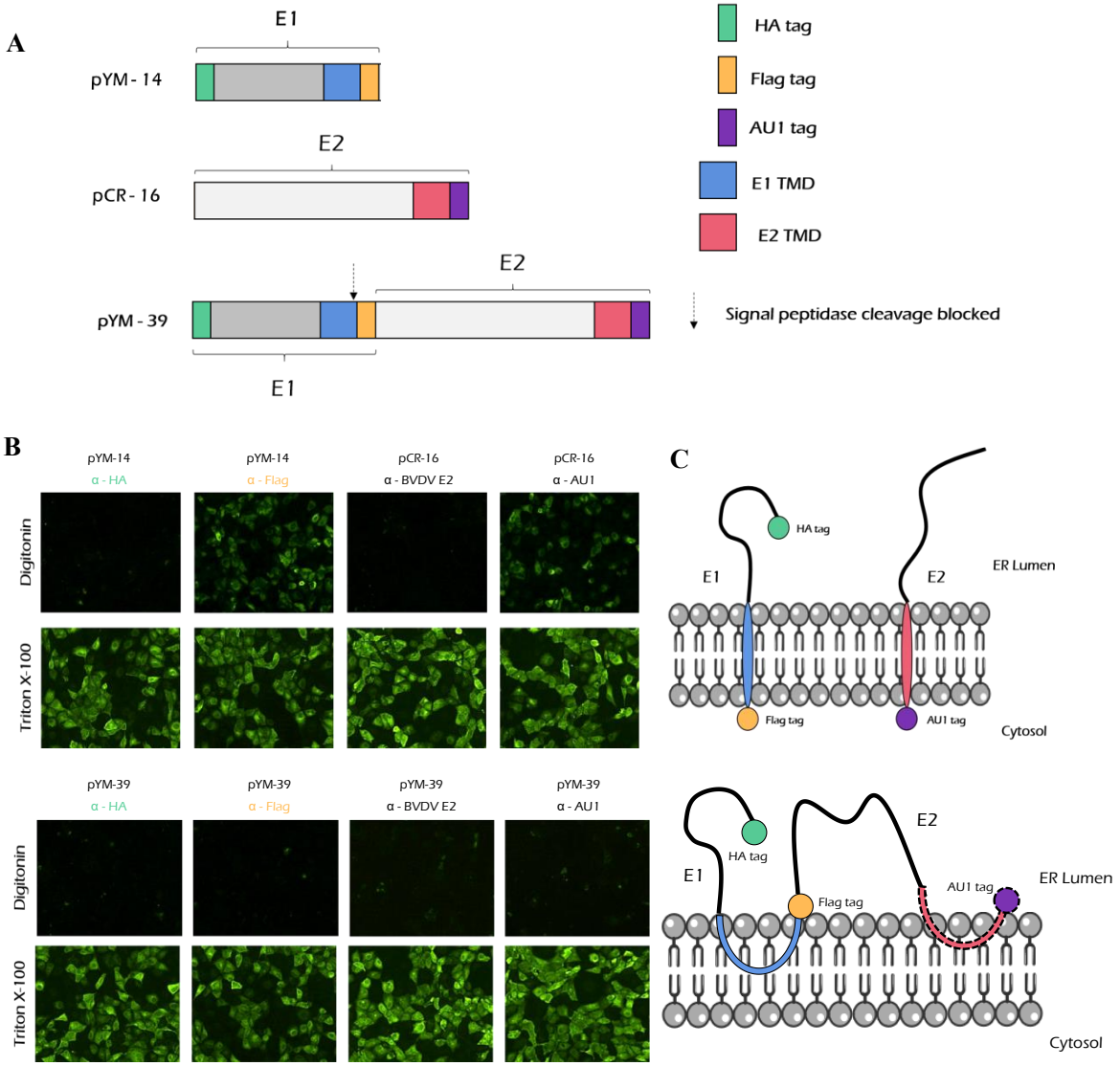


Figure 5.19 The TM region of pestiviral E1 form a hairpin structure before signal sequence cleavage

(A) Schematic representation of constructs used in this section

(B) RK-13 cells were transfected with the appropriate expression plasmids and fixed with 4% PFA on the following day. Either all the cell membranes were permeabilized with 0.05% Triton-X 100, or only the plasma membrane was permeabilized with a digitonin solution. E1: α-HA, α-FLAG, α-mouse Alexa-Fluor-488; E2: α-E2 (BVDV MIX anti E2 ectodomain), α-AU1, α-mouse-FITC;

(C) Schematic representation of the membrane topology of each corresponding samples beside.

5.5 The Prerequisites of E1 oligomerization and E1-E2 heterodimerization

5.5.1 Pestiviral glycoprotein E1 can form homo-oligomers independent of its membrane anchor

The glycoprotein E1 of the closely related HCV was shown to form non-covalently linked trimers on the virion. Additionally, the other two envelope proteins of pestiviruses, E^{rns} and E2, both can form homodimers (Tews et al., 2009b; Thiel et al., 1991; Tucakov et al., 2018; van Gennip et al., 2005). It is shown that E1E2 can form covalently linked E1-E2 heterodimers which were shown to be essential for viral entry (Ronecker et al., 2008). Because of the absence of specific antibodies against E1, it still unknown whether E1 of pestiviruses forms oligomers or not. Therefore, we expressed wild-type E1 and analysed it under non-reducing conditions. As shown in **Fig 5.20B** lane 1, interestingly, the overexpression of pestivirus E1 glycoprotein leads to the presence of three predominant bands which based on their electrophoretic mobility correspond to homo-trimer, homo-dimer and monomer of E1. This is the first time to observe that pestiviral E1 can form homo-oligomers. To determine whether the *C*-terminal membrane anchor has some effects on the oligomerization of E1, we constructed a series of E1 mutants with different length of truncations at the *C*-terminus in the E1 protein (**Fig 5.20A**). These mutants were expressed in RK-13 cells, and the expression of E1 mutants was analysed under non-reducing conditions by Western blot analysis with anti-HA antibody.

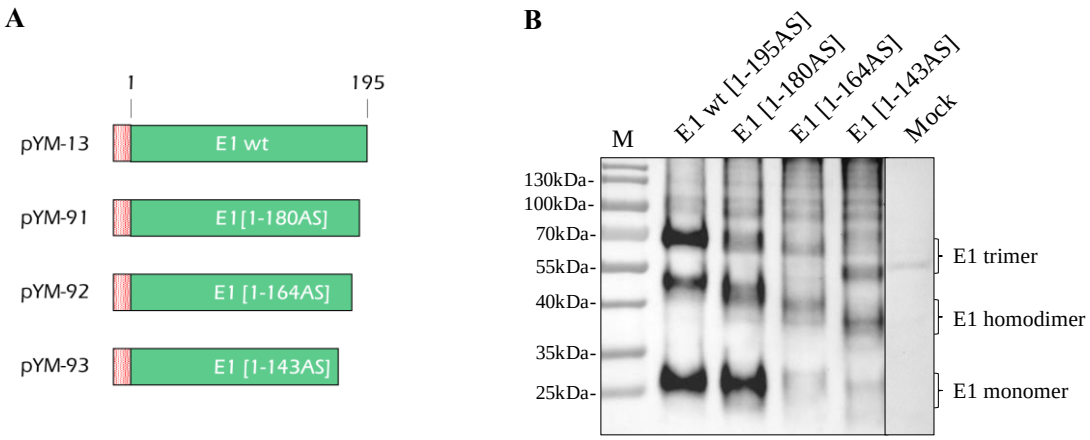


Figure 5.20 E1 can form homo-oligomer independent to its membrane anchor

(A) Schematic representation of constructs used in this section

(B) The HA-tagged E1 and its truncated variants were expressed in RK-13 cells, and expression products were analysed via Western blot on the following day. The cells were lysed with lysis buffer (-β-ME), then the samples were loaded onto SDS gels and separated electrophoretically. Western blot analysis was carried out using primary antibody α-HA and PO labeled 2nd antibodies for the detection of protein. Lane numbers refer to the order of the constructs numbers in **(A)**.

As shown in **Fig 5.20B**, surprisingly, for the E1 mutants with long truncation (pYM-92 and pYM-93) from which the putative membrane anchor was deleted, three clear bands were detected in the transfected-cell lysates. As expected due to the shortening with a decreasing molecular weights. This result indicated that pestivirus E1 can form homo-oligomer even in the absence of the carboxyterminal transmembrane region. It is worth noting that the signal from pYM-91 transfected cells is still comparable to that of the E1 wt. However, weaker bands were observed for the truncated samples with longer deletions (pYM-92 and pYM-93), which might be an indication for secretion of those E1 truncated constructs without membrane anchor.

5.5.2 Both Cys123 and Cys171 play an important role in E1 homo-oligomerization

Pestiviral envelope protein E^{rns} and E2 generate homodimers via intermolecular disulphide bonds established by their almost *C*-terminal cysteine residues (Li et al., 2013; Tews et al., 2009b; Wang et al., 2014). Accordingly, the homo-oligomerization of E1 is also supposed to be mediated via covalent linkage between pairs of cysteine residues. There are six cysteine residues distributed in E1 of BVDV CP7 strain. According to the alignment shown in **Fig 5.3**, the first 5 cysteine residues (at position 5, 20, 24, 94 and 123) are highly conserved throughout nearly all pestivirus species which was considered to be important for the stability and function of E1. In BVDV-2 strains, the 6th cysteine, which is located at position 171, is missing. This indicated that this site could be dispensable, however, based on some computational model prediction, it was suggested that this Cys171 should most likely be the critical site for E1E2 heterodimer formation (Wang et al., 2014). This point still await further experiential verification. To identify the potential determinants of E1 oligomerization, a series of E1 mutants containing single and double exchanges (Cys to Ser) were generated by using the site-directed mutagenesis. E1 wt and mutants were expressed in RK-13 cells, respectively, and the expression of E1 or E1 mutants was determined by Western blot analysis with anti-HA under non-reducing condition. As shown in **Fig 5.21**, compared to the E1 wt, the individual substitution at single exchanges affecting the first 4 Cys (position C5S, C20S, C24S and C94S) had no effect on the E1 oligomerization. Interestingly, both C123S and C171S reduced the presence of homo-trimer and homo-dimer of E1, C123 has wider influence on the oligomer formation of E1. If C123S/C171S double mutations were introduced into E1, only E1 monomer could be detected on the blot. These data also indicated that the first 4 cysteine residues most likely form intramolecular disulphide bonds, while the last two cysteines are involved in intermolecular linkage formation. Among them, Cys123 has more influence on the oligomerization of E1.

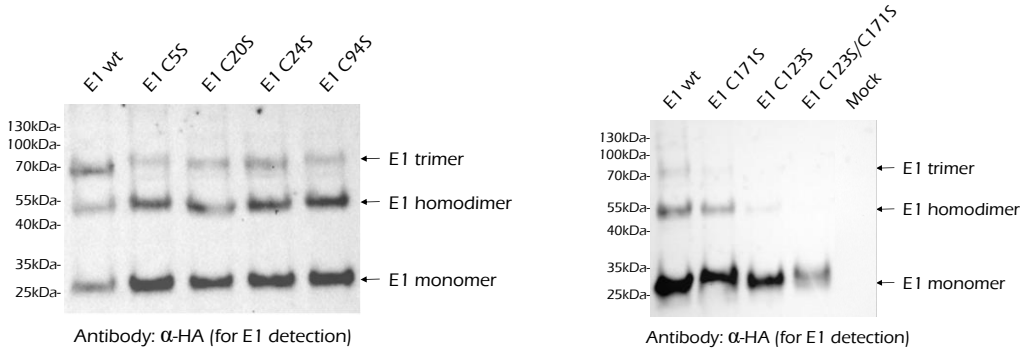


Figure 5.21 Both Cys123 and Cys171 play an important role in E1 homo-oligomerization

The given HA-tagged E1 wt and cysteine lacking mutants were transfected into RK-13 cells, and expression products were analysed via Western blot on the following day. The cells were lysed with lysis buffer (-β-ME), then the samples were loaded onto SDS gels and separated electrophoretically. Western blot analysis was carried out using primary antibody α-HA and secondary antibody anti-rabbit-PO for the detection of protein.

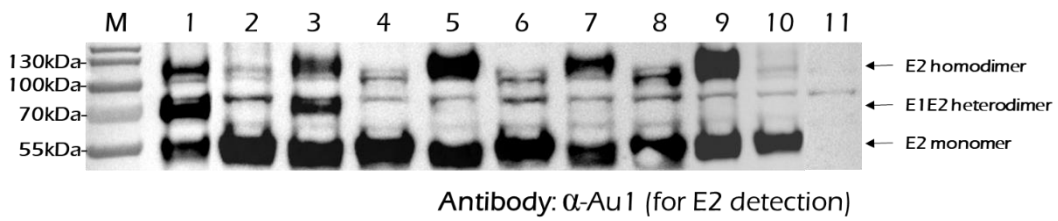
5.5.3 Critical sites for E1-E2 heterodimer formation and E2 homodimer formation

It is well known that covalently disulphide-linked E1-E2 heterodimers are needed for pestiviral infectivity (Ronecker et al., 2008). Publications (Li et al., 2013; Wang et al., 2014) suggested that except for cysteine residue at position 295 in E2, all the other cysteines of E2 form intramolecular disulphide bonds. Cys295 is the only free Cysteine residue in E2 that makes this

site the best candidate for the necessary disulphide linkage with E1 or with itself in homodimers. However, it is unclear which of the E1 cysteines is involved in E1-E2 heterodimer formation. To further investigate the prerequisite for E1-E2 crosslinking, AU1 tagged E2 or E2 mutants were used in this study. A substitution C295S was introduced into E2 for blocking the only unlinked cysteine residue in E2. E1 wt or E1 mutants containing single and double mutations (Cys to Ser) were co-transfected with E2 wt or E2 mutant (C295S) in RK-13 cells respectively, and the presence of E1-E2 heterodimer was analysed by Western blotting with anti-AU1 serum under non-reducing conditions. As shown in **Fig 5.22**, co-expression of E1 wt with E2 wt (shown in lane 1), resulted in a dominant band of ~70 kDa which according to size corresponds to E1-E2 heterodimer. In addition, bands of ~55 and ~110 kDa were visible that represent E2 monomer and homodimer, respectively. However, when E1 wt was co-expressed with E2 (C295S) mutant, both E1-E2 heterodimer and E2 homodimer were diminished (shown in lane 2). This situation is also true for other E1 variants when co-expressed with E2 (C295S) mutant (as shown in lane 4, 6 and 8). When E2 wt was expressed alone (shown in lane 9), we could see E2 homodimer with nearly equal amount as the E2 monomer. In contrast, when the only free cysteine in E2 was blocked, E2 homodimer completely disappeared (shown in lane 10). Those results indicated that the Cysteine residue at position 295 in E2 is critical for both E1-E2 heterodimerization and E2 homodimerization.

In addition, E1 mutants containing single and double mutations (Cys to Ser) were also tested in this study. The mutants E1 (C171S), E1 (C123S) and E1 double mutation (C171S and C123S) were co-transfected with E2 wt into RK-13 cells and analysed as described above. When E1 (C171S) was co-expressed with E2 wt, the E1-E2 heterodimer was still present (as shown in lane 3) indicating that Cys171 in E1 is not necessary for the heterodimer formation. However, when E2 wt was expressed in the presence of the E1 mutant (C123S), the amount of heterodimer was extremely reduced to about the level of the mock control and the amount of the E2 homodimer recovered to normal compared with the sample from E2 wt expressed alone (lane 5). These results clearly showed that Cys123, not Cys171, is the important site for E1-E2 heterodimer formation. This conclusion is contradictory to the computational prediction (Wang et al., 2014). This data also suggested that E2 prefer to form E1-E2 heterodimer in the presence of E1. The E2 homodimer might be an excess product for E2 not engaged in E1-E2 heterodimer formation. There are always some non-specific bands presenting in each group, which is most probably because of the background from the antibody.

	1	2	3	4	5	6	7	8	9	10	11
a	E1 wt	E1wt	E1 C171S	E1 C171S	E1 C123S	E1 C123S	E1 C123S/C171S	E1 C123S/C171S			Mock
b	E2 wt	E2 C295S	E2 wt	E2 C295S	E2 wt	E2 C295S	E2 wt	E2 C295S	E2 wt	E2 C295S	



Antibody: α-Au1 (for E2 detection)

Figure 5.22 Cys123 in E1 is the critical site for E1-E2 heterodimer formation; Cys295 in E2 is essential for not only E1-E2 heterodimerization but also for E2 homodimer formation

The given HA-tagged E1 wt or cysteine lacking mutants were co-transfected with AU1-tagged E2 or E2 mutant (C295S) into RK-13 cells, and expression products were analysed via Western blot on the following day. The cells were lysed with lysis buffer (-β-ME), then the samples were loaded onto SDS gels and separated electrophoretically. Western blot analysis was carried out using primary antibody α-AU1 and secondary antibody anti-rabbit-PO for the detection of protein. The table above of the gel provides the composition of the samples.

5.5.4 E1-E2 heterodimer formation is independent of the transmembrane region of E1

We showed that there is a dynamic membrane topology change of the transmembrane region of E1 after signal sequence cleavage with a relocation of the E1 *C*-terminus from ER to the cytosolic side. This is also true for E2 of pestiviruses (Radtke and Tews, 2017) and the closely related HCV envelope proteins (Cocquerel et al., 2002). It is supposed to be linked to multifunctionality of the membrane anchor, like E1-E2 heterodimer formation and ER retention in HCV. To determine whether the TM domain of E1 affects the described E1-E2 heterodimerization (Thiel et al., 1991), in this section, several *C*-terminally truncated E1 variants were tested for heterodimer formation ability.

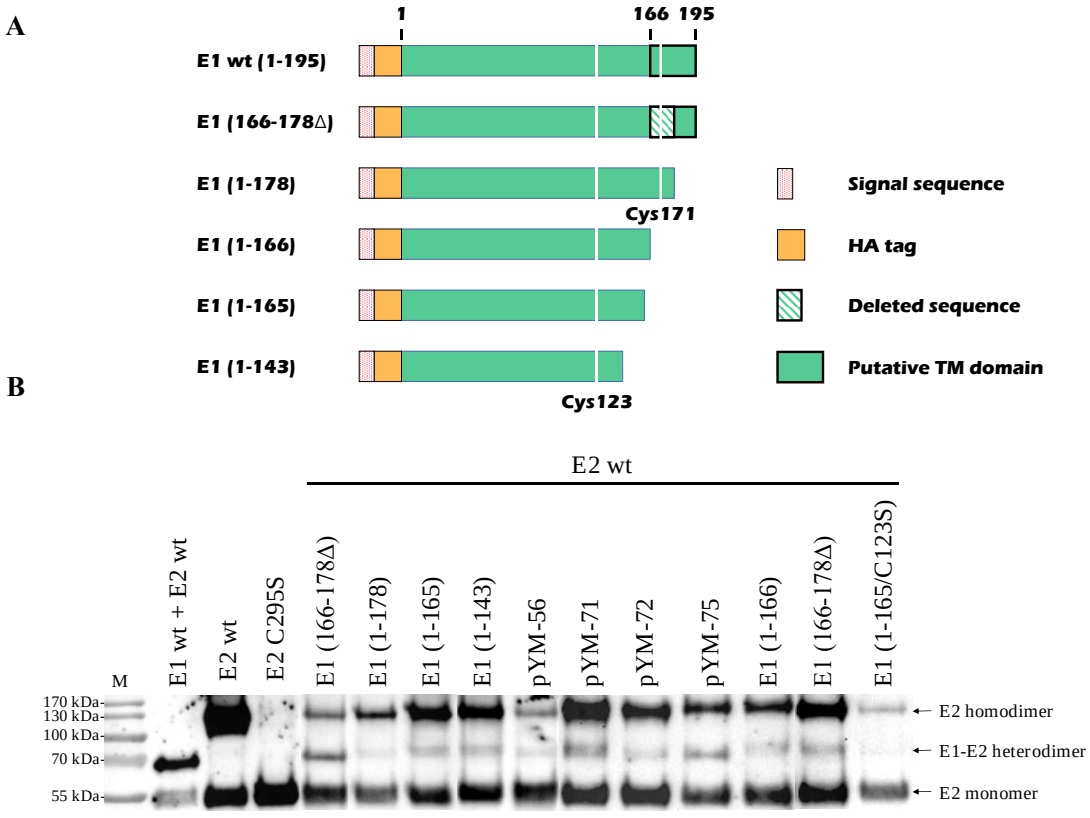


Figure 5.23 Heterodimer formation analysis

(A) Schematic representation of E1 wild type and truncated E1constructs used in this section. Cysteine residues Cys123 and Cys171 are presented as white lanes in the green bar presenting the proteins.

(B) The given *C*-terminally truncated E1 expression plasmids were co-transfected with E2 wt into RK-13 cells respectively, and co-expression products were analysed via Western blot on the following day. The cells

were lysed with lysis buffer (without β-ME, under non-reducing conditions), then the samples were loaded onto SDS-PAGE gels and separated electrophoretically. Western blot analysis was carried out using α-E2 (WB214) for the detection of E1-E2 heterodimer and E2 monomer/homodimer. Lane numbers refer to construct numbers shown above.

As shown in **Fig 5.23**, wild type E1 formed predominantly E1-E2 heterodimer in the presence of wild type E2 (**Fig 5.23B**, lane 1), whereas E2 formed large amount of E2 homodimer in the absence of E1 (lanes 2 and 3). Cys295Ser substitution in E2 can prevent E2 homodimer formation. Those three samples in lanes 1, 2 and 3 served as the positive control for the E1-E2 heterodimerization analysis. Different truncated variants of E1 were first analysed in this study together with E2. Western blot result of lane 4-7 in Fig **5.23B** showed that all the *C*-terminally truncated E1 still formed E1-E2 heterodimer in the presence of wild type E2. This is even also true for the construct in which the entire TM domain was deleted construct (lane 7). But the amount of heterodimer was dramatically reduced indicating that E1 *C*-terminus with the membrane anchor plays an important role in heterodimer formation. Next, we examined the effect of the replacement of TM domain of E1 for heterodimerization. Four representative constructs in which the TM domain of E1 was fully or partially replaced by that of VSV-g, furthermore were established. Those constructs were all shown before to be presented on the cell surface. Interestingly, even when the TM region of E1 was fully or partially exchanged by the corresponding sequence of VSV-g (shown in **Fig 5.12**), E1-E2 heterodimer could be detected (lane 8-11) but again only low amounts. It seems that the TM domain of E1 is dispensable for the dimerization but renders it much more efficient. The E1 ectodomain (1-166) was shown to be able to form E1-E2 heterodimer (lane 12), however, when a Cys123Ser mutation was introduced into this expression construct, only monomer and homodimer of E2 could be detected. Notably, this is also a further prove for our previous conclusion that Cys123, not Cys171, is important for the heterodimer formation of E1-E2.

5.5.5 E1 overrules the retention of E2 via intermolecular disulphide bond formation

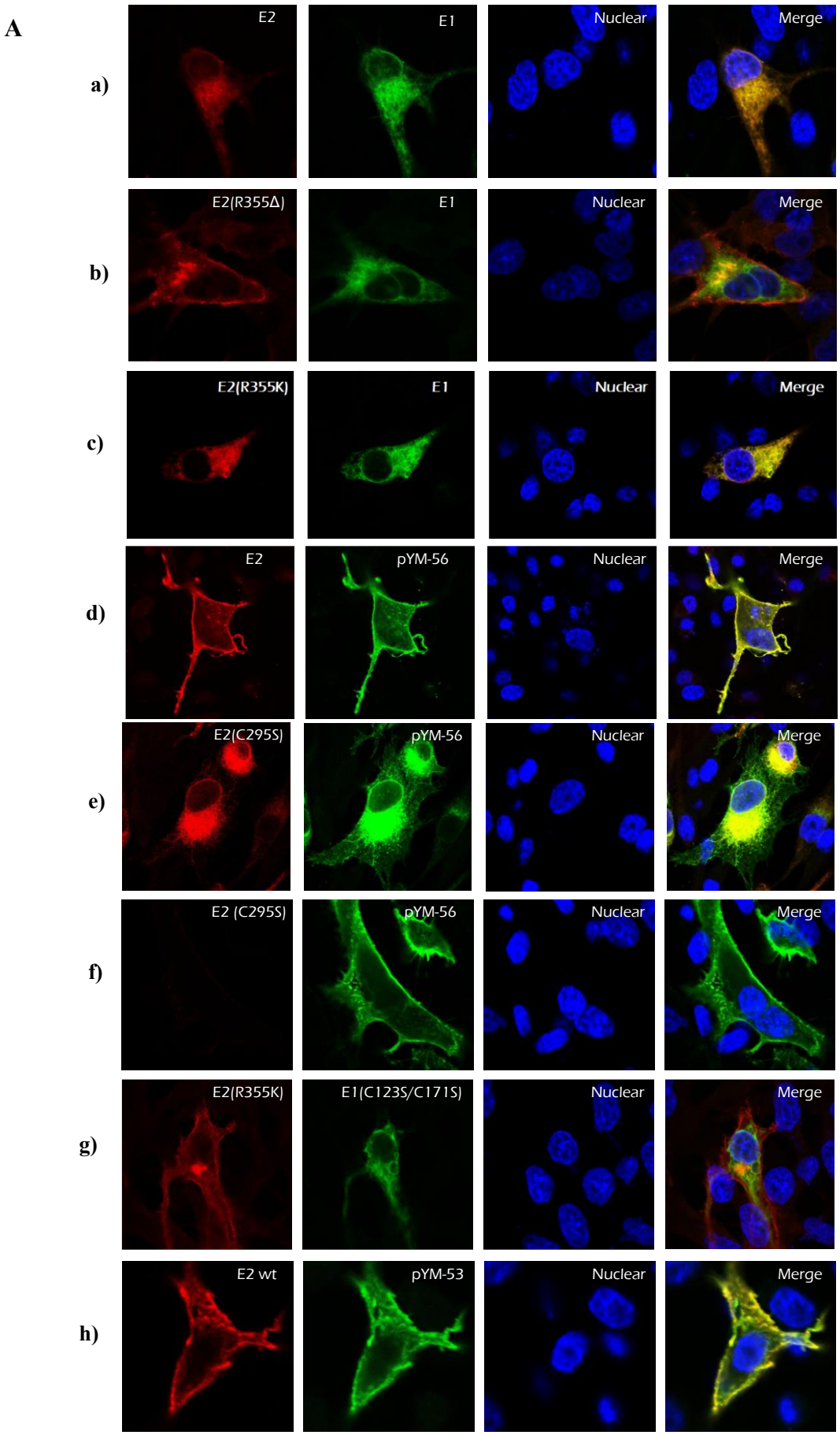
The initial work in our lab demonstrated that the arginine at position 355 in E2 has a big effect on the retention of E2. Several substitutions of R355 like R355A, R355K, R355E and R355Δ could had different effects on the cell surface presence of E2. Furthermore, published data also showed that co-expression of E1 can compensate for the reduced retention of some E2 mutants (Radtke and Tews, 2017). However, the retention deficit of E2 mutants E2 R355A and E2 R355Δ could not be compensated by E1. As already mentioned, R355 is important for the interaction of E2 and E1. Ronecker (Ronecker et al., 2008) showed that the E2 R355A mutation prevents generation of E1-E2 heterodimers. The exchange E2 R355K, however, had no influence on the formation of the E1-E2 heterodimers. These results suggest that heterodimerization between E1 and E2 could be essential to compensate for the reduced E2 retention. Moreover, the natural pestivirus isolate BVDV NewYork'93 (Meyer et al., 2002) contains a R355K substitution indicating that this mutation at position 355 is tolerable for the virus.

To investigate the mechanism of this interesting phenomenon, two representative E2 retention deficit mutants pCR-78 (E2 R355K) and pCR-79 (E2 R355Δ) were used in this study. First of all, we tested the cellular colocalization of wild type E1 with E2 or E2 mutants. The HA-tagged E1 expression plasmid was co-transfected with plasmids coding for E2 or E2 mutants into BHK-21 cells. On the following day, the cells were fixed with 4% PFA and permeabilized with 0.05% Triton-X100. Staining was done with HA-tag antibody plus anti-rabbit label and BVDV E2 mAb Mix plus anti-mouse label. As shown in Fig 5.20A, E1 wt showed a good intracellular

colocalization with E2 wt without any cell surface localization (**Fig 5.24A.a**). When the E2 R355Δ mutant was co-expressed with E1 wt, E1 and E2 showed a totally different cellular localization. Most of E2 R355Δ presents on the cell surface whereas E1 is still retained within the cell (**Fig 5.24A.b**). Interestingly, when R355 in E2 was replaced by lysine (R355K), E1 compensated the reduced E2 retention showing retention behaviour similar to E2 wt (**Fig 5.24A.c**). These results agree with the preliminary conclusions from our lab (Radtke and Tews, 2017). Then we examined those two E2 mutants for their ability to form heterodimers with E1 wt. Electrophoretic separation of the proteins under non-reducing conditions (followed by Western blot) was carried out. As shown in **Fig 5.24B**, E2 containing mutation R355K still can form E1-E2 heterodimers like E2 wt, however, when the arginine at position 355 in E2 was deleted, E1-E2 heterodimer can't be detected at all. Moreover, also the E2 homodimer was not detectable (lane 1, 2 and 3). To prevent the covalent linkage between E1 and E2 mutant R355K, the E1 mutant haboring C123S/C171S was used. When the covalent disulphide linkage between E1 and E2 was hampered (**Fig 5.24B**, lane 6), E1 lost the ability to compensate for the reduced E2 retention, leading to E2 present on the cell surface like the R355Δ mutant (**Fig 5.24A.g**). This data indicated that E1-E2 heterodimer formation is essential for the ability of E1 to compensate for the retention deficit of the E2 mutant.

Furthermore, the representative construct pYM-56 (HA tagged E1ecto + VSV-g TMD), which results in the cell surface presence of E1 was also tested in this study. Since E2 wild type contains a retention signal of its own, we wanted to know whether E2 still can keep a E1 retention deficient mutant within the cell. For this purpose, pCR-16 coding for AU1 tagged BVDV CP7 E2 was co-expressed with pYM-56 in BHK-21 cells. Surprisingly, both the E1 retention deficient mutant and E2 were present on the cell surface (as shown in **Fig 5.24A.e and f**). This indicates that E1, somehow, can overrule the retention signal of E2. Western blot results (**Fig 5.24B**, lane 4) showed that the E1 retention deficient mutant still form heterodimer with E2. However, when C295S was introduced into E2 for interrupting the covalent linkage between E1 and E2 (**Fig 5.24B**, lane 5), E2 can't follow the E1 mutant to go to the cell surface, but seems colocalized with E1 in Golgi apparatus. These results clearly demonstrated that E1 can overrule the retention of E2 when covalently disulphide linked E1 with E2, and it also suggested that E1 has a more dominant role in subcellular localization than E2.

As shown in **Fig 5.24A.h**, the pYM-53 construct which is code for HA-tagged E1 contains six mutations in its transmembrane region that results in the cell surface presence of E1 can overrule the retention single of E2 wt. Additionally, the E1-VSVg chimera plasmids which showed the cell surface expression (like pYM-71, pYM-72 and pYM-75) were also tested in this section. All of them make E2 wt follow them to go to the cell surface, furthermore, the E1-E2 heterodimerization still exist between those E1 variants and E2 wt (data not shown). This finding further proved the importance of E1-E2 heterodimer formation in the ability of E1 to compensate for the retention deficit of the E2 mutant.


A
a)
E2
E1
Nuclear
Merge
b)
E2(R355Δ)
E1
Nuclear
Merge
c)
E2(R355K)
E1
Nuclear
Merge
d)
E2
pYM-56
Nuclear
Merge
e)
E2(C295S)
pYM-56
Nuclear
Merge
f)
E2 (C295S)
pYM-56
Nuclear
Merge
g)
E2(R355K)
E1(C123S/C171S)
Nuclear
Merge
h)
E2 wt
pYM-53
Nuclear
Merge

B

	1	2	3	4	5	6
a	E1 wt	E1 wt	E1 wt	pYM-56	pYM-56	E1 (C123S/C171S)
b	E2 wt	E2 (R355K)	E2 (R355Δ)	E2 wt	E2 (C295S)	E2 (R355K)

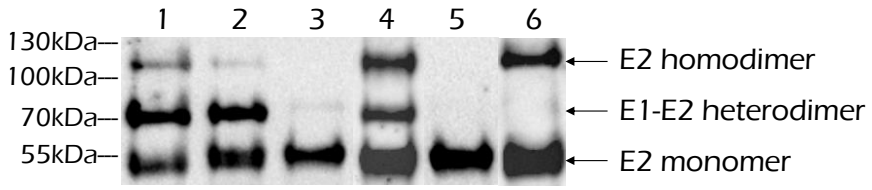


Figure 5.24 E1 overrules the retention signal of E2 via intermolecular disulphide bond between E1 and E2

(A) Co-localization analysis of E1 (or variants) and E2 (or variants) via confocal microscopy. BHK-21 cells were transfected with the appropriate expression plasmids and fixed with 4% PFA on the following day. All the cell membranes were permeabilized with 0.05% Triton-X100 (**d, f and h** samples were **not permeabilized**). E1: α-HA, α-mouse Alexa-Fluor-488;

(B) Western blot analysis for E1-E2 heterodimer formation of corresponding plasmid. α-E2 WB 214, α-mouse PO.

5.5.6 Both Cys123 in E1 and Cys295 in E2 are important for viral infectivity

Mutation analysis which has an influence on the E1-E2 heterodimerization were identified. In order to find out whether substitution of these residues interferes with the formation of infectious virus particles, selected mutations were introduced into the infectious clone for BVDV CP7 (clone p798).

In this section, the mutations C171S and C123S in E1 as well as C295S in E2 were introduced into BVDV CP7 full-length infectious clone 798 using common cloning techniques. Plasmid 798 contains the cDNA of the entire BVDV CP7 genome. Using T7 RNA polymerase transcription, viral genome like RNAs can be generated. These RNAs were used for electroporation (EP) of MDBK-B2 cells as described in **section 4.3.4**. Each sample was electroporated into two 30mm dishes, one for the duplicate. 24h after electroporation, the cells in one dish of each sample were tested for viral protein using indirect immunofluorescence. The viral protein NS3 was detected by the primary antibody Code4 which is the monoclonal antibody for pestivirus NS3 protein. After staining with anti-mouse FITC, samples were further checked with normal immuno-fluorescence microscopy.

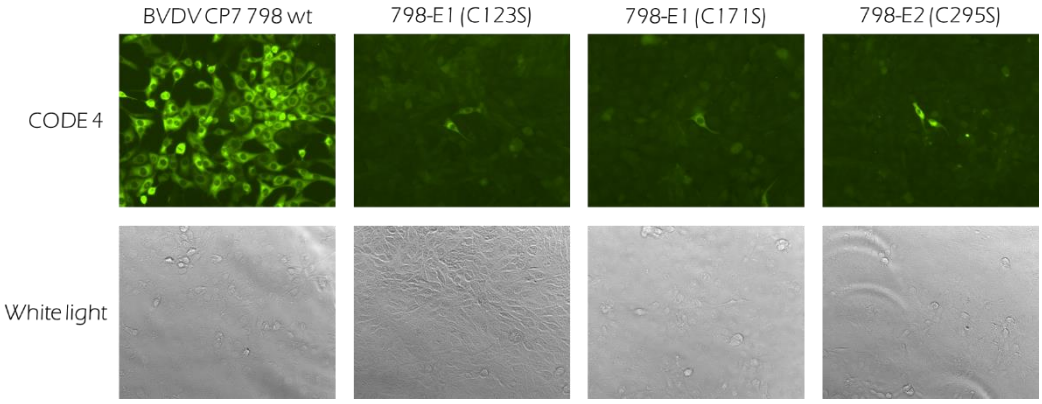


Figure 5.25 Indirect immunofluorescence analysis after electroporation

MDBK-B2 cells were first transfected by electroporation with the RNAs transcribed from the given plasmids. One day after EP, the cells were fixed with 4% PFA and permeabilized with 0.05% Triton X-100. The viral protein NS3 was detected with the primary antibody Code4 and α-mouse FITC.

As shown in **Fig 5.25**, for the positive control, nearly all the 798 electroporated cells were positive ensuring that the RNAs obtained from the *in-vitro* transcription. For the mutants, much lower number of positive cells were detected, indicating that the mutated RNA was functional as a replicon, though with decreased fitness. Two days after EP, the electroporated cells were splitted into a new 30 mm dishes to check whether the signal increased via virus spread. In addition, the supernatant from the electroporated cells was used for re-infection to determine whether newly generated viral particles were capable of infectious new cells. Both cells from the newly splitted plates and supernatant infected cells were checked again for the presence of NS3 by indirect immunofluorescence. Moreover, the samples were checked two days after split and SN infection for CPE.

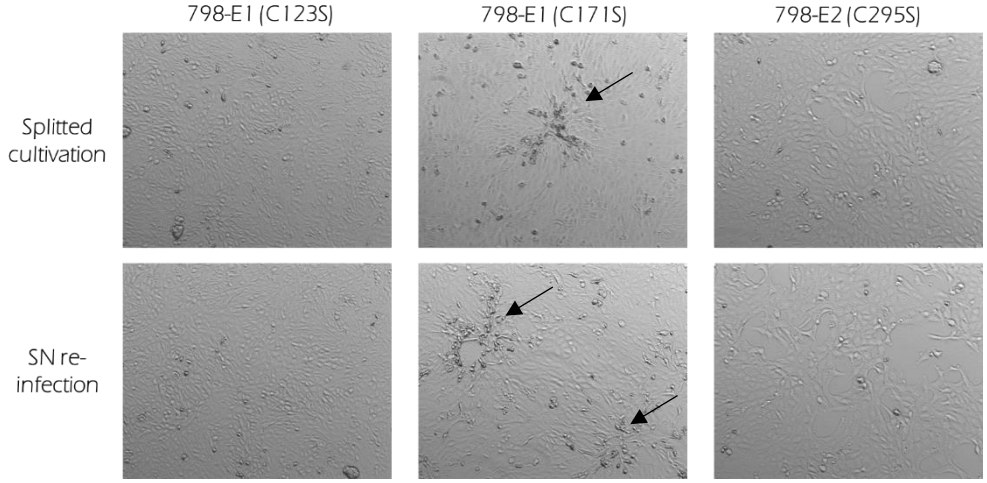


Figure 5.26 CPE observation after electroporation

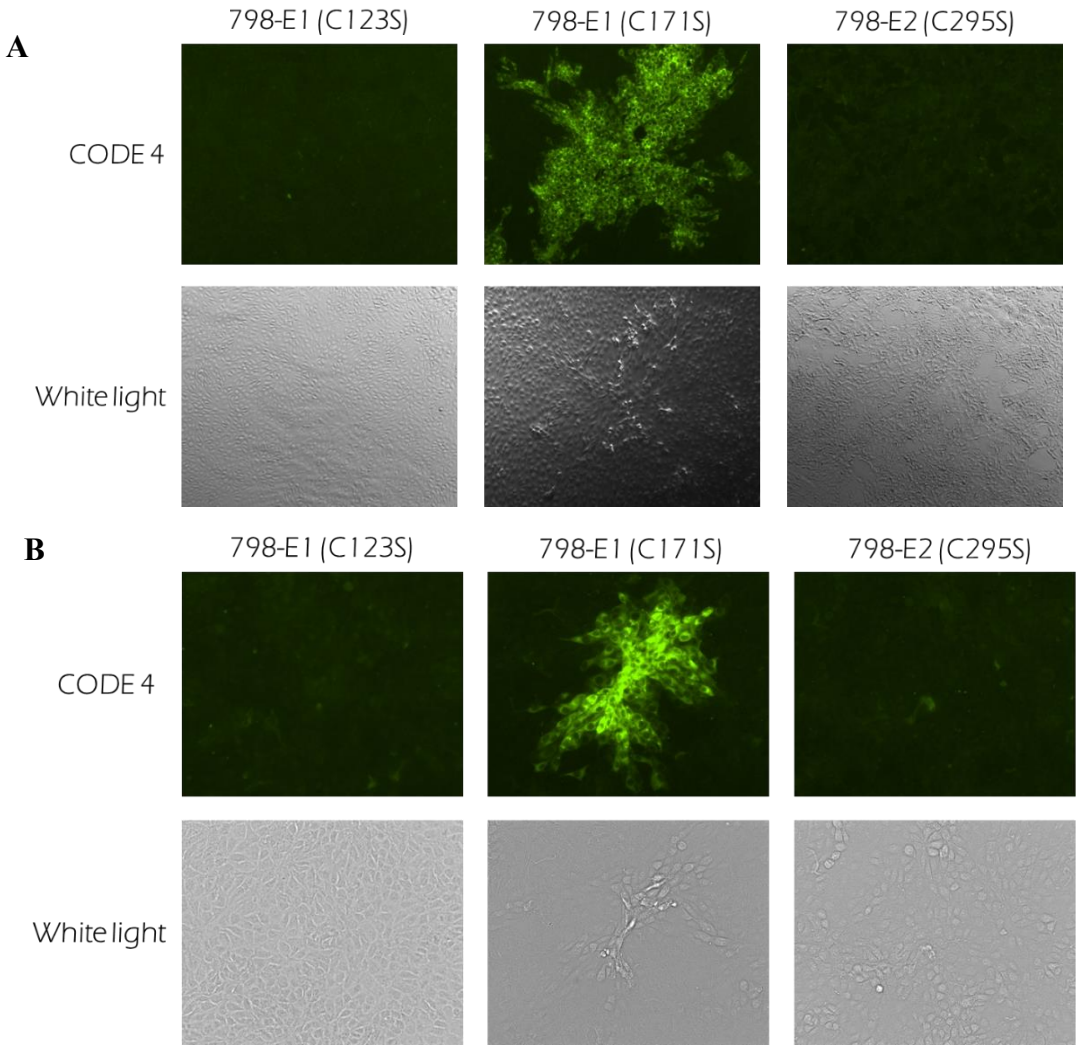


Figure 5.27 Indirect immunofluorescence analysis reinfection after electroporation

(A) Splitted cultivation after electroporation. The cells were fixed with 4% PFA and permeabilized with 0.05% Triton X-100. The viral protein NS3 was detected with the primary antibody Code4 and α-mouse FITC.

(B) Supernatant re-infection after electroporation. The treatment is same to **(A)**.

As shown in **Fig 5.26** and **Fig 5.27**, only 798-E1 C171S induced typical CPE in both splitted and SN infection dishes. After the RNA isolation and RT-PCR, the sequencing results confirmed that those recovered viruses were still carrying their corresponding substitution, no reversion arise. Interestingly, 798-E1 C123S and 798-E2 C295S were lost with the cell passage and could not be further enriched. These results demonstrated that BVDV CP7 798 infectious clones containing C123S in E1 or C295S in E2 can't generate infectious particles. These finding further clarified that those two Cysteine residues in E1 and E2 are critical for viral infectivity. It is worth noting that all the cells are positive in 798-E1 C171S splitted plate two days after splitted cultivation. However, if we use SN from 798-E1 C171S electroporated cells for the infection, two days later, only several viral plaques could be observed. This indicated that the infectivity of 798 containing the mutation C171S in E1 was also reduced. Nevertheless, the C171S mutation was stable in the recovered viruses for at least 3 generations.

5.6 The middle hydrophobic region (MHR) affects the secretion of E1

5.6.1 E1 is retained within the cell in the absence of the carboxyterminal membrane anchor

The hydrophobic region at the *C*-terminus of E1 was shown to be responsible for the ER retention. It is also supposed to be critical for membrane anchoring of E1, but the detailed knowledge which amino acids are in contact with the lipid bilayer remained unclear. It was suggested in a publication that the membrane anchor of pestiviral E1 harbors three helices including one perimembrane helix (pmH) and two transmembrane helices (tmH 1 and 2). This suggestion was based entirely on computational secondary structure prediction (sequences and corresponding helical wheel shown in **Fig 5.28**).

A

Name	Sequence	Hydrophobicity <H>	Hydrophobic moment <μH>
pmH	^{144}QVIKVVLRALKDLTRIWT^{161}	0.581	0.678
TMH1	^{164}TTTAFLVCLVKVVR^{177}	0.764	0.391
TMH2	^{179}QVLQGILWLMLI^{190}	1.222	0.104

B

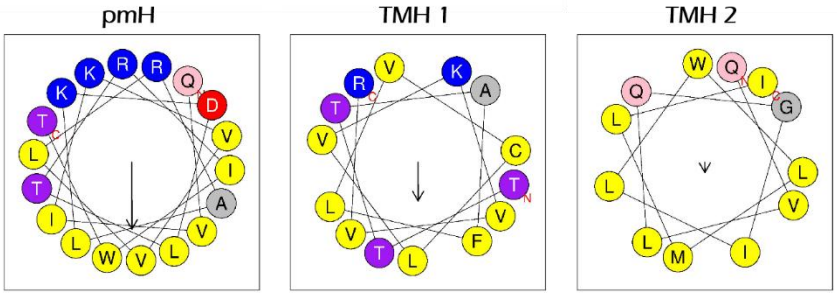


Figure 5.28 The sequence and helical wheel modeling of the putative membrane anchor of E1

(A) Hydrophobic region organization in E1. The conserved hydrophobic regions in E1 of pestiviruses. The data for hydrophobic moment <μH> and hydrophobicity <H> were all from the heliquest (http://heliquest.ipmc.cnrs.fr/).

(B) Helical wheel plots for the corresponding sequences shown in **(A)** which were generated by heliquest application. The arrow in the plots directs toward the hydrophobic face and the length of it corresponds to the hydrophobic moment <μH>.

The topology data presented above already demonstrated that it is highly unlikely that TmH 1 and TmH 2 represent two transmembrane domains, since the carboxylterminus is located on the cytoplasmic side. To determine experimentally which part of the E1 *C*-terminus served as membrane anchor, a series of *C*-terminally truncated E1 variants were generated (**Fig 5.29A**). The plasmids were expressed in BHK-21 cells by using the Vaccinia MVA T7 expression system and the newly synthesized proteins were labelled with ^{35}S amino acids. The supernatant (SN) and cell lysates (CL) of the transfected cells were harvested, from which proteins reacting with specific antibodies directed against the HA-tag were precipitated. The samples were then separated by SDS-PAGE and the labeled proteins were detected on imaging plates. All the samples were treated with PNGase to have concentrated bands for the following relative quantification.

As shown in **Fig 5.29B**, we could see no secretion at all for wild type E1. Additionally, the deletion of neither tmH 1 (E1 [166-178Δ]) nor tmH 2 (E1 [1-178]) resulted in the secretion of E1 indicating that neither of these two sequences alone is responsible for membrane binding of E1. Surprisingly, when the putative transmembrane region was completely removed (E1 [1-165/166]), only about 1% of E1 was secreted. Even when all the hydrophobic sequences located in the *C*-terminal region of E1 was deleted (E1 [1-143]), most of E1 is still found in the cell extract. Only approximately 5% could be detected in the supernatant. These data indicated that in the absence of the proposed membrane anchor, E1 somehow still was retained within the cell.

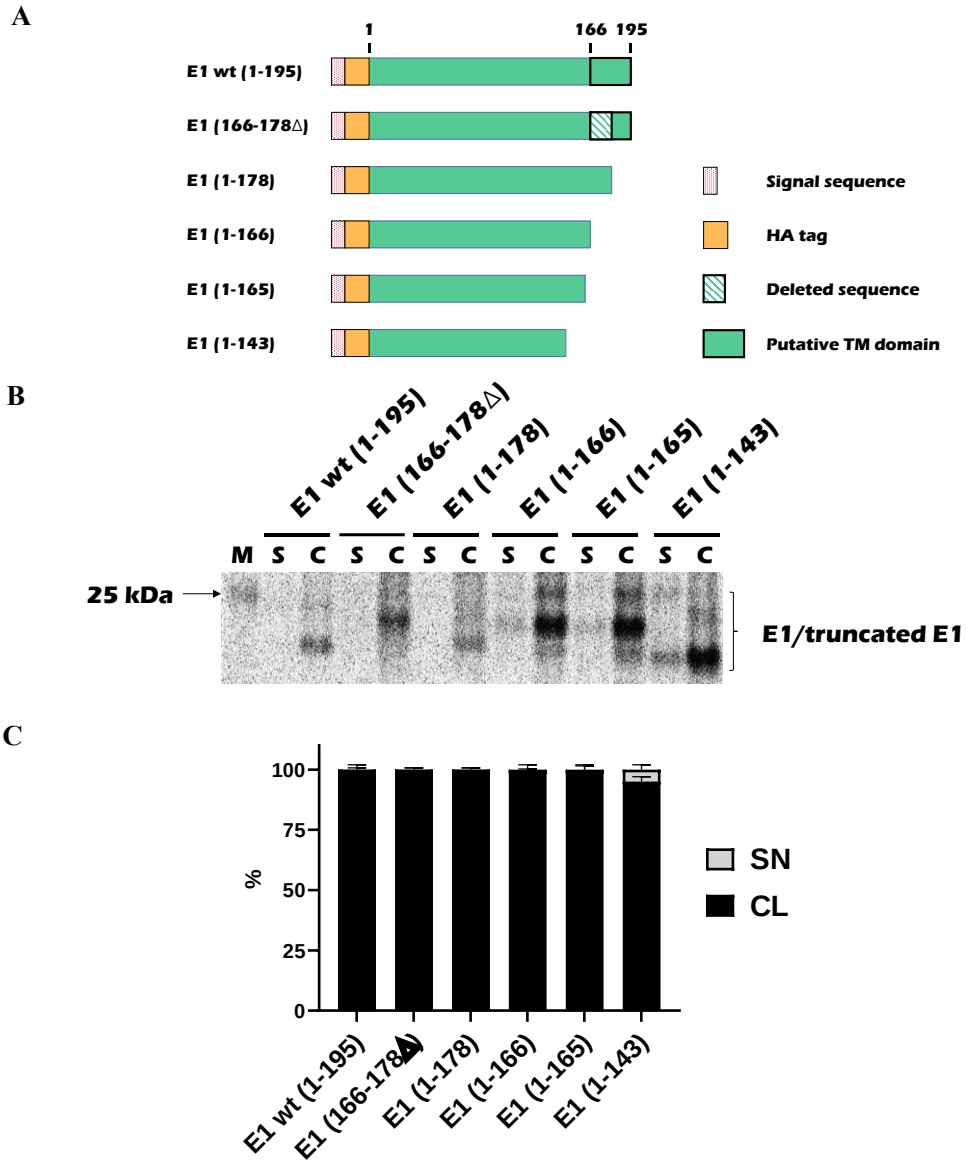


Figure 5.29 The secretion/retention analysis of E1/truncated E1 variants

(A) Schematic representation of wild type and truncated E1constructs used in this section.

(B) The immunoprecipitation results of different E1 constructs. The corresponding plasmids were expressed in BHK-21 cells by using Vaccinia virus MVA T7 expression system. Expressed proteins were labelled with ^{35}S radioactively. The supernatant (SN) and cell lysates (CL) of the transfected cells were produced, from which proteins reacting with a specific anti-serum directed against the HA tag were precipitated. The samples were pre-treated with PNGase, then separated using SDS-PAGE under reducing conditions and the labeled proteins were detected on imaging plates.

(C) Quantification of the secretion/retention rate for the E1/truncated E1 variants. The radioactivity of secretion (S) and cell proteins (C) was determined by Aida Image Analyzer 5.0. The relative percentage of each component was calculated with entire protein (S+C) added up to 100%. Results determined from 3 independent experiments are calculated as mean ± SD.

To determine the subcellular localization of those truncated E1 mutants, three representative E1 truncated constructs were further analysed by confocal microscopy. Briefly, the ER was fluorescently labeled by pDsRed-ER, which was transiently co-expressed with HA tagged E1. Similarly, the pDsRed-Golgi plasmid was used for Golgi-apparatus labeling. The results were analysed on a confocal fluorescence microscope.

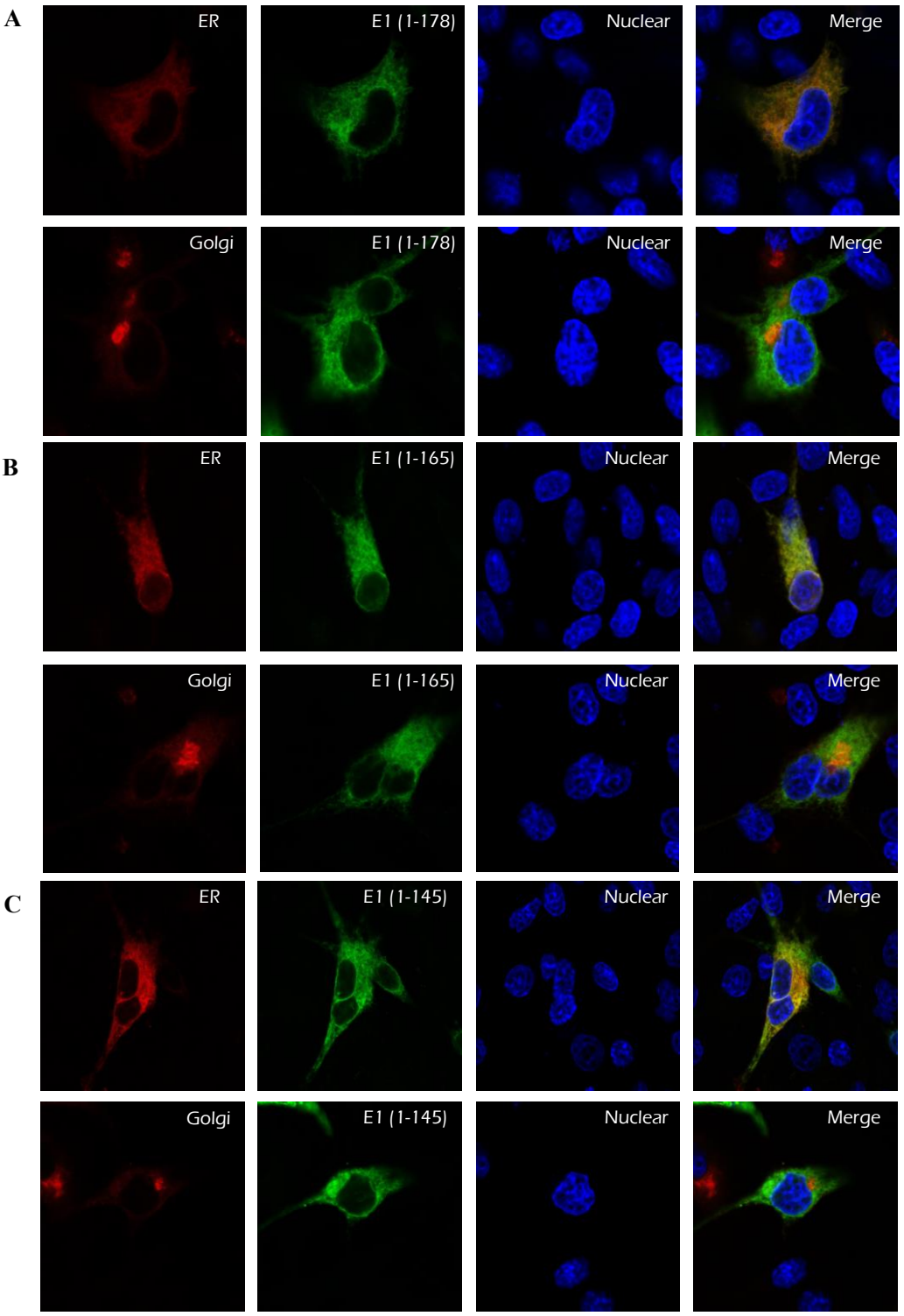


Figure 5.30 Subcellular localization of *C*-terminally truncated E1 variants

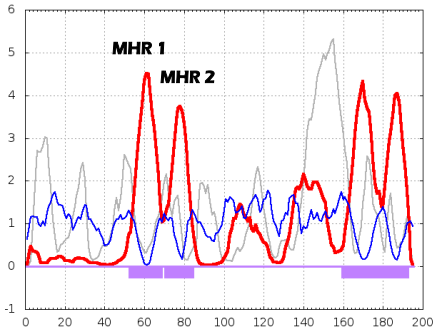
HA-tagged E1 with different length of *C*-terminal truncations (E1 [1-178], E1 [1-165] and E1 [1-145]) were co-expressed with pDsRed-ER/pDsRed-Golgi, respectively, for intracellular localization analysis. At 24h post-transfection, cells were fixed by 4% PFA permeabilized with 0.05% Triton X-100 and stained with specific antibodies against HA (in green). Compartments (ER or Golgi) are in red. Nuclei were stained with DAPI.

As shown in **Fig 5.30**, those *C*-terminally truncated E1 mutants present mainly ER localization. Even the longest deletion version (E1 [1-145]) was still predominantly found in the ER, while showing no localization in the Golgi compartment. This result also indicated that the ectodomain of E1, for some reason, was still retained in the ER preventing strong secretion E1.

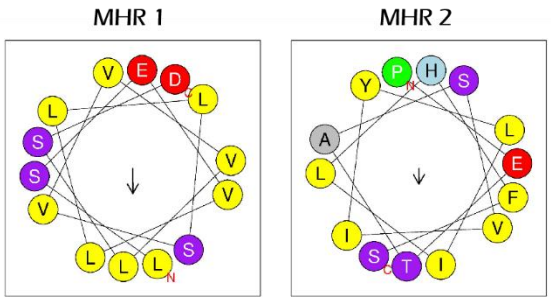
5.6.2 The middle hydrophobic region (MHR) affects the secretion of E1

Since there is a strong indication that there should be segment(s) in the ectodomain of E1 preventing strong secretion E1, we wanted to further investigate the hydrophobic region located in the middle part of E1 (h1, shown in **Fig 5.2**). By using bioinformatic prediction web application (TMHMM Server v. 2.0, http://www.cbs.dtu.dk/services/TMHMM/), we made a prediction of transmembrane helices in pestiviral glycoprotein E1. As shown in **Fig 5.31A**, the prediction suggested that the MHR 1 (54-67) and MHR 2 (70-83) (sequences and corresponding helical wheel shown in **Fig 5.31B and C**) probably form transmembrane helices, and thus could be the additional membrane binding region for E1. These structures could be involved in preventing secretion of the E1 ectodomain. To determine whether the middle hydrophobic region (MHR) of E1 affect the secretion of E1, a series of *C*-terminally truncated E1 variants in addition to MHR2 (70-83) deletion were generated in this section (shown in Fig **5.31D**). The expression plasmids were expressed in BHK-21 cells by using Vaccinia MVA T7 expression system and the newly synthesized proteins were labelled with ^{35}S amino acids. The supernatant (SN) and cell lysates (CL) of the transfected cells were harvested, and proteins reacting with specific antibodies directed against HA tag were precipitated. The samples were then separated by SDS-PAGE and the labeled proteins were detected on imaging plates. All the samples were treated with PNGase to have concentrated bands for the following quantification.

A



B



C

Name	Sequence	Hydrophobicity <H>	Hydrophobic moment <μH>
MHR1	^{53}LSEVLLLSVVVLSD^{67}	0.846	0.254
MHR2	^{70}PETASVIYLILHFS^{83}	0.834	0.159

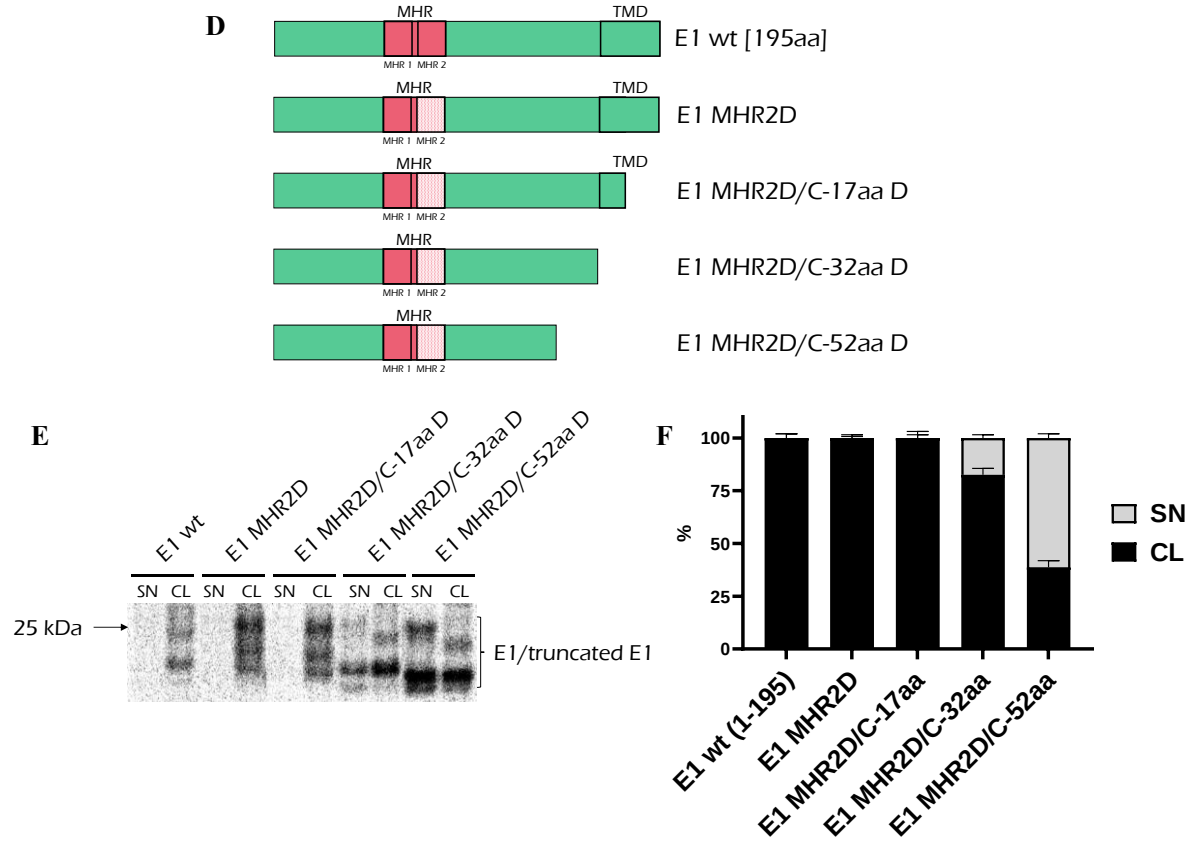


Figure 5.31 Secretion/retention analysis for the *C*-terminally truncated E1 MHR 2 deletion variants

(A) The prediction result of transmembrane helices in pestiviral E1 glycoprotein. Red line: Transmembrane helix preference. (THM index); Blue line: Beta preference. (BET index); Gray line: Modified hydrophobic moment index. (INDA index); Violet boxes (below abscisa): Predicted transmembrane helix position. (DIG index)

(B) Helical wheel plots for the corresponding sequences shown in **(C)** which were generated by heliquest application. The arrow in the plots directs toward the hydrophobic face and the length of it corresponds to the hydrophobic moment <μH>.

(C) Hydrophobic region organization in the ectodomain of E1. The conserved hydrophobic regions in E1 of pestiviruses. The data for hydrophobic moment <μH> and hydrophobicity <H> were all from the heliquest (http://heliquest.ipmc.cnrs.fr/).

(D) Schematic representation of wild type and truncated E1constructs used in this section. Middle hydrophobic region (MHR) presented in red box; Deletion sequence presented in light pink box.

(E) The immunoprecipitation results of different E1 constructs. The corresponding plasmids were expressed in BHK-21 cells by using Vaccinia virus MVA T7 expression system, and ^{35}S radioactively labelled in situ. The supernatant (SN) and cell lysates (CL) of the transfected cells were produced, from which proteins reacting with specific antibody direct against HA tag were precipitated. The samples were then separated using SDS-PAGE under reducing condition and the labeled proteins were detected on imaging plates.

(F) Quantification of the secretion/retention rate for the E1/truncated E1 variants. The radioactivity of secretion (S) and cell lysates (C) bands was determined by AIDA Image Analyzer 5.0. The relative percentage of each component was calculated with entire protein (S+C) added up to 100%. Results determined from 3 times independent experiments are calculated as mean ± SD.

As shown in **Fig 5.31E and F**, the deletion of MHR 2 alone had no effect on the retention of E1. Similarly, the construct which combines deletion of MHR 2 and deletion of the *C*-terminal 17aa presented 100% intracellular retention. Interestingly, the MHR 2 deletion construct with 32aa truncated from the *C*-terminus demonstrated about 20% secretion. Compared to the same truncated version which still contains MHR 2 (~ 1% secretion, shown in Fig 5.30E and F), the additional MHR 2 deletion results in significantly higher secretion (1% versus 20% respectively). Moreover, when the *C*-terminal truncated sequence was increased to 52aa, the MHR 2 deleted E1 showed about 55% secretion. The results showed that the hydrophobic sequence in the middle part of E1 partially prevents the secretion of the *C*-terminally truncated E1 proteins, indicating this region could be membrane associated. The upper additional bands of both SN and CL on the imaging plates are incompletely deglycosylated E1 or E1 variants.

To further investigate the effect of this part of E1, a series of entire MHR deletion constructs were generated. As shown in **Fig 5.32 B and C**, only removed of MHR alone can't lead to the release of E1. Moreover, this is also true for the entire MHR deletion combined with the 17aa truncation from the *C*-terminus. In contrast, both the *C*-terminal 32aa and 52aa deletions combined with whole MHR deletion constructs presented a clear secretion. Especially the latter, showed that nearly 88% of the proteins lost the intracellular retention. Since the deglycosylation via PNGase treatment sometime is incomplete, some upper additional bands of both SN and CL on the imaging plates are visible.

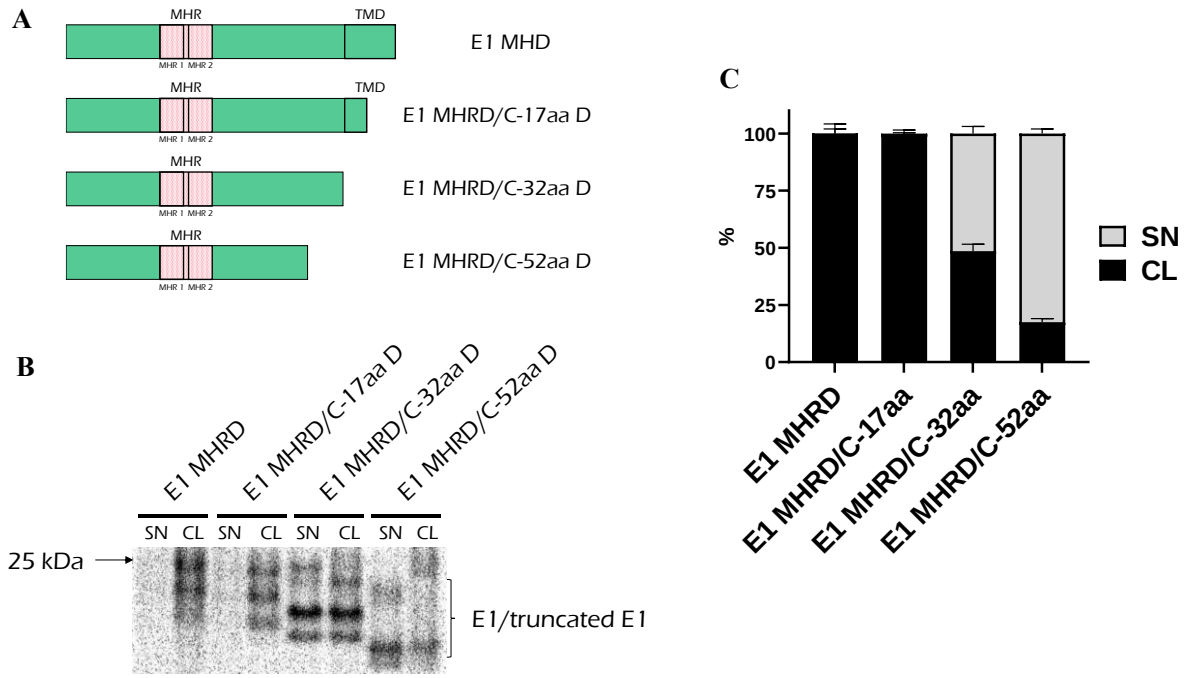


Figure 5.32 Secretion/retention analysis for the *C*-terminal truncated E1 entire MHR deletion variants

(A) Schematic representation of E1/truncated E1constructs used in this section. Middle hydrophobic region (MHR) presented in red box; Deletion sequence presented in light pink box.

(B) The immunoprecipitation results of different E1constructs. The corresponding plasmids were expressed in BHK-21 cells by using Vaccinia MVA T7 expression system, and ^{35}S radioactively labeled. The supernatant (SN) and cell lysates (CL) of the transfected cells were produced, from which proteins reacting with specific antibody direct against HA tag were precipitated. The samples were then separated using SDS-PAGE under reducing condition and the labeled proteins were detected on imaging plates.

(C) Quantification of the secretion/retention rate for the E1/truncated E1 variants. The radioactivity of secretion (S) and cell lysates (C) bands was determined by phosphorimager analysis software AIDA Image Analyzer 5.0. The relative percentage of each component was calculated with entire protein (S+C) added up to 100%. Results determined from 3 times independent experiments are calculated as mean ± SD.

Taken together, in this section, we provide a new view for the role of the hydrophobic region in the ectodomain of E1 in the morphogenesis of pestiviral envelope protein. There is a strong indication that pestiviral E1 is not readily secreted, even without the *C*-terminal membrane anchor, it still can be retained within cell, probably membrane associated. The data demonstrated that the MHR sequence is beneficial for the intracellular retention. Accordingly, we concluded that the hydrophobic sequence (53-83) in the middle part of E1 affect the secretion and intracellular retention of E1 probably by membrane binding.

Chapter 6: Discussions

The present study mainly focused on molecular characterization of the glycoprotein E1 of pestiviruses with a special emphasis on investigations with regard to its intracellular localization, retention signal and membrane topology in order to gain an initial insight into the molecular features that play a role in its engagement in virus assembly, budding and oligomerization of envelope proteins. In addition, we also tested E1 mutants with regard to their influence on infectious virus recovery.

6.1. The organization of domains in pestiviral E1

The E1 envelope protein of pestiviruses contains approximately 195 aa, which is only half the size of E2 (374 aa, BVDV CP7 as an example). There are nearly 40% hydrophobic residues in E1 indicating that E1 harbors several hydrophobic regions which are supposed to be functional in the viral life cycle. E2 has been already shown to be a type I transmembrane protein with a large ectodomian at the *N*-terminus residing in the ER lumen and a hydrophobic *C*-terminus achoring the protein in the ER membrane (Radtke and Tews, 2017). In this study, we showed that E1 adopts the same basic membrane topology. The alignment of multiple E1 sequences throughout all genotypes of pestiviruses (from A to K) combined with several bioinformatic predictions revealed the presence of some hydrophobic regions including the hydrophobic region in the middle (MHD, 54-83), a proposed perimembrane helix (pmH, 144-161) and the putative transmembrane region (TMD, 166-195) (as shown in **Fig 6.1**). According to the secondary structure prediction (**5.1.3.1**), MHD consists of two α-helices which are MHR 1 (53-67) and MHR 2 (70-83). Based on the conclusion from the published paper (Wang et al., 2014) and the prediction of JPred4, there are three α-helices located at the *C*-terminus of E1 which corresponding to the h3 and h4 region in the Kyte & Doolittle hydropathy plot in **Fig 5.2**. The perimembrane helix (pmH), which locates upstream of the putative transmembrane region, forms an ideal amphipathic helix (**Fig 5.28B**). The hydrophobic face of this helix is supposed to be the 'V, I, A, V, L, V, W, L, I' peptide. The putative transmembrane region (TMD, 166-195) actually is divided into two parts corresponding to the two peaks in h4 region in Fig 5.2, the first part is considered to be the membrane anchor of E1 whereas the second is thought to serve as a signal sequence for the translocation of the following E2. In our study, we showed that TMD is also responsible for the retention of E1.

Pestiviral E1 contains six cysteine residues, the first four located in the *N*-terminal or middle part most likely form intramolecular disulfide bonds, since they do not have any effect on both E1 oligomerization and E1-E2 heterodimerization when exchanged. The last two could form intermolecular disulfide bonds (shown in **Fig 5.20 and 5.21**). In all genotypes, pestiviral E1 possesses two conserved potentional *N*-linked glycosylation sites. In the *N*-terminal part of E1, one extra glycosylation site (N6) is CSFV/BDV-specific (as shown in **Fig 6.1**). The molecluar weight of mature E1 in those two species indicates that this site is also used for glycosylation (Risatti et al., 2007b). The glycosylation sites of E1 distribute at the *N*-terminus or middle part. All of them are used, indicating that this part of the protein is most likely exposed on the surface of E1 after correct folding. In our study, we showed that different glycan types are present at N19 and N100, which might be due to the interference from two neighbour cysteine residues (C20 and C24) to the N19 glycosylation site. It is known that the glycosylation of proteins is

closely related to the correct folding and biological activities. We also found that E1 variants lacking glycosylation site(s) still can form oligomers as well as E1-E2 heterodimers when co-expressed with E2 (data not shown). For HCV, mutations at N196 or N305 in E1 have strong effects on the E1-E2 heterodimer formation (Meunier et al., 1999). In CSFV, individual *N*-linked glycosylation sites in E1 are not essential for viral particle formation or virus infectivity. However, in the context of two or more putative glycosylation site modifications, residue N100 is critical for virus viability (Fernandez-Sainz et al., 2009).

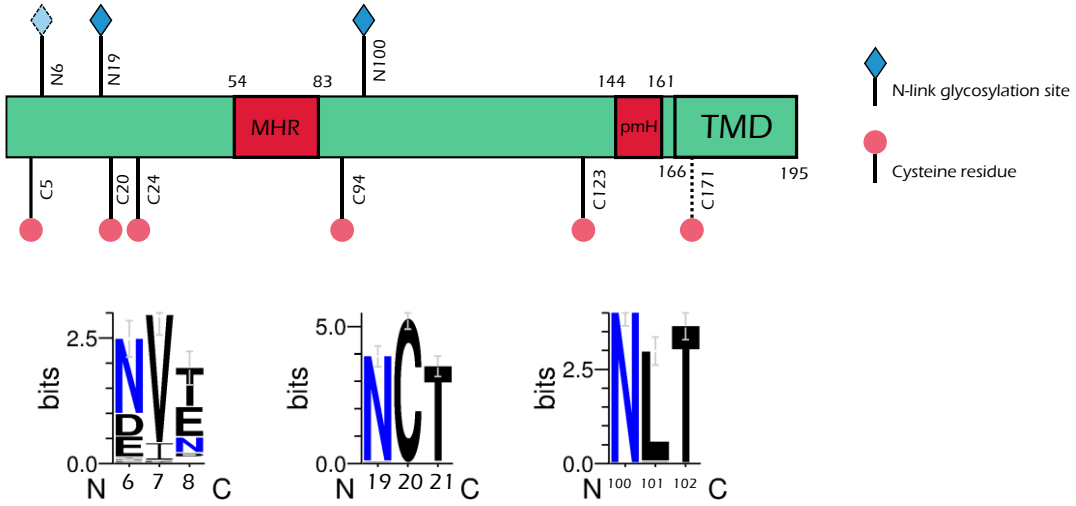


Figure 6.1 Schematic representation of pestiviral E1 envelope protein

Two *N*-glycosylation sites (N19 and N100) are conserved in all genotypes and one CSFV/BDV-specific site (N6) is labeled as light blue rhombus, and cysteine residues are marked with a pink ball (conserved cysteine residues with solid and non-conserved cysteine residues with dash line, respectively). MHR (hydrophobic region in the middle): 54-83, pmH (perimembrane helices): 144-161, TMD (transmembrane domain):166-195. Main hydrophobic regions are marked with red boxes. Sequence logos were generated based on 68 pestivirus E1 sequence throughout 11 ITCV-classified species (from pestivirus A to K) using WebLogo.

6.2 The retention of pestiviral E1 glycoprotein

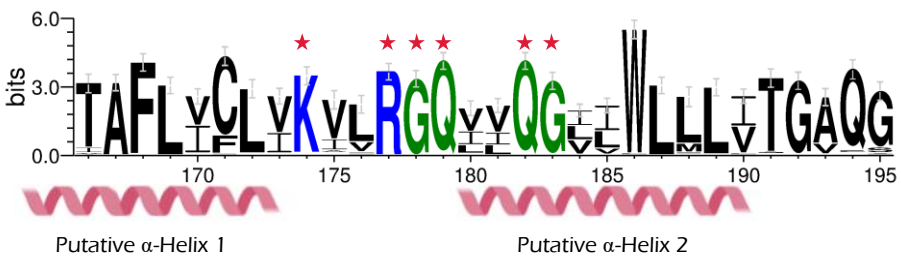
Viruses have to take advantage of the protein biosynthesis machinery of the host cells for their own protein synthesis and processing via the conventional protein modification and trafficking pathway. It means that proteins translocated into the ER after the co- or post-translational modification, are generally delivered to their destination via the secretory route or remain within defined intracellular compartment(s) when they contain a respective localization signal. Pestiviral E1 shows no secretion as well as no expression on the cell surface indicating that E1 has to accumulate at specific intracellular site(s) before viral budding. In our study, we showed that E1 is mainly concentrated in the ER regardless whether other viral proteins are present or not. For the other two pestiviral envelope proteins, it was shown that both E^{rns} and E2 also localize in the ER (Burrack et al., 2012; Grummer et al., 2001; Köhl et al., 2004; Radtke and Tews, 2017). In addition, ultrastructural study of pestivirus Giraffe-1 using electron microscopy also showed that ER is the initial cellular organelle for pestivirus assembly and viral budding (Schmeiser et al., 2014). These data strongly indicate that pestiviruses most likely bud at the ER site, so that the envelope is derived from the host ER membrane. Other viruses bud through the plasma membrane whereas several enveloped viruses also bud at intracellular organelle membranes, like the ER (e.g. rotaviruses and hepatitis C virus), the ER-Golgi intermediate

compartment (ERGIC) (e.g. coronaviruses and poxviruses) and the Golgi apparatus (e.g. bunyavirus) (Cocquerel et al., 1998; Griffiths G, 1992).

It is confirmed that E^{rns} and E2 contain retention signals of their own (Burrack et al., 2012; Radtke and Tews, 2017). Additionally, E2 can form covalently linked heterodimers with E1. It was therefore supposed that the ER retention signal of E2 could be sufficient to retain E1-E2 complexes within the ER (Cocquerel et al., 1998; Radtke and Tews, 2017). In our study, we showed that there is also an ER retention signal present in the transmembrane region of the pestiviral E1 glycoprotein. So this raises the question - why the E1-E2 envelope protein complexes of pestiviruses have two signals for ER retention? One reason could be that it takes time for E1 and E2 to form a complex. Thus, both proteins should have a signal ensuring accumulation at the same location to allow heterodimer formation.

The retention signals of both E1 and E2 are located in the TM domains. The TMDs of both envelope proteins are multifunctional. In addition to ER retention, they also ensure the membrane anchoring, serve as signal sequences for the protein downstream, and probably are involved in E1-E2 interactions. The TM regions of the envelope proteins of the members of *Flaviviridae* usually are composed of two hydrophobic sections separated by a short segment containing at least one fully conserved positively charged residue (as shown in **Fig 6.2 A and B**) (Cocquerel et al., 1999; Cocquerel et al., 1998; Cocquerel et al., 2000a).

A



B

BVDV CP7 (AAC55984):	QYWFDLEITDHHRDYFAESLLVIVVALLGGRYVLWLLVTYMILSEQMASG
HCV JFH 1 (AB047639):	VDVQYMYGLSPAITKYVVRWEWVVLLFLLLADARVCACLWMLILLGQAEA
TBEV (U27495):	GGVGFLPKLLLGVALAWLGLNMRNPTMSMSFLLAGGLVLAMTLGVGA
YFV (AY640589):	GGLNWITKVIMGAVLIWVGINTRNMTMSMSMILVGVIMMFLSLGVGA
DENV2 (M20558):	SGVSWTMKILIGVIITWIGMNSRSTSLSVSLVLVGVVTLYLGAMVQA
WNV (DQ211652):	GGMSWITQGLLGALLLWMGINARDRSIALTFLAVGGVLLFLSVNVHA
JEV (EF571853):	GGMSWITQGLMGALLLWMGVNARDRSIALAFLATGGVLVFLATNVHA

Figure 6.2 Comparison of putative TMD sequences of envelope proteins in the family *Flaviviridae*

(A) The sequence logo of the TM domain of pestiviral E1 was generated by WebLogo 3 web application (http://weblogo.threeplusone.com/create.cgi) demonstrates the alignment results of 68 pestivirus E1 sequences throughout all pestiviral species in one letter code. The size of the letters in the sequence logo corresponds to the degree of conservation over the 68 sequences (from pestivirus A to K). The fully conserved residues are highlighted with red stars. Two stretches of hydrophobic regions are shown below as pink helix. Polar charged residues in blue, polar non-charged residues in green.

(B) *C*-terminal transmembrane sequences of the envelope proteins (E2 for BVDV, E for others) of different members of the family *Flaviviridae*. The sequences that were predicted to form hydrophobic

α-helices are shown in red colour. Fully conserved positively charged residues are highlight with green stars. The corresponding GenBank accession numbers are also shown besides the virus name.

Sequence analysis of putative TMD sequences of envelope proteins in the family *Flaviviridae* revealed a quite similar organization for the representative members in this viral family (as shown in **Fig 6.2B**). The presence of at least one conserved positively charged residue was observed in a short linking segment connecting the two hydrophobic stretches. Publications showed that those conserved positively charged residues play an essential role in retention and heterodimerization of glycoproteins as well as assembly of the viral particle (Cocquerel et al., 2000a). In our study, we found that two fully conserved positively charged residues (K174 and R177) located in the middle segment have nearly no effect on E1 retention and on E1-E2 heterodimer formation (**Fig 5.14** and data not shown). This finding indicates that pestiviral E1 may adopt another different retention mechanism from other envelope proteins of family *Flaviviridae.* As shown in Fig 6.2B, compared to BVDV E2 or other envelope proteins of the members of family *Flaviviridae*, four fully conserved, but non-charged polar residues are found in the middle part of the E1 TM domain in addition to K174 and R177. We consider that since those polar residues are present in the connecting region, their concomitant mutation might also be required to alter the retention status of E1. In our study, surprisingly, only mutation of all the six fully conserved polar residues to Alanines could result in the presence of E1 on the cell surface (**Fig 5.14 and 5.15**). However, when those six fully conserved polar residues were replaced by Leucines, the retention signal of E1 was not affected (data not shown). It is possible that this newly made transmembrane region is too artificial to allow correct folding of E1. Due to the 'check mechanism' of the ER this misfolded E1 can still not leave the ER. According to the so-called "lipid-based" rule (Bretscher and Munro, 1993), membrane thickness could also play a role in the ER retention mediated by the TMD of E1. The lengths of the transmembrane regions of membrane proteins are strongly associated with their intracellular location in different organelles along the secretory pathways (Sharpe et al., 2010; Singh and Mittal, 2016). The artificial transmembrane anchor which is composed of 'pure leucine residues' might form a longer straight α-helix so that E1 changes its intracellular location, but still is located within the cell.

In our study, we tried a series of single/double/multiple mutations or deletions in the transmembrane region of E1. All the data indicated that the retention signal of E1, unlike that of E2, is super complex and stable. We consider that those six conserved polar residues in the TM region of E1 could serve as a functional group in the retention behaviour, that makes the retention of E1 resistant against changes. Moreover, introducing a diacidic 'DXE' export signal from the VSV-G cytoplasmic tail to the *C*-terminus of the E1 TM domain did not result in the presence of E1 on the cell surface. This finding suggested that the ER retention signal present in the E1 TM domain is dominant over the diacidic export signal, and thus proved the stability and strength of the retention signal of E1 (data not shown).

Interestingly, we found that E1 can overrule the retention signal of E2 when covalently linked to E2. In other words, the subcellular localization of E2 largely depends on that of E1 when the covalent linkage exists. In contrast, E2 is not able to overrule E1 in the covalently linked E1-E2 heterodimer. It could be supposed that E1 is the superior for the retention of E2. Only a single mutation at R355 or Q370 can suppress the retention of E2 (Kohl et al., 2004; Radtke and Tews, 2017), making E2 retention vulnerable to mutations. One could consider that pestiviruses use the complex and stable retention signal of E1 as a 'error correction' mechanism, ensuring that E2 still can be processed in the ER as usual when some unpredictable mutations

occur. A similar phenomenon was reported between pre-membrane (prME) and envelope (E) of YFV (yellow fever virus) (Ciczora et al., 2010).

Both glycoproteins E1 and E2 of pestiviruses were identified to be type I transmembrane proteins retained in the ER. Cellular type I transmembrane proteins located in the ER often contain a dilysine motif (e.g. -KKXX, and -KXKXX) in the cytosolic *C*-terminus (Munro and Pelham, 1987), while luminal ER proteins often have a KDEL sequence (Shin et al., 1991). The ER retention signals, in principle, can be divided into two types. One can make proteins resident in the ER at steady condition without cycling between the ER and the Golgi apparatus. HCV E1 and E2 contain this kind of retention signal (Cocquerel et al., 1999; Cocquerel et al., 1998; Cocquerel et al., 2000b; Duvet, 1998). The other is a so-called retrieval signal, which can return the target proteins from the Golgi complex to the ER via COPI vesicles. The retention signals in both E1 and E2 of pestiviruses most likely belong to the former. 'KDEL' as a retrieval signal returns the cargo from the Golgi to the ER via a well-characterized specific receptor (Capitani and Sallese, 2009; Jia et al., 2020; Jin et al., 2017; Yamamoto et al., 2001). As many viral envelope proteins, the pestiviral glycoproteins do not contain one of these known retention signals. Thus, the retention mechanism of pestiviral envelope proteins is still not clear at the molecular level.

It is worth noting that different cellular retention mechanisms in flaviviruses were reported. Yellow fever virus (YFV), another member of the genus *Flavivirus*, also contains retention signals in both pre-membrane (prM) and envelope (E) protein. The mechanism of ER retention of YFV mainly relies on the length of the transmembrane stretches (Ciczora et al., 2010). However, for HCV, the retention mechanisms primarily depends on the polar charged residues in the middle part of TM domain. The same is true for pestivirus E2. In our study, we showed that 6 fully conserved polar residues affect the ER retention of E1, indicating this mechanism could be different from the above mentioned types. It is striking that transmembrane domains of members of the family *Flavivirus* as retention signals for their envelope proteins take advantages of at least three different types of ER retention mechanisms in their morphogenesis, the exact molecular mechanism for the retention behaviour of pestiviral envelope proteins still awaits further investigation.

6.3 The membrane topology of the pestiviral E1 glycoprotein

Pestiviral envelope proteins are synthesized as a polyprotein. The signal peptidase is responsible for the cleavage at the E^{rns}/E1 and E1/E2 sites. E^{rns} uses an amphipathic helix as its membrane anchor which is arranged in plane to the membrane surface making the E^{rns}/E1 cleavage site really unusual. It is known that the processing of internal signal sequences is essential for the membrane topology of downstream following polypeptides. Therefore, the membrane topology before/after the signal peptidase cleavage was analysed in this study. Some paper suggested that one peptide composed of 16 Leucines is sufficient to form an α-helix to go through the membrane. However, the TM region of integral membrane proteins normally contain stretches of 20 to 25 hydrophobic residues (Ulmschneider and Sansom, 2001; von Heijne, 1995). As shown in Fig 6.2, there are two hydrophobic stretches in the TM domain of pestiviral E1. Both are about 10aa long, so each hydrophobic region is too short to form a single transmembrane spanning α-helix. Since the length of each hydrophobic region at the *C*-terminus of E1 does not fit to single membrane-spanning topology, it indicated that they should adopt an extended structure to go through the membrane. In our study, we found that pestiviral E1, like E2, adopts a typical type I transmembrane topology after the signal peptidase cleavage (schematic models were shown in **Fig 5.17 and 5.18**). However, when the E1/E2 cleavage site

was blocked, the TM domain of E1 forms a hairpin-like structure with the carboxyterminus exposed in the ER lumen (**Fig 6.3 A**). A dynamic change in the orientation of the *C*-terminus of E1 was also shown in HCV E1 and E2 (Cocquerel et al., 2002). A publication on HCV suggested that the extended 'hairpin-like' structure of the TM domain is thermodynamically not stable, since those exposed polar charged residues are not favourable in the hydrophilic membrane environment. (Cocquerel et al., 1998). After the reorientation of the *C*-terminus of the TM of E1 or E2 from the luminal side to the cytosol side, as a consequence, the *C*-terminus of those two envelope proteins are accessible from the cytosolic side, and thus form cytosolic tails. These regions could be functional in the interaction between the viral envelope proteins and the capsid/genome RNA complex to allow egress into the ER lumen by interacting with the cytosol tail with envelope proteins (Mettenleiter et al., 2013)

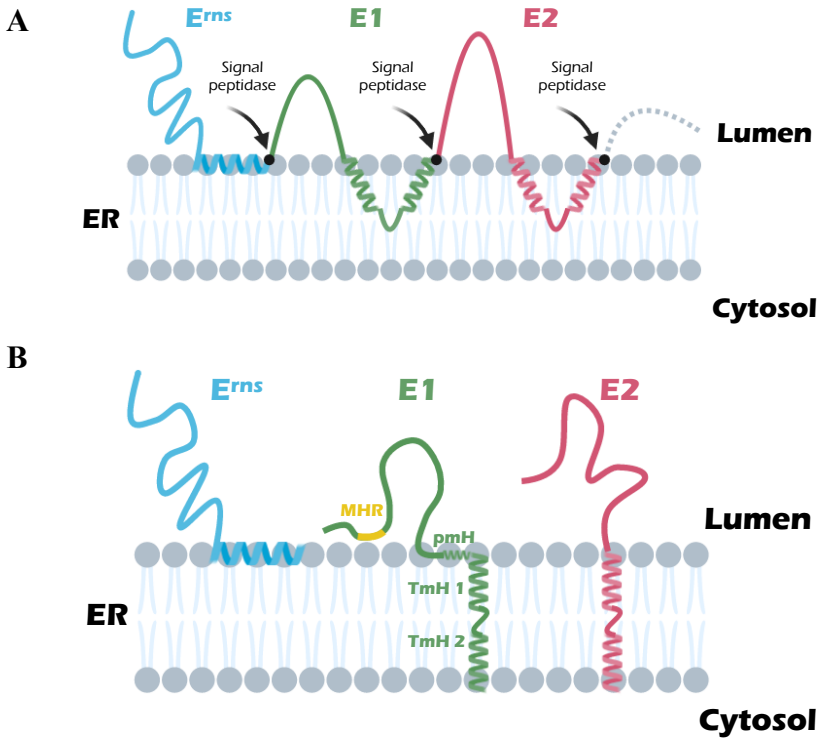


Figure 6.3 Schematic model of the reorientation behaviour of the TM domains of pestiviral envelope proteins during the early steps of their biogenesis

(A) The signal peptidase (shown as black arrow) is responsible for the cleavage at the E^{rns}/E1 and E1/E2 sites (shown as black dots). Before signal sequence cleavage between E1 and E2, the signal sequence present in the *C*-terminal half of the TM domain of E1 is oriented towards the ER lumen. As shown for the TM domain of E1, also the TM domain of E2 transiently adopts a hairpin structure to allow the translocation of the following proteins p7 and amino-terminus of NS2.

(B) After signal peptidase cleavage, the signal sequences present in the *C*-terminal half of the TM domain of E1 or E2 are reoriented towards the cytosol, establishing a transmembrane configuration spanning the lipid bilayer. The other hydrophobic regions in E1 were also shown.

As shown in **Fig 6.3 A and B**, the second half of the TM domain of E1, which is supposed to serve as a signal sequence for the downstream protein, remains linked to the *C*-terminus of E1 after the signal peptidase cleavage being part of a TM domain. After cleavage of the E1/E2 bond the *C*-terminus of E1 reorientates from the ER luminal side to the cytosol. The signal sequences at the second half of the *C*-terminus of E1 is supposed to contribute to several new

functions. After the reorientation, those two extended hydrophobic stretches can serve as a stable membrane anchor as well as retention signal for the cellular localization of E1. E2 was shown to have a similar reorientation behaviour at its *C*-terminus. In HCV, reorientation of the E1 and E2 *C*-terminus was also shown (Cocquerel et al., 2002). It was suggested that the reorientated transmembrane regions of E1 and E2 are also essential for heterodimerization. According to the data we have, for pestiviruses, it seems that the transmembrane region of E1 is also engaged in the formation of E1-E2 heterodimer (data shown in **Fig 5.23**).

In our studies, we observed an interesting phenomenon: if the cleavage at the E1/E2 site was hampered in the E^{rns}-E1-E2 precursor, the processing of the E^{rns}/E1 site was also inhibited. The results indicate that there is a processing hierarchy in E^{rns}-E1-E2 so that the E^{rns}/E1 site can only be processed after the E1/E2 site (Mu et al, under revision). Earlier studies already revealed that the E^{rns}-E1-E2 polyprotein is firstly cleaved into two parts: E^{rns}-E1 precursor and E2. Afterwards, the former would be further processed into E^{rns} and E1. It is known that signal peptidase (SPase) is responesible for the cleavage of the C/E^{rns}, E^{rns}/E1, E1/E2 and E2/p7 sites (Elbers et al., 1996; Rumenapf et al., 1993). SP cleavage is not dependent on a consensus sequence. Rather, substrate recogination relies on common structural characteristics. The signal peptide is about 20-30 amino acids long (but might be up to 80), with a typical tripartite structure: an n-region contains several basic amino acids, a hydrophobic h-region in the middle, and a slightly polar and rather unstructed c-region (von Heijne, 1990). Since E^{rns} does not contain an α-helical transmembrane domain but uses an unique amphipathic helix as the *C*-terminal membrane anchor, which is in plane to the membrane, the processing site of the E^{rns}/E1 site is quite unusual. An earlier study suggested that this special membrane anchor of E^{rns} plays an important role in the delicate balance between secretion and membrane association of E^{rns} (Tews and Meyers, 2007). The TM domain of E1 is supposed to be important for the cleavage at the E^{rns}/E1 most likely because this region represents the membrane anchor for the E^{rns}-E1 precursor. Binding of the precursor to the membrane and ensuring ER retention via the E1 signal gives time for cleavage to occur. Moreover, we also got strong indication that the full-length E1 sequence downstream of the E^{rns}/E1 cleavage site is important for efficient cleavage of E^{rns}/E1. All in all, a series of data shows that E1 affects the processing at the E^{rns}/E1 site by hooking the E^{rns}/E1 precursor to the membrane to give time for establishing the membrane contact of the amphipathic helix of the E^{rns}. Afterwards, E1 has to be folded properly to generate a cleavable structure at E^{rns}/E1 site (Mu et al, under revision).

In our study, we showed that before the cleavage at the E1/E2 site occurs, the *C*-terminus of E1 formed a hairpin like structure. This could prevent folding of E1 and impair the membrane anchoring of E^{rns}. This could be the reason why E^{rns}-E1 precursors are always detectable in the transfected or infected cells. However, the connection of the membrane topology change and the processing hierarchy in pestiviral polyproteins is still unclear and awaits further detailed chatacterization.

6.4 The oligomerization of E1 and the heterodimerization of E1-E2

In this study, we observed trimeric E1 under non-reducing condition. This is the first time to show the E1 glycoprotein of pestiviruses (at least of BVDV CP7) oligomerizes when overexpressed. Since the other two envelope proteins of pestiviruses can form covalently linked homodimers, there is a long-standing question whether pestiviral E1 does form oligomers, too. E1 contains six cysteine residues in most pestiviral species. Which cysteine residues are involved in intra- or intermolecular disulphide bond formation was not clear. The prediction from a computational model suggested that the last cysteine residue (Cys171) at the *C*-terminus

of E1 could be engaged in intermolecular disulphide linkage (Wang et al., 2014). However, the multiple sequence alignment throughout the whole species of pestiviruses (from species A to K) showed that Cys171 is not fully conserved. A C171F mutation is found in the species BVDV-2 and pronghorn antelope pestivirus (**Fig 5.3** and **Supplementary Material 1**). Since the E1-E2 heterodimer is critical for viral infectivity, it is unlikely that Cys171 is the important cysteine residue involved in E1-E2 heterodimer formation. Additionally, Cys171 is located in the putative transmembrane region and it is difficult to imagine how this residue could get contact with the free thiol at position 295 of E2. In our study, we showed that Cys123 plays the essential role in both E1-E2 heterodimer formation and the oligomerization of E1. However, this finding is contrasted by another E1-E2 heterodimer study which suggested that both Cys24 and Cys94 are important for E1-E2 heterodimerization (Fernandez-Sainz et al., 2011). We can't explain this discrepancy. Even though the respective study was done with CSFV and not BVDV, it is not probable that there is species dependent variation in this point. It could be due to misfolding of the mutated proteins in the CSFV study. In our hand, mutation of those residues did not interfere with oligomer formation (**Fig 5.21**). Moreover, at least Cys24 is located in the *N*-terminal glycosylated part of E1 that can be supposed to be exposed and is therefore not prove to interaction with Cys295 of E2 that is located close to the membrane. In HCV E1 forms a non-covalently linked trimeric form, which is thermally instable. The conserved 'GxxxG' motif in the transmembrane region of HCV E1 has been shown to be critical for this trimerization (Falson et al., 2015). Such a motif is missing in pestiviruses.

One paper about the function of CSFV E1 demonstrated that individual Cys to Ser mutations have no effect on E1-E2 heterodimerization in infected cells (Risatti et al., 2005), suggesting that there should be more than one disulphide linkage contributing to the heterodimer formation of E1-E2. However, the intramolecular disulphide bonds in pestiviral E2 (BVDV as an example) have been clarified via crystal structure analysis (El Omari et al., 2013). As shown in **Fig 6.4**, there are 17 cysteine residues distributed in the ectodomain of E2. Among them, the first 16 cysteine residues form 8 pairs of intramolecular disulphide linkages. Accordingly, the Cys295 is the only unpaired cysteine in E2 making it the logical candidate for both E2 homodimer and E1-E2 heterodimer formation. In this situation, it is unlikely that there is more than one intramolecular disulphide bond contributing to the heterodimer formation of E1-E2. In our study, we showed that Cys123, not Cys171, in E1 is the critical site for E1-E2 heterodimerization. The distance between Cys123 in E1 and the border of the putative TM domain of E1 is similar to that of Cys295 in E2, so that those two cysteine residues have a higher probability to get in contact, whereas Cys171 most likely is hidden in the membrane. Moreover, the published study also suggested that the individual substitutions of Cys residues in CSFV E1 are not essential for the infectivity *in vivo*, showing virulence features similar to those of parental virus. In contrast, we showed that the single mutation Cys123Ser in E1 triggered a defect of viral infectivity *in vitro*.

It is known that the E1-E2 heterodimer is involved in the attachment and entry step of pestiviruses (Ronecker et al., 2008). It is still unclear which of these two proteins contains the fusion relevant domain. Unexpectedly, the recently published crystal structure of BVDV E2 does not exhibit a class II fusion protein fold which was supposed to be present in pestiviruses in analogy to HCV. This indicated that the fusion machinery of pestiviruses could be totally different from so far reported examples (Li et al., 2013). It was considered that pestiviral E1 serves as a fusion protein. The hydrophobic residues from 54 to 83 (MHR) in the middle part of E1 have been proposed to be a putative fusion peptide. In our study, the MHR was shown to be closely related to the secretion/membrane association indicating that this segment with high probabilities binds to the membrane. The trimer formation is a typical feature of fusion protein, indicating E1 might be a good candidate to contain the fusion peptide. However, E1 seems to

be too short to harbor a typical class II or III fold of a normal fusion peptide. The crystal structure of E2 demonstrated that E2 has a very extended ectodomain (El Omari et al., 2013; Li et al., 2013). Consequently, it is hard to imagine how E1 could function as a fusogenic protein spanning the distance that the extended E2 establishes between viral and cellular membrane. The mechanism of the pestiviral fusion step still needs further detailed characterization. To answer this question, the crystal structure of pestiviral E1 is in urgently need.

In our study, the E1-E2 heterodimer was always detected with specific antibodies directed against E2. We tried several antibodies against the HA-tag. Unfortunately, none of them can be used for the detection of the E1-E2 heterodimer. It seems that when E1 is co-expressed with E2, the *N*-terminal HA-tags were shrouded by the structure of E2 after the folding of both proteins, so that the HA epitope is not detectable.

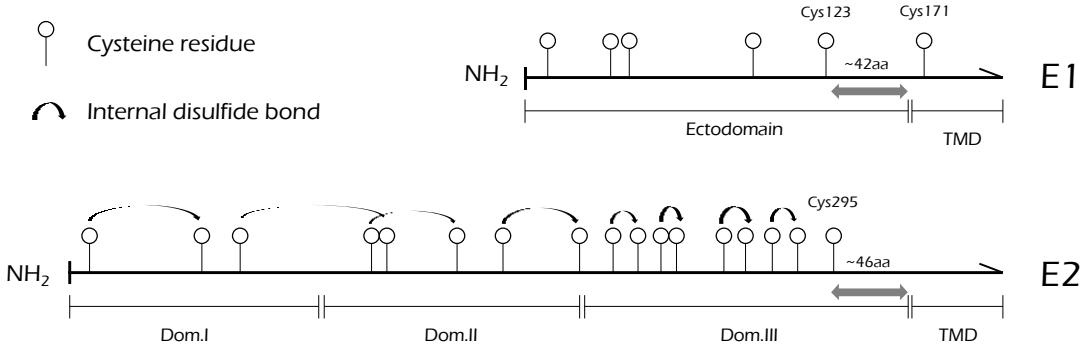


Figure 6.4 Schematic drawing of the cysteine residue distribution in native pestiviral E1 and E2 glycoproteins

Schematic representation of the distribution of cysteine residues in the E1 and E2 proteins of pestiviruses (BVDV CP7 as an example). Numbered circles represent the positions of the Cys residues in E1 and E2. Curved black arrows indicate internal disulphide bonds (El Omari et al., 2013). Straight grey arrows represent the distance away from the putative transmembrane region.

The primary interaction of the two partners in hetero- or homodimers of viral envelope proteins is often established by the TM regions (Cocquerel et al., 1999; Cocquerel et al., 1998; Op De et al., 2000). It has been shown that substitutions in positions K174 and R177 in E1 play a role in E1-E2 heterodimer formation (Ronecker et al., 2008). However, we observed that double K174A/R177A mutation in E1 has no significant effect on the E1-E2 heterodimerization. Even pYM-53, in which totally all the polar residues in TM region of E1 were removed, it can still form E1-E2 heterodimer when co-expressed with E2 wt. We accidentally found that the formation of the E1-E2 heterodimer is independent of the TM domain of E1, which is a rather surprising finding. The fact that E1 still forms E1 oligomers or E1-E2 heterodimers in the absence of its TM domain further proved that the Cys171, which is located in the putative TM region, is dispensable for the interaction between E1 and E2.

Normally, when membrane proteins lose their membrane anchors, they will be delivered to extracellular space via the conventional secretion pathway. Surprisingly, pestiviral E1 was still present in the ER in the absence of its membrane anchor. In our study, we also demonstrated that the hydrophobic sequence in the middle part (MHR) of E1 could be membrane associated. This region can affect the secretion of a series of *C*-terminally truncated E1 variants. Removal of all hydrophobic regions of E1 can result in ~88% secretion (shown in **Fig 5.31**). We hypothesize that this MHR could play an essential role in membrane binding so that E1 can still stay in the cell even without the membrane anchor. This finding is surprising since mutations of the *C*-terminal TM region can result in transport of E1 out of ER to the plasma membrane,

so that one would expect transport of a TM deleted E1 to the cell surface, which is not the case (data not shown). However, one has to keep in mind that the TM deleted E1 represents a highly artificial protein that might stay bound to ER chaperones because of misfolding. This would explain its retention in the ER despite the absence of its retention signal.

The results from VSV-g/E1 chimera analysed by FACS and IF assay clearly revealed that there is no retention associated signal located in the ectodomain of E1 (1-165) (as shown in **Fig 5.9**). In addition, the deletion of the entire MHR does not lead to the secretion or changes in intracellular location of E1. It can affect the secretion of E1 only under conditions of *C*-terminal truncation. In this study, we understand that the transmembrane regions of pestiviral envelope proteins could play other roles than simply membrane anchoring. HCV glycoproteins E1 an E2 are good examples of multifunctionality performed by their TMDs (Cocquerel et al., 1999; Cocquerel et al., 1998; Cocquerel et al., 2002; Cocquerel et al., 2000b). Normally, TM regions also play an essential role in the biological characteristics of membrane proteins, like translocation and folding. We speculated that the removal of the membrane anchor sequence results in the conformational changes of E1, the MHR might consequently be exposed to the water environment of the ER lumen. The exposed hydrophobic regions are not favourable under this condition, so that the MHR embeds into the membrane and interacts with the hydrophobic tails of the lipid molecules in the bilayer to make this region sequestered away from water (the schematic model shown in **Fig 6.3B**). We verified this MHR located on the ER luminal side by using selective permeabilization IF assay and Avi-tag biotinylation assay (data not shown). The polar charged residues Asp67 and Glu71 in this region may also play important roles in ER retention that makes E1 still retain within the ER even without the membrane anchor.

The experiment described in this thesis provide the first systematic analysis of the pestiviral E1 protein. The in past surprising results represent a signal can't step forward to understanding the biochemical and functional characteristics of this interesting protein but reveal also important questions that are still open and await further detailed investigation. Only with additional research on E1 will the mechanisms involved in pestiviral membrane fusion, assembly and budding be elucidated.

Zusammenfassung

Struktur- und Funktionsanalyse von pestiviralem E1-Glykoprotein

Pestiviren, Mitglieder der Familie *Flaviviridae*, gehören weltweit zu den wichtigsten Krankheitserregern von Nutztieren. Die Mitglieder der Gattung *Pestivirus* haben ein breites Wirtsspektrum (hauptsächlich Schweine und Wiederkäuer) und induzieren eine Vielzahl klinischer Manifestationen bei Nutztieren und Wildtieren. Obwohl mehrere gute Impfstoffe gegen die wichtigsten Pestiviren entwickelt wurden und seit langem eine Reihe strenger Biosicherheitsmaßnahmen wie Quarantäne- und Ausrottungsstrategien durchgeführt wurden, verursachen Pestiviren in der Tierhaltung erhebliche finanzielle Verluste.

Auf der Oberfläche von Pestiviruspartikeln befinden sich drei Hüllproteine. Unter diesen ist E1 am schlechtesten charakterisiert. Aufgrund des Fehlens spezifischer Antikörper gegen E1 sind sowohl funktionelle als auch strukturelle Informationen zu E1 immer noch unzureichend. E1 wurde nur im Zusammenhang mit den beiden anderen Hüllproteinen analysiert. In meiner Doktorarbeit konzentrierte ich mich auf die funktionelle und strukturelle Charakterisierung des Glykoproteins E1 der Pestiviren hinsichtlich seiner intrazellulären Lokalisation, des Retentionssignals, der Membrantopologie und der Oligomerisierung der Bildung von E1- und E1-E2-Heterodimeren, um einen ersten Einblick in die Molekular- und Zellbiologie dieses interessanten Proteins und die Voraussetzungen für die Interaktion der Pestivirus Hüllproteine und die Bildung von Pestiviruspartikeln zu erhalten.

Zunächst haben wir gezeigt, dass es keine Sekretion oder Zelloberflächenexpression von E1 gibt. Dies führte zu der grundlegenden Frage nach dem intrazellulären Kompartiment, in dem E1 hauptsächlich konzentriert ist. Mithilfe der Kolokalisationsanalyse mit Markerproteinen stellten wir fest, dass E1 überwiegend im ER und nicht im Golgi-Kompartiment lokalisiert ist. Da dieser Befund erhalten wurde, als E1 alleine exprimiert wurde, wurde damit nachgewiesen, dass E1 ein eigenes ER-Retentionssignal enthält.

Um die Determinanten für die ER-Retention von E1 zu charakterisieren, wurde eine Reihe von chimären und mutierten VSVg-E1-Proteinen analysiert. Es wurde gefunden, dass das intrazelluläre Retentionssignal in der mutmaßlichen TM-Domäne (letzte 30aa am *C*-Terminus) enthalten ist. Darüber hinaus konnte das Signal unter Verwendung einer Mutationsanalyse auf sechs vollständig konservierte polare Reste im mittleren Teil der TM Domäne des E1 eingegrenzt werden.

Dann wurde die Membrantopologie von E1 vor und nach der Signalpeptidspaltung bestimmt. Unter Verwendung von zwei unabhängigen biologischen Methoden kamen wir zu dem Schluss, dass E1 ein typisches Typ I-Transmembranprotein mit einem hydrophoben Membrananker am *C*-Terminus ist. Interessanterweise nimmt die Transmembrandomäne von E1, wenn die Situation vor der Signalpeptidabspaltung durch die Blockierung der Spaltstelle zwischen E1 und E2 nachgeahmt wird, eine haarnadelartige Struktur an, wobei sich der *C*-Terminus im ER-Lumen befindet.

Die Voraussetzungen für die Oligomerisierung von E1 und die Heterodimerisierung von E1-E2 wurden ebenfalls in unserer Studie untersucht. Überraschenderweise fanden wir, dass pestivirales E1 Homotrimere bildete, über die noch nie berichtet wurde. Sowohl Cys123 als auch Cys171 in E1 beeinflussen die Oligomerisierung in unterschiedlichem Maße. Eine

Koexpressionsanalyse mit E1/E2-Mutanten zeigte, dass Cys123 in E1 und, wie bereits bekannt, Cys295 in E2 die kritischen Reste für die Bildung von E1-E2-Heterodimeren sind. Cys295 in E2 wurde bereits als entscheidend für die E2-Homodimerisierung bestimmt. Um die Bedeutung der Bildung von E1-E2-Heterodimeren für die Lebensfähigkeit und Replikation von Pestiviren zu testen, analysierten wir auch Mutationen im infektiösen cDNA Klon von BVDV CP7, die keine E1-E2-Heterodimere bilden konnten. Es wurde gezeigt, dass diese BVDV-Mutanten ihre Infektiosität verlieren, was weiter beweist, dass diese beiden Stellen in E1 und E2 eine wesentliche Rolle im BVDV-Lebenszyklus spielen, höchstwahrscheinlich aufgrund ihrer Rolle bei der Heterodimerbildung.

In unserer Studie zeigte pestivirales E1 einige unerwartete Eigenschaften. Nach den vorläufigen Daten aus unserem Labor haben wir ein interessantes Phänomen beobachtet, dass E1 die Retention von E2 außer Kraft setzen kann, so dass das Fehlen des E1-Retentionssignals das Heterodimer zur Zelloberfläche lenkt, obwohl das E2-Retentionssignal noch vorhanden ist. Es wurde gezeigt, dass die kovalente Bindung zwischen E1 und E2 eine wesentliche Rolle für dieses Protein spielt. Weiterhin fanden wir, dass die E1-E2-Heterodimerbildung unabhängig von der TM-Domäne von E1 ist und damit einen völlig anderen Mechanismus zeigt als der des eng verwandten Hepatitis-C-Virus. Überraschenderweise führt die vollständige Deletion der TM-Region von E1 nicht zur Sekretion von Protein. Wir konnten zeigen, dass die hydrophobe Region im mittleren Teil von E1 höchstwahrscheinlich an die Membran bindet und die Sekretion von E1 in Abwesenheit der TM-Domäne verringert.

Summary

Structural and functional analysis of pestiviral E1 glycoprotein

Pestiviruses are grouped in the family *Flaviviridae*, are among the most important pathogens of farm animals worldwide. The members of the genus *Pestivirus* have a broad host range (mainly pigs and ruminants), and induce a variety of clinical manifestations in farm or wild animals. Even though several good vaccines against the most important pestiviruses have been developed and a series of strict bio-safety measures like quarantine and stamping-out strategies have long been carried out, pestiviruses cause severe financial losses in the animal farming industry.

Three envelope proteins are found on the surface of pestiviral virions. Among them, E1 is the least characterized. Due to the absence of specific antibodies directed against E1, both functional and structural information on E1 are still poor. E1 has only been analysed in context with the other two envelope proteins. In the present thesis work, I focused on the functional and structural characterization of the glycoprotein E1 of pestivirus with regard to its intracellular localization, the retention signal, the membrane topology and the oligomerization of E1 and E1-E2 heterodimer formation in order to gain an initial insight into the molecular and cellular biology of this interesting protein. The results of these analyses are also discussed the prerequisites for the steps leading to the assembly and budding of these viruses.

First of all, we showed that there is no secretion or cell surface expression of E1. This led to the basic question about the intracellular compartment where E1 is mainly concentrated. By using colocalization analysis with marker proteins, we determined that E1 localizes predominantly in the ER and not Golgi compartment. Since this finding was obtained when E1 was expressed alone, it proved that E1 contains an ER-retention signal of its own.

To characterize the determinants for ER retention of E1, a series of VSVg-E1 chimeric and mutated proteins were analysed. The intracellular retention signal was found to map to the putative TM domain (last 30aa at the *C*-terminal), furthermore, by using site direct mutagenesis analysis, the signal could be narrowed down to six fully conserved polar residues in the middle part of TM domain of E1.

Then, the membrane topology of E1 before and after the signal peptide cleavage were determined. By using two independent biological methods, we concluded that E1 is a typical type I transmembrane protein with a hydrophobic membrane anchor at its *C*-terminus. Interestingly, the pre-cleavage situation is mimicked by blocking the cleavage site between E1 and E2, the transmembrane domain of E1 adopt a hairpin-like structure with the *C*-terminus located in the ER lumen.

The prerequisites for the oligomerization of E1 and heterodimerization of E1-E2 were also explored in this study. Surprisingly, we found that pestiviral E1 formed homotrimers which has never been reported before. Both Cys123 and Cys171 in E1 affect the oligomerization in varying degrees. Co-expression analysis with E1/E2 mutants demonstrated that Cys123 in E1 and Cys295 in E2 are the critical sites for E1-E2 heterodimer formation. Meanwhile, Cys295 in E2 is also determinant for E2 homodimerization. To test for the importance of E1-E2 heterodimer formation for pestivirus viability and replication, we analysed the full-length infectious clone of BVDV CP7 bearing mutations that were not able to form E1-E2

heterodimers. Those BVDV mutants were shown to lose infectivity, further proving that those two sites in E1 and E2 play an essential role in the BVDV life cycle, most likely because of their role in heterodimer formation.

In our study, pestiviral E1 exhibited some unexpected characteristics. According to the preliminary data from our lab, we observed an interesting phenomenon that E1 can overrule the retention signal of E2, so that absence of the E1 retention signal directs the heterodimer to the cell surface even though the E2 retention signal is still present. It was shown that the covalent linkage between E1 and E2 plays an essential role for this process. Further, we found that the E1-E2 heterodimer formation is independent of the TM domain of E1, showing a totally different mechanism to that of the closely related Hepatitis C virus. Surprisingly, the complete deletion of the TM region of E1 does not result in the secretion of the protein. We were able to demonstrate that the hydrophobic region in the middle part of E1 most likely binds to the membrane and reduces the secretion of E1 in the absence of the TM domain.

References

Anelli, T., Sitia, R., 2008. Protein quality control in the early secretory pathway. The EMBO journal 27, 315-327.

Aoki, H., Ishikawa, K., Sakoda, Y., Sekiguchi, H., Kodama, M., Suzuki, S., Fukusho, A., 2001. Characterization of classical swine fever virus associated with defective interfering particles containing a cytopathogenic subgenomic RNA isolated from wild boar. The Journal of veterinary medical science / the Japanese Society of Veterinary Science 63, 751-758.

Aoki, H., Ishikawa, K., Sekiguchi, H., Suzuki, S., Fukusho, A., 2003. Pathogenicity and kinetics of virus propagation in swine infected with the cytopathogenic classical swine fever virus containing defective interfering particles. Arch Virol 148, 297-310.

Armengol, E., Wiesmuller, K.H., Wienhold, D., Buttner, M., Pfaff, E., Jung, G., Saalmuller, A., 2002. Identification of T-cell epitopes in the structural and non-structural proteins of classical swine fever virus. J Gen Virol 83, 551-560.

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

Bai, L., Wang, T., Zhao, G., Kovach, A., Li, H., 2018. The atomic structure of a eukaryotic oligosaccharyltransferase complex. Nature 555, 328-333.

Beltzer, J.P., Fiedler, K., Fuhrer, C., Geffen, I., Handschin, C., Wessels, H.P., Spiess, M., 1991. Charged residues are major determinants of the transmembrane orientation of a signal-anchor sequence. J Biol Chem 266, 973-978.

Bintintan, I., Meyers, G., 2010. A new type of signal peptidase cleavage site identified in an RNA virus polyprotein. J.Biol.Chem. 285, 8572-8584.

Blattner, C., Lee, J.H., Sliepen, K., Derking, R., Falkowska, E., de la Pena, A.T., Cupo, A., Julien, J.P., van Gils, M., Lee, P.S., Peng, W.J., Paulson, J.C., Poignard, P., Burton, D.R., Moore, J.P., Sanders, R.W., Wilson, I.A., Ward, A.B., 2014. Structural Delineation of a Quaternary, Cleavage-Dependent Epitope at the gp41-gp120 Interface on Intact HIV-1 Env Trimers. Immunity 40, 669-680.

Blome, S., Staubach, C., Henke, J., Carlson, J., Beer, M., 2017. Classical Swine Fever-An Updated Review. Viruses 9.

Bolin, S.R., McClurkin, A.W., Cutlip, R.C., Coria, M.F., 1985. Severe clinical disease induced in cattle persistently infected with noncytopathogenic bovine viral diarrhea virus by superinfection with cytopathogenic bovine viral diarrhea virus. Am J Vet Res 46, 573-576.

Borca, M.V., Holinka, L.G., Ramirez-Medina, E., Risatti, G.R., Vuono, E.A., Berggren, K.A., Gladue, D.P., 2019. Identification of structural glycoprotein E2 domain critical to mediate replication of Classical Swine Fever Virus in SK6 cells. Virology 526, 38-44.

Branza-Nichita, N., Durantel, D., Carrouee-Durantel, S., Dwek, R.A., Zitzmann, N., 2001. Antiviral effect of N-butyldeoxynojirimycin against bovine viral diarrhea virus correlates with misfolding of E2 envelope proteins and impairment of their association into E1-E2 heterodimers. J Virol 75, 3527-3536.

Brauer, P., Parker, J.L., Gerondopoulos, A., Zimmermann, I., Seeger, M.A., Barr, F.A., Newstead, S., 2019. Structural basis for pH-dependent retrieval of ER proteins from the Golgi by the KDEL receptor. Science 363, 1103-1107.

Bretscher, M.S., Munro, S., 1993. Cholesterol and the Golgi apparatus. Science 261, 1280-1281.

Brock, K.V., 2004. The many faces of bovine viral diarrhea virus. The Veterinary clinics of North America. Food animal practice 20, 1-3.

Brownlie, J., 1990. Pathogenesis of mucosal disease and molecular aspects of bovine virus diarrhoea virus. Vet.Microbiol. 23, 371-382.

Brownlie, J., Clarke, M.C., Howard, C.J., 1984. Experimental production of fatal mucosal disease in cattle. Vet Rec 114, 535-536.

Bulleid, N.J., 2012. Disulfide bond formation in the mammalian endoplasmic reticulum. . Cold Spring Harbor perspectives in biology 4(11), a013219.

Burrack, S., Aberle, D., Bürck, J., Ulrich, A.S., Meyers, G., 2012. A new type of intracellular retention signal identified in a pestivirus structural glycoprotein. FASEB journal : official publication of the Federation of American Societies for Experimental Biology 26, 3292-3305.

Cao, L., Yu, B., Kong, D., Cong, Q., Yu, T., Chen, Z., Hu, Z., Chang, H., Zhong, J., Baker, D., He, Y., 2019. Functional expression and characterization of the envelope glycoprotein E1E2 heterodimer of hepatitis C virus. PLoS Pathog 15, e1007759.

Cao, L.W., Diedrich, J.K., Ma, Y.H., Wang, N.S., Pauthner, M., Park, S.K.R., Delahunty, C.M., McLellan, J.S., Burton, D.R., Yates, J.R., Paulson, J.C., 2018. Global site-specific analysis of glycoprotein N-glycan processing. Nat Protoc 13, 1196-1212.

Capitani, M., Sallese, M., 2009. The KDEL receptor: New functions for an old protein. Febs Letters 583, 3863-3871.

Carbaugh, D.L., Baric, R.S., Lazear, H.M., 2019. Envelope Protein Glycosylation Mediates Zika Virus Pathogenesis. J Virol 93.

Castelli, M., Clementi, N., Pfaff, J., Sautto, G.A., Diotti, R.A., Burioni, R., Doranz, B.J., Dal Peraro, M., Clementi, M., Mancini, N., 2017. A Biologically-validated HCV E1E2 Heterodimer Structural Model. Scientific reports 7, 214.

Ceppi, M., de Bruin, M.G., Seuberlich, T., Balmelli, C., Pascolo, S., Ruggli, N., Wienhold, D., Tratschin, J.D., McCullough, K.C., Summerfield, A., 2005. Identification of classical swine fever virus protein E2 as a target for cytotoxic T cells by using mRNA-transfected antigen-presenting cells. J.Gen.Virol. 86, 2525-2534.

Chen, J., He, W.R., Shen, L., Dong, H., Yu, J., Wang, X., Yu, S., Li, Y., Li, S., Luo, Y., Sun, Y., Qiu, H.J., 2015. The laminin receptor is a cellular attachment receptor for classical Swine Fever virus. J Virol 89, 4894-4906.

Chirico, W.J., Waters, M.G., Blobel, G., 1988. 70k Heat-Shock Related Proteins Stimulate Protein Translocation into Microsomes. Nature 332, 805-810.

Ciczora, Y., Callens, N., Montpellier, C., Bartosch, B., Cosset, F.L., De Beeck, A.O., Dubuisson, J., 2005. Contribution of the charged residues of hepatitis C virus glycoprotein E2 transmembrane domain to the functions of the E1E2 heterodimer. J Gen Virol 86, 2793-2798.

Ciczora, Y., Callens, N., Seron, K., Rouille, Y., Dubuisson, J., 2010. Identification of a dominant endoplasmic reticulum-retention signal in yellow fever virus pre-membrane protein. J Gen Virol 91, 404-414.

Cocquerel, L., Duvet, S., Meunier, J.C., Pillez, A., Cacan, R., Wychowski, C., Dubuisson, J., 1999. The transmembrane domain of hepatitis C virus glycoprotein E1 is a signal for static retention in the endoplasmic reticulum. Journal of virology 73, 2641-2649.

Cocquerel, L., Meunier, J.C., Pillez, A., Wychowski, C., Dubuisson, J., 1998. A retention signal necessary and sufficient for endoplasmic reticulum localization maps to the transmembrane domain of hepatitis C virus glycoprotein E2. J Virol 72, 2183-2191.

Cocquerel, L., Op De, B.A., Lambot, M., Roussel, J., Delgrange, D., Pillez, A., Wychowski, C., Penin, F., Dubuisson, J., 2002. Topological changes in the transmembrane domains of hepatitis C virus envelope glycoproteins. EMBO J. 21, 2893-2902.

Cocquerel, L., Wychowski, C., Minner, F., Penin, F., Dubuisson, J., 2000a. Charged residues in the transmembrane domains of hepatitis C virus glycoproteins play a major role in the processing, subcellular localization, and assembly of these envelope proteins. Journal of virology 74, 3623-3633.

Cocquerel, L., Wychowski, C., Minner, F., Penin, F., Dubuisson, J., 2000b. Charged residues in the transmembrane domains of hepatitis C virus glycoproteins play a major role in the processing, subcellular localization, and assembly of these envelope proteins. J Virol 74, 3623-3633.

Crooks, G.E., Hon, G., Chandonia, J.M., Brenner, S.E., 2004. WebLogo: a sequence logo generator. Genome Res 14, 1188-1190.

Depner, K., Bauer, T., Liess, B., 1992. Thermal and pH stability of pestiviruses. Rev.Sci.Tech. 11, 885-893.

Deshaies, R.J., Sanders, S.L., Feldheim, D.A., Schekman, R., 1991. Assembly of Yeast Sec Proteins Involved in Translocation into the Endoplasmic-Reticulum into a Membrane-Bound Multisubunit Complex. Nature 349, 806-808.

Duvet, S., Cocquerel, L., Pillez, A., Cacan, R., Verbert, A., Moradpour, D., Wychowski, C., & Dubuisson, J., 1998. Hepatitis C virus glycoprotein complex localization in the endoplasmic reticulum involves a determinant for retention and not retrieval. The Journal of biological chemistry 273(48), 32088–32095.

Edwards, S., Fukusho, A., Lefevre, P.C., Lipowski, A., Pejsak, Z., Roehe, P., Westergaard, J., 2000. Classical swine fever: the global situation. Veterinary Microbiology 73, 103-119.

El Omari, K., Iourin, O., Harlos, K., Grimes, J.M., Stuart, D.I., 2013. Structure of a pestivirus envelope glycoprotein E2 clarifies its role in cell entry. Cell reports 3, 30-35.

Elbers, K., Tautz, N., Becher, P., Rümenapf, T., Thiel, H.J., 1996. Processing in the Pestivirus E2-NS2 region: identification of the nonstructural proteins p7 and E2p7. J Virol 70, 4131-4135.

Falkowska, E., Le, K.M., Ramos, A., Doores, K.J., Lee, J.H., Blattner, C., Ramirez, A., Derking, R., van Gils, M.J., Liang, C.H., McBride, R., von Bredow, B., Shivatare, S.S., Wu, C.Y., Chan-Hui, P.Y., Liu, Y., Feizi, T., Zwick, M.B., Koff, W.C., Seaman, M.S., Swiderek, K., Moore, J.P., Evans, D., Paulson, J.C., Wong, C.H., Ward, A.B., Wilson, I.A., Sanders, R.W., Poignard, P., Burton, D.R., 2014. Broadly neutralizing HIV antibodies define a glycan-dependent epitope on the prefusion conformation of gp41 on cleaved envelope trimers. Immunity 40, 657-668.

Falson, P., Bartosch, B., Alsaleh, K., Tews, B.A., Loquet, A., Ciczora, Y., Riva, L., Montigny, C., Montpellier, C., Duverlie, G., Pecheur, E.I., le Maire, M., Cosset, F.L., Dubuisson, J., Penin, F., 2015. Hepatitis C Virus Envelope Glycoprotein E1 Forms Trimers at the Surface of the Virion. J Virol 89, 10333-10346.

Farquhar, R., Honey, N., Murant, S.J., Bossier, P., Schultz, L., Montgomery, D., Ellis, R.W., Freedman, R.B., Tuite, M.F., 1991. Protein Disulfide Isomerase Is Essential for Viability in Saccharomyces-Cerevisiae. Gene 108, 81-89.

Fernandez-Sainz, I., Holinka, L.G., Gavrilov, B.K., Prarat, M.V., Gladue, D., Lu, Z., Jia, W., Risatti, G.R., Borca, M.V., 2009. Alteration of the N-linked glycosylation condition in E1 glycoprotein of Classical Swine Fever Virus strain Brescia alters virulence in swine. Virology 386, 210-216.

Fernandez-Sainz, I., Holinka, L.G., Gladue, D., O'Donnell, V., Lu, Z., Gavrilov, B.K., Risatti, G.R., Borca, M.V., 2011. Substitution of specific cysteine residues in the E1 glycoprotein of classical swine fever virus strain Brescia affects formation of E1-E2 heterodimers and alters virulence in swine. J Virol 85, 7264-7272.

Fernandez-Sainz, I.J., Largo, E., Gladue, D.P., Fletcher, P., O'Donnell, V., Holinka, L.G., Carey, L.B., Lu, X., Nieva, J.L., Borca, M.V., 2014. Effect of specific amino acid substitutions in the putative fusion peptide of structural glycoprotein E2 on Classical Swine Fever Virus replication. Virology 456-457, 121-130.

Firth, C., Bhat, M., Firth, M.A., Williams, S.H., Frye, M.J., Simmonds, P., Conte, J.M., Ng, J., Garcia, J., Bhuva, N.P., Lee, B., Che, X., Quan, P.L., Lipkin, W.I., 2014. Detection of zoonotic pathogens and characterization of novel viruses carried by commensal Rattus norvegicus in New York City. mBio 5, e01933-01914.

Fournillier, A., Wychowski, C., Boucreux, D., Baumert, T.F., Meunier, J.C., Jacobs, D., Muguet, S., Depla, E., Inchauspe, G., 2001. Induction of hepatitis C virus E1 envelope protein-specific immune response can be enhanced by mutation of N-glycosylation sites. Journal of virology 75, 12088-12097.

Franzoni, G., Kurkure, N.V., Essler, S.E., Pedrera, M., Everett, H.E., Bodman-Smith, K.B., Crooke, H.R., Graham, S.P., 2013. Proteome-wide screening reveals immunodominance in the CD8 T cell response against classical swine fever virus with antigen-specificity dependent on MHC class I haplotype expression. PloS one 8, e84246.

Garry, R.F., Dash, S., 2003. Proteomics computational analyses suggest that hepatitis C virus E1 and pestivirus E2 envelope glycoproteins are truncated class II fusion proteins. Virology 307, 255-265.

Goder, V., Spiess, M., 2001. Topogenesis of membrane proteins: determinants and dynamics. FEBS Lett 504, 87-93.

Gorlich, D., Prehn, S., Hartmann, E., Kalies, K.U., Rapoport, T.A., 1992. A Mammalian Homolog of Sec61p and Secyp Is Associated with Ribosomes and Nascent Polypeptides during Translocation. Cell 71, 489-503.

Gorlich, D., Rapoport, T.A., 1993. Protein Translocation into Proteoliposomes Reconstituted from Purified Components of the Endoplasmic-Reticulum Membrane. Cell 75, 615-630.

Griffiths G, R.P., 1992. Cell biology of viruses that assemble along the biosynthetic pathway. Semin Cell Biol. 3(5), 367-381.

Grummer, B., Beer, M., Liebler-Tenorio, E., Greiser-Wilke, I., 2001. Localization of viral proteins in cells infected with bovine viral diarrhoea virus. J.Gen.Virol. 82, 2597-2605.

Grummer, B., Grotha, S., Greiser-Wilke, I., 2004. Bovine viral diarrhoea virus is internalized by clathrin-dependent receptor-mediated endocytosis. J Vet Med B 51, 427-432.

Harada, T., Tautz, N., Thiel, H.J., 2000. E2-p7 region of the bovine viral diarrhea virus polyprotein: processing and functional studies. J Virol 74, 9498-9506.

Hartmann, E., Rapoport, T.A., Lodish, H.F., 1989. Predicting the Orientation of Eukaryotic Membrane-Spanning Proteins. Proceedings of the National Academy of Sciences of the United States of America 86, 5786-5790.

Heijne, G., 1986. The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology. The EMBO journal 5, 3021-3027.

Heimann, M., Roman-Sosa, G., Martoglio, B., Thiel, H.J., Rümenapf, T., 2006. Core protein of pestiviruses is processed at the C terminus by signal peptide peptidase. J.Virol. 80, 1915-1921.

Higy, M., Junne, T., Spiess, M., 2004. Topogenesis of membrane proteins at the endoplasmic reticulum. Biochemistry 43, 12716-12722.

Hughes, H., Stephens, D.J., 2008. Assembly, organization, and function of the COPII coat. Histochem Cell Biol 129, 129-151.

Hulst, M.M., Moormann, R.J., 1997. Inhibition of pestivirus infection in cell culture by envelope proteins E(rns) and E2 of classical swine fever virus: E(rns) and E2 interact with different receptors. J.Gen.Virol. 78 (Pt 11), 2779-2787.

J A Rothblatt , D.I.M., 1986. Secretion in yeast: translocation and glycosylation of prepro-alpha-factor in vitro can occur via an ATP-dependent post-translational mechanism. The EMBO journal 5, 1031-1036.

Jackson, M.R., Nilsson, T., Peterson, P.A., 1990. Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum. The EMBO journal 9, 3153-3162.

Jansens, A., van Duijn, E., Braakman, I., 2002. Coordinated nonvectorial folding in a newly synthesized multidomain protein. Science 298, 2401-2403.

Jia, J., Yue, X.H., Zhu, L.H., Jing, S.Y., Wang, Y.J., Gim, B., Qian, Y., Lee, I.T., 2020. KDEL receptor is a cell surface receptor that cycles between the plasma membrane and the Golgi via clathrin-mediated transport carriers. Cell Mol Life Sci.

Jin, H., Komita, M., Aoe, T., 2017. The Role of BiP Retrieval by the KDEL Receptor in the Early Secretory Pathway and its Effect on Protein Quality Control and Neurodegeneration. Front Mol Neurosci 10, 222.

Jo, W.K., van Elk, C., van de Bildt, M., van Run, P., Petry, M., Jesse, S.T., Jung, K., Ludlow, M., Kuiken, T., Osterhaus, A., 2019. An evolutionary divergent pestivirus lacking the N(pro) gene systemically infects a whale species. Emerg Microbes Infect 8, 1383-1392.

Kohl, W., Zimmer, G., Greiser-Wilke, I., Haas, L., Moennig, V., Herrler, G., 2004. The surface glycoprotein E2 of bovine viral diarrhoea virus contains an intracellular localization signal. J Gen Virol 85, 1101-1111.

Köhl, W., Zimmer, G., Greiser-Wilke, I., Haas, L., Moennig, V., Herrler, G., 2004. The surface glycoprotein E2 of bovine viral diarrhoea virus contains an intracellular localization signal. J.Gen.Virol. 85, 1101-1111.

Kolykhalov, A.A., Agapov, E.V., Blight, K.J., Mihalik, K., Feinstone, S.M., Rice, C.M., 1997. Transmission of hepatitis C by intrahepatic inoculation with transcribed RNA. Science 277, 570-574.

Kosmidou, A., Büttner, M., Meyers, G., 1998. Isolation and characterization of cytopathogenic classical swine fever virus (CSFV). Arch Virol 143, 1295-1309.

Kümmerer, B.M., Tautz, N., Becher, P., Thiel, H., Meyers, G., 2000. The genetic basis for cytopathogenicity of pestiviruses. Vet Microbiol 77, 117-128.

Kupfermann, H., Thiel, H.J., Dubovi, E.J., Meyers, G., 1996. Bovine viral diarrhea virus: Characterization of a cytopathogenic defective interfering particle with two internal deletions. J Virol 70, 8175-8181.

Kushima, Y., Wakita, T., Hijikata, M., 2010. A Disulfide-Bonded Dimer of the Core Protein of Hepatitis C Virus Is Important for Virus-Like Particle Production. Journal of virology 84, 9118-9127.

Laboissiere, M.C., Sturley, S. L., & Raines, R. T. , 1995. The essential function of protein-disulfide isomerase is to unscramble non-native disulfide bonds. Journal of Biological Chemistry 270(47), 28006-28009.

Lackner, T., Müller, A., König, M., Thiel, H.J., Tautz, N., 2005. Persistence of bovine viral diarrhea virus is determined by a cellular cofactor of a viral autoprotease. J.Virol. 79, 9746-9755.

Lackner, T., Müller, A., Pankraz, A., Becher, P., Thiel, H.J., Gorbalenya, A.E., Tautz, N., 2004. Temporal modulation of an autoprotease is crucial for replication and pathogenicity of an RNA virus. J.Virol. 78, 10765-10775.

Lackner, T., Thiel, H.J., Tautz, N., 2006. Dissection of a viral autoprotease elucidates a function of a cellular chaperone in proteolysis. Proc.Natl.Acad.Sci.U.S.A 103, 1510-1515.

Lamp, B., Riedel, C., Roman-Sosa, G., Heimann, M., Jacobi, S., Becher, P., Thiel, H.J., Rümenapf, T., 2011. Biosynthesis of classical swine fever virus nonstructural proteins. J.Virol. 85, 3607-3620.

Lamp, B., Riedel, C., Wentz, E., Tortorici, M.A., Rümenapf, T., 2013. Autocatalytic cleavage within classical swine fever virus NS3 leads to a functional separation of protease and helicase. J Virol 87, 11872-11883.

Lecot, S., Belouzard, S., Dubuisson, J., Rouille, Y., 2005. Bovine viral diarrhea virus entry is dependent on clathrin-mediated endocytosis. J Virol 79, 10826-10829.

Li, Y., Wang, J., Kanai, R., Modis, Y., 2013. Crystal structure of glycoprotein E2 from bovine viral diarrhea virus. Proc Natl Acad Sci U S A 110, 6805-6810.

Liang, D., Sainz, I.F., Ansari, I.H., Gil, L.H., Vassilev, V., Donis, R.O., 2003. The envelope glycoprotein E2 is a determinant of cell culture tropism in ruminant pestiviruses. J Gen Virol 84, 1269-1274.

Maurer, K., Krey, T., Moennig, V., Thiel, H.J., Rümenapf, T., 2004. CD46 is a cellular receptor for bovine viral diarrhea virus. J.Virol. 78, 1792-1799.

McCaffrey, K., Boo, I., Tewierek, K., Edmunds, M.L., Poumbourios, P., Drummer, H.E., 2012. Role of conserved cysteine residues in hepatitis C virus glycoprotein e2 folding and function. J Virol 86, 3961-3974.

Mettenleiter, T.C., Muller, F., Granzow, H., Klupp, B.G., 2013. The way out: what we know and do not know about herpesvirus nuclear egress. Cellular microbiology 15, 170-178.

Meunier, J.C., Fournillier, A., Choukhi, A., Cahour, A., Cocquerel, L., Dubuisson, J., Wychowski, C., 1999. Analysis of the glycosylation sites of hepatitis C virus (HCV) glycoprotein E1 and the influence of E1 glycans on the formation of the HCV glycoprotein complex. J Gen Virol 80 (Pt 4), 887-896.

Meyer, C., Von Freyburg, M., Elbers, K., Meyers, G., 2002. Recovery of virulent and RNase-negative attenuated type 2 bovine viral diarrhea viruses from infectious cDNA clones. J.Virol. 76, 8494-8503.

Meyers, G., Tautz, N., Becher, P., Thiel, H.J., Kümmerer, B., 1996a. Recovery of cytopathogenic and noncytopathogenic bovine viral diarrhea viruses from cDNA constructs. J Virol 70, 8606-8613.

Meyers, G., Tautz, N., Dubovi, E.J., Thiel, H.J., 1991. Viral cytopathogenicity correlated with integration of ubiquitin-coding sequences. Virology 180, 602-616.

Meyers, G., Thiel, H.J., 1995. Cytopathogenicity of classical swine fever virus caused by defective interfering particles. J Virol 69, 3683-3689.

Meyers, G., Thiel, H.J., 1996. Molecular characterization of pestiviruses. Advances in virus research 47, 53-117.

Meyers, G., Thiel, H.J., Rümenapf, T., 1996b. Classical swine fever virus: Recovery of infectious viruses from cDNA constructs and generation of recombinant cytopathogenic defective interfering particles. J Virol 70, 1588-1595.

Moennig, V., Floegel-Niesmann, G., Greiser-Wilke, I., 2003. Clinical Signs and Epidemiology of Classical Swine Fever: A Review of New Knowledge. The Veterinary Journal 165, 11-20.

Moennig, V., Frey, H.R., Liebler, E., Polenz, P., Liess, B., 1990. Reproduction of mucosal disease with cytopathogenic bovine viral diarrhoea virus selected in vitro. Vet Rec 127, 200-203.

Moriishi, K., Matsuura, Y., 2012. Exploitation of lipid components by viral and host proteins for hepatitis C virus infection. Front Microbiol 3, 54.

Munir Iqbal, H.F.-S.a.J.W.M., 2000. Interactions of bovine viral diarrhoea virus glycoprotein Erns with cell surface glycosaminoglycans. Journal of General Virology 81, 451–459.

Munro, S., Pelham, H.R., 1987. A C-terminal signal prevents secretion of luminal ER proteins. Cell 48, 899-907.

Neufeldt, C.J., Joyce, M.A., Van Buuren, N., Levin, A., Kirkegaard, K., Gale, M., Jr., Tyrrell, D.L., Wozniak, R.W., 2016. The Hepatitis C Virus-Induced Membranous Web and Associated Nuclear Transport Machinery Limit Access of Pattern Recognition Receptors to Viral Replication Sites. PLoS Pathog 12, e1005428.

Nilsson, T., Jackson, M., Peterson, P.A., 1989. Short Cytoplasmic Sequences Serve as Retention Signals for Transmembrane Proteins in the Endoplasmic-Reticulum. Cell 58, 707-718.

Nishimura, N., Balch, W.E., 1997. A di-acidic signal required for selective export from the endoplasmic reticulum. Science 277, 556-558.

Op De, B.A., Montserret, R., Duvet, S., Cocquerel, L., Cacan, R., Barberot, B., Le, M.M., Penin, F., Dubuisson, J., 2000. The transmembrane domains of hepatitis C virus envelope glycoproteins E1 and E2 play a major role in heterodimerization. J.Biol.Chem. 275, 31428-31437.

Ozcelik, D., Seto, A., Rakic, B., Farzam, A., Supek, F., Pezacki, J.P., 2018. Gene Expression Profiling of Endoplasmic Reticulum Stress in Hepatitis C Virus-Containing Cells Treated with an Inhibitor of Protein Disulfide Isomerases. ACS Omega 3, 17227-17235.

Postel, A., Letzel, T., Muller, F., Ehricht, R., Pourquier, P., Dauber, M., Grund, C., Beer, M., Harder, T.C., 2011. In vivo biotinylated recombinant influenza A virus hemagglutinin for use in subtype-specific serodiagnostic assays. Analytical Biochemistry 411, 22-31.

Radtke, C., Tews, B.A., 2017. Retention and topology of the bovine viral diarrhea virus glycoprotein E2. J Gen Virol 98, 2482-2494.

Rapoport, T.A., 2007. Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. Nature 450, 663-669.

Reimann, I., Depner, K., Trapp, S., Beer, M., 2004. An avirulent chimeric Pestivirus with altered cell tropism protects pigs against lethal infection with classical swine fever virus. Virology 322, 143-157.

Risatti, G.R., Holinka, L.G., Fernandez Sainz, I., Carrillo, C., Lu, Z., Borca, M.V., 2007a. N-linked glycosylation status of classical swine fever virus strain Brescia E2 glycoprotein influences virulence in swine. J Virol 81, 924-933.

Risatti, G.R., Holinka, L.G., Fernandez, S.I., Carrillo, C., Lu, Z., Borca, M.V., 2007b. N-linked glycosylation status of classical swine fever virus strain Brescia E2 glycoprotein influences virulence in swine. J.Virol. 81, 924-933.

Risatti, G.R., Holinka, L.G., Lu, Z., Kutish, G.F., Tulman, E.R., French, R.A., Sur, J.H., Rock, D.L., Borca, M.V., 2005. Mutation of E1 glycoprotein of classical swine fever virus affects viral virulence in swine. Virology 343, 116-127.

Ron, D., Walter, P., 2007. Signal integration in the endoplasmic reticulum unfolded protein response. Nat Rev Mol Cell Biol 8, 519-529.

Ronecker, S., Zimmer, G., Herrler, G., Greiser-Wilke, I., Grummer, B., 2008. Formation of bovine viral diarrhea virus E1-E2 heterodimers is essential for virus entry and depends on charged residues in the transmembrane domains. J Gen Virol 89, 2114-2121.

Rumenapf, T., Unger, G., Strauss, J.H., Thiel, H.J., 1993. Processing of the envelope glycoproteins of pestiviruses. J Virol 67, 3288-3294.

Rümenapf, T., Unger, G., Strauss, J.H., Thiel, H.J., 1993. Processing of the envelope glycoproteins of pestiviruses. J Virol 67, 3288-3295.

Saaranen, M.J., Ruddock, L.W., 2019. Applications of catalyzed cytoplasmic disulfide bond formation. Biochem Soc Trans 47, 1223-1231.

Sainz, I.F., Holinka, L.G., Lu, Z., Risatti, G.R., Borca, M.V., 2008. Removal of a N-linked glycosylation site of classical swine fever virus strain Brescia Erns glycoprotein affects virulence in swine. Virology 370, 122-129.

Schmeiser, S., Mast, J., Thiel, H.J., König, M., 2014. Morphogenesis of pestiviruses: new insights from ultrastructural studies of strain Giraffe-1. J Virol 88, 2717-2724.

Schneider, T.D., Stephens, R.M., 1990. Sequence logos: a new way to display consensus sequences. Nucleic Acids Res 18, 6097-6100.

Sethi, M.K., Fanayan, S., 2015. Mass Spectrometry-Based N-Glycomics of Colorectal Cancer. Int J Mol Sci 16, 29278-29304.

Sharpe, H.J., Stevens, T.J., Munro, S., 2010. A comprehensive comparison of transmembrane domains reveals organelle-specific properties. Cell 142, 158-169.

Shental-Bechor, D., Levy, Y., 2009. Folding of glycoproteins: toward understanding the biophysics of the glycosylation code. Curr Opin Struct Biol 19, 524-533.

Shi, H., Leng, C., Xu, Q., Shi, H., Sun, S., Kan, Y., Yao, L., 2018. Characterization of a Pestivirus H isolate originating from goats. Virus genes 54, 603-607.

Shin, J., Dunbrack, R.L., Jr., Lee, S., Strominger, J.L., 1991. Signals for retention of transmembrane proteins in the endoplasmic reticulum studied with CD4 truncation mutants. Proc Natl Acad Sci U S A 88, 1918-1922.

Simmonds, P., Becher, P., Bukh, J., Gould, E.A., Meyers, G., Monath, T., Muerhoff, S., Pletnev, A., Rico-Hesse, R., Smith, D.B., Stapleton, J.T., Ictv Report, C., 2017. ICTV Virus Taxonomy Profile: Flaviviridae. J Gen Virol 98, 2-3.

Singh, S., Mittal, A., 2016. Transmembrane Domain Lengths Serve as Signatures of Organismal Complexity and Viral Transport Mechanisms. Scientific reports 6.

Smith, D.B., Meyers, G., Bukh, J., Gould, E.A., Monath, T., Scott Muerhoff, A., Pletnev, A., Rico-Hesse, R., Stapleton, J.T., Simmonds, P., Becher, P., 2017. Proposed revision to the taxonomy of the genus Pestivirus, family Flaviviridae. Journal of General Virology 98, 2106-2112.

Stark, R., Meyers, G., Rümenapf, T., Thiel, H.J., 1993. Processing of pestivirus polyprotein: Cleavage site between autoprotease and nucleocapsid protein of classical swine fever virus. J Virol 67, 7088-7095.

Tautz, N., Elbers, K., Stoll, D., Meyers, G., Thiel, H.J., 1997. Serine protease of pestiviruses: Determination of cleavage sites. Journal of virology 71, 5415-5422.

Tautz, N., Meyers, G., Thiel, H.J., 1998. Pathogenesis of mucosal disease, a deadly disease of cattle caused by a pestivirus. Clin Diagn Virol 10, 121-127.

Tautz, N., Tews, B.A., Meyers, G., 2015. The Molecular Biology of Pestiviruses. Advances in virus research 93, 47-160.

Taylor, M.E., and Kurt Drickamer, 2011. Introduction to glycobiology Oxford university press.

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.

Tews, B.A., Schurmann, E.M., Meyers, G., 2009a. Mutation of cysteine 171 of pestivirus E rns RNase prevents homodimer formation and leads to attenuation of classical swine fever virus. J Virol 83, 4823-4834.

Tews, B.A., Schürmann, E.M., Meyers, G., 2009b. Mutation of cysteine 171 of pestivirus E rns RNase prevents homodimer formation and leads to attenuation of classical swine fever virus. J.Virol. 83, 4823-4834.

Thiel, H.J., Stark, R., Weiland, E., Rümenapf, T., Meyers, G., 1991. Hog cholera virus: molecular composition of virions from a pestivirus. J Virol 65, 4705-4712.

Tong, Y., Lavillette, D., Li, Q., Zhong, J., 2018. Role of Hepatitis C Virus Envelope Glycoprotein E1 in Virus Entry and Assembly. Front Immunol 9, 1411.

Tucakov, A.K., Yavuz, S., Schurmann, E.M., Mischler, M., Klingebeil, A., Meyers, G., 2018. Restoration of glycoprotein E(rns) dimerization via pseudoreversion partially restores virulence of classical swine fever virus. J Gen Virol 99, 86-96.

Ulmschneider, M.B., Sansom, M.S., 2001. Amino acid distributions in integral membrane protein structures. Biochim Biophys Acta 1512, 1-14.

van Gennip, H.G., Hesselink, A.T., Moormann, R.J., Hulst, M.M., 2005. Dimerization of glycoprotein E(rns) of classical swine fever virus is not essential for viral replication and infection. Arch.Virol. 150, 2271-2286.

von Heijne, G., 1990. The signal peptide. J Membr Biol 115, 195-201.

von Heijne, G., 1995. Membrane protein assembly: rules of the game. Bioessays 17, 25-30.

Wahid, A., Helle, F., Descamps, V., Duverlie, G., Penin, F., Dubuisson, J., 2013. Disulfide bonds in hepatitis C virus glycoprotein E1 control the assembly and entry functions of E2 glycoprotein. J Virol 87, 1605-1617.

Wang, J., Li, Y., Modis, Y., 2014. Structural models of the membrane anchors of envelope glycoproteins E1 and E2 from pestiviruses. Virology 454-455, 93-101.

Wang, Z., Nie, Y., Wang, P., Ding, M., Deng, H., 2004. Characterization of classical swine fever virus entry by using pseudotyped viruses: E1 and E2 are sufficient to mediate viral entry. Virology 330, 332-341.

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.

Weiland, E., Stark, R., Haas, B., Rumenapf, T., Meyers, G., Thiel, H.J., 1990. Pestivirus Glycoprotein Which Induces Neutralizing Antibodies Forms Part of a Disulfide-Linked Heterodimer. Journal of virology 64, 3563-3569.

Weiland, F., Weiland, E., Unger, G., Saalmüller, A., Thiel, H.J., 1999. Localization of pestiviral envelope proteins E(rns) and E2 at the cell surface and on isolated particles. J.Gen.Virol. 80 (Pt 5), 1157-1165.

Wiskerchen, M., Collett, M.S., 1991. Pestivirus gene expression: protein p80 of bovine viral diarrhea virus is a proteinase involved in polyprotein processing. Virology 184, 341-350.

Wolk, B., Buchele, B., Moradpour, D., Rice, C.M., 2008. A dynamic view of hepatitis C virus replication complexes. J Virol 82, 10519-10531.

Wu, Z., Ren, X., Yang, L., Hu, Y., Yang, J., He, G., Zhang, J., Dong, J., Sun, L., Du, J., Liu, L., Xue, Y., Wang, J., Yang, F., Zhang, S., Jin, Q., 2012. Virome analysis for identification of novel mammalian viruses in bat species from Chinese provinces. J Virol 86, 10999-11012.

Xu, J., Mendez, E., Caron, P.R., Lin, C., Murcko, M.A., Collett, M.S., Rice, C.M., 1997. Bovine viral diarrhea virus NS3 serine proteinase: polyprotein cleavage sites, cofactor requirements, and molecular model of an enzyme essential for pestivirus replication. J Virol 71, 5312-5322.

Yamamoto, K., Fujii, R., Toyofuku, Y., Saito, T., Koseki, H., Hsu, V.W., Aoe, T., 2001. The KDEL receptor mediates a retrieval mechanism that contributes to quality control at the endoplasmic reticulum. Embo Journal 20, 3082-3091.

Zhang, Y.N., Liu, Y.Y., Xiao, F.C., Liu, C.C., Liang, X.D., Chen, J., Zhou, J., Baloch, A.S., Kan, L., Zhou, B., Qiu, H.J., 2018. Rab5, Rab7, and Rab11 Are Required for Caveola-Dependent Endocytosis of Classical Swine Fever Virus in Porcine Alveolar Macrophages. Journal of virology 92.

Zheng, G., Li, L.F., Zhang, Y., Qu, L., Wang, W., Li, M., Yu, S., Zhou, M., Luo, Y., Sun, Y., Munir, M., Li, S., Qiu, H.J., 2020. MERTK is a host factor that promotes classical swine fever virus entry and antagonizes innate immune response in PK-15 cells. Emerg Microbes Infect 9, 571-581.

Zheng, J., Yamada, Y., Fung, T.S., Huang, M., Chia, R., Liu, D.X., 2018. Identification of N-linked glycosylation sites in the spike protein and their functional impact on the replication and infectivity of coronavirus infectious bronchitis virus in cell culture. Virology 513, 65-74.

Appendix

Supplementary Material 1:

```
E1|KJ950914.1|Pestivirus_J_NrPV/NYC-D23              NKIYCEKFHVLGDLVYVNSCLPMGLPTGARFVSKNVISLEPEKTAQIIPRLTHHLDSGIL      60
E1|EF100713.2|Porcine_Bungowannah                    -SPYCPVAKRVFNIIYTNNCTPLGLPDKSKIIGPGTFDISGR-DEFIFPKLPYHVDDFIL      58
E1|AY781152.3|Pronghorn_antelope_pestivirus          SEVYCKVEKRVGSLWYTRNCTPACLPGHTEILGAGVFDTNPQ-GRSLIPRLPGHITEAVI      59
E1|FJ527854.1|BVDV_XJ-04                             ASPYCDVERKIGYIWYTKNCTPACLPRNTRIIGPGKFDTNAE-DGKILREMGGHLSEFVL      59
E1|LC006970.1|BVDV_2_KZ-91-CP                        ASPYCDVERKIGYIWYTKNCTPACLPRNTKIIGPGKFDTNAE-DGKILHEMGGHLSEFVL      59
E1|GQ888686.2|BVDV_2_JZ05-1                          ASPYCEVERKIGYIWYTKNCTPACLPRNTKIIGPGKFDTNAE-DGKILHEMGGHLSEFAL      59
E1|KX096718.1|BVDV_2_HB-1511                         ASPYCDVERKIGYIWYTKNCTPACLPRNTRIIGPGKFDTNAE-DGKILHEMGGHLSEFVL      59
E1|JF714967.1|BVDV_2_HLJ-10                          ASPYCNVERKIGYIWYTKNCTPACLPRNTRIIGPGKFDTNAE-DGKILHEMGGHLSEFVL      59
E1|KT875169.1|BVDV_2_91W                             ASPYCDVERKIGYIWYTKNCTPACLPRNTRIIGPGKFDTNAE-DGKILHEMGGHLSEFAL      59
E1|HQ258810.1|BVDV_SH-28                             ASPYCDVERKIGYIWYTKNCTPACLPRNTRIIGPGKFDTNAE-DGKILHEMGGHLSEFVL      59
E1|KT832818.1|BVDV_2_USMARC-60765                    ASPYCDVERKIGYIWYTKNCTPACLPKNTRIIGPGKFDTNAE-DGKILHEMGGHLSEFVL      59
E1|AB567658.1|BVDV_Hokudai-Lab/09                    ASPYCDVEKKIGYVWYTKNCTPACLPKNTKIIGPGKFDTNAE-DGKIIHEMGGHLSEFVL      59
E1|KJ000672.1|BVDV_2_SD1301                          ASPYCNVEKKIGYIWYTRNCTPACLPKNTKIIGPGKFDTNAE-DGKIIHEMGGHLSEFAL      59
E1|JQ799141.1|BVDV_1_sichuan                         ASPYCNVERKLGYVWYTKNCTPACLPKNTKIIGPGMFDTNAE-DGKILHEMGGHLSEFIL      59
E1|AB078950.1|BVDV_1_KS86-1ncp                       ASPYCEVKRKIGYIWYTQNCTPACLPKNTKIIGPGKFDTNAD-DGKILYEMGGHLSEVLL      59
E1|JN400273.1|BVDV_1_SD0803                          ASPYCAVERKIGYIWFTKNCTPACLPKNTRIVGPGKFDTNAD-DGKILHEMGGHLSEVLL      59
E1|KX987157.1|BVDV_1_SLO/1170/2000                   ASPYCDVERKIGYIWYTKNCTPACLPRNTKIVGPGKFDTNGD-DGKILHEMGGHLSEVLL      59
E1|LT631725.1|BVDV_1_UM/126/07                       ASPYCNVERKIGYIWFTKNCTPACLPKNTKIVGPGKFDTNAD-DGKILHEMGGHLSEVLL      59
E1|KF896608.1|BVDV_1_Bega-like                       ASPYCDVERKLGYIWFTKNCTPACLPKNTKIVGPGMFDTNGD-DDKILHEMGGHLSEVLL      59
E1|M96751.1|BVDV_1_SD1                               ASPYCDVDRKIGYIWFTKNCTPACLPKNTKIIGPGKFDTNAE-DGKILHEMGGHLSEVLL      59
E1|KC757383.1|BVDV_1_10JJ-SKR                        SSPYCDVERKLGYIWFTKNCTPACLPKNTKIVGPGKFDTNAE-DGKIMHEMGGHLSEVLL      59
E1|AF526381.3|BVDV_1_ZM-95                           ASPYCEVERKIGYIWYTKNCTPACLPKNTKIVGPGRFDTNAE-DGKILHEMGGHLSEVLL      59
E1|KP941591.1|BVDV_1_USMARC-55925                    ASPYCEVERRLGYVWYTKNCTPACLPRNTKIIGPGRFDTNAE-DGKILHEMGGHLSEVLL      59
E1|LC089876.1|BVDV_1_Shitara/02/06                   ASPYCDVDRKVGYIWYTKNCTPACLPKNTMIIGPGKFDTNAE-DGKILHEMGGHLSEMLL      59
E1|KC853441.1|BVDV_1_SuwaCp                          ATPYCEVEQKIGYIWITKNCTPACLPKNTKIVGPGKFDTNAE-DGKILHEMGGHLSEVLL      59
E1|KX577637.1|BVDV_1_SLO/2407/2006                   ATPYCEVDRKLGYIWYTKNCTPACLPGNTKIVGPGKFDTNAE-DGKIIHEMGGHVSEVLL      59
E1|KP313732.1|BVDV_1_Carlito                         ATPYCEVERKLGYVWYTKNCTPACLPGNTKIVGPGKFDTNAE-DGKILHEMGGHLSEVLL      59
E1|FJ040215.1|BVDV_3_Th/04_KhonKaen                  ATPYCNVSRKIGYVWYTNNCTPACLPMNTRIVGPGKFDTNGD-DGKILHEMGGHLSELAI      59
E1|KC297709.1|BVDV_3_LVRI/cont-1                     ATPYCNINKRIGYVWYTNNCTPACLPKNPRIVGPGKFDTNGD-DGKILHEMGSHLSELTI      59
E1|KC788748.1|BVDV_3_Italy-129/07                    ATPYCNINKKIGYVWYTNNCTPACLPKNTRIVGPGKFDTNGD-DGKILHEMGSHLSELAI      59
E1|JX469119.1|BVDV_3_JS12/01                         ATPYCNINKKIGYVWYTNNCTPACLPKNTRIMGPGKFDTNGE-DGKILHEMGSHLSELAI      59
E1|JX985409.1|BVDV_3_CH-KaHo/cont                    ATPYCNINKKIGYVWYTNNCTPACLPKNTRIVGPGKFDTNGD-DGKILHEMGSHLSELAI      59
E1|AB871953.1|BVDV_3_D32/00_'HoBi'like               ATPYCNINKKIGYVWYTNNCTPACLPKNTRIVGPGKFDTNGD-DGKILHEMGSHLSELAI      59
E1|HQ231763.1|BVDV_3_Italy-1/10-1                    ATPYCNINKKIGYVWYTNNCTPACLPKNTRIVGPGKFDTNGD-DGKILHEMGSHLSELAI      59
E1|JQ612704.1|BVDV_3_Italy-83/10-ncp                 ATPYCNINKKIGYVWYTNNCTPACLPKNTRIVGPGKFDTNGD-DGKILHEMGSHLSELAI      59
E1|NC_018713.1|BVDV_3_LV03/12                        ATPYCNINKKIGYVWYTNNCTPACLPKNTRIVGPGKFDTNGD-DGKILHEMGSHLSELAI      59
E1|KJ660072.1|Pestivirus_PG-2                        -SPYCEVNRKLGYIWYTNNCTPACLPGNTKIIGPGKFDTNAE-DGKILHELGGHISEFLL      58
E1|MH410816.1|Pestivirus_PG-2_GIRAFFE                -SPYCEVNRKLGYIWYTNNCTPACLPGNTKIIGPGKFDTNAE-DGKILHELGGHISEFLL      58
E1|AF144617.2|Pestivirus_giraffe-1_H138              ASPYCEVNKKLGYIWYTNNCTPACLPGNTKIIGPGKFDTNAE-DGKILHELGGHLSEFIL      59
E1|KJ463422.1|BDV_FNK2012-1                          QSPYCNVTRKIGYIWYTNNCTPACLPRNTKIIGPGKFDTNAE-DGKILHEMKGHISEFIL      59
E1|U70263.1|BDV_BD31                                 QSPYCNVTRKIGYIWYTNNCTPACLPKNTKIIGPGKFDTNAE-DGKILHEMKGHISEFIL      59
E1|KF925348.1|BDV_Coos_Bay-5_nc                      QSPYCNVTRKIGYIWYTNNCTPACLPKNTKIIGPGKFDTNAE-DGKILHEMKGHISEFIL      59
E1|AB897785.1|BDV_X818                               QSPYCNVTRKIGYVWYTNNCTPACLPKNTKIIGPGKFDTNAE-DGKILHEMRGHISEFIL      59
E1|AF144618.2|Pestivirus_reindeer-1_V60-Krefeld      QSPYCNVTKKLGYIWYTNNCTPACLPRNTKIVGPGKFDTNAE-DGKILHEMRGHISEFLL      59
E1|KC963426.1|BDV_JSLS12-01                          QSPYCNVTRKIGYIWYTNNCTPACLPGNTKIIGPGKFDTNAE-DGKILHEMRGHISEFLL      59
E1|AF407339.1|BDV_Aveyron                            QSPYCNVTKKLGYIWYTNNCTPACLPKNTKIIGPGKFDTNAE-DGKILHEMRGHISEFLL      59
E1|AY646427.1|CSFV_94.4/IL/94/TWN                    MSPYCNVTRKVGYIWYNNNCTPACLPKNTKIIGPGKFDTNAE-DGKILHEMGGHLSEFLL      59
E1|KP233070.1|CSFV_GXF29/2013                        MSPYCNVTRKIGYIWYTNNCTPACLPKNTKIIGPGKFDTNAE-DGKILHEMGGHLSEFLL      59
E1|KC851953.1|CSFV_IND/UK/LAL-290                    MSPYCNVTRKIGYIWYTNNCTPACLPKNTKIIGPGKFDTNAE-DGKILHEMGGHLSEFLL      59
E1|AF407339.1|CSFV_39                                MSPYRNVTRKIGYIWYTNNCTPACLPRNTKIIGPGKFDTNAE-DGKILHEMGGHLSEFLL      59
E1|FJ529205.1|CSFV_Zj0801                            MSPYCNVTRKIGYIWYTNNCTPACLPTNTKIIGPGKFDTNAE-DGKILHEMGGHLSEFLL      59
E1|KJ619377.1|CSFV_Bergen                            MSPYCNVTRKIGYIWYTNNCTPACLPKNTKIIGPGKFDTNAE-DGKILHEMGGHLSEFLL      59
E1|KU504339.1|CSFV_GD19/2011                         MSPYCNVTRKVGYIWYTNNCTPACLPKNTKIIGPGKFDTNAE-DGKILHEMGGHLSEFLL      59
E1|KM362426.1|CSFV_IND/AS/GHY/G4                     MSPYCNVTRKIGYIWYTNNCTPACLPKNTKIIGPGKFDTNAE-DGKILHEMGGHLSEFLL      59
E1|GQ923951.1|CSFV_SXCDK                             MSPYCNVTRKIGYIWYTNNCTPACLPKNTKIIGPGKFDTNAE-DGKILHEMGGHLSEFLL      59
E1|J04358.2|CSFV_Alfort/Tuebingen                    MSPYCNVTRKIGYIWYTNNCTPACLPKNTKIIGPGKFDTNAE-DGKILHEMGGHLSEFLL      59
E1|AY259122|CSFV_Alfort/Tuebingen                    MSPYCNVTRKIGYIWYTNNCTPACLPKNTKIIGPGKFDTNAE-DGKILHEMGGHLSEFLL      59
E1|KF669877.1|CSFV_JJ9811                            LSPYCNVTSKIGYIWYTNNCTPACLPKNTRIIGPGKFDTNAE-DGKILHEMGGHLSEFLL      59
E1|X87939.1|CSFV_Alfort/187                          LSPYCNVTSKIGYIWYTNNCTPACLPKNTKIIGPGKFDTNAE-DGKILHEMGGHLSEFLL      59
E1|X87939|CSFV_Alfort/187                            LSPYCNVTSKIGYIWYTNNCTPACLPKNTKIIGPGKFDTNAE-DGKILHEMGGHLSEFLL      59
E1|KJ660072.1|CSFV_Riems                             LSPYCNVTSKIGYIWYTNNCTPACLPKNTKIIGPGKFDTNAE-DGKILHEMGGHLSEFLL      59
E1|GU270877.1|BDV_H2121_Chamois-1                    QSPYCNVTKKIGYVWYTSNCTPACLPKNTKIIGPGKFDTNAE-DGKILHEMKGHVSEFLL      59
E1|KF918753.1|BDV_Gifhorn_genotype-3                 QSPYCNVTRKIGYIWYTNNCTPACLPGNTKIIGPGKFDTNAE-DGKILHEMRGHISEFLL      59
E1|JX428945.1|Pestivirus_Aydin/04-TR                 QSPYCNVTRRIGYIWYTNNCTPACLPRNTKIIGPGKFDTNAE-DGKILHEMGGHISEFLL      59
E1|AF037405.1|Pestivirus_Aydin/04-TR                 QSPYCNVTRRIGYIWYTNNCTPACLPRNTKIIGPGKFDTNAE-DGKILHEMGGHISEFLL      59
E1|KM408491.1|Pestivirus_Burdur/05-TR                QSPYCNVTRKIGYIWYTNNCTPACLPKNTKIIGPGKFDTNAE-DGKILHEMGGHISEFLL      59
                                                        *      :  :    .* *  **    ::. . :. .      :: .:  *: .  :

E1|KJ950914.1|Pestivirus_J_NrPV/NYC-D23              LVLVAMSDFMPETSSALYLILHFMIPNSRHRTISEEGLTMALNLTSTEPVSSVIPTSVYV      126
E1|EF100713.2|Porcine_Bungowannah                    LSLIAMSDFAPETSSIIYLALHYLMPSNDNRDFVMDLDPNKLNLTATKSVASVVPTSVNV      118
E1|AY781152.3|Pronghorn_antelope_pestivirus          LSLVALSEVMPETSSALYIALHYFLHPMN--ETIGYCDKNQLNLTITTTVDKVIPNSVYV      117
E1|FJ527854.1|BVDV_XJ-04                             LSLVVLSDFAPETASAIYLVLHFAIPQSH--ISVDTCDKNQLNITVAITVAEVIPGSVWN      117
E1|LC006970.1|BVDV_2_KZ-91-CP                        LSLVVLSDFAPETASVIYLVLHFTIPQSY--VSVDTCDKNQLNLTVATTVAEVIPGSVWN      117
E1|GQ888686.2|BVDV_2_JZ05-1                          LSLVVLSDFAPETASVIYLVLHFAIPQSH--INVDTCDKSQLNLTVATTVAEVIPGTVWN      117
E1|KX096718.1|BVDV_2_HB-1511                         LSLVVLSDFAPETSSAIYLVLHFAIPQSH--VNVDTCDKNQLNLTVATTVAEVIPGSVWN      117
```

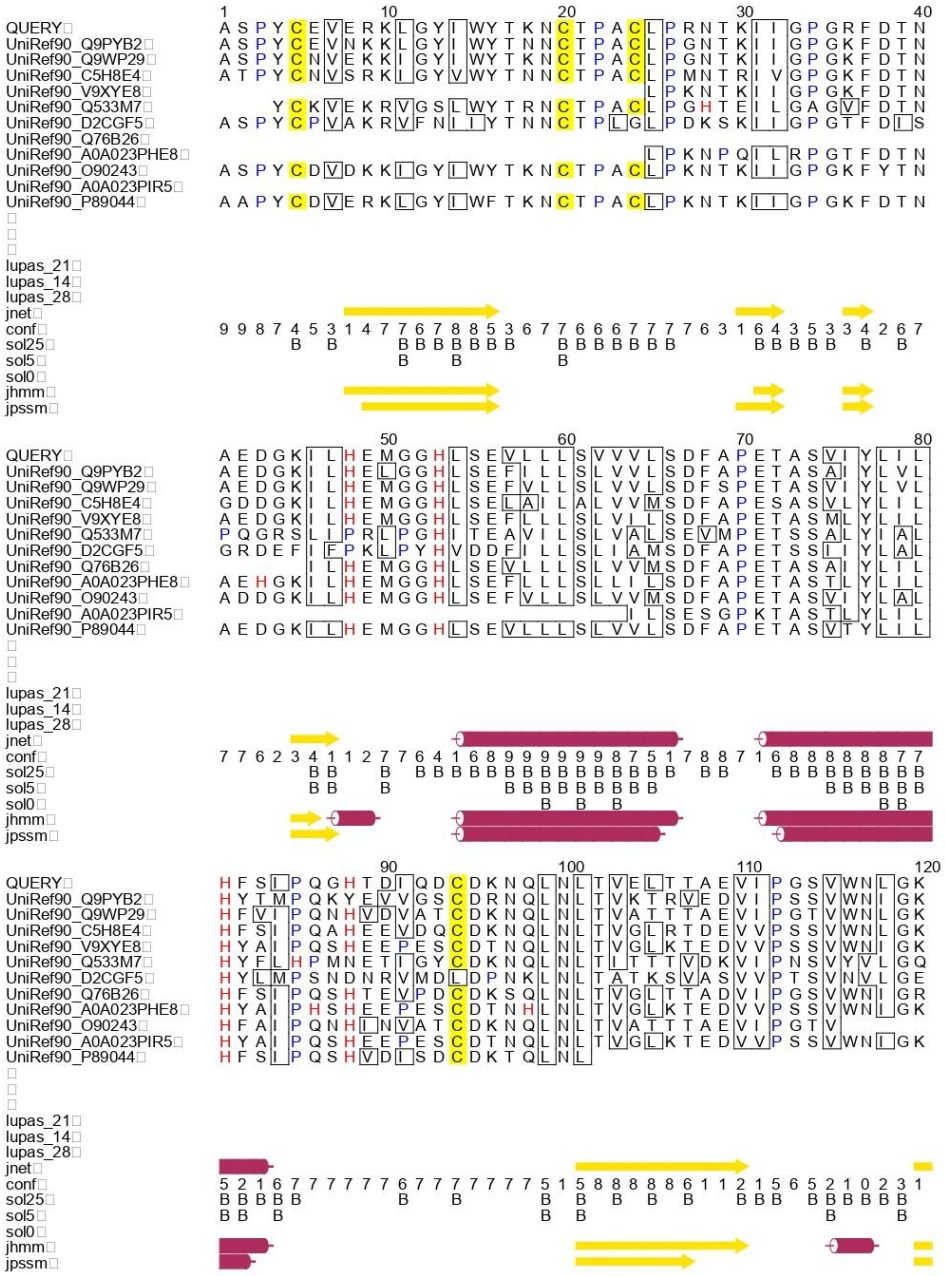
```
E1|JF714967.1|BVDV_2_HLJ-10                          LSLVVLSDFAPETASVIYLVLHFAIPQSH--VSVDTCDKNQLNLTVATTVAEVIPGTVWN      117
E1|KT875169.1|BVDV_2_91W                             LSLVVLSDFAPETASAIYLVLHFVIPQSH--VSVDTCDKNQLNLTVATTVAEVIPGTVWN      117
E1|HQ258810.1|BVDV_SH-28                             LSLVVLSDFAPETASVIYLVLHFAIPQNH--VNVDTCDKNQLNLTVATTVAEVIPGSVWN      117
E1|KT832818.1|BVDV_2_USMARC-60765                    LSLVVLSDFAPETASVIYLVLHFAIPQNH--INVDTCDKNQLNLTVATTVAEVIPGSVWN      117
E1|AB567658.1|BVDV_Hokudai-Lab/09                    LSLVVLSDFAPESASVIYLVLHYAIPQNH--AEVTTCDKNQLNLTVTTTTADVIPGSVWN      117
E1|KJ000672.1|BVDV_2_SD1301                          LSLVVLSDFSPETASVIYLVLHYVIPQNH--VEVTTCDKNQLNLTIATTTADVIPGSVWN      117
E1|JQ799141.1|BVDV_1_sichuan                         LSLVVLSDFAPETASALYLVLHFTIPQSH--TPVMDCDKSQLNLTVGLTTADVVPSSVWN      117
E1|AB078950.1|BVDV_1_KS86-1ncp                       LSLVVMSDFAPETASAIYLILHFSIPQSH--TEVSDCDKSQLNLTVGLTTADVIPGSVWN      117
E1|JN400273.1|BVDV_1_SD0803                          MSLVVLSDFAPETASVVYLILHFSIPQGH--TDVLDCDKGQLNLTVGLRTEDVIPGSVWN      117
E1|KX987157.1|BVDV_1_SLO/1170/2000                   LSLVVLSDFAPETASAVYLILHFSIPQSH--SDILDCDKNQLNLTVELTTEQVIPGSVWN      117
E1|LT631725.1|BVDV_1_UM/126/07                       LSLVVLSDFAPETASVMYIILHFFIPQSH--IDVSDCDKNQLNLTIGLKTDDVVPNSVWN      117
E1|KF896608.1|BVDV_1_Bega-like                       LSLVVLSDFAPETASAMYLILHFSIPQSH--TDITDCDKNQLNLTIGLTTADMVPGSVWN      117
E1|M96751.1|BVDV_1_SD1                               LSLVVLSDFAPETASAMYLILHFSIPQSH--VDITDCDKTQLNLTIELTTADVIPGSVWN      117
E1|KC757383.1|BVDV_1_10JJ-SKR                        LSLVVLSDFAPETASTLYLVLHFSIPQRH--TDILDCDKSQLNLTVGVTTADVIPGSVWN      117
E1|AF526381.3|BVDV_1_ZM-95                           LSLVVLSDFAPETASVAYLALHFGIPQRH--TEVLDCDKNQLNLTVGVTTEDVIPGSVWN      117
E1|KP941591.1|BVDV_1_USMARC-55925                    LSVVVLSDFAPETASLIYLILHFSIPQGH--TDIHDCDRNQLNLTVELTTAEVIPGSVWN      117
E1|LC089876.1|BVDV_1_Shitara/02/06                   LSLVVMSDFAPETASVIYLILHFSIPQGH--TEVLDCDKNQLNLTVGLTTADVIPGSVWN      117
E1|KC853441.1|BVDV_1_SuwaCp                          LSLVVLSDFAPETASVIYLVLHFTVPQSH--TDVLDCDKNQLNLTIGVTTADVIPGSVWN      117
E1|KX577637.1|BVDV_1_SLO/2407/2006                   LSLVVLSDFAPETASVVYLILHFSIPQRH--TDILDCDKNQLNLTVELTTADVIPGSVWN      117
E1|KP313732.1|BVDV_1_Carlito                         LSLVVLSDFAPETASVMYLVLHFSIPQSH--TDILDCDKDQLNLTVKLTTADVIPGSVWN      117
E1|FJ040215.1|BVDV_3_Th/04_KhonKaen                  LALVVMSDFAPESASVLYLILHFSIPQAH--EEVDQCDKNQLNLTVGLRTDEVVPSSVWN      117
E1|KC297709.1|BVDV_3_LVRI/cont-1                     LALVVMSDFAPESASVLYLILHFSIPQAH--EEVDQCDRNQLNLTVSLRTDEVVPSSVWN      117
E1|KC788748.1|BVDV_3_Italy-129/07                    LALVVMSDFTPESTSVLYLILHFSIPQAH--EEVDQCDRNQLNLTVSLRTDEVVPSSVWN      117
E1|JX469119.1|BVDV_3_JS12/01                         LALVVMSDFAPESASVLYLILHFSIPQAH--EEVDQCDRNQLNLTVSLRTDEVVPSSVWN      117
E1|JX985409.1|BVDV_3_CH-KaHo/cont                    LALVVMSDFAPESASVLYLILHFSIPQAH--EEVDQCDRNQLNLTVSLRTDEVVPSSVWN      117
E1|AB871953.1|BVDV_3_D32/00_'HoBi'like               LALVVMSDFAPESASVLYLILHFSIPQAH--EEVDQCDRNQLNLTVGLRTDEVVPSSVWN      117
E1|HQ231763.1|BVDV_3_Italy-1/10-1                    LALVVMSDFAPESASVLYLILHFSIPQAH--EEVDQCDRNQLNLTVSLRTDEVVPSSVWN      117
E1|JQ612704.1|BVDV_3_Italy-83/10-ncp                 LALVVMSDFAPESASVLYLILHFSIPQAH--EEVDQCDRNQLNLTVSLRTDEVVPSSVWN      117
E1|NC_018713.1|BVDV_3_LV03/12                        LALVVMSDFAPESASVLYLILHFSIPQAH--EEVDQCDRNQLNLTVSLRTDEVVPSSVWN      117
E1|KJ660072.1|Pestivirus_PG-2                        LSLVVLSDFAPETASAIYLVLHYTIPQKH--EVVENCDMNQLNLTVTKRVKDVIPSSVWN      116
E1|MH410816.1|Pestivirus_PG-2_GIRAFFE                LSLVVLSDFAPETASAIYLVLHYTIPQKH--EVVENCDMNQLNLTVTKRVKDVIPSSVWN      116
E1|AF144617.2|Pestivirus_giraffe-1_H138              LSLVVLSDFAPETASAIYLVLHYTMPQKY--EVVGSCDRNQLNLTVKTRVEDVIPSSVWN      117
E1|KJ463422.1|BDV_FNK2012-1                          LSLVVLSDFAPETASTLYLVLHFTLPQTY--EVPSGCDTNQLNLTVDLRVDDVIPSSVWN      117
E1|U70263.1|BDV_BD31                                 LSLVVLSDFAPETASTLYLVLHFTLPQTY--EVPSECDTNQLNLTVGLRVDDVIPSSVWN      117
E1|KF925348.1|BDV_Coos_Bay-5_nc                      LSLVVLSDFAPETASTLYLVLHFTLPQTY--EVPNECDTSQLNLTVGLRVDDVIPSSVWN      117
E1|AB897785.1|BDV_X818                               LSLVVLSDFAPETASTLYLVLHFALPQTH--EVPSVCDTNQLNLTVSLRVDDVIPSSVWN      117
E1|AF144618.2|Pestivirus_reindeer-1_V60-Krefeld      LSLVVLSDFAPETASTLYLILHFTIPQSH--EAPSECDTNQLNLTIGLRVDEVVPSSVWN      117
E1|KC963426.1|BDV_JSLS12-01                          LSLVILSDFAPETASSLYLILHFAIPQSH--EIPDGCDTNQLNLTVNLRVEDVIPSSVWN      117
E1|AF407339.1|BDV_Aveyron                            LSLVVLSDFAPETASALYLLLHFAIPQSH--ENPSSCDTNQLNLTTGLKVEDVIPSSVWN      117
E1|AY646427.1|CSFV_94.4/IL/94/TWN                    LSLVVLSDFAPETASALYLIFHYTIPQIH--EVPEDCDTNQLNLTVELRTEDVVPSSVWN      117
E1|KP233070.1|CSFV_GXF29/2013                        LSLVILSDFAPETASTLYLILHYAVPQSY--GEPEGCDTNQLNLTVGLRTQDVVPSSVWN      117
E1|KC851953.1|CSFV_IND/UK/LAL-290                    LSLVILSDFAPETASTLYLILHYAIPQSH--EEPESCDTNQLNLTVGLKTEDVVPSSVWN      117
E1|AF407339.1|CSFV_39                                LSLVILSDFAPETASTLYLILHYAIPQSH--EEPEGCDTNQLNLTVGLKTEDVVPSSVWN      117
E1|FJ529205.1|CSFV_Zj0801                            LSLVILSDFAPETASTLYLILHYAIPQSH--EEPEGCDTNQLNLTVGLRTEDVVPSSVWN      117
E1|KJ619377.1|CSFV_Bergen                            LSLVILSDFAPETASTLYLILHYAIPQSH--DEPEGCDTNQLNLTVGLRTEDVVPSSVWN      117
E1|KU504339.1|CSFV_GD19/2011                         LSLVILSDFAPETASTLYLILHYAIPQSH--EEPEGCDTNQLNLTVGLRTEDVVPSSVWN      117
E1|KM362426.1|CSFV_IND/AS/GHY/G4                     LSLVILSDFAPETASTLYLILHYAIPQPH--KEPEGCDTNQLNLTVGLRTEDVVPSSVWN      117
E1|GQ923951.1|CSFV_SXCDK                             LSLVILSDFAPETASTLYLILHYAIPQSH--EEPEGCDTNQLNLTVGLRTEDVIPSSVWN      117
E1|J04358.2|CSFV_Alfort/Tuebingen                    LSLVILSDFAPETASTLYLILHYAIPQSH--EEPEGCDTNQLNLTVKLRTEDVVPSSVWN      117
E1|AY259122|CSFV_Alfort/Tuebingen                    LSLVILSDFAPETASTLYLILHYAIPQSH--EEPEGCDTNQLNLTVKLRTEDVVPSSVWN      117
E1|KF669877.1|CSFV_JJ9811                            LSLVVLSDFAPETASTLYLILHYAIPQVH--KEPDGCDTNQLNLTVELRTEDVVPSSVWN      117
E1|X87939.1|CSFV_Alfort/187                          LSLVVLSDFAPETASALYLILHYVIPQSH--EEPEGCDTNQLNLTVELRTEDVIPSSVWN      117
E1|X87939|CSFV_Alfort/187                            LSLVVLSDFAPETASALYLILHYVIPQSH--EEPEGCDTNQLNLTVELRTEDVIPSSVWN      117
E1|KJ660072.1|CSFV_Riems                             LSLVVLSDFAPETASALYLIFHYVIPQPH--DEPEGCDTNQLNLTVELRTEDVIPSSVWN      117
E1|GU270877.1|BDV_H2121_Chamois-1                    LSLVILSDFAPETASTLYLILHFTIPQSY--ESPDDCDTNQLNLTVSLKVEDVIPSSVWN      117
E1|KF918753.1|BDV_Gifhorn_genotype-3                 LSLVVLSDFAPETASVLYLILHFSIPQGY--ESPDQCDMNQLNLTVDLKVEDVIPSSVWN      117
E1|JX428945.1|Pestivirus_Aydin/04-TR                 LSLVVLSDFAPETASAIYLVLHFTIPQSY--ENPKDCDKNQLNLTIGLRTEDVVPSSVWN      117
E1|AF037405.1|Pestivirus_Aydin/04-TR                 LSLVVLSDFAPETASAIYLVLHFTIPQSY--ENPKDCDKNQLNLTIGLRTEDVVPSSVWN      117
E1|KM408491.1|Pestivirus_Burdur/05-TR                LSLVVLSDFAPETASALYLILHFTIPQSY--ENPADCDKNQLNLTVGLKTEDVVPSSVWN      117
                                                     : :: :*:. **::*  *: :*: :               **:*    . .::* :*


E1|KJ950914.1|Pestivirus_J_NrPV/NYC-D23              EGQWTCWKPSWWPYNADIALFFEGAFEMLELIARAVGDLMKVWTEATAVAFLCFLIKAFR      180
E1|EF100713.2|Porcine_Bungowannah                    LGEWVCVKPSWWPYSAEITNLIGGVITVADLVIKTIEELLNLWTEATAVAFLAALIKIFR      178
E1|AY781152.3|Pronghorn_antelope_pestivirus          LGQWVCVKPGWWPYDSEVTLVVNEVINVLDIGGRAARVLLQVWDAATAIAVLIFIMKVAR      177
E1|FJ527854.1|BVDV_XJ-04                             LGKYVCIRPNWWPYETTAVFVLEEAGQVVKLGLRAIRDLTRIWNAATTTAFLVFLVKVLR      177
E1|LC006970.1|BVDV_2_KZ-91-CP                        LGKYVCIRPDWWPYETATVFVLEEAGQVVKLGLRAIRDLTRIWNAATTTAFLVFLVKALR      177
E1|GQ888686.2|BVDV_2_JZ05-1                          LGKYVCIRPDWWPYETTTVFVLEEAGQVIKLGLRAIRDLTRIWNAATTTAFLVFLVKVLR      177
E1|KX096718.1|BVDV_2_HB-1511                         LGKYVCIRPDWWPYETTTVLVLEEAGQVVKLGLRAIRDLTRIWNAATTTAFLVFLVKVLR      177
E1|JF714967.1|BVDV_2_HLJ-10                          LGKYVCIRPDWWPYETTTVFVLEEAGQVVKLGLRAIRDLTRIWNAATTTAFLVFLVKVLR      177
E1|KT875169.1|BVDV_2_91W                             LGKYVCIRPDWWPYETTTVFVLEEAGQVVKLGLRAIRDLTRIWNAATTTAFLVFLVKVLR      177
E1|HQ258810.1|BVDV_SH-28                             LGKYVCIRPDWWPYETTTVFVLEEAGQVVKLGLRAIRDLTRIWNAATTIAFLVFLVGALR      177
E1|KT832818.1|BVDV_2_USMARC-60765                    LGRYVCIRPDWWPYETTTVFVLEEAGQVVKLGLRAIRDLTRIWNAATTTAFLVFLVKVLR      177
E1|AB567658.1|BVDV_Hokudai-Lab/09                    LGKYVCIRPDWWPYETATVFALEEAGQVVKLGLRAIRDLTRIWNAATTTAFLVFLVKVLR      177
E1|KJ000672.1|BVDV_2_SD1301                          LGKYVCIRPDWWPYETTVVFALEEAGQVVKLGMRAIRDLTRIWNAATTIAFLVFLVKVLR      177
E1|JQ799141.1|BVDV_1_sichuan                         LGKWVCVRPDWWPYETATVLLFEEAGQVVKIALRAVRDITRIWNAATTTAFLVCLVRVAR      177
E1|AB078950.1|BVDV_1_KS86-1ncp                       IGRYVCIRPDWWPYETTTVLVVEEVSQVVKLVLRAIRDLTRIWDAATSTAFLVCLVKVVR      177
E1|JN400273.1|BVDV_1_SD0803                          LGRYVCIRPNWWPYETTTVLAFEEIGQVIKIALRALRDLTRIWNAATATAFLVCLVKVVR      177
E1|KX987157.1|BVDV_1_SLO/1170/2000                   LGRYVCIRPNWWPYETAPVLALEEIGQVINIVLRALRDLTRIWNAATTTAFLVCLVKVVR      177
E1|LT631725.1|BVDV_1_UM/126/07                       MGKYVCIRPNWWPYETTAVLAFEEVSQVVKIVLRALRDLTRIWNAATTTAFLVCLVKVVR      177
E1|KF896608.1|BVDV_1_Bega-like                       LGKYVCIRPNWWPYETVTVLAFEEVGQVIKIVLRALRDLTRIWNAATTTAFLVCLVKVVR      177
E1|M96751.1|BVDV_1_SD1                               LGKYVCIRPDWWPYETAAVLAFEEVGQVVKIVLRALRDLTRIWNAATTTAFLVCLIKMVR      177
```

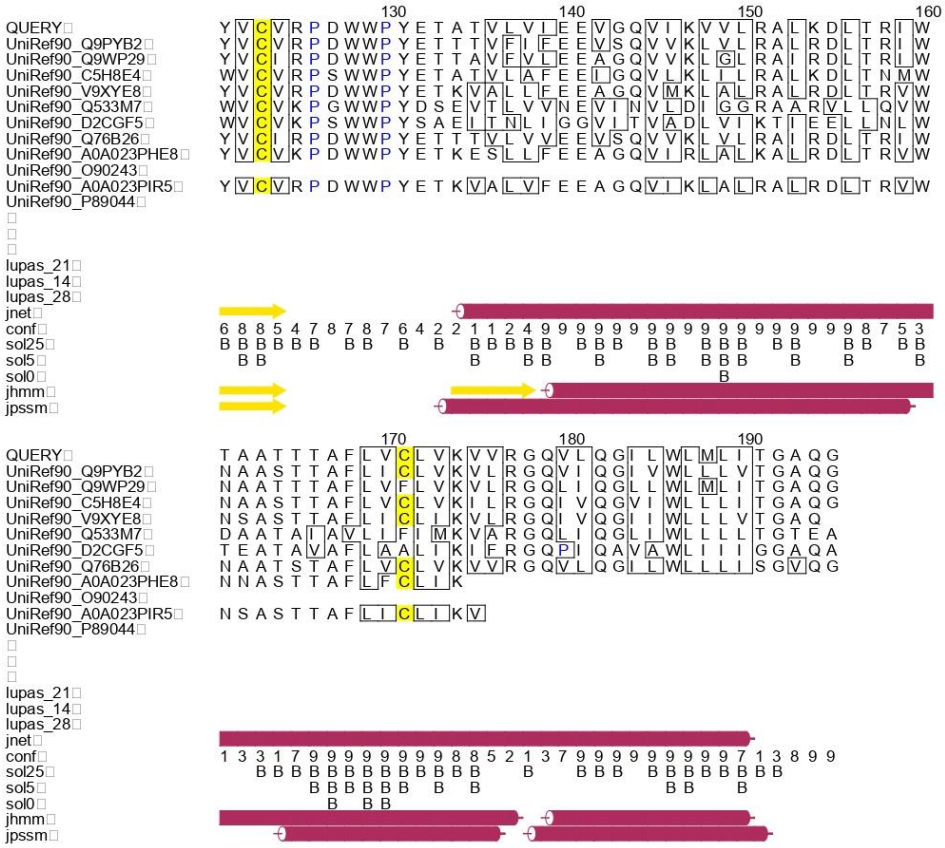
```
E1|KC757383.1|BVDV_1_10JJ-SKR                    LGKWVCIRPDWWPYETAPVLAFEEAGQVVRIALRALRDLTRIWNAATTTAFLVCLVKVVR  177
E1|AF526381.3|BVDV_1_ZM-95                       LGRYVCIRPDWWPYETATVLVFEEISQVIKLVLRALRDLMRIWNAATTTAFLVCLVKVVR  177
E1|KP941591.1|BVDV_1_USMARC-55925                LGKYVCIRPDWWPYETATVLVFEEVGQVIKIVLRALRDLTRIWTAATTTAFLVCLVKVVR  177
E1|LC089876.1|BVDV_1_Shitara/02/06               LGRYVCIRPDWWPYETATVMAFEEVGQVIKIALRALKDLTRIWNAATTTAFLVCLVKVVR  177
E1|KC853441.1|BVDV_1_SuwaCp                      LGKYVCIRPDWWPYETAAVLVFEEVGQVIKIALRALRDLTRIWNAATTTAFLVCLVKVVR  177
E1|KX577637.1|BVDV_1_SLO/2407/2006               LGKYVCIRPDWWPYETATVLAFEEIGQVIKIALRALRDLTRIWNAATTTAFLVCMVKVVR  177
E1|KP313732.1|BVDV_1_Carlito                     LGKYVCIRPDWWPYETATVLAFEEVGQVIKIALRALRDLTRIWNAATTTAFLVCLVKVVR  177
E1|FJ040215.1|BVDV_3_Th/04_KhonKaen              LGKWVCVRPSWWPYETATVLAFEEIGQVLKLILRALKDLTNMWNAASTTAFLVCLVKILR  177
E1|KC297709.1|BVDV_3_LVRI/cont-1                 LGKWVCVRPPWWPYETATVLAFEEIGQVLKLILRALKDLTNMWNAASTTAFLVCLVKILR  177
E1|KC788748.1|BVDV_3_Italy-129/07                LGKWVCVRPPWWPYETATVLAFEEIGQVLKLILRALKDLTNMWNAASTTAFLVCLVKILR  177
E1|JX469119.1|BVDV_3_JS12/01                     LGKWVCVRPSWWPYETATVLAFEEIGQVLKLILRALKDLTNMWNAASTTAFLVCLVKILR  177
E1|JX985409.1|BVDV_3_CH-KaHo/cont                LGKWVCVRPSWWPYETATVLAFEEIGQVLKLILRALKDLTNMWNAASTTAFLVCLVKILR  177
E1|AB871953.1|BVDV_3_D32/00_'HoBi'like           LGKWVCVRPPWWPYETATVLAFEEIGQVLKLILRALKDLTNMWNAASTTAFLVCLVKILR  177
E1|HQ231763.1|BVDV_3_Italy-1/10-1                LGKWVCVRPPWWPYETATVLAFEEIGQVLKLILRALKDLTNMWNAASTTAFLVCLVKILR  177
E1|JQ612704.1|BVDV_3_Italy-83/10-ncp             LGKWVCVRPPWWPYETATVLAFEEIGQVLKLILRALKDLTNMWNAASTTAFLVCLVKILR  177
E1|NC_018713.1|BVDV_3_LV03/12                    LGKWVCVRPPWWPYETATVLAFEEIGQVLKLILRALKDLTNMWNAASTTAFLVCLVKILR  177
E1|KJ660072.1|Pestivirus_PG-2                    IGKYVCVRPDWWPYETTTVLIFEEISHVVKLVLRALRDLTRIWNAASTTAFLICLVKVLR  176
E1|MH410816.1|Pestivirus_PG-2_GIRAFFE            IGKYVCVRPDWWPYETTTVLIFEEISHVVKLVLRALRDLTRIWNAASTTAFLICLVKVLR  176
E1|AF144617.2|Pestivirus_giraffe-1_H138          IGKYVCVRPDWWPYETTTVFIFEEVSQVVKLVLRALRDLTRIWNAASTTAFLICLVKVLR  177
E1|KJ463422.1|BDV_FNK2012-1                      LGKYVCVRPDWWPYETTIVLLFEEAGQVVKLVLRAIRDLTRIWNSASTTAFLICLIKVLR  177
E1|U70263.1|BDV_BD31                             LGKYVCVRPDWWPYETTIVLLFEEAEQVVKLVLRAIRDLTRVWNSASTTAFLICLAKVLR  177
E1|KF925348.1|BDV_Coos_Bay-5_nc                  LGKYVCVRPDWWPYETTIVLLFEEAEQVVKLVLRAIRDLTRVWNSASTTAFLICLVKVLR  177
E1|AB897785.1|BDV_X818                           LGKYVCVRPDWWPYETTMVLLFEEAGQVVKLVLRAIRDLTRVWNSASTTAFLICLVKVLR  177
E1|AF144618.2|Pestivirus_reindeer-1_V60-Krefeld  LGKYVCVRPDWWPYETTVVLLCEEAGQVIKLVLRAIRDLTRVWNSASTTAFLICLIKVLR  177
E1|KC963426.1|BDV_JSLS12-01                      LGKYVCVRPDWWPYETTAVLLFEEAGQIVKLILRAIRDLTRVWNSASTTAFLICLVKVLR  177
E1|AF407339.1|BDV_Aveyron                        LGKYVCVRPDWWPYETTVVLLFEEAGQVVKLILRAIRDLTRVWNSASTTAFLICLVKVLR  177
E1|AY646427.1|CSFV_94.4/IL/94/TWN                VGEYVCVRPDWWPYETKVVLLFEEAGQVIKLALRAMRDLTRVWSSASTIAFLICLIKVLR  177
E1|KP233070.1|CSFV_GXF29/2013                    IGKYVCVRPDWWPYETKVALLFEEAGQVIKLALRALRDLTRVWNSASTTAFLICLIKVLR  177
E1|KC851953.1|CSFV_IND/UK/LAL-290                IGKYVCVRPDWWPYETKVALLFEEVGQVIKLALRALRDLTRVWNSASTTAFLICLIKVLR  177
E1|AF407339.1|CSFV_39                            IGKYVCVRPDWWPYETKVALLFEEAGQVIKLALRALRDLTRVWNSASTTAFLICLIKILR  177
E1|FJ529205.1|CSFV_Zj0801                        IGKYVCVRPDWWPYETKVALLFEEAGQVVKLALRALRDLTRVWNSASTTAFLICLIKILR  177
E1|KJ619377.1|CSFV_Bergen                        IGKYVCVRPDWWPYETKVALLFEEAGQVIKLALRALRDLTRVWNSASTTAFLICLIKTLR  177
E1|KU504339.1|CSFV_GD19/2011                     IGKYVCVRPDWWPYETKVALLFEEAGQVIKLALRALRDLTRVWNSASTTAFLICLIKILR  177
E1|KM362426.1|CSFV_IND/AS/GHY/G4                 IGKYVCVRPDWWPYETKVALLFEEAGQVIKLVLRALRDLTRVWNSASTTAFLICLIKVLR  177
E1|GQ923951.1|CSFV_SXCDK                         IGKYVCVRPDWWPYETKVALLFEEAGQVVKLVLRALRDLTRVWNSASTTAFLICLIKVLR  177
E1|J04358.2|CSFV_Alfort/Tuebingen                IGKYVCVRPDWWPYETKVALLFEEAGQVIKLVLRALRDLTRVWNSASTTAFLICLIKVLR  177
E1|AY259122|CSFV_Alfort/Tuebingen                IGKYVCVRPDWWPYETKVALLFEEAGQVIKLVLRALRDLTRVWNSASTTAFLICLIKVLR  177
E1|KF669877.1|CSFV_JJ9811                        VGKYVCVRPDWWPYETKVALLFEEAGQVIKLALRALRDLTRVWNSASTTAFLICLIKVLR  177
E1|X87939.1|CSFV_Alfort/187                      VGKYVCVRPDWWPYETKVALLFEEAGQVVKLALRALRDLTRVWNSASTTAFLICLIKVLR  177
E1|X87939|CSFV_Alfort/187                        VGKYVCVRPDWWPYETKVALLFEEAGQVVKLALRALRDLTRVWNSASTTAFLICLIKVLR  177
E1|KJ660072.1|CSFV_Riems                         VGKYVCVRPDWWPYETEVALLFEEVGQVVKLALRALRDLTRVWNSASTIAFLICLIKVLR  177
E1|GU270877.1|BDV_H2121_Chamois-1                LEKYVCIRPYWWPYETTVALLFEEAGQVLKLALRAIRDLTRVWNSASTTAFLICLVKLLR  177
E1|KF918753.1|BDV_Gifhorn_genotype-3             LGKYVCIRPDWWPYETTVVLLFEEAGQVIKLALRAIRDLTRVWSSASTTAFLICLVKVLR  177
E1|JX428945.1|Pestivirus_Aydin/04-TR             IGKYVCIRPDWWPYETTVVLLFEEVGQVIKLALRALRDLTRVWNSASTTAFLICLIKVLR  177
E1|AF037405.1|Pestivirus_Aydin/04-TR             IGKYVCIRPDWWPYETTVVLLFEEVGQVIKLALRALRDLTRVWNSASTTAFLICLIKVLR  177
E1|KM408491.1|Pestivirus_Burdur/05-TR            LGKYVCIRPDWWPYETTVVLLFEEVGQVIKLALRALRDLTRVWNSASTTAFLICLIKVLR  177
                                                   .:.* :* ****.:  .       :  :  ::   : .:*  *:: *.*  :    *

E1|KJ950914.1|Pestivirus_J_NrPV/NYC-D23          GQILQGVILLLLLSSAEG 198
E1|EF100713.2|Porcine_Bungowannah                GQPIQAVAWLIIIGGAQA 196
E1|AY781152.3|Pronghorn_antelope_pestivirus      GQLIQGLIWLLLLTGTEA 195
E1|FJ527854.1|BVDV_XJ-04                         GQLIQGLLWLMLITGAQS 195
E1|LC006970.1|BVDV_2_KZ-91-CP                    GQLIQGLLWLMLITGAQG 195
E1|GQ888686.2|BVDV_2_JZ05-1                      GQLIQGLLWLMLITGAQG 195
E1|KX096718.1|BVDV_2_HB-1511                     GQLIQGLLWLMLITGAQS 195
E1|JF714967.1|BVDV_2_HLJ-10                      GQLIQGLLWLMLITGAQS 195
E1|KT875169.1|BVDV_2_91W                         GQLIQGLLWLMLITGAQS 195
E1|HQ258810.1|BVDV_SH-28                         GQLIQGLLWLMLITGAQG 195
E1|KT832818.1|BVDV_2_USMARC-60765                GQLIQGLLWLMLITGAQS 195
E1|AB567658.1|BVDV_Hokudai-Lab/09                GQLIQGLLWLMLITGAQG 195
E1|KJ000672.1|BVDV_2_SD1301                      GQLIQGLLWLMLITGVQS 195
E1|JQ799141.1|BVDV_1_sichuan                     GQVLQGLIWLLLITGAQG 195
E1|AB078950.1|BVDV_1_KS86-1ncp                   GQVLQGILWLLLISGVQG 195
E1|JN400273.1|BVDV_1_SD0803                      GQVLQGILWLLLITGVQG 195
E1|KX987157.1|BVDV_1_SLO/1170/2000               GQVLQGILWLLLITGVQG 195
E1|LT631725.1|BVDV_1_UM/126/07                   GQVLQGILWLLLITGVQG 195
E1|KF896608.1|BVDV_1_Bega-like                   GQVLQGILWLLLITGVQG 195
E1|M96751.1|BVDV_1_SD1                           GQVVQGILWLLLITGVQG 195
E1|KC757383.1|BVDV_1_10JJ-SKR                    GQVVQGVIWLLLITGVQG 195
E1|AF526381.3|BVDV_1_ZM-95                       GQVLQGVLWLLLITGAQG 195
E1|KP941591.1|BVDV_1_USMARC-55925                GQVLQGILWLILITGAQG 195
E1|LC089876.1|BVDV_1_Shitara/02/06               GQVLQGILWLLLITGVQG 195
E1|KC853441.1|BVDV_1_SuwaCp                      GQILQGILWLLLITGVQG 195
E1|KX577637.1|BVDV_1_SLO/2407/2006               GQVLQGVLWLLLITGVQG 195
E1|KP313732.1|BVDV_1_Carlito                     GQMLQGILWLLLITGVQG 195
E1|FJ040215.1|BVDV_3_Th/04_KhonKaen              GQIVQGVIWLLLITGAQG 195
E1|KC297709.1|BVDV_3_LVRI/cont-1                 GQIVQGVIWLLLITGAQG 195
E1|KC788748.1|BVDV_3_Italy-129/07                GQIVQGVIWLLLITGAQG 195
E1|JX469119.1|BVDV_3_JS12/01                     GQIVQGVIWLLLITGAQG 195
E1|JX985409.1|BVDV_3_CH-KaHo/cont                GQIVQGVIWLLLITGAQG 195
E1|AB871953.1|BVDV_3_D32/00_'HoBi'like           GQIVQGVIWLLLITGAQG 195
```

S1: Multiple sequences alignment of pestiviral E1 glycoprotein

Multiple sequences alignment (68 sequences of E1 throughout the whole species of pestiviruses were used in this analysis) was carried out with Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/). An * (asterisk) represents position which contain 100% conserved residue, a : (colon) indicates conservation between amino acids with highly similar properties and a . (dot) with low similar characteristics.

Supplementary Material 2:





S2: Multiple sequences alignment of pestiviral E1 glycoprotein

Illustration of a JPred4 secondary structure prediction of pestiviral glycoprotein E1. The E1 aa sequence of BVDV CP7 serves as a query sequence. JPred provides several annotation tracks for visualization. The final JNet prediction followed by a confidence score for the prediction (0-9, from least to highest confidence). Red cylindrical represents these regions form α-helices with high probabilities, yellow arrows indicated β-strand formation. Cysteines that are highly conserved are marked in yellow.

Publications

Mu, Y., Bintintan, I., & Meyers, G. (2020). Downstream Sequences Control the Processing of the Pestivirus E^{rns}-E1 Precursor. ***Journal of virology***, 95(1), e01905-20. https://doi.org/10.1128/JVI.01905-20

Yu Mu, Christina Radtke, Birke Andrea Tews, and Gregor Meyers. 2021. Characterization of membrane topology and retention signal of pestiviral glycoprotein E1. ***Journal of virology*** (Accepted)

Yu Mu, Christina Radtke, Birke Andrea Tews, and Gregor Meyers. 2021. Contribution of cysteine residues for the formation of E1-E2 heterodimer from pestiviruses. Submitted to ***Journal of virology***. (Under revision)

Acknowledgements

Looking backwards, four years studying life in Germany is like a sudden meteor that vanishes in a second. It feels like that I just came to Greifswald yesterday, how time flies! At the time that my studying journey is coming to an end, I am full of emotional thoughts. Along this way, I am so grateful to the people who offered so much help, support and encouragement to me.

My deepest gratitude goes first and foremost to my supervisor, Professor Dr. Gregor Meyers, Head of the laboratory for immunology and molecular biology of positive strand RNA viruses, Institute of Immunology (IfI), Friedrich-Loeffler Institute (FLI), Insel of Riems, Germany. Many thanks for providing me a great opportunity that I could join his excellent group. I have learned a lot from him, not only about the scientific research, but also about the attitudes toward life. No matter when and where I met problems with the study or life, he always encouraged me and offered helpful support and suggestions. Under his guidance, I finished the program and thesis for my PhD study smoothly. I really have learned a lot from him and gained much knowledge over four years study in his lab.

I would like to express my heartfelt gratitude to my supervisory committee members, Univ.-Prof. Dr. Klaus Osterrieder and PD Dr. rer. nat. Michael Veit, Institute of Virology, Faculty of Veterinary Medicine, Freie-Universität Berlin, Germany for allowing me to be under their supervision and giving me valuable advice and unlimited support for my thesis.

I am also deeply indebted to my mentors Dr. Birke Andrea Tews and Dr. Christine Luttermann for their unconditional support, enthusiasm in each discussion and valuable suggestions. They are rigorous in work and meticulous in details. They are role models whom I should chase for in my future.

I deeply appreciate the help and support from all the members in our lab. Many thanks to Dr. Anna Katharina Tucakov, Dr. Jolene Carlson, PhD students Juliane Kühn and Anne Klingebeil for all the discussions in each lab meeting and daily lab work. Many thanks to technician Gaby Stooß, Kristin Trippler and Rico Jahnke for their excellent supportive work in the lab.

I would also like to express my appreciation to all the members of Institute of Immunology (IFI) for their suggestions and assistance in every IFI seminars and cooperation lab work.

Moreover, my thanks would go to my beloved family for their endless love, encouragement, support and concern in accomplishment of this doctoral degree all through these years. I also owe my sincere gratitude to my fiancée Dandan Li. We have been in love for 10 years, many thanks for her always be there, standing by me when I was in hard time of my life.

Last but not least, I would like to thank the China Scholarship Council (CSC) for their financial support, without it, I could not be possible to afford the opportunity to study in Germany. I am very proud of my motherland.

Declaration of authorship/Selbstständigkeitserklärung

I hereby confirm that the present work was solely composed by my own. I certify that I have used only the specified sources and aids.

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

Ort, Datum	Unterschrift
Berlin, 19.04.2021	**Yu Mu**