Cover Page

Universiteit Leiden

Author: Treffers, Emma Elisabeth (Emmely)
Title: A +RNA virus diptych : Chikungunya virus-host interactions and arteriviral programmed ribosomal frameshifting
Issue Date: 2015-12-16

A +RNA virus diptych

**Chikungunya virus-host interactions
and
Arteriviral programmed ribosomal frameshifting**

Emmely Treffers

A +RNA virus diptych

Chikungunya virus-host interactions
and
Arteriviral programmed ribosomal frameshifting

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 16 december 2015
klokke 16.15 uur

door

Emma Elisabeth Treffers
geboren te Leiden
in 1985

Promotor

Prof. dr. E.J. Snijder

Co-promotores

Dr. M.J. van Hemert

Dr. P.A. van Veelen

Leden promotiecommissie

Prof. dr. A.E. Gorbalenya

Prof. dr. M. Wuhrer

Dr. R.C. Olsthoorn (Leiden Institute of Chemistry, Leiden University)

Dr. C.R. Jimenez (VU University Medical Center, Amsterdam)

TABLE OF CONTENTS

PART 1
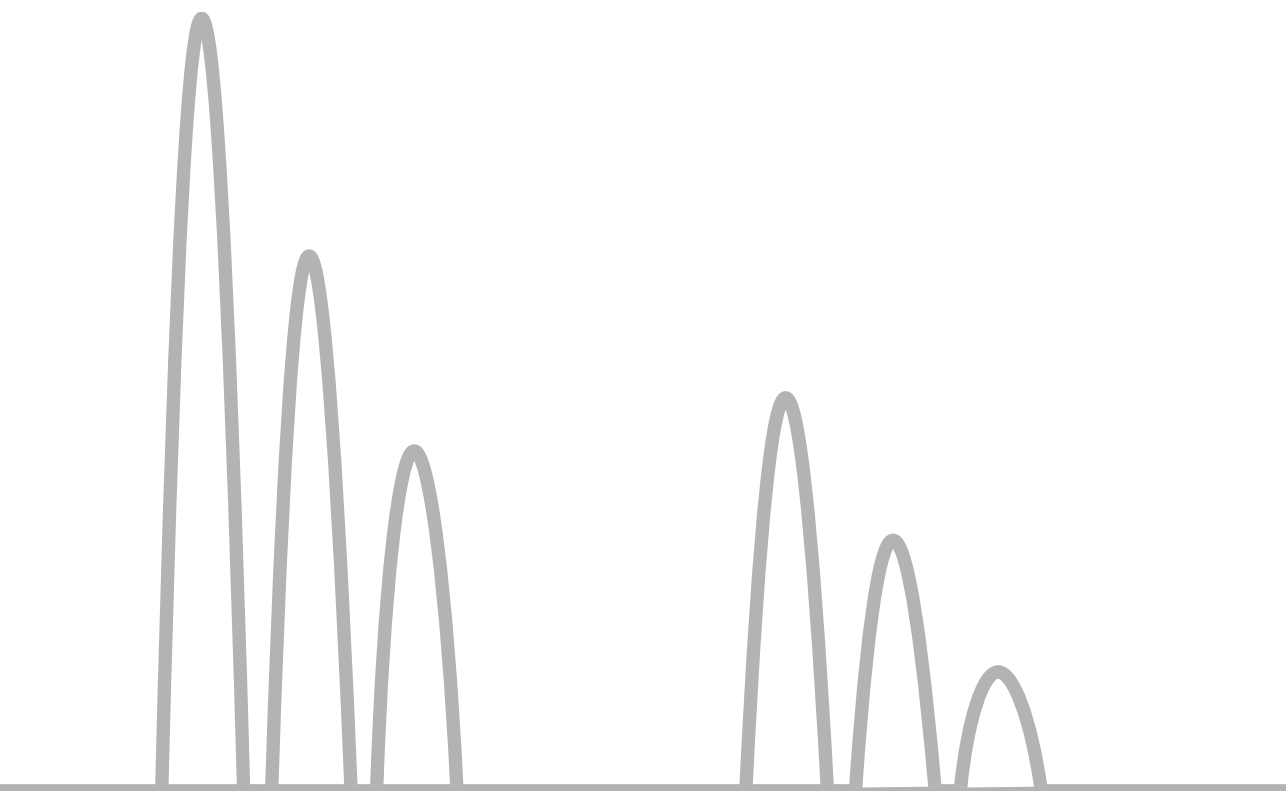
Quantitative proteomics of chikungunya virus

Chapter 1

General introduction part 1

+RNA VIRUSES

Viruses are obligate intracellular parasites that are highly dependent on many cellular components and processes. Using genome type and replication strategy, the Baltimore classification system distinguishes double-stranded (ds)DNA, single-stranded (ss)DNA, dsRNA, positive (+)ssRNA, and negative (-)ssRNA genomes. Additionally, there are +ssRNA viruses that replicate via a DNA intermediate and dsDNA viruses that replicate through a ssRNA intermediate [1]. The +RNA viruses are the largest virus group, which contains many important human and animal pathogens, such as SARS coronavirus [2-4], poliovirus [5], hepatitis C virus [6], dengue virus [7], chikungunya virus [8], foot and mouth disease virus [9] and porcine reproductive and respiratory syndrome virus [10, 11].

All +RNA viruses encode an RNA-dependent RNA polymerase (RdRp) to replicate their genomes and several virus groups transcribe subgenomic RNAs to express the genes encoding their structural and accessory proteins [12]. +RNA viral genome replication is associated with modified cellular membranes to which components of the viral replication complex are anchored. The origin of these membranes differs between virus groups, but membrane-associated replication might serve the same fundamental purpose(s), as it is a general feature of all eukaryotic +RNA viruses [13].

Chikungunya virus

The work described in this first part of my thesis focuses on chikungunya virus (CHIKV). CHIKV is a reemerging arthropod-borne human pathogen that was first discovered in Tanzania in 1952, during an outbreak on the Makonde Plateau in the Southern region of Tanganyika [8]. Before 2005 only occasional small-scale outbreaks were reported. CHIKV reemerged around 2005 on the east coast of Africa and several Indian Ocean islands, and subsequently spread across the Asian continent while infecting millions of people [14, 15]. This outbreak was closely associated with the occurrence of a single point mutation in the viral genome, resulting in an A226V substitution in the CHIKV envelope protein E1. This mutation increased the epidemic potential of CHIKV, as it allowed the virus to be more efficiently spread by a new vector, *Aedes albopictus* (Asian tiger mosquito). The global distribution of *Ae. albopictus*, which includes urban areas in southern Europe and the USA, is wider than for *Ae. aegypti*, which was the main vector before 2005 [16]. CHIKV spread to the Americas around the end of 2013 and has since caused an explosive outbreak in the Caribbean and South/Central America [17]. Infected travelers have returned to Europe, Australia, the USA and Canada [18-25]. In countries with *Ae. aegypti* or *Ae. albopictus* populations this poses the risk of establishing new CHIKV reservoirs [26, 27], and locally-transmitted infections have already occurred in e.g.

Italy, France and the USA in 2007, 2010 and 2014 [28-30]. During the outbreak in Italy over 200 confirmed cases were reported [28].

In the Makonde language Chikungunya means "that which bends up", which refers to the posture of people suffering from the disease due to the severe and persisting poly-arthralgia that characterizes the disease [8]. Other symptoms include a short episode of high fever, myalgia, headache and rash. Asymptomatic infections range from 3-25% [31-34]. Therapy is limited to supportive care as antiviral treatments and vaccines for CHIKV are still being developed [35, 36]. The development of antiviral therapies and vaccines is hampered by the rapidly occurring resistance, caused by the generally high mutation rate in RNA viruses [37-40].

Alphavirus genome organization and expression

CHIKV has a single stranded +RNA genome of 12kb and belongs to the alphavirus genus of the Togavirus family [41]. The alphavirus genome contains two large open reading frames (ORFs) that are translated into polyproteins. The genome also serves as the mRNA template for translation of the nonstructural proteins nsP1-4 that are encoded by the first ORF. Translation of nsP4 is dependent on read-through of an opal stop codon at the end of the nsP3-coding region [42]. The polyprotein is proteolytically processed by a protease domain residing in nsP2 [43, 44]. The structural polyprotein is encoded by the second ORF and is expressed from a subgenomic mRNA. The structural polyprotein is processed into the capsid (C) protein, the envelope proteins E1, E2 and E3, and the 6k ion channel protein by a protease domain in C and host proteases in the exocytic pathway [45, 46]. A third small ORF is embedded within the sequence encoding 6k and is translated after a −1 ribosomal frameshift resulting in the synthesis of a transframe (TF) protein that shares the N-terminal sequence with 6k and with the C-terminal amino acids encoded by the -1 ORF [47].

Alphavirus replicative cycle in mammalian cells

The alphavirus nucleocapsid contains a single copy of genomic RNA and is enveloped in a lipid bilayer derived from the host plasma membrane. This bilayer contains glycosyl-ated E1 and E2 heterodimers that are assembled into 80 trimeric spikes [48]. It is still unknown which cellular receptor(s) is required for CHIKV attachment to the host cell. Alphavirus cell entry is initiated by interaction of E2 with cellular proteins. Generally, alphavirus entry is dependent on clathrin-mediated endocytosis and E1-mediated fu-sion with endosomal membranes [49]. After fusion, the nucleocapsid is released into the cytoplasm and is almost immediately uncoated to release the viral RNA for translation of the nonstructural polyprotein [50].

The nonstructural polyprotein is processed by nsP2 in a sequential manner. The nsP3/nsP4 junction is cleaved first, followed by the nsP1/nsP2 junction and finally the nsP2/

nsP3 junction. This sequential processing regulates the activity of RNA synthesis by the replication and transcription complex (RTC) consisting of the nonstructural proteins [51, 52]. The precursor P123 and nsP4 form a complex that synthesizes -strands exclusively, while the nsP1/nsP2 cleavage converts the complex into one that also synthesizes +strands. After the nsP2/nsP3 junction has been cleaved, the complex can only synthesize +strand genomes and the 26S subgenomic mRNA [53]. Viral replication takes place in spherules at the plasma membrane and in cytoplasmic vacuoles that are derived from modified lysosomal and endosomal membranes [54, 55].

nsP1 has guanyl transferase and guanine-7-methyltransferase enzymatic activities required for capping of viral genomes and subgenomic mRNAs [56-58]. This protein is membrane-associated and is thought to anchor the RTC to the modified membrane structures [59]. Besides the protease activity that resides in a C-terminal domain, nsP2 also has RNA helicase [60] and RNA triphosphatase/nucleoside triphosphatase activities [61, 62] residing in the N-terminal part of the protein. nsP2 is also responsible for the transcriptional and translational shut-off of the host cell during alphavirus infection [63]. nsP2 localizes to cytoplasmic foci and the nucleus [64, 65]. The exact functions of nsP3 still remain unclear but the protein is required for RNA synthesis [43, 66] and the macro domain of CHIKV nsP3 has ADP-ribose 1"-phosphate phosphatase and RNA binding activity [67]. During infection nsP3 localizes to cytoplasmic foci where it might inhibit stress granule formation [65, 68, 69]. nsP4 is the RNA dependent RNA polymerase (RdRp) that synthesizes both + and –stranded RNA [43, 70]. This protein might also have terminal adenylyltransferase (TATase) activity which could be involved in the maintenance of the poly(A)tail [71]. Compared to the other nsPs, nsP4 is very short lived due to ubiquitin-mediated degradation, which is one of the ways by which the nsP4 concentration is regulated during infection [72].

Processing of the alphavirus structural polyprotein C-E3-E2-6k-E1 starts by release of C through autoproteolysis [45, 73]. Subsequently, a signal sequence in the new N-terminus of the polyprotein initiates translocation of pE2 (the E3-E2 precursor [74]) across the endoplasmatic reticulum (ER) membrane [75]. A hydrophobic stretch near the C-terminus of E2 anchors the protein in the ER membrane [76]. The insertion signals for 6k and E1 are located after the E2 anchor sequence and in the C-terminal domain of 6k, respectively. The E2-6k and 6k-E1 junctions are cleaved by signalases in the ER lumen after insertion [46]. In the ER lumen and during transport through the Golgi apparatus pE2 and E1 become glycosylated and palmitoylated [77-79]. pE2 and E1 form a heterodimer in the ER and during transport of the pE2-E1 complex through the *trans* Golgi complex to the plasma membrane pE2 is cleaved into E3 and E2 by furin, a host protease [80, 81]. CHIKV E3 is not incorporated into virions [82]. The alphavirus genomic RNA contains a packaging signal that is recognized by C, which results in the subsequent oligomerization of capsid proteins to generate the nucleocapsid [83, 84]. Budding occurs at the plasma
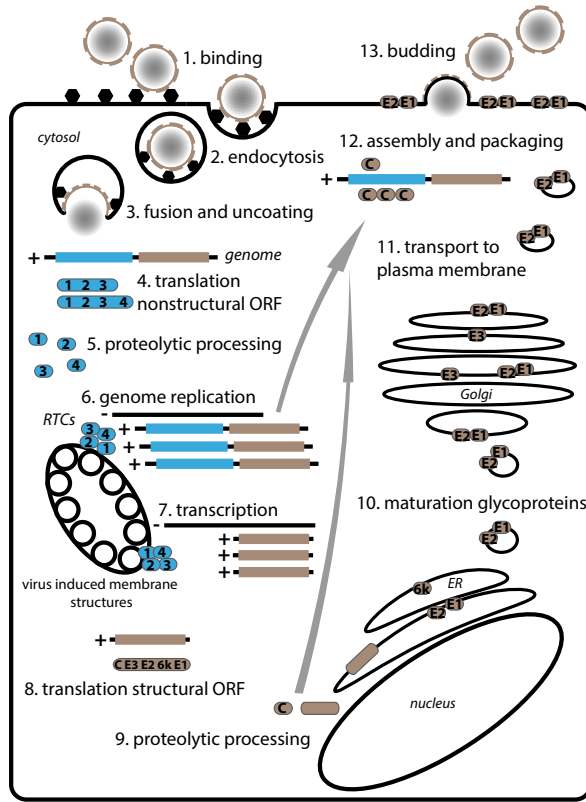
Figure 1: Alphavirus replicative cycle. 1. Alphavirus particles consist of a nucleocapsid that is enveloped in a host plasma membrane-derived lipid bilayer. Alphavirus entry is initiated by binding of the E2 glycoprotein to cellular membrane proteins. **2.** Alphavirus particles generally enter the cell through clathrin-mediated endocytosis. **3.** The E1 glycoprotein mediates fusion with endosomal membranes. The nucleocapsid is released into the cytosol and uncoated to release the +RNA genome. **4.** The nonstructural ORF of the genome is translated by cellular ribosomes into two polyproteins. Translation of nsP4 depends on stop codon readthrough. **5.** nsP2 proteolytically processes the nonstructural polyproteins into the individual subunits. The nsPs assemble into replication and transcription complexes (RTCs) at the plasma membrane and on cytoplasmic vacuoles that are derived from modified lysosomal and endosomal membranes. **6.** During genome replication, nsP4, the RNA-dependent RNA polymerase (RdRp) first synthesizes minus strands which are later used as templates to synthesize new + strand genomes. **7.** – strands are used to transcribe + strand subgenomic mRNAs. **8.** The structural ORF is translated from the +strand subgenomic mRNA to produce the structural polyprotein. **9.** Processing of the structural polyprotein starts by release of C through autoproteolysis. The remaining polyprotein is translocated across the endoplasmatic reticulum (ER) membrane. The E2-6k and 6k-E1 junctions are cleaved by signalases in the ER lumen. **10.** pE2 and E1 form a heterodimer and are glycosylated and palmitoylated in the ER lumen and during transport through the Golgi apparatus. **11.** During transport through the *trans* Golgi complex to the plasma membrane pE2 is cleaved into E3 and E2 by furin. At the plasma membrane E2 and E1 are incorporated into new virions. **12.** The alphavirus genomic RNA contains a packaging signal that is recognized by C, which results in oligomerization of capsid proteins to generate the nucleocapsid. **13.** Budding occurs at the plasma membrane through interaction of the nucleocapsid with the cytoplasmic domain of E2.

membrane through interaction of the nucleocapsid with the cytoplasmic domain of E2 [85]. A schematic representation of the alphavirus replicative cycle in mammalian cells is depicted in Figure 1.

Chikungunya virus-host interactions
CHIKV is, like all viruses, very dependent on host factors during its replicative cycle. Assuming cell entry of CHIKV occurs in a similar fashion as for other alphaviruses, it requires binding of virus particles to host receptors [86, 87]. An active endocytic pathway, including endosomal acidification, is required for fusion of the viral and host membranes and release of the nucleocapsid into the cytosol [88]. Uncoating of the nucleocapsid requires interaction with the large ribosomal subunit [89]. Translation of viral mRNA requires the cellular translational machinery since viruses do not encode their own ribosomes [90]. RTCs are associated with modified membrane structures [54] and several host factors have been reported to be recruited to these nsP-containing complexes [65, 91-94]. Viral proteins can be phosphorylated and glycosylated by cellular enzymes [77-79, 95, 96]. During budding the nucleocapsid becomes enwrapped in a host-derived membrane envelope [97]. Several steps in the replication cycle require the presence of host membrane lipids [98]. Alphaviruses inhibit host transcription and translation and suppress (innate) immune responses to promote efficient replication and translation of viral genomes [99, 100].

A better understanding of these complex interactions between a virus and its host can facilitate the development of antiviral therapies. Instead of targeting viral enzymes, host factors may also be targeted to inhibit viral replication. With this strategy resistance is expected to occur less frequently, since cellular antiviral drug targets are unlikely to mutate during infection [101, 102].

Several different methods have been used to study CHIKV-host interactions in both mammalian and insect hosts. RNA interference (RNAi) screens [103], transcriptomics [104-106], yeast two-hybrid assays [91, 107], co-immunoprecipitation with viral proteins [108], computational methods [109, 110] and proteomics [111-120]. In these studies a great number of (potential) interaction partners of viral proteins and cellular processes that are affected by CHIKV infections have been identified. Several CHIKV proteomics studies have been performed so far, in which infected cells, infected mice and patient sera were analyzed. However, a large-scale quantitative proteomics study in which several time points post infection during the first round of replication on synchronously infected cells was still missing. Additionally, a systematic analysis of the posttranslational modifications on host proteins during CHIKV infection has never been performed. The aim of the research in this thesis was to track changes in host protein abundance and phosphorylation status during CHIKV infection in well-characterized cells during a single round of replication before cytopathic effects become apparent.

QUANTITATIVE PROTEOMICS

In this first part of my thesis mass spectrometry-based quantitative proteomics was used to study the consequences of CHIKV infection in mammalian cells. Proteomics is the systematic study of the total set of proteins produced from the genome present in the cell at a given time point [121]. Proteomics is a hypothesis-generating approach and is often used as a starting point to identify proteins and pathways that might be involved in a certain process (such as viral infection). To determine the exact role of identified host factors it is usually necessary to do follow-up experiments. With quantitative proteomics the changes in protein abundance in the cell during viral infection can be studied. Relative changes in host protein abundance can provide novel insights into which host proteins and pathways are important during viral infection. An increase in the abundance of certain proteins can, for example, indicate activation of an innate immune pathway, while a decrease could indicate that the virus targets this protein for degradation [122-125].

In this thesis, the stable isotope labeling by amino acids in cell culture (SILAC) method (Figure 2) was used to study the full spectrum of host proteins during CHIKV infection. Cells are grown in media lacking certain essential amino acids, which are supplemented with non-radioactive, isotopically labeled forms of these amino acids [126]. In the heavy state, several ^{12}C and/or ^{14}N atoms are replaced with ^{13}C and ^{15}N isotopes to increase the mass of the amino acid. Cells are grown in SILAC media for at least five cell-doublings to allow >96% incorporation of the heavy amino acids. When labeled arginine and lysine are used in combination with trypsin digestion, all peptides, except for the C-terminal one, will contain at least one labeled amino acid [127]. There is no chemical difference between the heavy isotope-labeled amino acids and their natural counterparts and as a result cell growth and behavior in the light and heavy isotope-labeled conditions is the same. Light- and heavy-labeled peptides co-elute from the high performance liquid chromatography (HPLC) column and are analyzed together in the mass spectrometer [126]. Consequently, the relative quantification of peptide and protein ratios with SILAC is very accurate [128].

An advantage of SILAC over other quantitative proteomics methods such as iTRAQ [129], ICAT [130], dimethyl labeling [131] or label-free approaches [132-139] is that experimental and control samples can be mixed directly after harvesting, so before further sample processing, because every protein in the sample has already been labeled. This reduces the amount of variability between experimental and control conditions caused by sample handling [126] and increases reproducibility [140].

SILAC is most often applied in cell culture but has also been used successfully to label complete animals such as mouse [141], fly [142] and zebrafish [143].
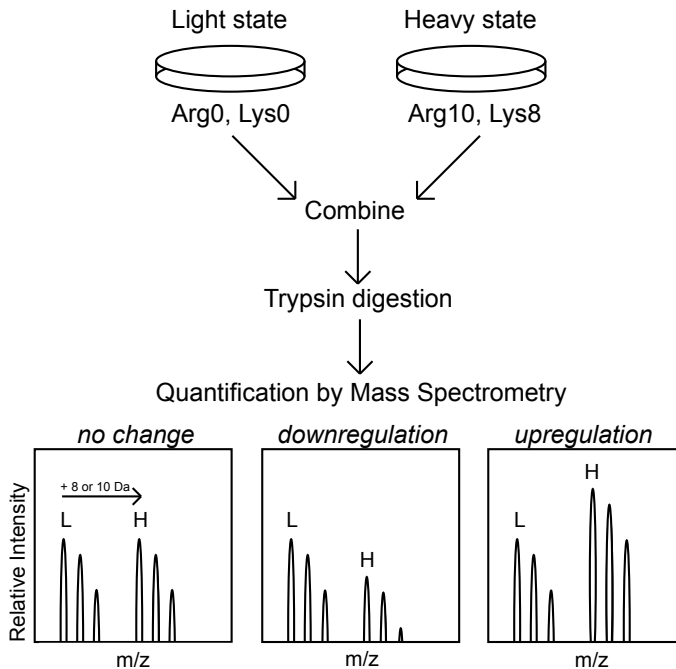
Figure 2: SILAC labeling strategy. Cells are grown in media containing stable isotope-encoded amino acids until the amino acids have been fully incorporated into the cells. After the experiment (e.g. a viral infection) has been performed, the light and heavy samples are mixed and further processed as a single sample. Peptides originating from the light and heavy experimental conditions can be distinguished through their mass difference by mass spectrometry and the relative intensities are used to determine whether the abundance of a protein did not change, was downregulated or upregulated as a result of the experimental treatment.

Quantification of post translational modifications

Most proteins undergo post-translational modifications (PTMs). These modifications play an important role in regulating biological processes and add another layer of complexity to the proteome. PTMs allow the cell to respond rapidly to extra- and intracellular stimuli by changing protein activity, subcellular localization, binding partners and stability [144-146]. Notably, the function of a certain protein can change significantly due to the addition of a specific PTM, while the abundance of that protein does not change at all. Over 400 different PTMs have been described, the most common being phosphorylation, acetylation, N-linked glycosylation and amidation [147]. A single protein can often be modified by distinct PTMs at the same or multiple residues and crosstalk between different PTMs generates an even more complex proteome [146, 148].

The analysis of PTMs poses several technical challenges. Modifications are frequently transient, which makes the timing of sampling during an experiment crucial. The abun-

dance of modified proteins is often very low compared to their unmodified counterparts, which makes it essential to enrich for modified proteins or peptides prior to analysis [149]. Modified peptides can be lost during sample preparation, either due to instability of the modification or because modified peptides preferentially adsorb to metal or plastics [150, 151]. The identification and quantification of a modified site is often based on a single peptide, while multiple peptide identifications are used when protein abundance is quantified. This makes PTM quantification less accurate than the quantification of relative protein abundance. The software search engine that is used to identify peptides and proteins from mass spectrometry data, tests each possible arrangement of amino acids with and without the variable modifications of interest (such as phosphorylation) to find the best peptide match in the database that is used for the search. The addition of each extra variable modification during the search greatly increases search complexity and generates more false assignments [151, 152]. For peptides that contain multiple residues that could contain the modification it is not always possible to identify the exact modified site [152]. When a change in modified peptide abundance is observed, it is important to verify that the total protein abundance did not change as well [151].

In this thesis, SILAC was used to study changes in host protein phosphorylation during CHIKV infection. One third of eukaryotic proteins is estimated to become phosphorylated [153]. Proteins are phosphorylated by the transfer of a phosphoryl group from ATP or GTP to either a side chain of a serine or threonine residue by protein serine/threonine kinases or to a tyrosine residue by protein tyrosine kinases [153]. Humans express ±520 kinases of which around 90 are tyrosine kinases [154, 155]. Protein phosphorylation is a reversible modification and can be removed by protein phosphatases. The human genome contains genes for 107 protein tyrosine phosphatases [156] and around 30 protein serine/threonine phosphatases [157]. The distribution of phosphotyrosine, phosphothreonine and phosphoserine sites in the cell is ± 2%, 12%, and 86% [158].

Several features of phosphorylated peptides aid their identification by mass spectrometry. The addition of HPO_3 results in an increase in amino acid residue mass of 80 Da. Sites on phosphopeptides can be identified from mass shifts in fragment ions generated by MS/MS [151]. Peptides containing phosphotyrosine can often be identified by a fragment ion of 216 Da which derives from peptide bond cleavage on either side of the phosphotyrosine residue [151, 159]. Additionally, peptides containing phosphoserine and phosphothreonine often undergo a 98 Da neutral loss corresponding to the loss of H_3PO_4 [151].

OUTLINE PART 1

In this first part of my thesis, changes in the host cellular proteome following CHIKV infection were studied to better understand the interplay between viral replication and host cell infrastructure and metabolism. **Chapter 2** describes a SILAC-based quantitative proteomics study to analyze temporal changes in the proteome of CHIKV-infected cells. This study revealed that changes in protein abundance during CHIKV infection are relatively small and that most of the proteins that showed significantly changed abundance were downregulated. Four proteins that were significantly downregulated during CHIKV infection, Rnd3, DDX56, UbcH10 and Plk1, were overexpressed from plasmids in host cells and this inhibited CHIKV infection. **Chapter 3** describes a SILAC-based quantitative phosphoproteomics study of CHIKV-infected cells. Eukaryotic elongation factor 2 (eEF2) was identified as a factor that becomes phosphorylated in various cell lines early during infection with CHIKV, Semliki forest virus (SFV) or Sindbis virus (SINV). Infection with coxsackie virus B3, a picornavirus, also triggered eEF2 phosphorylation. eEF2 phosphorylation might reflect part of the antiviral response of the cell, but it is not activated via one of the 'classical' pathways that are generally involved in the innate immune response to virus infections. In **chapter 4** the findings of part 1 are summarized.
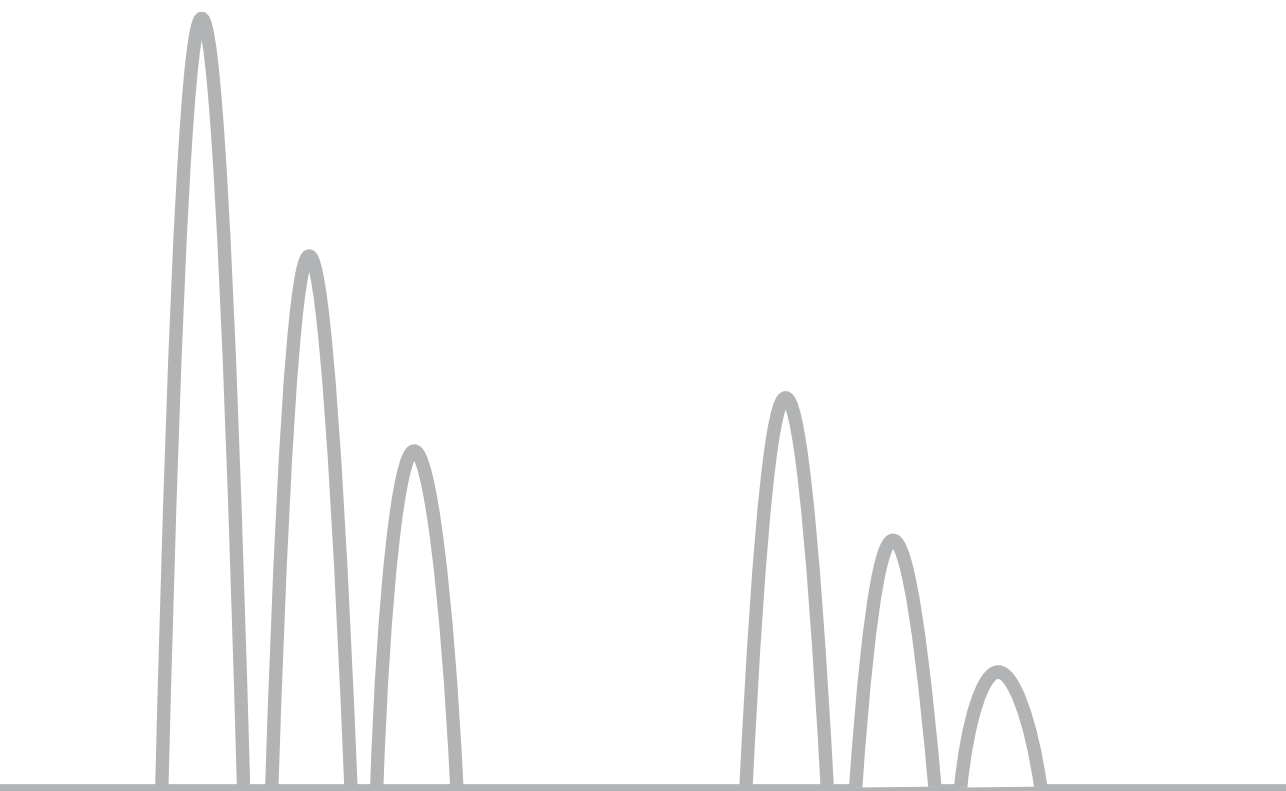
Chapter 2

Temporal SILAC-based quantitative proteomics identifies host factors involved in chikungunya virus replication

Emmely E. Treffers[1,2], Ali Tas[1], Florine E.M. Scholte[1], Myrthe N. Van[1], Matthias T. Heemskerk[1], Arnoud H. de Ru[2], Eric J. Snijder[1], Martijn J. van Hemert[1#], and Peter A. van Veelen[2#]

[1]Molecular Virology Laboratory, Department of Medical Microbiology, Leiden University Medical Center, ZA, Leiden, The Netherlands, [2]Department of Immunohematology and Blood transfusion, Leiden University Medical Center, ZA, Leiden, The Netherlands
[#]These authors contributed equally

ABSTRACT

Chikungunya virus (CHIKV) is an arthropod-borne reemerging human pathogen that generally causes a severe persisting arthritis. Since 2005, the virus has infected millions of people during outbreaks in Africa, Indian Ocean Islands, Asia, and South/Central America. Many steps of the replication and expression of CHIKV's 12-kb RNA genome are highly dependent on cellular factors, which thus constitute potential therapeutic targets. SILAC and LC-MS/MS were used to define the temporal dynamics of the cellular response to infection. Using samples harvested at 8, 10, and 12 h post infection, over 4700 proteins were identified and per time point 2800-3500 proteins could be quantified in both biological replicates. At 8, 10, and 12 h post infection, 13, 38, and 106 proteins, respectively, were differentially expressed. The majority of these proteins showed decreased abundance. Most subunits of the RNA polymerase II complex were progressively degraded, which likely contributes to the transcriptional host shut-off observed during CHIKV infection. Overexpression of four proteins that were significantly downregulated (Rho family GTPase 3 (Rnd3), DEAD box helicase 56 (DDX56), polo-like kinase 1 (Plk1), and ubiquitin-conjugating enzyme E2C (UbcH10) reduced susceptibility of cells to CHIKV infection, suggesting that infection-induced downregulation of these proteins is beneficial for CHIKV replication. All MS data have been deposited in the ProteomeXchange with identifier PXD001330 (http://proteomecentral.proteomexchange. org/dataset/PXD001330).

INTRODUCTION

Chikungunya virus (CHIKV) is an arthropod-borne RNA virus and its transmission to humans by mosquitoes generally leads to a short episode of high fever, rash and joint pains. The arthritis can be debilitating and may last for many months [14]. The virus was first isolated during an outbreak in Tanzania in 1952/1953 and large outbreaks used to be rare until the virus reemerged in 2005-2006 on the east coast of Africa and several Indian Ocean islands. Since then it has spread across the Asian continent, causing several large outbreaks that affected millions of people [14]. Hundreds of infected travelers have returned to Europe and locally transmitted infections have occurred in Italy and France in 2007, 2010, and 2014 [28, 29]. In the autumn of 2013, the virus was introduced to the Caribbean island of Saint Martin, which – for the first time – led to a massive outbreak in the Caribbean and South/Central America [17]. By November 2014, the number of estimated cases in the Americas already exceeded 900,000. and in July 2014 the first locally transmitted infections in the USA were reported [30].

CHIKV belongs to the alphavirus genus of the togavirus family and has a 12 kb single-stranded positive polarity (+) RNA genome that contains two open reading frames (ORFs). The first ORF of the genome is directly translated into a large polyprotein that is proteolytically processed into the nonstructural proteins (nsP) 1 to 4 by a protease residing in nsP2 [41, 76]. The four nsPs, probably together with host proteins, form a replication complex that first synthesizes a negative-stranded (-) RNA. Next, this –RNA serves as the template for the synthesis of genomic RNA and a subgenomic RNA. The latter is translated into a polyprotein that is proteolytically processed into the structural proteins C (capsid), envelope glycoproteins E1, E2 and E3 and the 6K ion channel protein by the joint action of a protease domain in C and host proteases in the exocytic pathway [41, 76]. For a schematic overview of CHIKV genome organization and expression, see Figure S1. Due to its small genome size, CHIKV is dependent on host factors during many steps of its replication cycle. For example, cellular receptors are required for CHIKV entry, ribosomes are required for translation of the viral mRNA, cellular membrane-derived spherules form a platform for viral RNA synthesis, and the secretory pathway is required for the maturation of the viral glycoproteins and virion biogenesis [76]. In addition to exploiting cellular resources, CHIKV suppresses cellular transcription, translation and (innate) immune responses to facilitate efficient virus replication [99, 100].

Currently, there is no licensed vaccine or antiviral treatment on the market to prevent or treat CHIKV infection [35]. Efforts to develop antiviral drugs have traditionally focused on viral targets (e.g. replicative enzymes), but drug resistance often emerges due to the high mutation rate of RNA viruses. Consequently, there is a growing interest in targeting host factors involved in viral replication as an alternative strategy, since host genes are unlikely to mutate during antiviral therapy [101, 102].

To gain more insight into the complex interaction between CHIKV and its host we set out to study temporal changes in the cellular proteome during CHIKV infection through stable isotope labeling by amino acids in cell culture (SILAC) and subsequent LC-MS/MS analysis [126]. This was done from the moment that CHIKV +RNA became clearly detectable (8 h p.i.) until the moment that viral RNA and protein levels reached their maximum (10-12 h p.i.), but before serious cytopathic effects were observed.

Quantitative proteomics strategies have been used previously to study the effects of CHIKV infection in cell culture, mice and patient serum (for a recent review see [160]). 2D-DGE and 2D-DIGE approaches [112, 113, 115, 119] identified between 15 and 51 differentially-expressed proteins. A GeLC-MS/MS approach identified 1455 proteins, of which 90 were significantly downregulated [111]. Another study identified 3505 proteins of which only 514 were present in all samples and none of these were differentially expressed [120]. An iTRAQ approach resulted in the identification of 569 proteins of which 63 were differentially expressed [116]. A combination of 2D-DIGE and iTRAQ resulted in the identification of 177 differentially-expressed proteins out of a total of 2686 identified proteins [114]. Advantages of our approach over these previously published studies are that we analyzed synchronously infected cells during the first round of replication in a well-characterized system and tracked changes in abundance over time. This approach resulted in the identification of over 4700 cellular proteins, of which 13, 38 and 106 were differentially expressed at 8, 10 and 12 hours post infection (h p.i.), respectively. The majority of these proteins were downregulated. Peptides derived from ten of the twelve subunits of the RNA polymerase II complex were identified and 7 of these were significantly degraded by 12 h p.i., likely contributing to the well-known CHIKV-induced shut-off of cellular transcription [99, 161]. Overexpression (prior to infection) of four host factors found to be significantly downregulated during CHIKV-infection (Rho family GTPase 3 (Rnd3), DEAD (Asp-Glu-Ala-Asp) box helicase 56 (DDX56), polo-like kinase 1 (Plk1), and ubiquitin-conjugation enzyme E2C (UbcH10)) reduced the percentage of cells that could be productively infected with the virus. This suggests that the infection-induced downregulation of these and presumably also other host proteins is not merely "collateral damage", but is beneficial for CHIKV replication.

MATERIALS AND METHODS

Cells

The previously described 293/ACE2 cell line [162] was obtained from dr. Shinji Makino (UTMB, Texas). These cells were chosen as they could be efficiently transfected with plasmids and infected with CHIKV (and other viruses), while adhering better to tissue culture plastics than standard HEK293 cells. This cell line was originally described to be

derived from human HEK293 cells, but cloning and sequencing of genes from 293/ACE2 cells for overexpression studies that were done after we had completed our proteomics studies, suggested this was not the case. BLAST analysis [163] of the sequences of cDNAs of a variety of host factors obtained from 293/ACE2 cells revealed that these had greater similarity to sequences from several non-human primates, like *Papio anubis* (olive baboon) and *Macaca mulatta* (rhesus macaque), than to those of *Homo sapiens*. Subsequent analysis of the cytidine monophospho-N-acetylneuraminic acid hydroxy-lase gene also revealed the absence of a *Homo sapiens*-specific 92-bp deletion [164]. Therefore, we now assume that the parental cells used to produce the 293/ACE2 cell line were actually of nonhuman primate origin. Cells were cultured in DMEM (Lonza, Switzerland), 10% FCS (PAA, Austria), 2 mM L-glutamine, 100 IU/ml penicillin and 100 μg/ml streptomycin at 37 °C and 5% CO_2.

Plasmids and oligonucleotides

cDNA was made from total cellular RNA using random hexamers (Promega, Madison, WI, USA, C118A) and RevertAid H Minus M-MuLV RT (Fermentas, Lithuania).

pCMV-FLAG-Rnd3 was constructed by amplifying the human Rnd3 gene from MRC-5 cell cDNA (all oligonucleotide sequences are listed in Table S1) using oligonucleotides RND3-fw and RND3-rev and cloning it into pCMV-Tag2 (Stratagene, United States).

pCDNA3-DDX56-FLAG was constructed by amplifying the DDX56 gene from 293/ACE2 cDNA using oligonucleotides DDX56-fw1 and DDX56-rev1 and cloning it into pCDNA3.0 (Life Technologies Europe, The Netherlands) after a second PCR with oligonucleotides DDX56-fw2 and DDX56–rev2 .

pCMV-FLAG-Plk1-WT and pCMV-FLAG-Plk1-KM (kinase domain mutant) [165] were a kind gift from dr. Hyungshin Yim (Hanyang University, Korea). pCMV-FLAG-Plk1-FAA (polo-box domain mutant) was constructed by site-directed mutagenesis of pCMV-FLAG-plk1-WT using oligonucleotides FAA-QC1 and FAA-QC2 to introduce W414F and V415A, followed by the introduction of the L427A mutation using oligonucleotides FAA-QC3 and FAA-QC4. These Polo-box domain mutations were previously described [165].

pCDNA3-FLAG-UbcH10 [166] was a kind gift from prof. Akira Nakagawara (Chiba Cancer Centre, Japan). All constructs were verified by sequencing.

pCMV-FLAG-Ub [167] expresses FLAG-tagged ubiquitin.

SILAC labeling & CHIKV infection

293/ACE2 cells were cultured in SILAC DMEM (PAA, Austria), 10% dialyzed FBS (Gibco, UK), 0.280 mM arginine, 0.398 mM lysine, 0.5 mM proline, 10 mM HEPES, 2 mM L-glutamine and 100 IU penicillin and 100 μg/ml streptomycin for > 5 cell doublings to ensure complete incorporation of the labeled amino acids. Arginine to proline conversion was not observed under these labeling conditions. The SILAC light medium was supplemented

with Arg $^{12}C_6$ $^{14}N_4$ and Lys $^{12}C_6$ $^{14}N_2$ (Sigma, The Netherlands), the SILAC heavy medium was supplemented with Arg $^{13}C_6$ $^{15}N_4$ and Lys $^{13}C_6$ $^{15}N_2$ (Cambridge Isotope Laboratories, Massachusetts, USA).

SILAC-labeled cells were seeded in 10 cm^2 dishes 1 day before infection with CHIKV strain LS3 at a multiplicity of infection (MOI) of 5, as previously described [168]. At 1 h p.i. the inoculum was removed and replaced with SILAC DMEM, 2% FBS, 0.280 mM arginine, 0.398 mM lysine, 0.5 mM proline, 25 mM HEPES, 2 mM L-glutamine, 100 IU/ml penicillin and 100 µg/ml streptomycin. Infected and mock-infected cells were lysed at 8, 10 and 12 h p.i. in 4% SDS, 0.1 M Tris pH 7.6, followed by incubation at 96 °C for 10 min to ensure complete inactivation of the virus. The experiment was performed in duplicate with a label swap (Figure 1A).
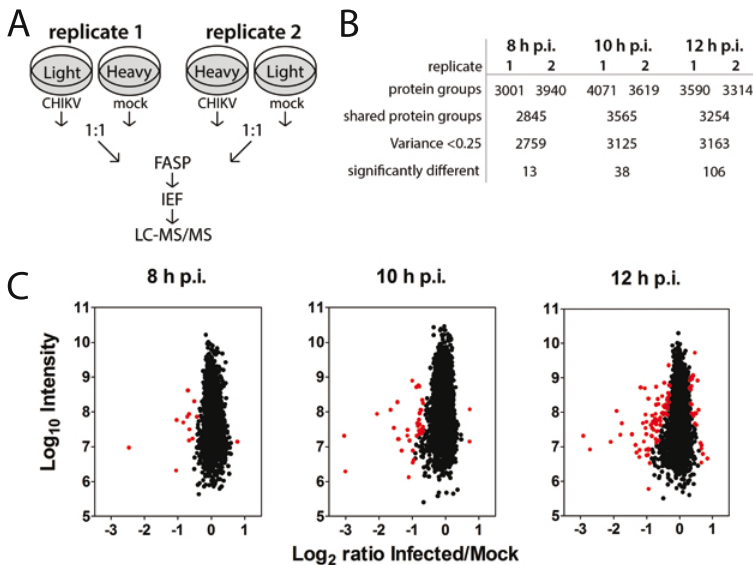
Figure 1: SILAC experiment (*A*) Experimental set-up of the SILAC-based proteomics analysis of virus-induced changes in the cellular proteome during the course of CHIKV infection. At each time point post infection 2 biological replicates were used, in which the SILAC label was swapped between mock-infected and infected samples. Infected and mock-infected cells were lysed and equal amounts of protein were mixed and subsequently digested into peptides using the FASP procedure. Peptides were separated into 12 fractions using IEF and analyzed using LC-MS/MS. (*B*) Number of identified proteins at each time point post infection. The total number of identified proteins with ≥2 razor peptides of which ≥1 unique was 4763. The number of identified proteins with ≥ 2 evidence counts for each biological replicate and the total number of proteins that were identified in both replicates with ≥ 2 evidence counts at each time point post infection are shown. For follow-up analysis only protein groups with a variance <0.25 were selected. (*C*) Proteome-wide quantification at 8, 10 and 12 hours post CHIKV infection. Average normalized $\log_2$ protein ratios are plotted against the $\log_{10}$ of summed peptide intensities. Each dot represents a protein group. Red dots indicate protein groups of which the abundance differed significantly between infected and mock-infected control cells (Benjamini-Hochberg FDR of 0.05).

Protein digestion and IEF

Equal amounts (200 µg) of mock- and CHIKV-infected lysates were mixed, followed by the addition of DTT to a final concentration of 0.1 M and 5 min incubation at 70 ºC. Protein digestion was performed using the Filter Aided Sample Preparation (FASP) method [169] and 100 µg protein was loaded per 0.5-ml 30 kDa Microcon filter devices (Millipore, Massachusetts, USA). Filter devices were washed twice with 8 M urea, 0.1 M Tris pH 8.5 and cysteines were alkylated with 50 mM iodoacetamide in the same buffer. Samples were washed 3 times with 8 M urea, 0.1 M Tris pH 8 and proteins were digested by overnight incubation at room temperature (RT) in the same buffer with 1 µg endo Lys-C (Wako Pure Chemical Industries, Japan). The sample was diluted fourfold with 50 mM ammonium bicarbonate pH 8.4 containing 1 µg trypsin (Worthington Chemical Corporation, New Jersey, USA) and digested for 4 h at RT. Peptides were collected by centrifugation, acidified by addition of TFA to 1% and desalted using SPE. Peptides were separated into 12 fractions on 13-cm Immobiline DryStrips pH 3-10 (GE Healthcare, UK) in a 3100 Offgel Fractionator (Agilent Technologies, California, USA) that was run up to 20 kVh [170, 171]. Peptide fractions were desalted using SPE, frozen in liquid nitrogen and lyophilized in a CHRIST RVC 2-18.

Mass spectrometry

Lyophilized peptide fractions were dissolved in 95/3/0.1 (v/v/v) water/ACN/formic acid. The samples were analyzed by nanoflow LC-MS. The 1100 HPLC system (Agilent Technologies, California, USA) operating in split-flow configuration as described by Meiring et al [172] was coupled on line to a 7-tesla LTQ-FT Ultra MS (Thermo, Massachusetts, USA). HPLC columns were made in house. The samples were trapped on a 15-mm ReproSil-Pur C18, 3-µm pre-column (inner diameter 100-µm) and eluted to a 20-cm ReproSil-Pur C18, 3-µm analytical column (inner diameter 50-µm) by a 2.5 h 0-50% solvent B (10/90/0.1 v/v/v water/ACN/formic acid) gradient. The eluent was emitted via a customized nano-electrospray probe into the mass spectrometer performing in data-dependent mode, automatically switching between MS and MS/MS mode. Full scan MS spectra were acquired in the FT-ICR. The top 5 most intense ions were fragmented with collision-induced dissociation in the linear ion trap. The data were acquired with Xcalibur, version 2.0 SR2 (Thermo, Massachusetts, USA). Each fraction was measured twice.

Data analysis

Raw data files were analyzed with MaxQuant 1.2.2.5 [173] using the Andromeda search engine [174], databases used for the main search were ipi.HUMAN.v3.72 (86,392 entries) and a custom-made database containing the protein sequences of CHIKV-LS3 (11 entries) using the GenBank sequence (accession KC149888) [168]. A smaller database, human.first.search (15,612 entries) containing a subset of human protein sequences was

used for the first search. A concatenated reversed database (KR special amino acids) was used to reach a FDR of 0.01 and a list of common contaminants was included in the search. FDR at the peptide level was also 0.01. Enzyme specificity for the search was Trypsin/P. Variable modifications included in the search were oxidation (M) and acetylation (protein N-term) and carbamidomethyl (C) was included as a fixed modification. Up to 2 missed cleavages and a maximum of 5 modifications per peptide were allowed. Mass tolerance for precursor ions was 7 ppm and for fragment ions 0.5 Da. MaxQuant results were further analyzed with Perseus version 1.2.0.17, Microsoft Excel 2010 and GraphPad Prism version 5.

Contaminants, protein groups identified with the concatenated reversed database and protein groups that were only identified by site were removed from further analysis. Proteins that were identified with ≥2 razor peptides, of which ≥1 unique peptides were included for further analysis. For each time point only proteins with a ratio count ≥2 for both biological replicates were selected. Normalized ratios from one of the biological replicates were inverted to ensure all ratios were displayed as infected/mock. Normalized ratios were $\log_2$ transformed and averages and variances were calculated. Proteins with a variance <0.25 were included in the analysis. The list of excluded proteins was manually inspected and some of these proteins were still included in the analysis when a large change was observed in both replicates in the same direction but with relatively large variation. The Significance B function in Perseus 1.2.0.17 was used separately for each time point, both sided, with Benjamini-Hochberg FDR 0.05 to determine which protein groups displayed significantly changed abundance during CHIKV infection.

To estimate the percentage of protein originating from CHIKV, the total intensity of all identified CHIKV proteins was divided by the total intensity of all protein groups (including CHIKV proteins) identified per biological replicate. This was also done separately for the nonstructural and structural polyprotein.

Previously published protein half-lives [175] were used to predict $\log_2$ protein ratios for a hypothetical situation in which a block in protein translation occurs at 8 h p.i. These theoretical changes in protein abundance were compared to the observed protein degradation rates. Proteins were only selected for follow-up if their observed degradation was faster than published.

Homo sapiens, *Macaca mulatta*, *Papio anubis*, *Pan troglodytes*, *Gorilla gorilla* and *Pongo abelii* protein sequences of Rnd3, DDX56, UbcH10 and Plk1 were aligned using Geneious 7.1.5 [176] to determine cross species sequence conservation of the peptides identified during LC-MS/MS.

Western Blot Analysis
CHIKV- or mock-infected 293/ACE2 cells in 10 cm² dishes were washed with PBS and then lysed in 0.5 ml of 4× Laemmli sample buffer (100 mM Tris-HCl, pH 6.8, 40% glycerol,

8% SDS, 40 mM DTT, 0,04 mg/ml bromophenol blue). Proteins were separated by SDS-PAGE in a 10% polyacrylamide gel and were transferred to Hybond LFP-membranes (GE Healthcare, UK) by semi-dry blotting with a Transblot SD semi-dry transfer cell (BioRad, California, USA). Membranes were blocked with 1% casein in PBST and incubated o/n with rabbit antisera against CHIKV nsP1 [168], nsP2 [177], nsP3 [178], C, kind gifts from prof. Andres Merits (University of Tartu, Estonia), CHIKV E1, E2 [179], mouse monoclonal antibody H68.4 against the transferrin receptor (Invitrogen) or mouse monoclonal antibody against β-actin (Sigma) diluted in PBST with 1% casein. Biotin-conjugated swine-a-rabbit (DAKO) or goat-a-mouse (DAKO), and Cy3-conjugated mouse-a-biotin (Jackson, Pennsylvania, USA) diluted in PBST with 0.5% casein were used for fluorescent detection with a Typhoon-9410 scanner (GE Healthcare, UK).

Immunofluorescence Microscopy

293/ACE2 cells grown on coverslips were fixed o/n in 3% paraformaldehyde (PFA) in PBS. Quenching was done with 10 mM Glycine in PBS and cells were permeabilized with 0.1% Triton X-100 in PBS for 10 min. Coverslips were incubated for 1 h with CHIKV E2 antiserum [179] and mouse monoclonal antibody J2 against dsRNA (English & Scientific Consulting, Hungary) diluted in PBS with 0.5% BSA. Primary antibodies were detected with donkey-a-rabbit-Cy3 (Jackson, Pennsylvania, USA) and goat-a-mouse-Alexa488 (Jackson, Pennsylvania, USA). Nuclei were stained with Hoechst-33342. Coverslips were mounted with Prolong (Invitrogen) and examined using a Zeiss Axioskop2 fluorescence microscope with Axiocam HRc camera and AxioVision software (Zeiss, Germany).

FACS

One day prior to transfection with 0.5 μg plasmid and 2.5 μl Lipofectamine2000 (Life Technologies Europe, The Netherlands) in antibiotics-free medium, 7.5×10^4 293/ACE2 cells were seeded in 12-well clusters. At 4 hours post transfection (h p.t.) the medium was replaced by antibiotics-containing medium. At 10 h p.t. cells were either mock infected or infected with CHIKV-LS3-GFP [168] (Figure S1A) or GFP-expressing adenovirus HAdV-GFP/LUC [180] at an MOI of 5. At 10 h p.i. cells were harvested by trypsinization and fixed o/n at 4 °C in 3% PFA in PBS. Cells were washed twice with 1% FCS in PBS and once with 0.1% Saponin, 0.5% BSA, and 0.02% NaN_3 in PBS. Cells were permeabilized for 10 min on ice with 0.1% Saponin, 0.5% BSA, and 0.02% NaN_3 in PBS and stained on ice for 30 min with anti-FLAG-APC (APC anti-DYKDDDDK tag, clone L5, Biolegend, California, USA) diluted in the same buffer. Cells were washed 3 times in 0.5% BSA in PBS and stored at 4 °C in 1% FCS, 1% PFA in PBS until they were analyzed using a FACSCalibur (BD Biosciences, California, USA) and CellQuestPro software (BD Biosciences, California, USA). Intact single cells were gated. Contour plots were made using CellQuestPro with the following settings: log density 50%, smoothing 2, 1% threshold. Each contour

plot contains approximately 1×10^4 cell events. Each plot was divided into quadrants to differentiate between non-transfected, uninfected cells (lower left quadrant (LL)), non-transfected, infected cells (lower right quadrant (LR)), transfected, uninfected cells (upper left quadrant (UL)) and transfected, infected cells (upper right quadrant (UR)). To determine the effect of overexpression of host factors on the virus-driven expression of a GFP reporter gene, for each plot the percentage of transfected cells that also became infected ("double positive") was calculated as follows: UR/(UL+UR)*100. This number was compared to the percentage of non-transfected cells that became infected within the same sample, which was calculated as: LR/(LL+LR)*100.

RESULTS

Proteomics

Quality of proteomics dataset
To analyze changes in the cellular proteome in response to CHIKV infection we have performed a time course experiment in which we compared cell lysates of infected and mock infected SILAC-labeled mammalian cells harvested at 8, 10 and 12 h p.i. To minimize background signal from uninfected cells, an MOI of 5 was chosen to ensure that >99% of cells became infected. A CHIKV-induced cytopathic effect in 293/ACE2 cells became visible by microscopy around 14 h p.i., and, therefore, 12 h p.i. was chosen as the latest time point for proteomic analysis, to avoid a high proportion of false positive hits, e.g. through proteins that leaked out of these cells. At 8 h p.i. the changes in the proteome were limited and we, therefore, did not analyze earlier samples, but focused on samples with a high expression of viral proteins and viral RNA. At each time point, CHIKV proteins were identified by mass spectrometry (Table S2). To estimate how much of the total protein in infected cells was of viral origin, the sum of the intensities of all protein groups that originated from CHIKV was divided by the sum of all intensities for a given biological replicate (Table S3). CHIKV proteins accounted for roughly 0.6% to 1.1% of total protein, depending on the time-point analyzed. Based on the peptides analyzed, we estimated that about 70-80% of CHIKV protein in the cell originates from the structural polyprotein.

Successful CHIKV infection of the SILAC-labeled cells was also confirmed by western blot (WB) and immunofluorescence microscopy (IFA) (Figure 2A). At 4 h p.i., a weak labeling for double-stranded RNA (dsRNA), a marker for viral replication, was observed in a small number of cells, while at 8 h p.i. nearly all cells stained positive for dsRNA and the viral envelope protein E2 (Figure 2A). This confirmed that virtually all cells were CHIKV-infected under our experimental conditions. Viral proteins nsP1 and E2 were also
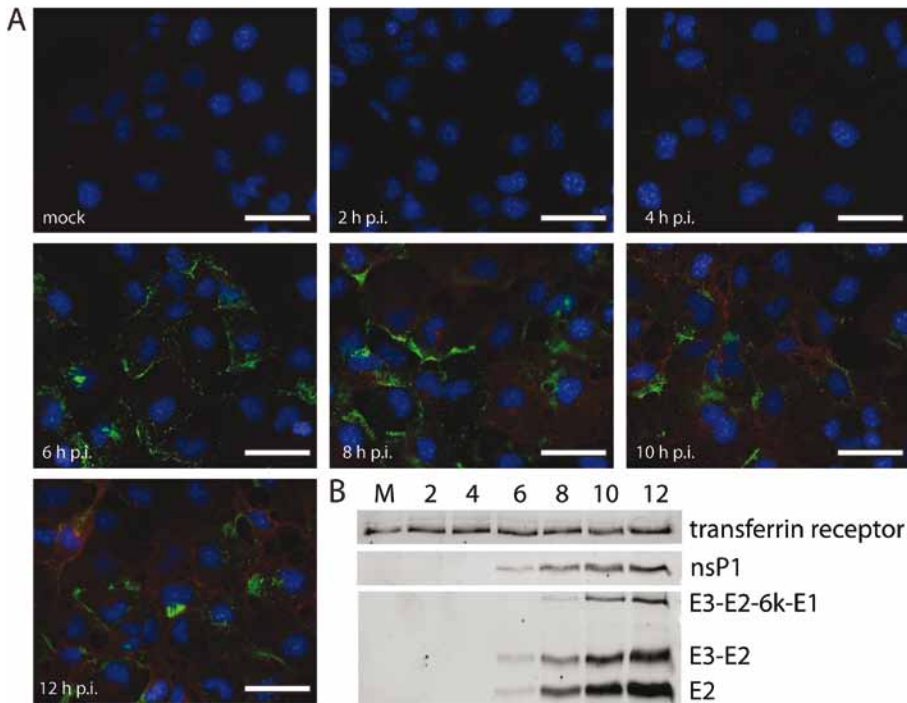
Figure 2: Immunofluorescence microscopy and Western blot analysis of CHIKV-infected SILAC-labeled cells at different time points post infection. (*A*) Cells were immunolabeled for dsRNA (green) and the CHIKV E2 protein (red), and nuclear DNA (blue) was stained with Hoechst-33342. Scale bar corresponds to 50 μm. (*B*) Protein lysates were separated by SDS-PAGE and viral proteins nsP1 and E2 were detected by Western blotting. The precursors E3-E2 and E3-E2-6k-E1 could also be detected with the E2 antibody. The transferrin receptor was used as a loading control.

detected by WB at 6 h p.i. and their abundance increased during the course of infection (Figure 2B).

A total number of 4763 protein groups was identified with ≥2 razor peptides of which at least 1 peptide was unique. In each biological replicate 3000-4000 protein groups were quantified with ≥2 evidence counts and at each time point 2800-3550 protein groups were quantified in both biological replicates (Figure 1B and Table S2).

Proteins displaying significantly changed abundance during CHIKV infection
The relative abundance of the majority of the identified proteins did not change during the course of infection. At 8 h p.i., the abundance of only 13 proteins had significantly changed with a Benjamini-Hochberg FDR of 0.05. At 10 and 12 h p.i., this number had increased to 38 and 106 proteins, respectively (Figure 1B and C and Table S4). The major-

ity of proteins with a significantly changed abundance showed a decrease during the course of CHIKV infection.

Nine proteins were found to be significantly downregulated at all investigated time points (Table 1) and for most of these there was a clear correlation between their decreasing abundance and the progression of the infection. One of these proteins, Rpb1, is a subunit of the RNA polymerase II complex (POLR2). Six other subunits from this complex were also significantly downregulated at 12 hpi (Table 2).

Virus-specific downregulation versus normal turn-over of downregulated proteins
Alphaviruses, including CHIKV, are known to induce a host translational shut-off [100]. In CHIKV-infected 293/ACE2 cells this process starts between 8 and 9 h p.i. [168]. Therefore, a decrease in the abundance of any host protein might merely be the result of its normal turn-over (degradation in the absence of synthesis). For a subset of proteins the observed degradation rate was compared to their predicted decay based on published half-lives in uninfected cells [175]. Indeed, several proteins were identified that showed no relative change in abundance at 8 h p.i., while their abundance had decreased significantly by 10 and/or 12 h p.i. (Figure S2A) with a degradation rate that was close to the predicted one. These downregulated proteins were not selected for follow-up analysis. Other proteins were specifically downregulated as a result of the infection, as their abundance decreased much more rapidly than predicted (Figure S2B).

Effect of overexpression of host proteins that were downregulated during CHIKV infection

Four significantly downregulated proteins, Rnd3, DDX56, Plk1 and UbcH10, were chosen for follow-up experiments to more extensively assess their role in CHIKV replication. Rnd3 and UbcH10 showed significant decreased abundance at all three time points that were studied (Table 1). DDX56 showed significant decreased abundance at 10 ($\log_2$ ratio -0.71) and 12 h p.i. ($\log_2$ ratio -0.64) and Plk1 at 12 h p.i. (log2 ratio -1.01) (Table S4). Plasmids expressing FLAG-tagged versions of these four proteins were transfected into 293/ACE2 cells that were subsequently infected with a GFP-expressing CHIKV reporter virus (Figure S1A). This virus was previously characterized and showed a growth phenotype similar to the parent CHIKV strain without the reporter [168]. The abundance of the FLAG-tagged host factor and GFP reporter protein was analyzed by FACS. FLAG-tagged ubiquitin (Ub) was used as a negative control. Abundance of this protein did not change during CHIKV infection as $\log_2$ ratios for this protein were -0.10, -0.23 and -0.33 at 8, 10 and 12 h p.i., respectively, which was expected based on its half-life of 11h [175]. A GFP-expressing adenovirus was used as an additional negative control, as we anticipated only a very limited overlap between the host factors involved in the replication of this DNA virus and those involved in CHIKV replication.

Protein	Accession number	8 h p.i.		10 h p.i.		12 h p.i.	
		Average log$_2$ ratio (var)	Counts/ replicate	Average log$_2$ ratio (var)	Counts/ replicate	Average log$_2$ ratio (var)	Counts/ replicate
CCN2/CTGF	P29279	**-2.47** (0.24)	2;8	**-3.04** (0.42)	3;9	**-2.92** (0.13)	6;5
THBS1	P07996	**-0.72** (0.03)	22;37	**-1.45** (0.65)	31;37	**-1.91** (0.50)	23;23
NSA2	O95478	**-0.65** (0.06)	3;17	**-1.56** (0.09)	12;16	**-1.79** (0.04)	9;14
RPB1/POLR2	P24928	**-0.83** (0.02)	16;34	**-1.65** (0.12)	33;23	**-1.75** (0.13)	19;22
RND3/ARHE	P61587	**-0.67** (0.01)	2;11	**-1.14** (0.05)	5;11	**-1.44** (1.12)	4;3
CCN1/CYR61	O00622	**-0.70** (0.02)	59;87	**-1.00** (0.03)	64;92	**-1.18** (0.01)	52;50
AMOTL2	Q9Y2J4	**-0.51** (0.00)	41;42	**-0.81** (0.02)	53;48	**-1.05** (0.01)	51;47
UBCH10/UBE2C	O00762	**-0.43** (0.00)	13;15	**-0.70** (0.06)	18;22	**-0.94** (0.03)	25;20
HMOX1	P09601	**-0.66** (0.04)	8;10	**-0.88** (0.21)	8;7	**-0.58** (0.00)	12;6

Table 1: Proteins that were significantly downregulated at all time points. Var stands for variance. Counts/replicate represents the number of ratio counts that was used from each biological replicate to determine the median ratio.

Subunit	Accession number	8 h p.i. Log$_2$ ratio (var)	10 h p.i. Log$_2$ ratio (var)	12 h p.i. Log$_2$ ratio (var)
RPB1 (POLR2A)	P24928	**-0.83*** (0.02)	**-1.65*** (0.12)	**-1.75*** (0.13)
RPB2 (POLR2B)	P30876	**-0.04** (0.15)	**-0.77** (0.58)	**-0.66*** (0.00)
RPB3 (POLR2C)	P19387	**0** (0.23)	**-0.34** (0.17)	**-0.76*** (0.02)
RPB4 (POLR2D)	O15514	n.d.	n.d.	n.d.
RPB5 (POLR2E)	P19388	**-0.41** (0.05)	**-0.73*** (0.03)	**-0.78*** (0.04)
RPB6 (POLR2F)	P61218	n.d.	n.d.	n.d.
RPB7 (POLR2G)	P62487	**-0.18** (0.00)	**-0.15** (0.01)	**-0.34** (0.10)
RPB8 (POLR2H)	P52434	**-0.09** (0.00)	**-0.17** (0.04)	**-0.14** (0.03)
RPB9 (POLR2I)	P36954	**-0.46** (n/a)	**-0.80*** (0.09)	**-0.83*** (0.04)
RPB10 (POLR2L)	P62875	**-0.24** (0.04)	**-0.68*** (0.10)	**-0.71*** (0.08)
RPB11 (POLR2J, POLR2J2, POLR2J3)	Q9GZM3	**-0.04** (0.19)	**-0.53** (0.00)	**-0.57*** (0.14)
RPB12 (POLR2K)	P53803	**-0.11** (0.00)	**0.07** (n/a)	**-0.55** (0.06)

Table 2: Relative abundance of RNA polymerase II subunits during CHIKV infection. Average log2 ratios for the individual subunits of the RNA polymerase II complex are given for each time point post infection. Var stands for variance. * indicates that the subunit was significantly downregulated (Benjamini-Hochberg FDR 0.05). n.d. indicates that this subunit was not identified during the experiment. n/a indicates that this subunit was only quantified in one biological replicate.
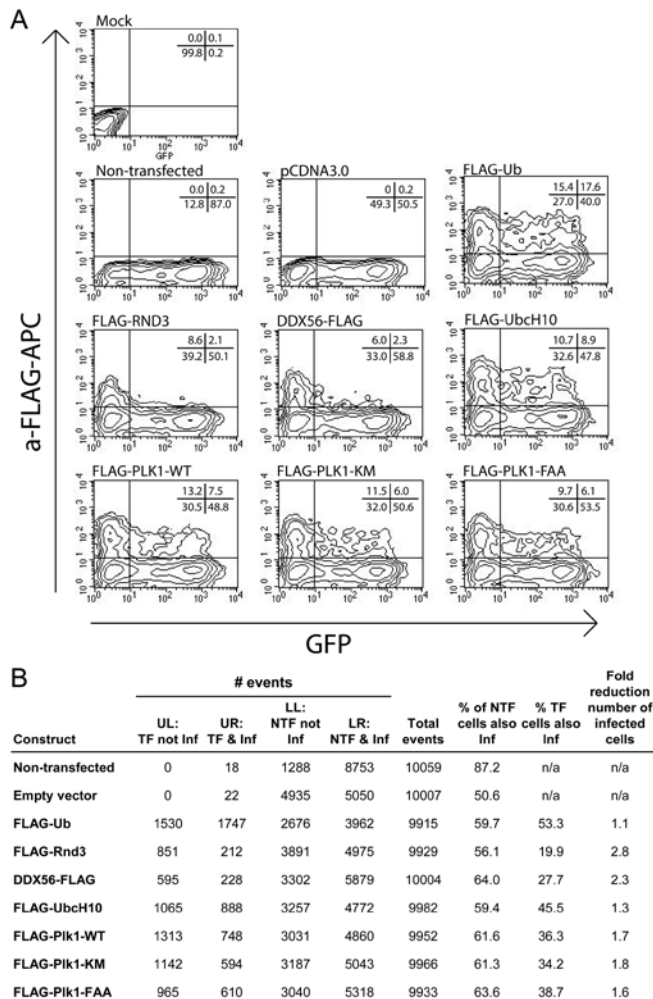
2

A Mock, Non-transfected, pCDNA3.0, FLAG-Ub, FLAG-RND3, DDX56-FLAG, FLAG-UbcH10, FLAG-PLK1-WT, FLAG-PLK1-KM, FLAG-PLK1-FAA

a-FLAG-APC (y-axis), GFP (x-axis)

B

Construct	# events				Total events	% of NTF cells also Inf	% TF cells also Inf	Fold reduction number of infected cells
	UL: TF not Inf	UR: TF & Inf	LL: NTF not Inf	LR: NTF & Inf				
Non-transfected	0	18	1288	8753	10059	87.2	n/a	n/a
Empty vector	0	22	4935	5050	10007	50.6	n/a	n/a
FLAG-Ub	1530	1747	2676	3962	9915	59.7	53.3	1.1
FLAG-Rnd3	851	212	3891	4975	9929	56.1	19.9	2.8
DDX56-FLAG	595	228	3302	5879	10004	64.0	27.7	2.3
FLAG-UbcH10	1065	888	3257	4772	9982	59.4	45.5	1.3
FLAG-Plk1-WT	1313	748	3031	4860	9952	61.6	36.3	1.7
FLAG-Plk1-KM	1142	594	3187	5043	9966	61.3	34.2	1.8
FLAG-Plk1-FAA	965	610	3040	5318	9933	63.6	38.7	1.6

Figure 3: (Over)expression of different FLAG-tagged host factors, followed by infection with a GFP-expressing CHIKV at 10 h p.t. at an MOI of 5. At 10 h p.i., the cells were fixed, permeabilized and stained with anti-FLAG-APC. APC staining and GFP expression were quantified by FACS. (*A*) FACS contour plots. The mock plot displays non-transfected cells that were not infected. In all other plots cells were infected with reporter virus CHIKV-LS3-GFP. pCDNA3.0 was used as an empty vector control and FLAG-Ub as a negative control. Each plot contains approximately 10,000 events. The x-axis shows the level of GFP expression, the y-axis shows the intensity of APC staining. The numbers in the upper right of each panel indicate the percentage of cells in each quadrant. Contour plots of control cells that were transfected with FLAG-tagged host factors, but were not infected with CHIKV can be found in Figure S4. (*B*) Summary of FACS data obtained with CHIKV-infected cells transfected with plasmids to overexpress FLAG-tagged host factors. The number of events that was recorded in each quadrant during FACS analysis is shown. The percentage of CHIKV infected cells was calculated using the total number of transfected cells and total number of nontransfected cells. The fold reduction in number of infected cells is shown in the last column. TF, Transfected, NTF, nontransfected, Inf, infected.

Transfection of cells in general had a (nonspecific) negative effect on the susceptibility of cells to CHIKV infection, as only 50.6% of cells transfected with an empty vector became infected with CHIKV, compared to 87.2% of nontransfected cells (Figure 3). This (nonspecific) effect on susceptibility was not observed when cells were infected with the adenovirus (Figure S3). Overexpression of the negative control FLAG-Ub had no effect on the number of GFP-positive cells during either CHIKV (Figure 3) or adenovirus (Figure S3) infection.

Overexpression of FLAG-Rnd3 and DDX56-FLAG inhibited CHIKV replication, resulting in a 2.8- and 2.3-fold reduction in the number of GFP-positive cells, respectively. Overexpression of FLAG-Plk1-WT and FLAG-UbcH10 had a more modest effect, resulting in a 1.7- and 1.3-fold reduction in the number of GFP-expressing cells, respectively. Mutations in the kinase domain and Polo-box domain of Plk1 did not change the effect that overexpression of this protein had on CHIKV infection (Figure 3 and Figure S4). The FACS analysis revealed that especially cells expressing high levels of the FLAG-tagged host proteins were less likely to become infected with CHIKV. The replication of adenovirus, determined by quantifying GFP reporter abundance, was not affected by the overexpression of FLAG-Plk1-WT, FLAG-UbcH10 and DDX56-FLAG. Compared to nontransfected cells in the same sample, FLAG-Rnd3 overexpression did reduce the infection rate of adenovirus-infected cells by 1.5 fold, and there was a correlation between the level of FLAG-Rnd3 abundance and the reduction in GFP reporter abundance (Figure S3).

DISCUSSION

Minimal changes in the cellular proteome in response to CHIKV infection

To obtain more insight into the complex interactions between CHIKV and its host cell, we have applied SILAC-based quantitative proteomics to analyze changes in abundance in the cellular proteome during the stages of infection with high viral RNA and protein expression. At 8 h p.i. almost 3000 proteins could be quantified and, surprisingly, the cellular proteome had remained largely unchanged. Mass spectrometry and WB analysis revealed that cells at this time point already contained substantial quantities of viral protein (0.6% of total protein) and microscopy showed that almost all cells had been successfully infected. This demonstrated that the limited response of the cells at this time point was not due to a low percentage of infected cells or slow progression of the infection. The number of affected proteins increased from only 13 at 8 h p.i. to 106 at 12 h p.i. Strikingly, the majority of these proteins was downregulated and the abundance of none of the upregulated proteins increased by more than 2-fold.

The innate immune response is rapidly activated upon detection of dsRNA in the cytoplasm [181] and we thus expected to observe upregulation of at least some in-

nate immunity-related proteins, but this was not the case. However, it was previously shown that CHIKV infection leads to a strong transcriptional upregulation of interferon regulatory factor 3 (IRF3)-dependent genes, such as interferon-β and Viperin in human fibroblasts, while this increase in mRNA levels did not result in increased protein abundance [182]. This suggests that CHIKV very efficiently evades or blocks the innate immune response and our data support these earlier observations. The report by White et al. and our current study emphasize the importance of directly measuring protein abundance instead of mRNA levels in virus-infected cells, as changes in the latter might not translate into altered protein levels. Alphaviruses are known to induce a host translational shut-off [100] and CHIKV-LS3 does so in 293/ACE2 cells from 8 h p.i. onward [168]. By considering a block of translation starting at 8 h p.i., predicting the decay of a protein in the following hours based on its published half-life, and comparing this to its real decrease in abundance, we could distinguish between proteins that merely decreased in abundance due to normal turnover and those that were degraded faster than predicted during CHIKV infection. Many other viruses also cause a host shut-off, and our study shows that it is important to take this into consideration when analyzing proteomics data to avoid follow-up research on false positive hits.

During the follow-up studies we discovered that the 293/ACE2 cell line was not of human but of non-human primate origin, most likely from an Old-World monkey related to *Papio anubis* and *Macaca mulatta*. The *M. mulatta* genome has been sequenced and this revealed that 89% of human-macaque orthologs differ at the amino acid level and the average human gene differs by 12 nonsynonymous and 22 synonymous substitutions from its macaque ortholog [183]. Since a single amino acid change in a peptide already prevents matching to the corresponding protein sequence in the search database, this could theoretically have lowered the number of identified proteins resulting from our search against a human proteome database. To exclude this, a second search against the uniprot *M. mulatta* database (69,931 entries, reviewed:358 (Swiss-Prot), unreviewed: 69,573 (TrEMBL); data not shown) was performed. This resulted in a similar number of identifications (4818) and proteins with significantly changed abundance (9, 42, 91, at 8, 10 and 12 h p.i., respectively), the majority of those being the same as in the original search against the human database. An advantage of using a human proteome database for the search, however, is that it is much better annotated than the non-human primate databases and results in less identifications of uncharacterized proteins. For the four proteins that were selected for our overexpression studies we additionally aligned sequences from *Homo sapiens, Macaca mulatta, Papio anubis, Pan troglodytes, Gorilla gorilla* and *Pongo abelii* (Figure S5). Mapping of the peptides that were identified with LC-MS/MS to these alignments indicated that each of these proteins was identified with

at least 4 peptides that are 100% conserved across these six species. From these findings we conclude that the initial identifications and quantifications are valid.

Downregulation of host proteins is beneficial for CHIKV infection

Although the overall effect of CHIKV infection on the cellular proteome was rather limited, several proteins were specifically downregulated. This could be mediated by the virus or might be a (antiviral) response of the cell. Four of these downregulated proteins were chosen for follow-up studies and the effect of their overexpression on CHIKV infection was analyzed. Figure S6 summarizes the possible roles these proteins play during CHIKV infection.

Rnd3

Of the four host factors, overexpression of FLAG-RND3 had the strongest negative effect on CHIKV-LS3-GFP infection, leading to 2.8-fold reduction in the number of GFP positive cells, compared to the non-transfected cells in the same sample. Rnd3 was also one of the most downregulated proteins during CHIKV infection. Expression of Rnd3 appears to be unfavorable for a broader range of viruses, as its overexpression also reduced adenovirus infection. The protein belongs to a subfamily of Rho GTPases that constitutively bind GTP. Its activity does not depend on a GDP/GTP switch, but is regulated by expression, localization and phosphorylation [184]. Many cellular functions have been ascribed to Rnd3 and the protein could inhibit CHIKV infection in several ways.

Rnd3 might limit CHIKV through its effect on the cytoskeleton since it is an important antagonist of RhoA induced stress fibers [184] and alphavirus replication complexes interact with cytoskeleton filaments [55, 185].

Another function of Rnd3 is the regulation of cap-dependent translation of mRNAs with highly structured 5′untranslated regions (UTRs) [186]. Rnd3 prevents the phosphorylation of 4E binding proteins (4E-BP1), which limits the release of bound eukaryotic initiation factor 4E (eIF4E). Low availability of eIF4E affects the activity of the eIF4F translation initiation complex, resulting in the translation of mainly mRNAs with 5′ UTRs containing little secondary structure. Consequently, Rnd3 would inhibit the translation of mRNAs with complex structured 5′ UTRs. +RNA virus genomes, including that of CHIKV, generally have highly structured 5′ UTRs and therefore Rnd3-mediated inhibition of translation of such mRNAs could be responsible for its negative effect on CHIKV replication. Degradation of Rnd3 during CHIKV infection might be a strategy by which the virus ensures efficient translation of its genome.

DDX56

Overexpression of DDX56-FLAG led to a 2.3-fold reduction in the percentage of CHIKV-infected cells, but had no effect on adenovirus infection, suggesting this was not due

to e.g. a general effect on cellular homeostasis. DDX56 is a nucleolar ATP-dependent helicase from the DEAD-box family and is involved in ribosome biogenesis [187, 188]. During CHIKV infection, DDX56 appears to have an antiviral role, as its overexpression decreased the number of infected cells. DDX56 degradation during infection might be induced by CHIKV nsP2 or C since a fraction of these proteins localizes to the nucleolus through nucleolar import signals [76, 108].

DDX56 has previously been shown to be a proviral host factor for West Nile virus (WNV) and human immunodeficiency virus type 1 (HIV-1) [189-191]. During WNV infection DDX56 relocalizes from the nucleolus to the cytosol where it interacts with WNV capsid protein and its helicase domain is essential for the assembly of infectious virions [189, 190]. DDX56 is one of several DEAD-box helicases that interact with the HIV-1 Rev protein and synergistically enhance Rev-dependent HIV-1 RNA export from the nucleus to the cytoplasm. DDX56 promotes relocalization of Rev from the nucleolus to the nucleus [191].

Plk1

Compared to Rnd3 and DDX56, overexpression of FLAG-Plk1-WT had only a modest effect on CHIKV-driven GFP reporter protein expression, but still reduced the percentage of infected cells by 1.7-fold. Overexpression of this protein had no effect on adenovirus infection. In an attempt to pinpoint which of the known domains is responsible for the antiviral effect of (overexpressed) Plk1, we analyzed the effect of a kinase defective and a polo-box domain Plk1 mutant. Overexpression of these variants had the same negative effect on CHIKV replication as WT Plk1, suggesting that the kinase- and polo-box domain are not responsible for the antiviral properties of this protein.

Plk1 was previously described as a host factor for other viruses and was therefore selected for follow-up studies [192-195]. It is a serine/threonine mitotic kinase that is expressed during the S, G2 and M phase of the cell cycle and is a key regulator of several steps during mitosis [196]. The C-terminal polo-box domain of Plk1 targets the N-terminal kinase domain to specific targets and subcellular locations through binding of phosphorylated residues on a specific motif [196]. Plk1 has also been described as a regulator of interferon induction by mitochondrial antiviral signaling protein (MAVS), which plays a crucial role in the signal transduction that leads to induction of interferon β in response to viral infection. Overexpression of Plk1 has been shown to inhibit the induction of interferon β [197] and therefore overexpression of Plk1 could be expected to have a proviral effect. However, during CHIKV infection this was not the case, conceivably because CHIKV employs alternative mechanisms to evade the interferon response, e.g. through inhibition of JAK-STAT signaling by nsP2 [198].

UbcH10

UbcH10 was chosen for follow-up studies because it was also significantly downregulated during infection with equine arteritis virus, an unrelated +RNA virus (unpublished data). However, overexpression of FLAG-UbcH10 had a limited impact on CHIKV infection, leading to a 1.3-fold reduction of the number of CHIKV-infected cells. UbcH10 is the E2 ubiquitin-conjugating enzyme of the anaphase-promoting complex (APC), an E3 ubiquitin ligase responsible for the ubiquitination of cyclins [199]. It is an essential factor for cell cycle progression and plays a role at several cell cycle checkpoints [200]. Its effect on CHIKV replication might be indirect and through its general effect on the cell (cycle), although it did not affect adenovirus infection.

POLR2 complex is specifically downregulated during CHIKV infection

We observed that several subunits of the POLR2 complex were downregulated during CHIKV infection. This confirmed the results of a previous study that showed that nsP2 of CHIKV, Sindbis virus and Semliki Forest virus induced degradation of Rpb1, the catalytic subunit of the POLR2 complex, through a proteasome-dependent pathway. This degradation results in a host transcriptional shut-off, which likely serves to inhibit cellular antiviral responses [63]. We identified 10 of the 12 subunits of the POLR2 complex and observed the progressive downregulation of most of the components during CHIKV infection (Table 2). By 12 h p.i., 7 subunits (Rpb1-3, Rpb5 and Rpb9-11) had been significantly downregulated (1.5- to 3.4-fold). The whole complex appears to be a target for degradation during CHIKV infection. Degradation of Rpb1 occurs most rapidly and is, therefore, most likely the first subunit to be degraded, closely followed by Rpb5 and Rpb9. The loss of Rpb2, Rpb10 and Rpb11 started at 10 h p.i., followed by that of Rpb3 and Rpb12 at 12 h p.i. Whether these proteins are directly targeted by the virus or are degraded because Rpb1 is lacking is not clear at the moment. We have previously shown that the host transcriptional shut-off triggered by CHIKV-LS3 in 293/ACE2 cells can be observed as early as 6 h p.i. and increases until there is almost no host transcription left at 12 h p.i. [168]. These findings are supported by the progressive degradation of the POLR2 subunits observed here. Rpb 5, 6, 8, 10 and 12 are also part of the POLR1 complex, which synthesizes ribosomal RNA, and the POLR3 complex, which produces tRNAs [201]. Consequently, the downregulation of these subunits might result in a very broad effect on cellular transcription and (indirectly on) translation.

Comparison with previously published CHIKV proteomics data sets

Previously conducted proteomics studies showed a very limited overlap in the lists of proteins with significantly changed abundance during CHIKV infection [111-116, 119, 120]. This is most likely due to differences in the experimental systems (mice, patient serum, cell culture), experimental conditions (virus strain, MOI, time post infection), and

the methods (database, software, statistics) used to analyze the data. The *in vivo* infection studies (e.g. with mice or serum) are based on relatively heterogeneous biological material (various cell types, infected and non-infected cells) and are complicated by the fact that not only the direct effects of CHIKV infection are observed, but also all kinds of indirect effects such as tissue damage and the response of uninfected cells to cytokines produced elsewhere in the infected organism. In addition, the applied proteomics techniques might be a factor in the limited overlap between the different studies. For example, in the study that combined iTRAQ and 2D-DIGE, the two datasets only had a few proteins in common, despite the fact that they were generated using the same biological samples [114]. The low number of identifications generated with the 2D-DIGE approach (32 proteins) was probably the main contributor to this limited overlap.

A comparison of our results with previously reported datasets identified a few proteins that were also differentially expressed in other studies. Synthenin-1 was down regulated 2-fold at 12 h p.i. in our study and was 5-fold downregulated 2 days post infection in a human microglial cell line in the study by Abere et al.[111]. CD63 and HTRA1 were both downregulated, while HYRC and DNAJC13 were both upregulated in our study and in the study by Wikan et al., while agrin was downregulated in our study but upregulated in their study [120]. HKQ and LAMB1 were downregulated in our study and 2 days post infection in the study by Frasier et al., while SNW1 was downregulated in our study but up regulated in their study [114]. In line with our results, three other studies also observed that the majority of significantly changed proteins were downregulated [111, 114, 119].

We set out to analyze the direct effects of CHIKV-infection on cells by specifically analyzing synchronously infected cells during the first round of replication in a well-defined and characterized system. We used a SILAC-based time course so we could track abundance during the course of infection. This increases the chance of identifying differentially expressed proteins and decreases the chance of finding false positives, which are unlikely to be the same at multiple time points. A similar approach was recently used to study temporal changes during CMV infection [202]. An additional advantage of our time course study is that kinetics of the differentially regulated proteins could be compared to published half-lives. Several of the other studies analyzed samples at much later time points post infection, probably long after the onset of the translational shut-off. Therefore, the downregulation of many of the proteins identified in these studies might have merely been the result of normal turnover (in the absence of translation).

CONCLUSION

The data presented in this study show that the immediate impact of CHIKV infection on host cell protein abundance is rather limited. This might be because the cellular response to the infection is limited at the level of protein abundance, or because it is effectively suppressed by CHIKV, e.g. through the virus-induced translational and transcriptional shut-off, rendering the cell unable to upregulate protein expression. The previously described CHIKV nsP2-induced degradation of an RNA polymerase subunit [63] and the downregulation of most components of the POLR2 complex that we observed in our study, likely contribute importantly to the host transcriptional shut-off and suggests that also other proteome changes that we describe are relevant.

The majority of the differentially expressed proteins were downregulated during the infection, likely to manipulate the intracellular environment in a way that is beneficial for CHIKV replication. In line with this assumption, we discovered that overexpression of four of these downregulated proteins had a negative effect on the ability of CHIKV to establish a productive infection in these cells. How these four proteins affect CHIKV replication, directly or indirectly, and whether they are specifically targeted for degradation by CHIKV, as previously described for Rpb1 [63], remain interesting questions for future studies.

Although CHIKV infection only led to a limited cellular response at the level of protein abundance, it remains very well possible that the cell responds to infection in ways that cannot be detected with a quantitative proteomics approach that only studies changes in total protein abundance. For example, this approach will not detect changes in subcellular localization or PTMs, such as phosphorylation, that can occur rapidly during (antiviral) signaling. Our findings advance the understanding of the response to CHIKV infection at the cellular level and this information might be used in future studies aimed at developing 'host-directed' antiviral strategies.

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository [203] with the dataset identifier PXD001330.

Centre, Japan) for generously supplying plasmids and the PRIDE Team for making our data publicly available.

2

SUPPORTING INFORMATION

Supporting information tables

RND3-fw	5'-TGCAGGAATTCCTAAGGAGAGAAGAGCCAGCCAG
RND3-rev	5'-CCAGAATTCTCACATCACAGTGCAGCTCTTCG
DDX56-fw1	5'-CGGCAATTGAGCGCCATGGAGGACTCTGAAGCACTG
DD56-rev1	5'-CCGCCTGCAGGTCAGGAGGGCTTGGCTGTGGGTCTG
DDX56-fw2	5'-TGGGATCCAGCGCCATGGAGGACTCTGAAGCACTG
DDX56-rev2	5'-CCAGATATCCGAGCCTGCAGGCTTGTCATC
FAA-QC1	5'-ATCTTCTGGGTCAGCAAGTTCGCGGACTATTCGGACAAG
FAA-QC2	5'-CTTGTCCGAATAGTCCGCGAACTTGCTGACCCAGAAGAT
FAA-QC3	5'-TATGGCCTTGGGTATCAGGCGTGTGACAACAGTGTGGG
FAA-QC4	5'-CCCACACTGTTGTCACACGCCTGATACCCAAGGCCATA

**Supporting Information Table S1: List of oligonucleotides**

Supporting Information Table S2: Excel table with protein groups identified during the SILAC time course. Contaminants, proteins identified with the concatenated reversed database, proteins that were only identified by site and proteins that were identified with <2 peptides were removed. Protein groups that were used for further analysis (variance <0.25) were placed in different sheets for each time point. Identified CHIKV proteins were placed in a separate sheet. This table can be downloaded on the website of the journal: http://onlinelibrary.wiley.com/doi/10.1002/pmic.201400581/suppinfo.

	8 h p.i.		10 h p.i.		12 h p.i.	
	replicate 1	**replicate 2**	**replicate 1**	**replicate 2**	**replicate 1**	**replicate 2**
Nonstructural polyprotein (%)	0.10	0.24	0.34	0.23	0.16	0.19
Structural polyprotein (%)	0.32	0.55	0.92	0.73	0.74	0.54
Total CHIKV protein (%)	0.43	0.79	1.26	0.95	0.90	0.72
Average total CHIKV protein (%)	0.61		1.11		0.81	

Supporting Information Table S3: Relative protein intensity in percentages originating from CHIKV proteins. To estimate what percentage of the total protein in infected cells was of viral origin, the sum of the intensities of all protein groups that originated from the CHIKV nonstructural polyprotein, structural polyprotein and total CHIKV protein was divided by the sum of all intensities for a given biological replicate. The average of the total amount of CHIKV protein in the 2 biological replicates at each time point is also shown. The amount of structural proteins was about 2 to 4-fold higher than that of the of the nsPs at the various time points post infection.

Supporting Information Table S4: Excel table with the proteins that showed significantly different abundance during CHIKV infection. Each time point is placed in a different sheet. This table can be downloaded on the website of the journal: http://onlinelibrary.wiley.com/doi/10.1002/pmic.201400581/suppinfo.
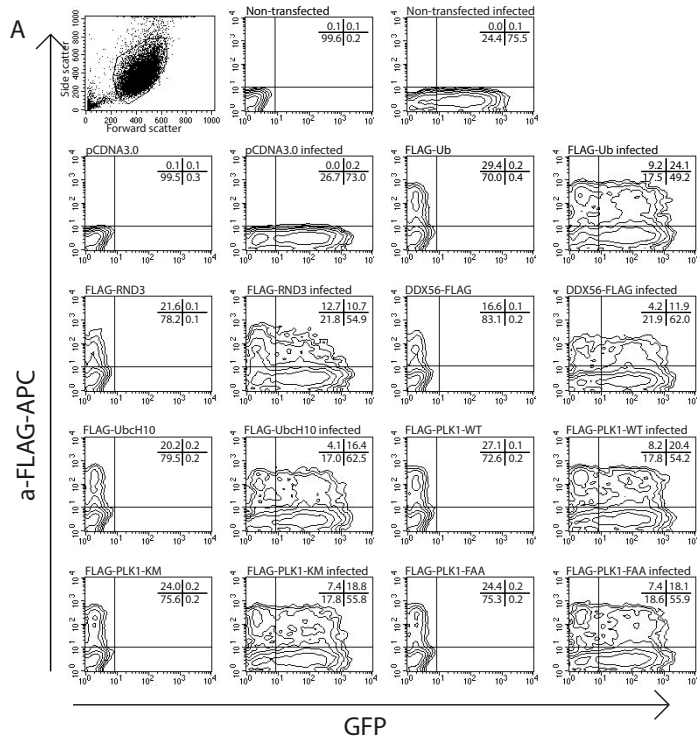
Supporting information figures
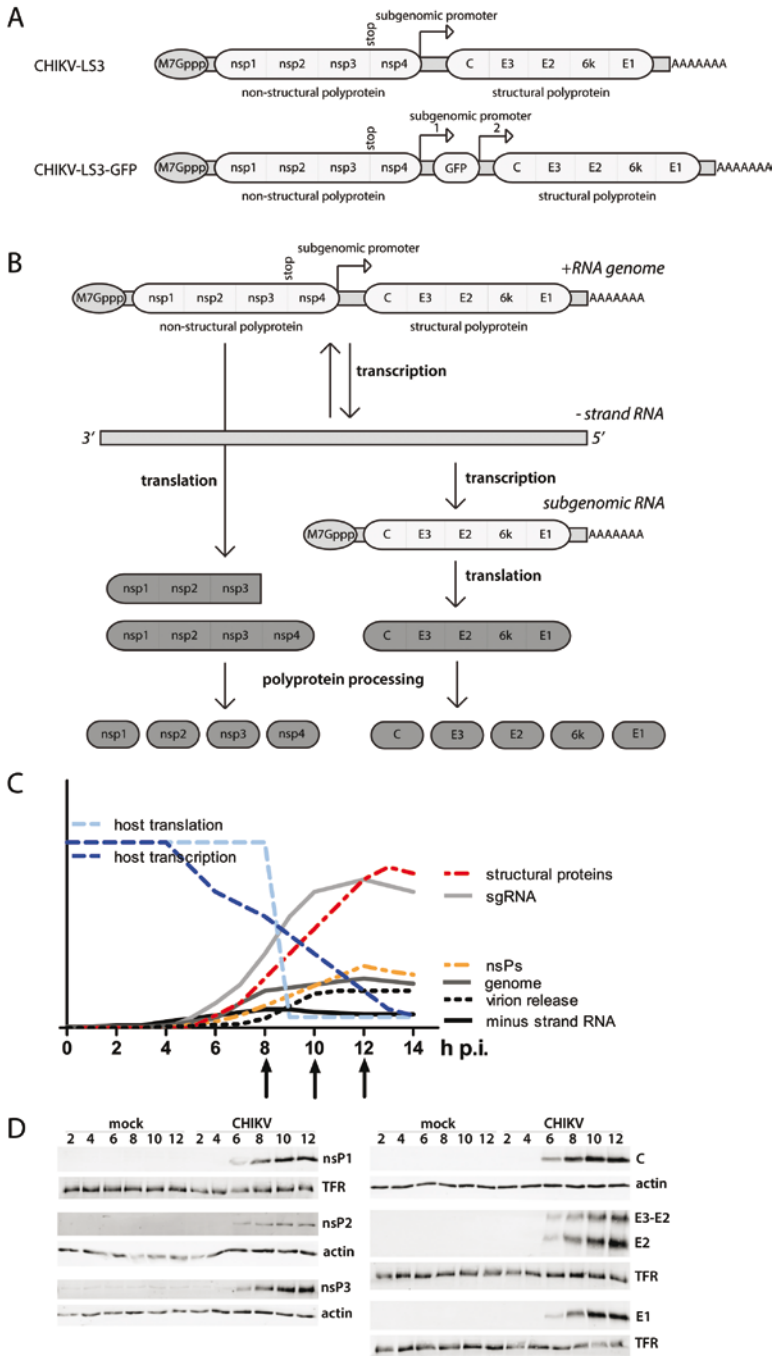
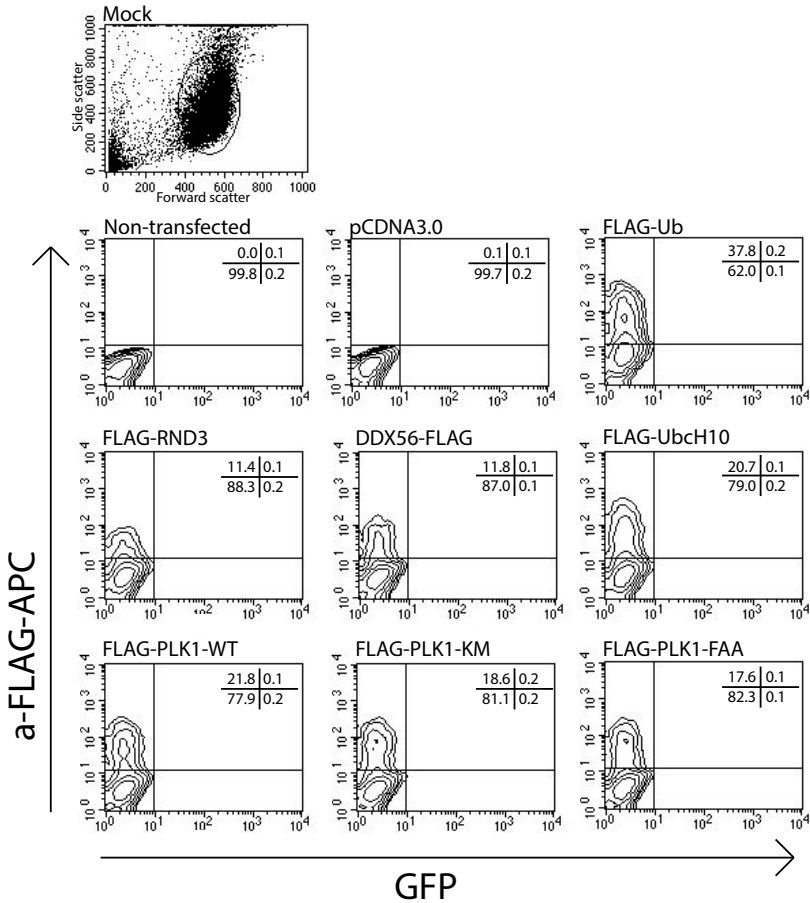**Supporting Information Figure S1: Schematic overview of CHIKV genome organization and protein expression**. (*A*) Schematic representations of CHIKV constructs used in this paper. CHIKV has a single-stranded positive polarity RNA genome that encodes for two polyproteins. The genome is 5' capped and has a 3' poly-A tail. The CHIKV-LS3-GFP genome contains a second subgenomic promotor from which GFP is expressed. (*B*) Schematic overview of viral protein expression. Upon entering the cell the CHIKV genome is directly translated by cellular ribosomes into the nonstructural polyprotein. The RNA dependent RNA polymerase nsP4 is only expressed upon translational readthrough of the stop codon between nsP3 and nsP4. The protease domain in nsP2 processes the polyprotein into the different subunits. The viral replication complex transcribes minus (-) strands from the +RNA genome. These are used as templates to transcribe new CHIKV genomes. A subgenomic promotor on the –strand is used to transcribe a subgenomic RNA that is translated into the structural polyprotein, which is proteolytically processed into individual subunits by the protease domain in C and host proteases. (*C*) Schematic overview of the accumulation of negative strand, genomic RNA, subgenomic RNA, nonstructural and structural proteins and the release of infectious virions, during the replication of CHIKV in 293/ACE2 cells. In addition the progression of the transcriptional and translational host shut-off are shown. Plots are based on data from Scholte et al. (RNA, viral titers, host shut-off) [168], and this paper (protein; see panel D and Supplemental Table 3). Arrows indicate the time points that were analyzed in our proteomics experiment. (*D*) Western blot analysis of samples from CHIKV-LS3 infected 293/ACE2 cells showing the expression of nsP1, nsP2, nsP3, C, E1, E3-E2 and E2 from 2-12 h p.i. Mock-infected cells were included as controls to demonstrate the specificity of the antibodies used.
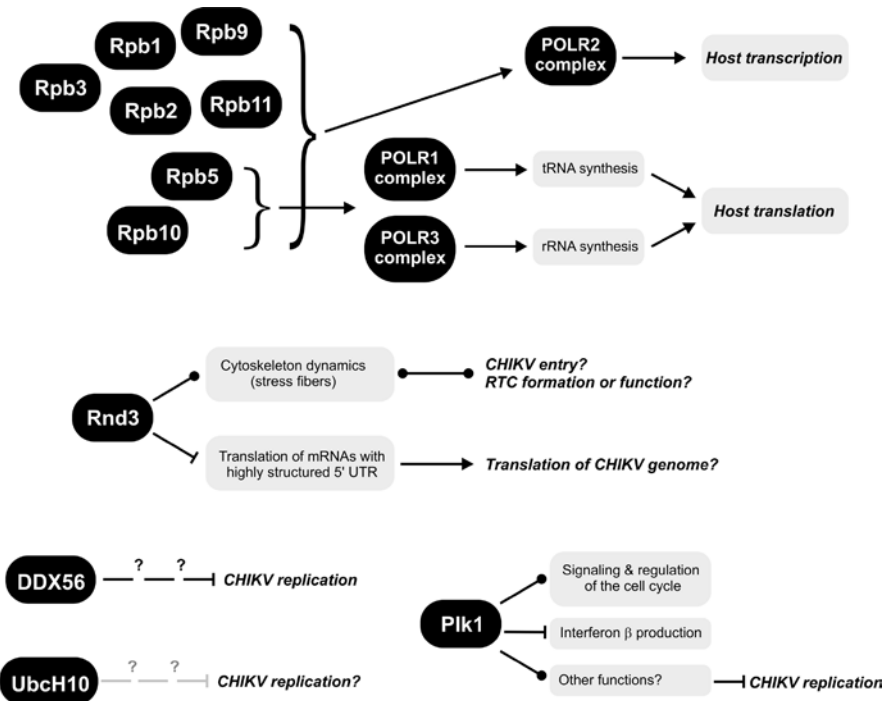
A

COL4A1

LAMB1

NOSIP

B

MDA9

HKQ

rpb10

- ● theoretical decay
- ■ observed decay
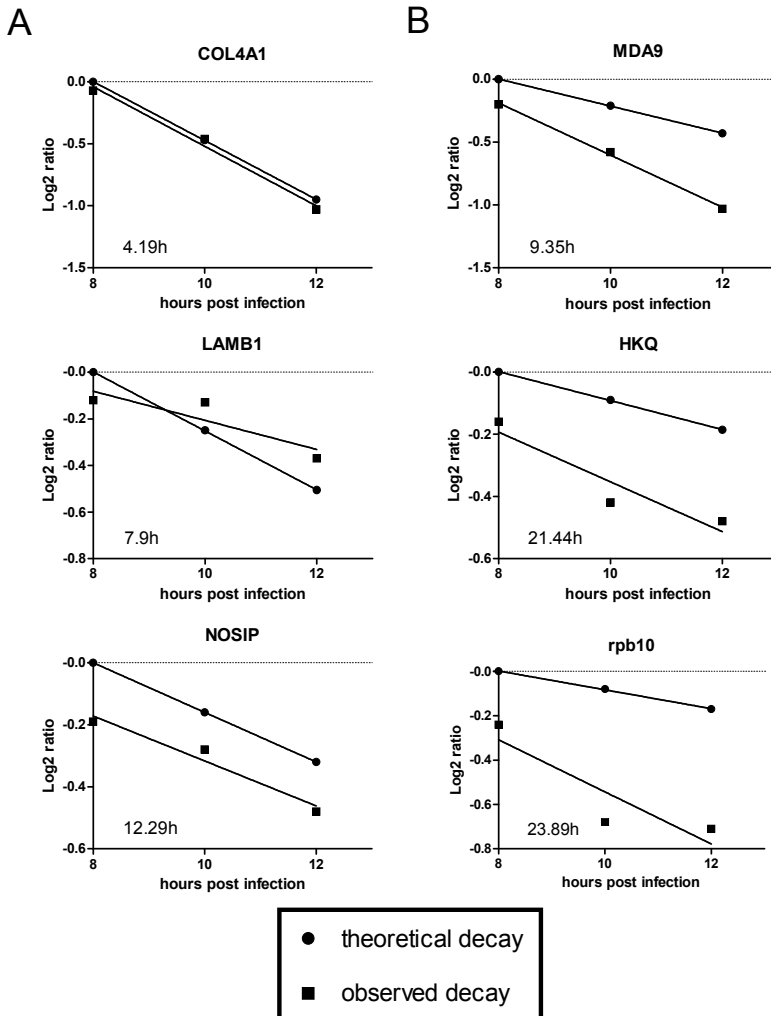
Supporting Information Figure S2: Comparison of observed and predicted log2 ratios over time of 6 proteins that were hardly affected at 8 h p.i. but were significantly downregulated later during CHIKV infection. The predicted change in relative abundance was calculated using published half-lives (indicated in each graph) [175] and assuming virus-induced host translational shut-off starts 8 h p.i. (*A*) Proteins of which the observed decay was very similar to the predicted curve. (*B*) Proteins that were degraded faster than predicted during CHIKV infection.

B

| | # events | | | | | | | Fold reduction |
Construct	UL: TF not Inf	UR: TF & Inf	LL: NTF not Inf	LR: NTF & Inf	Total events	% of NTF cells also Inf	% TF cells also Inf	number of infected cells
Non-transfected	1	13	2538	7870	10422	75.6	n/a	n/a
Empty vector	2	22	2767	7561	10352	73.2	n/a	n/a
FLAG-Ub	953	2506	1821	5119	10399	73.8	72.4	1.0
FLAG-Rnd3	1321	1112	2270	5719	10422	71.6	45.7	1.6
DDX56-FLAG	439	1229	2267	6425	10360	73.9	73.7	1.0
FLAG-UbcH10	430	1706	1761	6500	10397	78.7	79.9	1.0
FLAG-Plk1-WT	849	2110	1776	5607	10342	75.9	71.3	1.1
FLAG-Plk1-KM	769	1962	1841	5773	10345	75.8	71.8	1.1
FLAG-Plk1-FAA	763	1876	1929	5791	10359	75.0	71.1	1.1

Supporting Information Figure S3: (Over)expression of different FLAG-tagged host factors, followed by infection with a GFP-expressing AdenoV at 10 h p.t. and an MOI of 5. At 10 h p.i., the cells were fixed, permeabilized and stained with anti-FLAG-APC. APC staining and GFP expression were quantified by FACS. (*A*) The first scatter plot shows side- and forward scatter and the gate that was used to select cells to display in contour plots. The contour plots show a-FLAG-APC staining and GFP expression. For each host factor a mock- and adenovirus-infected plot is shown. pCDNA3.0 was used as an empty vector control and FLAG-Ub as a negative control. Each plot contains approximately 10,000 events. The x-axis shows the level of GFP expression, the y-axis shows the intensity of APC staining. The numbers in the upper right of each panel indicate the percentage of cells in each quadrant. (*B*) Summary of FACS data. The number of events that was recorded in each quadrant during FACS analysis is shown. The percentage of adenovirus-infected cells was calculated using the total number of transfected cells and the total number of non-transfected cells. The fold reduction in number of infected cells is shown in the last column. TF= Transfected, NTF= non-transfected, Inf= infected.

Supporting Information Figure S4: FACS contour plots of transfected uninfected control cells. 293/ACE2 cells were transfected with FLAG-tagged host factors and at 10 h p.t. mock-infected. Cells were fixed 10 h later, permeabilized and stained with anti-FLAG-APC. APC staining and GFP expression were measured by FACS. The first plot shows side- and forward scatter and the gate that was used to select cells to display in the contour plots. Each plot contains approximately 10,000 events. X-axis shows GFP expression, y-axis shows APC staining. Numbers in upper right of each panel indicate the percentage of cells in each quadrant.

Supporting Information Figure S5: Sequence alignments of the proteins that were used in the follow-up experiments of our proteomics study. Alignments were made for each of the four proteins that were selected for follow-up (overexpression) studies, using sequences from *Homo sapiens*, *Macaca mulatta*, *Papio anubis*, *Pan troglodytes*, *Gorilla gorilla* and *Pongo abelii*. Alignments were made with Geneious 7.1.5 [176]. The uniprot identifier for each sequence that was used for the alignment is shown after the species name. The level of sequence identity is shown in the bar above the sequence. The peptides identified in the LC-MS/MS analysis are indicated with bars underneath the sequence. (*A*) Rnd3, (*B*) DDX56, (*C*) UbcH10, (*D*) Plk1.

C Identity — UbcH10

1. homo sapiens_O00762
2. macaca mulatta_I2CW74
3. papio anubis_A0A096NX38
4. pan troglodytes_H2QKH0
5. gorilla gorilla_G3S6I6
6. pongo abelii_H2P240

D Identity — Plk1

1. homo sapiens_P53350
2. macaca mulatta_F7HSC1
3. papio anubis_A0A096NHF8
4. pan troglodytes_G2HFG8
5. gorilla gorilla_G3RAQ3
6. pongo abelii_H2NQF2

Supporting Information Figure S5 (continued)

Supporting Information Figure S6: Schematic overview of possible functions of significantly down-regulated proteins during CHIKV infection. The POLR2 complex is required for host transcription and degradation of its subunits results in a decrease in host transcription. The POLR1 and POLR3 complexes are involved in tRNA synthesis and rRNA synthesis, respectively, and degradation of their subunits affects host translation. Rnd3 is an antagonist of RhoA induced stress fibers and the protein might affect CHIKV entry or the replication and transcription complex (RTC). Rnd3 also inhibits translation of mRNAs with highly structured 5'UTRs and could affect translation of CHIKV genomes through this function. Overexpression of DDX56 inhibits CHIKV infection but how it affects replication is still unclear. The effect of overexpression of UbcH10 on CHIKV infection was minimal and it is still unknown how inhibition would take place. Plk1 has several functions in regulation of the cell cycle and interferon-β production. Since mutations in the kinase- and polobox-domain of this protein had no effect on the inhibitory effect that overexpression of Plk1 had on CHIKV infection it might inhibit CHIKV through yet another function.
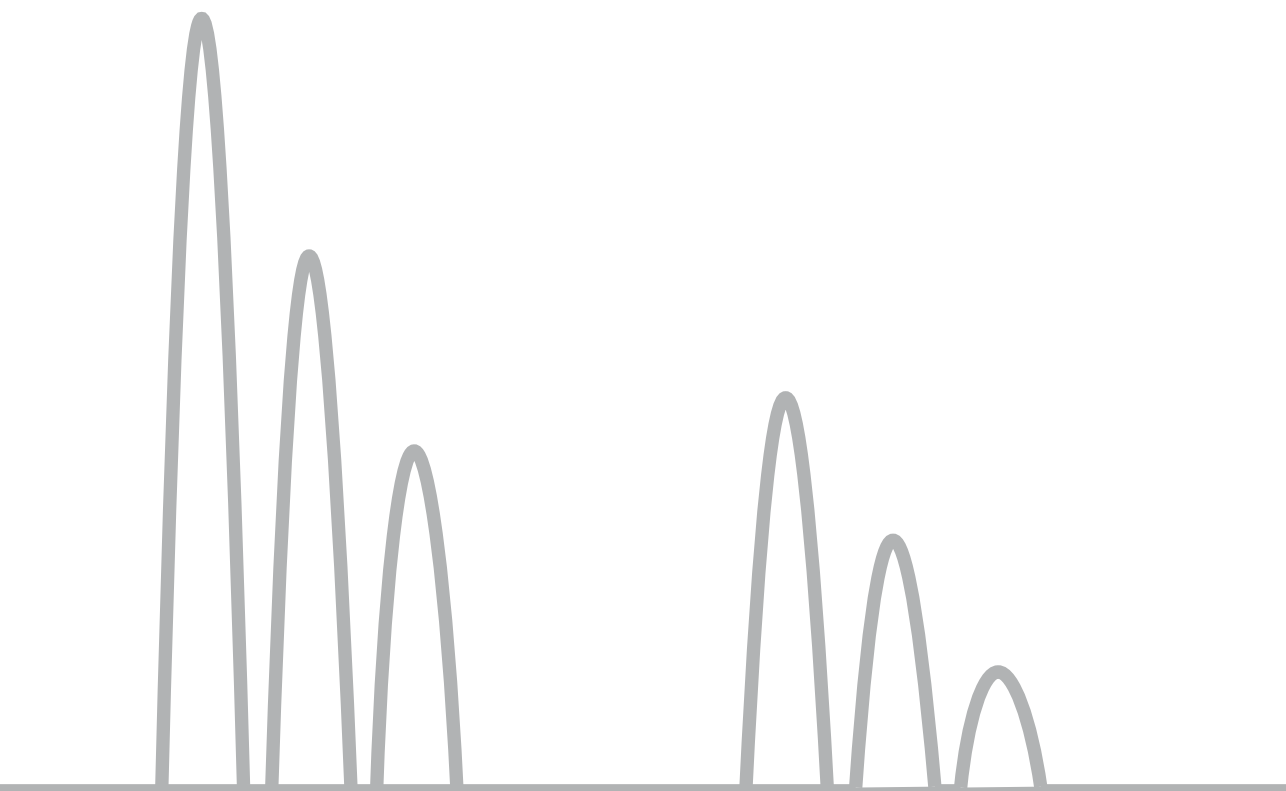
Chapter 3

Rapid and strong induction of eEF2 phosphorylation during alphavirus infection

Emmely E. Treffers[1,2], Ali Tas[1], Florine E.M. Scholte[1], Arnoud H. de Ru[2], Eric J. Snijder[1], Peter A. van Veelen[2#], Martijn J. van Hemert[1#]

[1]Molecular Virology Laboratory, Department of Medical Microbiology, Leiden University Medical Center, 2333 ZA, Leiden, The Netherlands, [2]Department of Immunohematology and Blood transfusion, Leiden University Medical Center, 2333 ZA, Leiden, The Netherlands
[#]These authors contributed equally

ABSTRACT

Chikungunya virus (CHIKV) is a reemerging alphavirus that causes a high fever, rash and severe persisting arthralgia. Since 2005, millions of people have been infected during outbreaks in Africa, Asia, and recently also in South/Central America. CHIKV is highly dependent on cellular factors during many steps of its replication cycle. Here stable isotope labeling with amino acids in cell culture (SILAC) and liquid chromatography-tandem mass spectrometry (LC-MS/MS) were used to assess temporal changes in the cellular phosphoproteome in response to CHIKV infection. Almost 3000 unique phosphorylation sites were identified and per time point the phosphorylation status of 800–1200 sites was quantified. At 2, 8, and 12 h p.i., we measured significant modulation of 10, 71, and 136 phosphorylation sites, respectively. The largest change in phosphorylation status was measured on residue T56 of eukaryotic elongation factor 2 (eEF2), which showed a >50-fold increase at 8 and 12 h p.i. while an increase in phosphorylation could already be detected by 2 h p.i. Infection with other alphaviruses (Semliki Forest virus and Sindbis virus) and a picornavirus triggered a similar strong eEF2 phosphorylation, which could not be prevented by inhibition of PKA or AMPK or by activation of mTORC1, known regulators of eEF2 phosphorylation. Activation of the RIG-I-mediated innate immune response was also not responsible for eEF2 phosphorylation. Transfection with a CHIKV replicon RNA also induced a strong eEF2 phosphorylation, which demonstrated that neither the viral entry process nor the expression of CHIKV structural proteins were responsible. We hypothesize that the increase in phosphorylation of eEF2 is triggered by a specific feature of the genome (e.g. RNA structure) or of (alpha)virus genome RNA translation or transcription and may constitute a novel cellular antiviral/stress response to reduce translation in infected cells.

INTRODUCTION

Chikungunya virus (CHIKV) is a reemerging human pathogen that has affected the lives of millions over the past decade. It has a 12-kb RNA genome of positive polarity and belongs to the alphavirus genus of the togavirus family [15]. CHIKV is an arthropod-borne virus and can be transmitted by mosquitos from the *Aedes* species [26]. Clinical symptoms include a high fever, rash and severe persisting polyarthralgia [14]. The virus was first discovered in 1952 in Tanzania [8] and for a long time mainly caused outbreaks of a limited magnitude. In 2005-2006, however, the virus reemerged in an epidemic form on the east coast of Africa and several Indian ocean islands, infecting millions of people as it spread across the Asian continent [204]. In the autumn of 2013, the virus reached the Caribbean island of Saint-Martin, which was the start of a massive outbreak in the Caribbean and South/Central America [205]. Within a year more than 1 million suspected cases were reported [206]. In the past decade, hundreds of infected travelers have returned to non-endemic countries and as a result small locally-transmitted outbreaks have already occurred in e.g. Italy, France, and the USA [20, 29, 30, 207].

The CHIKV replicative cycle depends on a wide range of interactions with the host, while the infection is also expected to trigger a variety of antiviral and stress responses. To gain more insight into the changes that occur in the host cell during the course of infection, we have previously performed a temporal quantitative proteomics analysis of CHIKV-infected cells [208]. Surprisingly, this study revealed that overall changes in protein abundance were minimal and that the majority of significantly changed proteins showed decreased abundance during infection. An important reason for the lack of protein upregulation during CHIKV infection is the host translational shut-off induced by alphaviruses, which renders the cell unable to respond to the infection at the level of protein synthesis after a certain time point [100]. However, cellular responses to infection might also involve modulation of the activity of proteins and pathways, e.g. by changes in the localization and post-translational modifications (PTM) of proteins, such as phosphorylation and ubiquitination. Compared to changes in protein expression levels, these processes, in general, occur on a different time scale and allow much more rapid responses than those that depend on the induction of protein synthesis. Such PTM-based responses would not be detected by the commonly used quantitative proteomics approaches that only analyze changes in protein abundance. For example, a protein could relocate to a different organelle resulting in an altered function, while the total abundance of this protein could remain unchanged. A change in its PTM status could also alter a protein's function, activity, interaction partners, or stability without affecting its abundance [145, 146, 209]. Especially PTMs can quickly (de)activate signaling pathways, leading to major changes within the cell [158]. Considering the limited changes in protein abundance in CHIKV-infected cells that we observed previously, we

hypothesized that the cell may mainly respond to infection at the level of PTMs. We, therefore, set out to study changes in host protein phosphorylation status during CHIKV infection, because the regulation of protein function through phosphorylation is one of the most widespread regulatory mechanisms in eukaryotes [209]. A temporal quantitative phosphoproteomics approach using stable isotope labeling by amino acids in cell culture (SILAC) [126] was used, and to our knowledge, this is the first in-depth analysis of the temporal changes in host protein phosphorylation status during alphavirus infection.

We have quantified changes in phosphorylation status of 2988 phosphorylation sites on 1038 phosphoproteins, of which 167 unique sites on 113 proteins were modulated significantly on at least one of the time points analyzed. In 110 instances a significant increase in phosphorylation was measured, whereas a significant decrease was observed in 107 instances. The largest change was measured on residue T56 of eukaryotic elongation factor 2 (eEF2). Phosphorylation of this threonine residue increased more than 50-fold after CHIKV infection, an observation that was also made early after infection with two other alphaviruses and a picornavirus. During translation eEF2 is required for the translocation step of peptide-chain elongation and phosphorylation of T56 blocks protein synthesis in the elongation phase [221]. siRNA knockdown of eEF2 resulted in decreased expression of CHIKV proteins, suggesting that the presence of (non-phosphorylated) eEF2 is important for CHIKV replication. Our study provides a first glimpse into the phosphorylation-based regulation of cellular protein function following CHIKV infection and provides a starting point for the more detailed exploration of host processes that restrict or promote alphavirus replication.

MATERIAL & METHODS

Cells and viruses

MRC-5 cells were cultured in EMEM (Lonza), 10% fetal calf serum (FCS) (PAA), 2 mM L-Glutamine, and 1% non-essential amino acids (NEAA) (GE Healthcare) at 37°C and 5% CO_2. Vero E6 cells were cultured in DMEM (Lonza) containing 8% FCS. 293/ACE2 cells [162] were cultured in DMEM (Lonza) containing 10% FCS and 2 mM L-Glutamine. BHK-21 cells were cultured in BHK-21 medium (Glasgow MEM; Invitrogen), containing 5% FCS, 10% tryptose phosphate broth, and 10 mM HEPES, pH 7.4. All culture media contained 100 IU/ml penicillin and 100 ug/ml streptomycin unless otherwise specified. All cells were grown at 37°C and 5% CO_2.

Infections with CHIKV LS3 and LS3-GFP were performed essentially as described previously [168]. The construction of a Chikungunya replicon that was derived from CHIKV LS3 by replacing the structural genes with a puromycin resistance/foot-and-mouth

disease virus 2A oligopeptide/green fluorescent protein (PAC-2A-GFP) reporter gene will be published elsewhere (details available upon request). The sequence of the 2A oligopeptide contains an amino acid motif that prevents formation of the peptide bond between glycine and the final proline of the sequence which allows expression of multiple proteins from a single ORF [210, 211].

Vero E6 cells were infected with the Sindbis virus (SINV) HR-strain or Semliki Forest virus (SFV) strain SFV4 [212] at a multiplicity of infection (MOI) of 5. Vero E6 cells were infected with a GFP-expressing recombinant human adenovirus (HAdV-GFP/LUC; [180]) at an MOI of 5 and with a GFP-expressing recombinant coxsackie B3 virus (CVB3) (a kind gift from prof. dr. Frank van Kuppeveld, Utrecht University) at an MOI of 1. BHK-21 cells were infected with the equine arteritis virus (EAV) Bucyrus strain at MOI 5 at 39.5°C essentially as described previously [213].

SILAC labeling

MRC-5 cells were cultured in SILAC DMEM (PAA) containing 10% dialyzed Fetal Bovine Serum (FBS) (Gibco), 0.280 mM arginine, 0.384 mM lysine, 0.5 mM proline, 10 mM HEPES, 2 mM L-Glutamine and 1% NEAA for >5 cell doublings to ensure complete incorporation of labeled amino acids. Arginine to proline conversion was not observed under these conditions. The 'light SILAC labeling' was performed using Arg $^{12}C_6$ $^{14}N_4$ and Lys $^{12}C_6$ $^{14}N_2$, whereas the 'heavy SILAC sample' was labeled with Arg $^{13}C_6$ $^{15}N_4$ and Lys $^{13}C_6$ $^{15}N_2$ (Cambridge Isotope Laboratories).

CHIKV infection of SILAC labeled cells

SILAC-labeled MRC-5 cells were seeded in 75-cm^2 flasks 1 day before infection with CHIKV-LS3 [168] at MOI 5. One hour post infection (h p.i.), the inoculum was removed and replaced with SILAC DMEM containing 2% dialyzed FBS, 0.280 mM arginine, 0.384 mM lysine, 0.5 mM proline, 25 mM HEPES, 2 mM L-Glutamine and 1% NEAA. At 2, 8, and 12 h p.i., infected and mock-infected cells were harvested for phosphoproteomics analysis by lysis in 4% SDS, 0.1M Tris pH 7.6, followed by heating to 96°C for 10 min. At 12 h p.i., protein lysates for western blot (WB) analysis were harvested in 4× Laemmli sample buffer (LSB) (100 mM Tris-HCl, pH 6.8, 40% glycerol, 8% SDS, 40 mM DTT, 0.04 mg/ml bromophenol blue) and cells grown on coverslips were fixed in 3% PFA in PBS . The experiment was performed in duplicate with a label swap (Figure 1A).
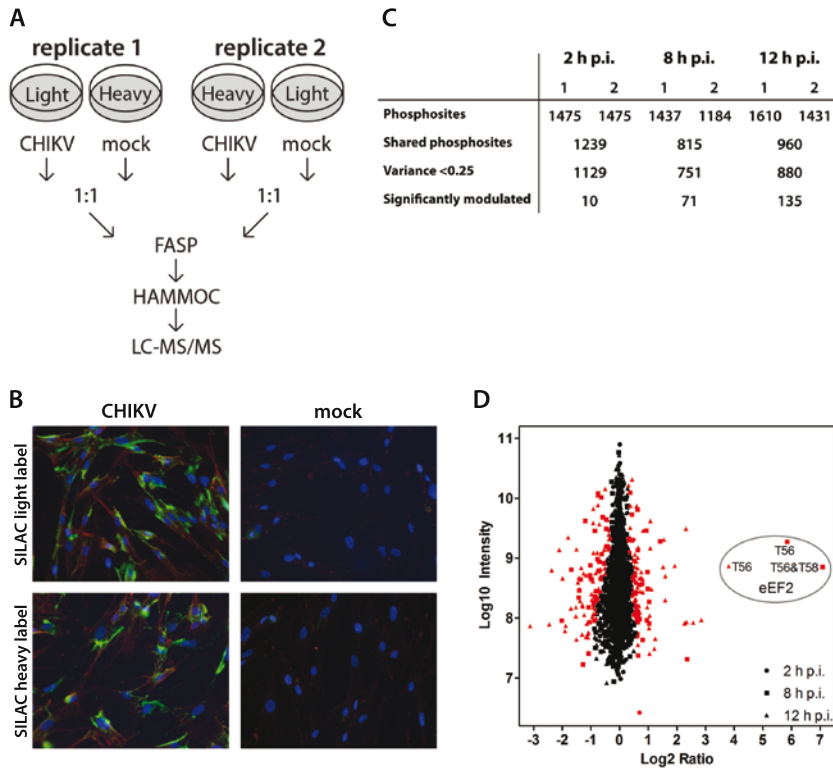
Figure 1: SILAC experiment. (*A*) Experimental set-up of the SILAC-based phosphoproteomics study of virus-induced changes in the host proteome during the course of CHIKV infection. B) At each time point two biological replicates were used, in which a SILAC label-swap was performed between infected and mock-infected samples. Infected and mock-infected cells were lysed at 2, 8 and 12 h p.i. and equal amounts of protein were mixed and subsequently digested into peptides using the filter-aided sample preparation procedure. The samples were enriched for phosphopeptides using the HAMMOC procedure and analyzed by LC-MS/MS. (*B*) Number of identified phosphorylation sites at each time point p.i. The number of identified phosphorylation sites for each biological replicate and the total number of phosphorylation sites that were identified in both replicates at each time point p.i. are shown. For follow-up analysis only phosphorylation sites with a variance <0.25 were selected. (*C*) Immunofluoresence microscopy analysis of SILAC-labeled CHIKV-infected cells at 12 h p.i. Cells were immunolabeled for dsRNA (green) and the CHIKV E2 protein (red), and nuclear DNA (blue) was stained with Hoechst-33342. (*D*) Proteome-wide quantification of phosphorylation sites at 2 (dots), 8 (squares), and 12 (triangles) h post-CHIKV infection. Average normalized $\log_2$ ratios are plotted against the $\log_{10}$ of peptide intensities. Each mark represents a protein group. Red marks indicate phosphorylation sites that were significantly modulated (Benjamini-Hochberg FDR of 0.05). The phosphorylation sites identified for eEF2 at 8 and 12 h p.i. are indicated.

Protein digestion

The protein concentration of the SILAC cell lysates was determined using the bicin chonicic acid assay (Pierce). Digestion of the proteins was performed using the Filter Aided Sample Preparation (FASP) method [169], for which equal amounts (900 μg) of mock and infected lysates were mixed and DTT was added to a final concentration of 50 mM, followed by a 5-min incubation at 70°C. Samples were loaded on two 15-ml 30 kDa Microcon filter devices (Millipore), which were washed twice with 8 M urea, 0.1 M Tris pH 8.5, while cysteines were alkylated with 50 mM iodoacetamide in the same buffer. Samples were washed 3 times with 8 M urea, 0.1 M Tris pH 8. Proteins were digested overnight at room temperature using 20 ug endoLysC (Wako Pure Chemical Industries) per filter device in the same buffer. The sample was diluted fourfold with 50mM ammonium bicarbonate pH 8.4 containing 20 ug trypsin (Worthington Chemical Corporation), and digested for 4h at room temperature. Peptides were collected by centrifugation, acidified to a final percentage of 1% TFA, and desalted using solid phase extraction. Peptides were eluted in 20/80/0.1 milliQ/acetonitrile (ACN) (Actu-All Chemicals)/trifluoric acid (TFA) (Sigma-Aldrich) (v/v/v).

Phosphopeptide enrichment

The samples were enriched for phosphopeptides by Hydroxy Acid Modified Metal Oxide Chromatography (HAMMOC) [214]. Per 500 ug peptide digest a 200 ul tip with an Octyl C8 membrane (Empore) and 2.5 mg Titansphere TiO_2 10 μm (GL Sciences) were used. The tips were preconditioned with 20/80/0.1 milliQ/ACN/TFA (v/v/v) (solution A) and equilibrated with 300 mg/ml DL Lactic acid (Fluka Analytical) in solution A. Peptide samples were mixed 1:1 with 300 mg/ml DL Lactic acid in solution A and loaded on the tips. The tips were washed with 300 mg/ml DL Lactic acid in solution A and solution A. 100 ul 20% phosphoric acid (Sigma-Aldrich) was put in collection tubes and phosphopeptides were eluted with 50 ul 0.5% piperidine (Actu-All Chemicals) followed by 50 ul 5% piperidine. Peptides were desalted on 200 ul tips with an SDB-XC membrane (Empore). Tips were preconditioned with solution A and equilibrated with 0.1% TFA. Samples were loaded on the tips and the tips were washed with 0.1 %TFA. Peptides were eluted with solution A and lyophilized in a CHRIST RVC-2-18 CDplus.

Mass spectrometry

Phosphopeptide-enriched samples were analyzed via on-line C18-nano-HPLC-MS with a system consisting of an Easy nLC 1000 gradient HPLC system (Thermo, Bremen, Germany), and a Q-Exactive mass spectrometer (Thermo). Fractions were injected onto a homemade precolumn (100 μm × 15 mm; Reprosil-Pur C18-AQ 3 μm, Dr. Maisch, Ammerbuch, Germany) and eluted via a homemade analytical nano-HPLC column (15 cm × 50 μm; Reprosil-Pur C18-AQ 3 um). The gradient was run from 0% to 30% solvent B

(10/90/0.1 water/ACN/FA v/v/v) in 120 min. The nano-HPLC column was drawn to a tip of ~5 μm and acted as the electrospray needle of the MS source. The Q-Exactive mass spectrometer was operated in top10-mode. Parameters were resolution 70,500 at an AGC target value of 3,000,000, maximum fill time of 250 ms (full scan), and resolution 17,500 at an AGC target value of 200,000/maximum fill time of 80 ms for MS/MS at an intensity threshold of 2,500. Apex trigger was set to 1 to 15 seconds, and allowed charges were 2-6. Each sample was analyzed in duplo.

Data analysis

Raw data files were analyzed using Maxquant 1.4.0.3 [173] using the Andromeda search engine [174]. Databases used for the main search were UNIPROT/KB_Human (88,665 entries) and a custom-made database containing the protein sequences of CHIKV-LS3 (11 entries) using the GenBank sequence (accession KC149888) [168]. For the first search a smaller database, human.first.search (15,612 entries) containing a subset of human protein sequences was used. A list of common contaminants was included in the search. To reach a false discovery rate (FDR) of 0.01 a concatenated reversed database (KR special amino acids) was used, FDR at the peptide level was also 0.01. Enzyme specificity was Trypsin/P. Variable modifications included in the search were oxidation (M), acetylation (protein N-term) and phospho (STY), whereas carbamidomethyl (C) was included as a fixed modification. Up to 3 missed cleavages and a maximum of 5 modifications per peptide were allowed. The minimum score for modified peptides was set to 40 and the minimum delta score for modified peptides was set to 17. Match between runs was turned on with a matching time window of 1 minute. MaxQuant results were further analyzed with Perseus version 1.2.0.17 and Microsoft Excel 2010 and GraphPad Prism version 5.

From the list of identified phosphorylation sites, contaminants and peptides identified with the concatenated reversed database were removed before further analysis. Normalized ratios from one of the biological replicates were inverted to ensure all ratios were displayed as infected/mock. Normalized ratios were $\log_2$ transformed and averages and variances were calculated for each experiment. Phosphorylation sites with a variance <0.25 were included in the analysis. The list of excluded sites was manually inspected and some of these were still included in the analysis when a large change was observed in both replicates in the same direction but with relatively large variation. Significance B with a Benjamini-Hochberg FDR of 0.05 (both-sided) was calculated in Perseus 1.2.0.17 separately for each time point to determine which phosphorylation sites were significantly modulated during CHIKV infection.

String 10 was used to determine enrichment with medium confidence for protein-protein interactions and GO terms [215].

Western blot analysis

Western blot (WB) analysis was performed essentially as described previously [208]. Primary antibodies used were rabbit antisera against CHIKV nsP1 [168], CHIKV capsid protein, SFV capsid protein (both kind gifts from prof. Andres Merits, University of Tartu, Estonia), rabbit polyclonal against eEF2 #2332, mouse monoclonal against STAT-1 (9H2) #9176 (both Cell Signaling Technology), mouse monoclonal H68.4 against the transferrin receptor (Invitrogen), or mouse monoclonal against b-actin (A5316; Sigma) diluted in 1% casein in phosphate buffered saline containing 0.1% Tween-20 (PBST). Rabbit polyclonal phospho-eEF2 (Thr56) #2331 (Cell Signaling Technology) was diluted in 1% BSA in tris buffered saline containing 0.1% Tween-20 (TBST). Biotin-conjugated swine-a-rabbit diluted 1:2000 (DAKO) or goat-a-mouse (DAKO) diluted 1:1000, and Cy3-conjugated mouse-a-biotin diluted 1:2500 (Jackson) were used for fluorescent detection with a Typhoon-9410 scanner (GE Healthcare).

Immunofluorescence microscopy

Immunofluorescence microscopy (IFA) was performed essentially as described previously [208]. Primary antibodies used were rabbit polyclonal against CHIKV E2 [179] and mouse monoclonal antibody J2 against dsRNA (English & Scientific Consulting) diluted in 0.5% BSA in PBS. Primary antibodies were detected with donkey-a-rabbit-Cy3 or goat-a-mouse-Alexa488 (Jackson). Nuclei were stained with Hoechst33342. Coverslips were mounted with Prolong (Invitrogen) and examined using a Zeiss Axioskop2 fluorescence microscope with Axiocam HRc camera and AxioVision software.

siRNA-mediated knockdown of eEF2

eEF2 protein levels were knocked down in Vero E6 and MRC-5 cells using ON-TARGETplus siRNA SMARTpool against human eEF2 (1938; Thermo Scientific) at a final concentration of 25 nM and 2 ul Dharmafect1 in 1 ml per 3.8 cm^2 well. A pool of non-targeting (scrambled) siRNAs (NTP; cat. nr. D-001810-10; Dharmacon) was used as a negative control. The cells were grown in medium without antibiotics 1 day prior to siRNA transfection until 1 day p.t. At 2 days p.t., cells were infected with CHIKV LS3 at an MOI of 1 and 16 h p.i. cells were harvested in 4× LSB. Protein levels for eEF2, CHIKV nsP1, CHIKV capsid and actin were determined by WB analysis.

Modulating eEF2 phosphorylation with compounds

The AMPK inhibitor dorsomorphin (Sigma-Aldrich) and the PKA inhibitor KT5720 (Abcam) were dissolved in anhydrous DMSO (Sigma-Aldrich). mTORC1 activator L-Leucine (Sigma-Aldrich) was dissolved in DMEM. Vero E6 cells were pretreated with compound for 1 h prior to infection with SFV at an MOI of 5 in the presence of the same concentration of compound. As a control no compound and 0.1% DMSO were used. Cells were

harvested at 5 h p.i. in 4× LSB and protein levels for peEF2, SFV capsid and actin were determined by WB analysis.

Treatment of cells with 5'pppRNA

The *in vitro* synthesis of 5'pppRNA representing sequences from the 5' and 3' untranslated regions of the vesicular stomatitis virus (VSV) genome was described previously [216]. MRC-5 cells were transfected with 0.1, 1, or 10 ng/ml 5'pppRNA or control RNA (same sequence but lacking the 5'ppp moiety; Integrated DNA Technologies Inc., IA, USA) 1 h prior to infection or mock infection with CHIKV-LS3-GFP at MOI 0.1, as described previously [216]. At 24 h p.i., cells were lysed in 4× LSB and protein levels for pEEF2, STAT-1, and actin were determined by WB analysis.

Replicon RNA transfection

BHK-21 cells were transfected by electroporation using $4x10^6$ cells in 400 ul PBS and 4 ug of *in vitro* transcribed capped or uncapped CHIKV replicon RNA per cuvette. After 2 pulses with an Eurogentec Easyjet Plus instrument set at 850 V and 25 µF, cells were taken up in pre-warmed medium and seeded at a density of $6x10^5$ cells/$10cm^2$ dish and incubated at 37°C. Cells were harvested 6, 8, and 10 h p.i. in 4× LSB for WB analysis or in 20 mM Tris-HCl (pH 7.4), 100 mM LiCl, 2 mM EDTA, 5 mM DTT, 5% (w/v) lithium dodecyl sulfate, and 100 mg/ml proteinase K for in-gel hybridization analysis.

RNA isolation, denaturing agarose electrophoresis and in-gel hybridization

Total RNA was isolated from cells lysed in 20 mM Tris-HCl (pH 7.4), 100 mM LiCl, 2 mM EDTA, 5 mM DTT, 5% (w/v) lithium dodecyl sulfate, and 100 mg/ml proteinase K as described previously [168]. RNA was separated in 1.5% denaturing formaldehyde-agarose gels using the MOPS buffer system as described previously [217]. RNA molecules were detected by direct hybridization of the dried gel with ^{32}P-labeled oligonucleotides essentially as described previously [218]. CHIKV positive- or negative-stranded RNAs were visualized as described previously [168] using probe CHIKV-hyb4 (5'-TGTGGGTTC-GGAGAATCGTGGAAGAGTT-3') or probe CHIKV-hyb2 (5'-AACCCATCATGGATCCTGTGTAC-GTGGA-3'), which are complementary to the 3' end of the genome (detects genome and subgenomic mRNA) and anti-genome (detects negative-stranded RNA), respectively. 18S ribosomal RNA (loading control) was detected with the oligonucleotide probe 5'-AT-GCCCCCGGCCGTCCCTCT-3'. Hybridized gels were exposed to Storage Phosphor screens and scanned with a Typhoon-9410 scanner (GE Healthcare).

3

RESULTS

To identify proteins of which the phosphorylation status is modified during CHIKV infection, we performed a SILAC time course experiment in which protein lysates of mock- and CHIKV-infected MRC-5 cells were compared. Samples harvested at 2, 8 and 12 h p.i. were enriched for phosphopeptides, followed by LC-MS/MS analysis and phosphorylation site quantification (Figure 1A). Our earlier study [208] demonstrated that until 8 h p.i. there were minimal changes in the host proteome at the level of protein abundance, but we expected that changes in phosphorylation status could occur much more rapidly and, therefore, decided to also analyze samples taken at 2 h p.i. to monitor early responses to CHIKV infection. At 8 h p.i. the viral proteins are clearly detectable by WB analysis in this cell line (Figure 2A) and by 12 h p.i. high levels of viral protein and RNA are present but cytopathic effects are not yet observed.
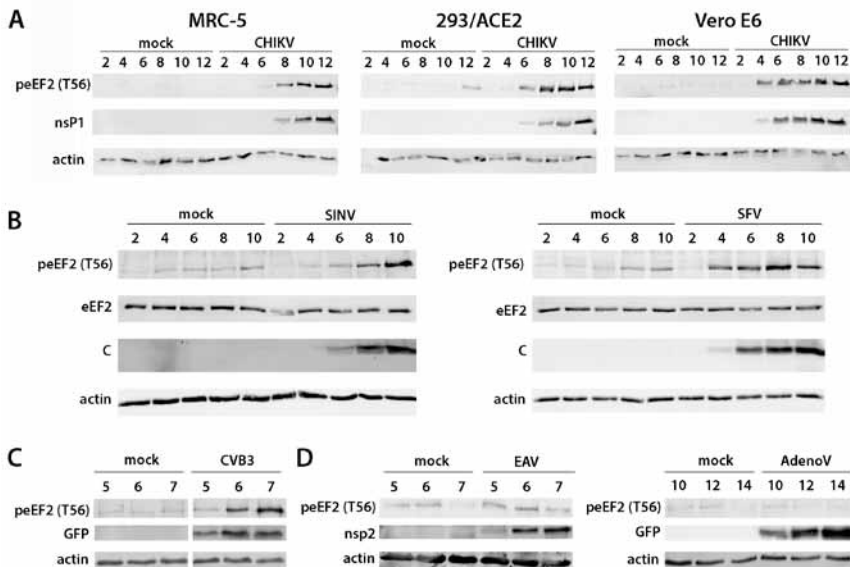
Figure 2: Alphavirus infection triggers early phosphorylation of eEF on T56. (*A*) WB analysis of CHIKV MOI 5 infected MRC-5, 293/ACE2 and Vero E6 cells at different time points post infection. Protein lysates were separated by SDS-PAGE and viral protein nsP1 and peEF2 (T56) were detected by WB. Actin was used as a loading control. (*B*) WB analysis of SINV and SFV MOI 5 infected Vero E6 cells at different time points post infection. Protein lysates were separated by SDS-PAGE and viral protein C, eEF2 and peEF2 (T56) were detected by WB. Actin was used as a loading control. (*C*) WB analysis of a GFP-expressing CVB3 MOI 1 infected VeroE6 cells at different h p.i. Protein lysates were separated by SDS-PAGE and GFP (marker for viral protein synthesis) and peEF2 (T56) were detected by WB analysis. Actin was used as a loading control. (*D*) WB analysis of EAV MOI 5 infected BHK-21 cells and GFP-expressing AdenoV infected Vero E6 cells at different time points p.i. Protein lysates were separated by SDS-PAGE and viral protein nsP2 or GFP, and peEF2 (T56) were detected by WB analysis. Actin was used as a loading control.

General proteomics results

Successful CHIKV infection of SILAC-labeled cells was confirmed by immunofluorescence microscopy at 12 h p.i. All cells stained positive for double-stranded RNA (a marker for viral replication) and E2, one of the CHIKV structural proteins (Figure 1B).

In total, 2988 phosphorylation sites were identified on 1060 unique proteins (Table S1). Per time point 800-1200 peptides containing phosphorylation sites were quantified in both biological replicates. (Figure 1C and D).

Phosphorylated CHIKV peptides were also identified. Two phosphorylation sites were identified on the capsid protein (C), two on envelope protein E2 and four on nonstructural protein nsP3. (Table S1)

During CHIKV infection 167 unique phosphorylation sites on 113 host proteins were modulated significantly on at least one time point with a Benjamini-Hochberg FDR of 0.05. Several sites were significantly modulated at multiple time points (Figure 1D and Table S2).

At 2 h p.i., changes in phosphorylation status were minimal, although 10 sites were significantly modulated, with four of these also being significantly modulated at the two later time points: T56 on eEF2 (Log$_2$ ratio 0.70), S430 on Vimentin (Log$_2$ ratio 0.48), and T185 and Y187 on MAPK (Log$_2$ ratio -0.48 and 0.29, respectively). At 8 h p.i., the phosphorylation status of 71 sites had changed significantly and at 12 h p.i. this number increased to 136 sites. 42 sites were significantly modulated at both 8 and 12 h p.i. (Table S2). Both increases (110 occurrences) and decreases (107 occurrences) in phosphoryla-

Protein	Accession number	Peptide sequence	Modified residue(s) #	Log2 ratio		
				2 h p.i.	8 h p.i.	12 h p.i.
EEF2	P13639	_AGETRFT(ph)DTR_	T56	0.70 *	5.85 *	3.81 *
		AGETRFT(ph)DT(ph)RKDEQER	T56, T58	NaN	7.08 *	NaN
HSPB1	P04792	_ALS(ph)RQLS(ph)SGVSEIR_	S78, S82	0.05	0.35	2.31 *
LARP1	Q6PKG0	_S(ph)LPTTVPES(ph)PNYR_	S769, S774	-0.02	-1.37 *	-2.11 *
MTDH	E5RJU9	_SETSWES(ph)PK_	S568	-0.06	-2.02 *	-3.12 *
NES	P48681	_S(ph)LGEEIQESLK_	S548	NaN	1.03 *	2.27 *
		SLRS(ph)LEEQDQETLR	S746	NaN	NaN	2.85 *
NUP50	Q9UKX7-2	_VAAETQS(ph)PSLFGSTK_	S221	NaN	1.30	2.34 *
PRRC2C	Q9Y520-2	_AFGSGIDIKPGT(ph)PPIAGR_	T2673	-0.14	-1.62	-2.45 *
RALY	Q9UKM9-2	_TRDDGDEEGLLTHS(ph)EEELEHS(ph)QDTDADDGALQ_	S288, S295	-0.27	-2.17	-2.37 *
SH3KBP1	B7Z6E8	_ANS(ph)PSLFGTEGKPK_	S587	NaN	2.35 *	2.57 *

Table 1: Peptides that contain a phosphorylation site that changed phosphorylation status >4 fold on at least one of the time points analyzed. (ph) indicates the preceding amino acid residue was modified by phosphorylation. # Amino acid residue number according to PhosphoSitePlus [283]. *Significantly changed phosphorylation status at this time point with Benjamini-Hochberg FDR 0.05. NaN: peptide was not identified at this time point.

tion were observed. On 11 peptides we identified sites with a larger than 4-fold change in phosphorylation status on at least one of the time points studied (Table 1).

The 113 proteins that contained phosphorylation sites that were significantly modulated were analyzed with medium confidence using String 10 [215]. The dataset was enriched for protein-protein interactions with a p-value of 2.22e-16. This set of proteins were also strongly enriched for the Gene Onthology (GO) [219, 220] molecular functions poly(A)-binding RNA binding (Bonferroni corrected p-value $1.10*10^{-20}$) and RNA binding protein (Bonferroni corrected p-value $1.48*10^{-19}$), the GO biological processes intracellular transport (Bonferroni corrected p-value $7.66*10^{-4}$) and the GO cellular component adherens junction (Bonferroni corrected p-value $7.48*10^{-20}$) and cell-substrate adherens junction (Bonferroni corrected p-value $1.12*10^{\wedge-19}$) (Table S3).

eEF2 becomes strongly phosphorylated during alphavirus infection

The protein that showed the largest increase in phosphorylation, i.e. a >50 fold increase at 8 and 12 h p.i., was eEF2. This protein became phosphorylated on T56 and T58 (Figure 1D). Because of this strong effect, it was selected for follow-up analysis. Phosphorylation of eEF2 on T56 during CHIKV infection was confirmed with WB on samples from independent time course infection experiments in MRC-5, Vero E6, and 293/ACE2 cells. In the proteomics experiment, a significant increase in phosphorylation (Log_2 ratio 0.70) was already observed at 2 h p.i. Also in Vero E6 cells strong phosphorylation at this residue could be detected early in infection (at 4 h p.i) by WB analysis. (Figure 2A).

Other proteins involved in translation, eEF1D, ribosomal protein S6 (rps6), ribosomal protein, large, P1 (rplp1) and rplp2, also contain phosphorylation sites that were significantly modulated during CHIKV infection suggesting that the slowing down of translation might have a broader basis than only through inhibition of eEF2.

Phosphorylation of eEF2 in Vero E6 cells was also observed during infection with two other alphaviruses, Sindbis virus (SINV) and Semliki Forest virus (SFV) (Figure 2B). We subsequently also tested three unrelated viruses, a GFP expressing CVB3 (+RNA virus, picornavirus family), EAV (+RNA virus, arterivirus family, order Nidovirales) and a GFP-expressing HAdV (dsDNA virus, adenovirus family). Infection with CVB3 resulted in early eEF2 phosphorylation (Figure 2C) but infections with EAV and HAdV did not (Figure 2D).

eEF2 is a GTPase that is required for the translocation step of peptide-chain elongation during translation. Its activity is regulated by the eEF2 kinase (eEF2K) that phosphorylates two threonine residues in an ordered process with T56 phosphorylation preceding T58 phosphorylation [221], but phosphorylation on T56 is already sufficient to inactivate eEF2 [222]. In our study, we only identified phosphorylated T58 on a peptide that was also phosphorylated on T56 (Table 1).

To obtain more insight into the role of eEF2 during CHIKV infection, the effect of siRNA-mediated depletion of eEF2 on CHIKV LS3-GFP replication in MRC-5 and Vero E6
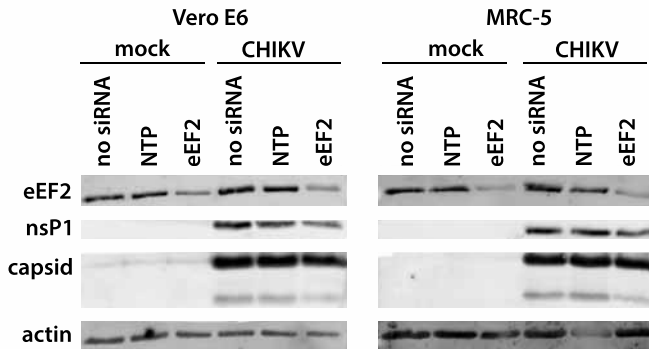
Figure 3: siRNA-mediated knockdown of eEF2 inhibits viral protein synthesis. Vero E6and MRC-5 cells were transfected with 25 nM siRNA targeting eEF2, a non-targeting pool of control siRNAs or no siRNA. 2 days post infection the cells were infected with CHIKV LS3 MOI 1 and protein lysates were harvested at 16 h p.i. Proteins were separated by SDS-PAGE and viral proteins nsP1 and C, eEF2 were detected by WB. Actin was used as a loading control.

cells was analyzed (Figure 3). The siRNA-mediated knockdown had no significant effect on cell viability, while the synthesis of CHIKV proteins was reduced, although this was a modest effect.

Preliminary studies into the mechanism that triggers eEF2 phosphorylation

Modulating pathways involved in regulating eEF2 activity
eEF2 phosphorylation is modulated by several pathways, including the AMPK, PKA, mTORC1 and MAP kinase pathways [225]. To investigate which cellular pathways could be responsible for eEF2 phosphorylation, and to understand how this PTM affects CHIKV replication, several inhibitors and agonists were used. Activation of PKA results in activation of eEF2K [223, 224]. However, inhibition of PKA with KT5720 only mildly inhibited phosphorylation of eEF2 on T56 during CHIKV infection and had no effect on the synthesis of capsid protein (Figure 4). Activation of AMPK can result in increased phosphorylation of eEF2 [225]. However, inhibition of AMPK with dorsomorphin increased phosphorylation on eEF2 T56 during CHIKV infection but did not appear to have an effect on capsid protein expression (Figure 4). Inhibition of mTORC1 promotes translation elongation [225]. However, activation of mTORC1 with L-Leucine increased eEF2 phosphorylation during CHIKV infection and at the same time reduced the amount of capsid protein accumulating in CHIKV-infected cells (Figure 4). These preliminary compound studies suggest that the pathways that were tested so far are not involved in the induction of eEF2 phosphorylation, but additional studies including more positive and negative controls are required to exclude their involvement for certain.
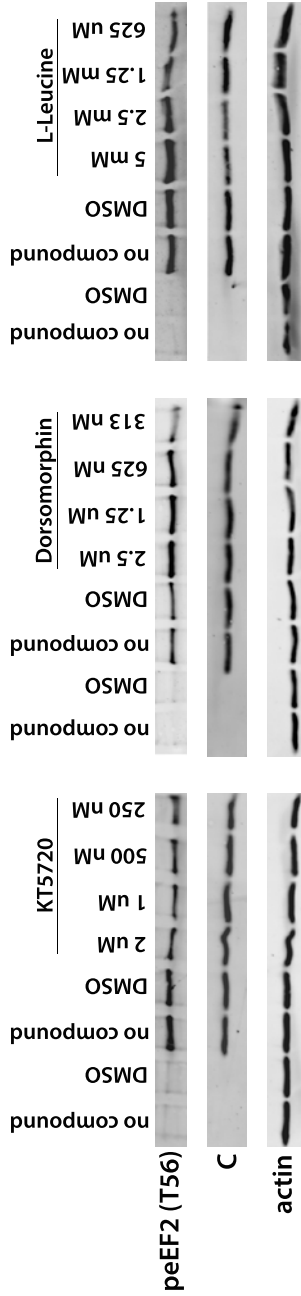
3

Figure 4: eEF2 phosphorylation is not mediated through PKA or AMPK activation or inhibition of mTORC1. Vero E6 cells were pretreated with different concentrations of KT5720 (PKA inhibitor), Dorsomorphin (AMPK inhibitor) or L-Leucine (mTORC1 activator) 1 hour prior to infection with SFV at an MOI of 5 in the presence of the same concentration of compound. As a control no compound and 0.1% DMSO were used. Protein lysates were harvested 5 h p.i. and separated by SDS-PAGE. Protein levels for peEF2, SFV Capsid were determined by WB. Actin was used as a loading control.

Innate immune response

We next hypothesized that phosphorylation of eEF2 might be part of an innate immune response that is triggered by PAMPs, e.g. via the detection by the RIG-I/MDA5-MAVS pathways. To test this possibility, RIG-I was activated by transfection of a non-cytotoxic dose of 5' triphoshorylated RNA (5'pppRNA) representing sequences from the 5'and 3' untranslated regions of the VSV genome into MRC-5 cells, which has been shown to induce a good antiviral response against CHIKV infection [216]. This treatment did not induce phosphorylation of eEF2 T56, while it did lead to the induction of STAT1 synthesis, indicating that the 5'pppRNA triggered an innate immune response in these cells. In cells that were infected with CHIKV one hour after transfection of 5'pppRNA, eEF2 phosphorylation was prevented in a dose-dependent manner, likely because CHIKV replication was inhibited in these cells by the 5'pppRNA treatment (Figure 5).
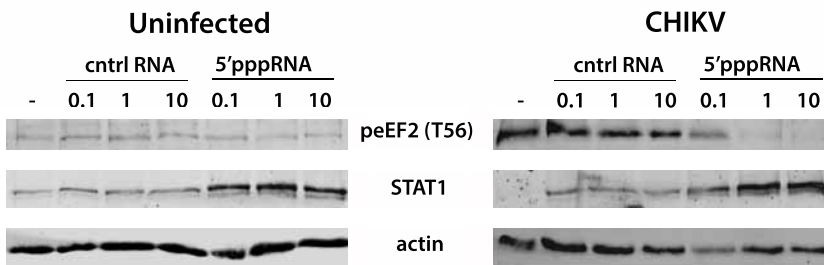
Figure 5: RIG-I activation does not trigger eEF2 phosphorylation. MRC-5 cells were transfected with 0.1, 1 or 10 ng/ml 5'pppRNA or a control RNA 1 hour prior to infection with CHIKV-LS3-GFP at an MOI of 0.1. Protein lysates were harvested 24 h p.i. and separated by SDS-PAGE. STAT1 and peEF2 (T56) were detected by WB analysis. Actin was used as a loading control.

No involvement of CHIKV structural proteins and virus entry

To determine whether any of the structural proteins play a role in triggering eEF2 phosphorylation, BHK-21 cells were transfected with a CHIKV RNA replicon lacking expression of all structural proteins and eEF2 phosphorylation was analyzed by WB (Figure 6A). Also in the absence of the structural proteins the CHIKV replicon triggered eEF2 phosphorylation. Uncapped replicon RNA was transfected as a control since it will not be translated and not be replicated and this construct did not induce phosphorylation. Since entry was bypassed by direct electroporation of the RNA into the cytosol, we conclude that the entry process did not trigger phosphorylation either. An increase in genome, subgenomic mRNA (sgRNA) and negative-stranded RNA could be detected
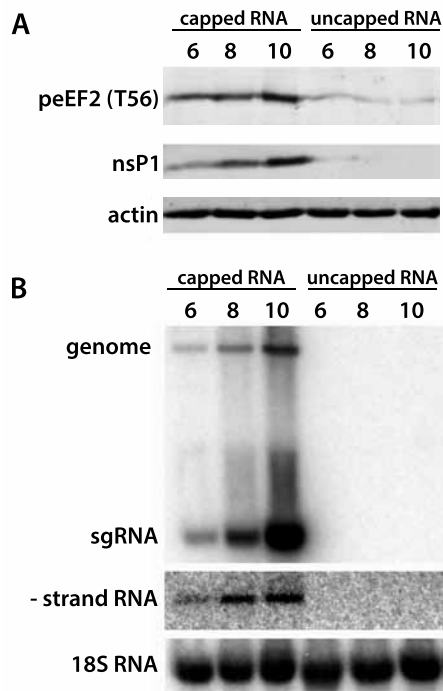
Figure 6: Transfection of a capped CHIKV replicon triggers eEF2 phosphorylation. BHK-21 cells were electroporated with *in vitro* transcribed capped or uncapped CHIKV LS3 replicon RNA. Protein lysates were harvested and RNA was isolated at 6, 8 and 10 h p.i. (*A*) Protein lysates were separated by SDS-PAGE and viral protein nsP1 and peEF2 (T56) were detected by WB. Actin was used as a loading control. (*B*) RNA was isolated and separated on a denaturing formaldehyde-agarose gel. Positive- and negative strands of CHIKV RNA were detected by in-gel hybridization with radioactively labeled oligonucleotides complementary to the 3' end of either negative- or positive-strand CHIKV RNA. The cellular 18S ribosomal RNA was used as a loading control.

for the capped replicon but the uncapped RNA was most likely quickly degraded after electroporation as it could not be detected at 6 h p.t. (Figure 6B).

DISCUSSION

Proteomics techniques to analyze changes in host protein abundance have already been applied to many different viruses. These changes are often relatively slow and many viruses cause a host transcriptional and/or translational shut-off that prevents the synthesis of new proteins after a certain time point. To gain more insight in signaling pathways that are triggered or inhibited during CHIKV infection, we have performed a phosphoproteomics study and quantified changes in phosphorylation on a large

number of phosphorylation sites on cellular proteins at 2, 8 and 12 h p.i. In our previous study, we studied protein abundance and only observed minimal changes at 8 h p.i. [208], but changes in phosphorylation status can occur very rapidly and may also be triggered during the earliest stages of viral replication, which is why we also included the 2 h p.i. time point in the current study.

Numerous proteins showed regulation of phosphorylation status during CHIKV infection. Several of these factors were previously implicated in viral infection, although not always through their phosphorylation status. For example, SH3KBP1 (S587; Log_2 ratio 2.35 and 2.65 at 8 and 12 h p.i. in our study) interacts with the ORF3 protein of hepatitis E virus which results in prolongation of endomembrane growth factor signaling and promotes cell survival during infection [226]. SH3KBP1 was also shown to inhibit herpes simplex virus replication [227]. Hyper-phosphorylation of nuclear pore proteins (Nups) induced by the leader protein inhibits nucleocytoplasmic trafficking during picornavirus infection [228, 229]. Although Nup50 (S221; Log_2 ratio 2.34 at 12 h p.i. in our study) was not implicated in picornavirus infection, Nup153 (S687; Log_2 ratio 0.64 and 1.24 at 8 and 12 h p.i. in our study) was [228-231].

For several proteins interactions with one of the alphavirus proteins have been described. For vimentin several sites significantly changed phosphorylation status during CHIKV infection and this protein was reported to be an interaction partner of nsP2 and nsP3 in a yeast two-hybrid assay [91]. However, another study, which used immunoprecipitation, reported that the interaction between these proteins was due to aspecific binding [94] so the changes in phosphorylation status on vimentin may also prove to be irrelevant to the CHIKV infection. Also Ybx1 (S165; Log_2 ratio -1.03 and -1.17 at 8 and 12 h p.i. in our study) and hnRNPK (S116; Log_2 ratio 0.31 at 8 h p.i. in our study) were reported to interact with nsP2 and nsP3 in the yeast two-hybrid assay [91].

eEF2 showed the largest change in phosphorylation status during CHIKV infection. The increase in phosphorylation could already be detected by 2 h p.i. and is thus an early event during infection. eEF2 is required for successful virus replication since knockdown resulted in a (modest) decrease in viral protein synthesis. Interestingly, several other proteins involved in translation, eEF1D, rps6, rplp1 and rplp2, showed significant changes in phosphorylation as well during CHIKV infection suggesting that the inhibition of protein synthesis might have a broader basis than only through inactivation of eEF2. The sites on rplp1 (S101 and S104) and rplp2 (S102 and S105) are part of the elongation factor binding site of the large ribosomal subunit and phosphorylation of these proteins is necessary for ribosome activity [232, 233] and at 8 h p.i. phosphorylation on these sites was significantly decreased (Log_2 ratio -0.75). The phosphorylated form of rplp2 was also shown to interact with eEF2 [232]. For rps6, an interaction with nsP2 of Venezuelan equine encephalitis virus (VEEV), SINV, and SFV has been reported [234]. This study reported a decrease in phosphorylation for rps6 during alphavirus infection. After 2 h

p.i., there might have been a decrease (S235 and S236; Log_2 ratio -0.62) during CHIKV infection, but the peptide was only identified in one of the biological replicates. At later time points we observed an increase (Log_2 ratio 1.02 and 0.96 at 8 and 12 h p.i.). For several elongation factors, including eEF1D, a role as cofactor during virus replication has been described (Li et al., 2013).

The strong induction of eEF2 phosphorylation was also observed during infections with two other alphaviruses, SFV and SINV, and picornavirus CVB3, and was previously reported to occur during Rift Valley Fever virus (RVFV) infection [235], suggesting that the mechanism triggering phosphorylation during viral infection may be more universal. However, phosphorylation was not triggered during infections with two other viruses, EAV and HAdV. It will be interesting to determine whether more viruses from different virus families trigger eEF2 phosphorylation as this may help to determine what the trigger is for phosphorylation induction. In the four phosphoproteomics studies that have been published, using cells infected with human immunodeficiency virus (HIV)-1, Sendai Virus, porcine reproductive and respiratory syndrome virus (PRRSV), and influenza A virus, the peptide containing eEF2 T56 was not identified so it is possible that phosphorylation was also not induced during infections with these viruses [236-239]. For PRRSV this is not surprising, since it is an arterivirus related to EAV. In the HIV-1 study, samples were already analyzed at 1 min p.i. which was most likely too early to observe an increase in eEF2 phosphorylation [239].

Regulation of eEF2 activity

eEF2 is a GTPase that is required for the translocation of the nascent polypeptide chain from the P and A site to the E and P site of the translating ribosome [240]. eEF2 can be modified on multiple residues, of which phosphorylation on T56 is physiologically the most important. Phosphorylation on this residue results in a structural change in eEF2 that reduces its affinity for the ribosome. This arrests protein translation at the translocation step of elongation [241]. During CHIKV infection, eEF2 became phosphorylated on T56 and T58. Both T56 and T58 are phosphorylated by eEF2 kinase (eEF2K) and T56 phosphorylation always precedes phosphorylation of T58 [221, 242]. In our experiment we only identified T58 phosphorylation on a peptide that was also modified on T56, which is in line with earlier findings. eEF2K belongs to a small group of atypical α-kinases and eEF2 is its only known substrate [225]. eEF2K activity is Ca^{2+}/calmodulin dependent [243, 244] and the kinase is additionally positively and negatively regulated by phosphorylation at several sites [225]. We did not identify any of these sites in our proteomics analysis. Besides an increase in Ca2+ levels, two other pathways positively regulate eEF2K activity. Cyclic AMP (cAMP) is a catabolic signal that inhibits protein synthesis by stimulating cAMP-dependent protein kinase (PKA). Phosphorylation of eEF2K on S500 by PKA makes it partially Ca2+/calmodulin independent [223, 224]. AMP-activated
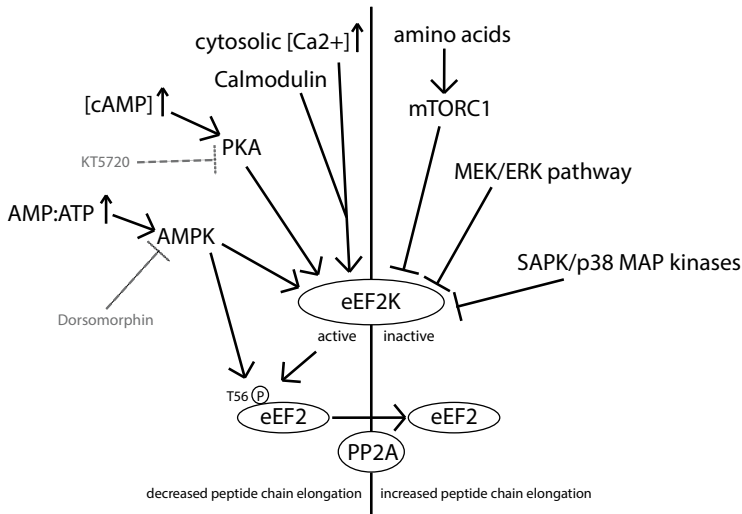
Figure 7: Schematic overview of different pathways involved in regulating eEF2 phosphorylation. An increase in the cytosolic Ca2+ concentration activates eEF2K. Activation of PKA by an increased concentration of cAMP activates eEF2K and can be inhibited by KT5720. Activation of AMPK by an increase in the AMP:ATP ratio activates eEF2K and AMPK possibly also directly phosphorylates eEF2. AMPK can be inhibited by Dorsomorphin. eEF2K can be inhibited by mTORC1, the MEK/ERK pathway and SAPK/p38 MAP kinases. eEF2 T56 phosphorylation is reversible through dephosphorylation by PP2A.

protein kinase (AMPK), a cellular energy sensor [245], also stimulates eEF2K activity, although it is still unclear whether AMPK regulates eEF2K directly through phosphorylation or indirectly through the inhibition of mammalian target of rapamycin complex 1 (mTORC1) signaling [225]. It is also possible that AMPK directly phosphorylates eEF2 to decrease protein synthesis [246]. mTORC1 negatively regulates eEF2K activity by inducing phosphorylation on S78, S359 and S366 and as a result promotes translation elongation [225]. There is also some evidence that eEF2K is phosphorylated on S366 and as a result is inactivated in response to MAP kinase pathways, SAPK/p38δ and MEK/ERK, either directly or through downstream effectors [247-249]. Low pH values also activate eEF2K which might inhibit protein synthesis during tissue acidosis [250]. Phosphorylation of eEF2 is reversible and dephosphorylation is done by protein phosphatase 2 A (PP2A) [251]. A schematic overview of the various pathways involved in modulating eEF2 activity is depicted in Figure 7.

Nutrient sensing as an antiviral mechanism

One of our hypotheses was that eEF2 becomes phosphorylated during alphavirus infection to reduce translation as part of an innate immune response. Viral infection triggers

innate immune responses through sensing of pathogen-associated molecular patterns (PAMPs) by pathogen recognition receptors such as Toll-like receptors (TLRs), RIG-I like receptors (RLRs), nucleotide-binding and oligomerization domain (NOD)-like receptors (NLRs) and DNA-dependent activator of IRFs (DAIs) [181, 252-254]. The activation of these innate sensors can induce rapid immune responses. Our data suggests that phosphorylation of eEF2 during alphavirus infection is not mediated through RIG-I since transfection of 5'pppRNA did not induce phosphorylation on T56. However, next to sensing of viral products such as dsRNA, nutrient sensing in metabolic stress pathways also plays a role in the innate immune response to viruses [255]. During nutrient deprivation, several pathways are activated to conserve cellular resources [256, 257]. Protein synthesis consumes a large part of the cellular energy and is one of the processes that is inhibited to save energy. There are several energy sensors that could mediate an increase in eEF2 phosphorylation.

During cellular stress an increase in the AMP:ATP ratio activates the AMPK pathway. Activation of AMPK results in the general inhibition of ATP-consuming (anabolic) processes, including protein synthesis, and the activation of ATP-generating (catabolic) processes. AMPK has already been described as an antiviral sensor [258]. In our CHIKV infection experiment, phosphorylation on S108 of the beta subunit of the AMPK complex (PRKAB1) was significantly increased at 8 and 12 h p.i. ($\log_2$ ratio 0.95 and 0.69) and phosphorylation on this residue is known to stimulate AMPK activity [259]. Other phosphorylation sites were not identified for this complex. Despite the likely increase in AMPK activity during alphavirus infection, our data indicate that phosphorylation of eEF2 during alphavirus infection is not mediated via AMPK activation alone. Inhibition of AMPK with dorsomorphin, an AMPK inhibitor even resulted in a small increase in phosphorylation during SFV infection and had no effect on translation of viral capsid protein. It has been shown that eEF2 becomes phosphorylated during RVFV infection, a –ssRNA virus of the bunyavirus family. In this study phosphorylation on eEF2 still occurred in AMPK knockout cells and was thus also independent of AMPK activation [235], although at this point we cannot prove that the mechanism of phosphorylation induction in these diverse viruses is triggered by the same stimulus.

Intracellular cyclic AMP (cAMP) is sensed by PKA. An increase in cAMP during alphavirus infection could result in increased activation of eEF2K. Inhibition of PKA with KT5720 resulted in a small decrease in eEF2 phosphorylation during SFV infection but if PKA is involved in triggering the increase in eEF2 phosphorylation it is most likely not the only pathway that is involved since inhibition of PKA should then have resulted in a stronger inhibition of phosphorylation.

The amino acid availability is rate-limiting in protein synthesis. mTORC1 functions as a nutrient sensor and its promotion of translation is directed by the intracellular amino acid supply [260]. mTORC1 inhibition due to a decrease in intracellular L-leucine levels

could result in an increase in phosphorylation of eEF2 [261]. However, stimulation of mTORC1 with L-leucine did not prevent phosphorylation of eEF2 during SFV infection. The highest concentrations even resulted in a small increase in phosphorylation and inhibited SFV as the amount of capsid that was produced during infection was decreased.

The preliminary compound studies suggest that the pathways that were tested so far do not appear to be involved in inducing eEF2 phosphorylation during alphavirus infection. The very early triggering of eEF2 phosphorylation may also argue against an energy sensor as the initiator because the level of replication and translation of viral genomes is not yet so high after 2 h p.i. that it is expected to deplete nutrient pools. The small effect that inhibition of PKA with KT5720 had on eEF2 phosphorylation suggests that this pathway may enhance the stimulation later during infection, when energy resources do become exhausted by massive replication and translation of viral RNAs, but that the initial trigger is another stimulus. Additional studies which include more controls are required because at this point we cannot prove that the compounds truly inhibited or activated the pathways that they were supposed to target in our experiments. The inclusion of known inducers of eEF2 phosphorylation through these pathways, out of the context of alphavirus infection, should help us to determine whether inhibition was truly successful. We have not yet tested all the pathways that could be responsible for eEF2 phosphorylation as it is also possible that one of the MAP kinase pathways is involved, but this requires additional studies. Our data suggests that activation of MAPK1 and MAPK3 is probably minimal during CHIKV infection. Phosphorylation of Y187 on MAPK1 and Y204 on MAPK3 increased significantly during CHIKV infection, but phosphorylation of T185 on MAPK1decreased and the status of T202 on MAPK3 did not change. Phosphorylation on both residues is required for full activation of the kinases [262].

No involvement of structural proteins

The alphavirus 6k protein is cotranslationally translocated across the ER membrane and has been reported to trigger depletion of reticular calcium stores, resulting in an increase in cytosolic Ca2+ levels [263]. Since eEF2K activity is Ca2+/Calmodulin dependent this led us to hypothesize that expression of 6k during alphavirus infection causes an increase in cytosolic calcium levels which results in eEF2K activation to slow down translation. However, transfection of a replicon that lacks all structural proteins into BHK-21 cells still induced high levels of eEF2 phosphorylation, so none of the structural proteins is responsible for triggering eEF2 phosphorylation. The early induction of eEF2 phosphorylation also argues against the involvement of 6k, since at 2 h p.i. the abundance of viral proteins in the cells is still very low and most likely not enough to cause massive depletion of calcium stores.

Other mechanisms that could trigger eEF2 phosphorylation

The transfection of viral RNA into cells circumvents clathrin-mediated endocytosis. Since the transfected replicon RNA was still able to induce strong eEF2 phosphorylation, it is not likely that one of the steps during CHIKV entry is responsible for triggering phosphorylation. Perhaps RNA structures in the viral genome are sensed in the cytosol and this triggers an innate immune response that slows down protein synthesis. Inhibition of CHIKV replication by pretreatment with 5'pppRNA of cells prior to infection also prevented the induction of eEF2 phosphorylation suggesting that the presence of a certain level of e.g. viral RNA or protein is required before phosphorylation is induced. In different cell lines infected with CHIKV, infection does not progress at the same rate. In Vero E6 cells, in which viral proteins could be detected with WB by 4 h p.i., the induction of phosphorylation could also be observed at that time point. In MRC-5 cells, in which infection progresses slower, viral proteins could not be detected until 8 h p.i. while only a small amount of phosphorylated eEF2 could be detected at 6 h p.i. This also suggests that a certain level of e.g. viral RNA or protein has to be reached before the induction of phosphorylation is fully triggered.

It was recently reported that eEF2 is required for the Gag-mediated block of stress granule assembly during HIV-1 infection [264]. Although this paper suggests that eEF2 phosphorylation might decrease during HIV-1 infection, this stress granule connection might be interesting to explore for CHIKV as well since CHIKV also exhibits complex interactions with stress granules and some of their components, such as the G3BPs [65].

An increase of cytosolic Ca2+ early during alphavirus infection could trigger activation of eEF2K, but there is conflicting evidence as to whether alphaviruses induce an increase in cytosolic calcium levels during infection [265, 266].

eEF2 abundance does not change during alphavirus infection

Other groups have reported a decrease in eEF2 abundance during CHIKV infection [111, 112, 119]. However, in our previous proteomics study the abundance of eEF2 did not change [208] and we have not been able to reproduce this finding in our WB analysis at the time points that we studied either. In the other studies, samples were analyzed at much later time points in infection (24 h p.i. or later) than we used and eEF2 might be degraded during this stage of infection. This might not be an eEF2-specific property, however, since alphaviruses induce a host translational shut-off [100]. Two of the studies that quantified a decrease in eEF2 levels used 2D-DIGE and most likely the large increase of phosphorylated eEF2 during CHIKV infection significantly decreased the amount of eEF2 in the spot containing the unmodified form of eEF2, which could explain the observed decrease in abundance in these studies [112, 119].

It has been shown that ribosomes and eEF2 co-localize with aggregation sites of the C protein during SINV infection [267]. eEF2 was also identified as one of the interaction

partners of SINV nsP3 during infection with a nsP3-GFP expressing SINV mutant [94]. In these studies the phosphorylation status of eEF2 was not determined, but it would be interesting to determine whether these interactions occur with the modified or unmodified form of eEF2.

CONCLUSION

In our previous proteomics analysis of CHIKV-infected cells, there were only minimal changes in protein abundance and most of the proteins that did show a significantly changed abundance were downregulated [208]. The current study shows that more and larger changes occur on the level of phosphorylation and that both increased phosphorylation and dephosphorylation can be observed. Our data suggest that, at least for alphaviruses, cellular responses to viral infection may mainly occur on the level of PTMs and it would certainly be interesting to study other PTMs in this context. However, it continues to be important to analyze changes in total protein abundance as well, since a two-fold change in protein abundance can also result in a two-fold change in phosphorylation status, which would falsely indicate that the phosphorylation status is regulated while it is in fact the protein abundance.

Our study revealed that early during infection with several alphaviruses and a picornavirus, eEF2 becomes phosphorylated. The phosphorylation status of sites of other proteins involved in translation also changed significantly during CHIKV infection and possibly also contribute to translation inhibition. The phosphorylation of eEF2 is unfavorable for CHIKV replication as a reduction in the eEF2 levels resulted in a decreased expression of CHIKV proteins. The trigger for eEF2 phosphorylation during CHIKV infection is not yet known, but it does not appear to be mediated through one of the energy-sensing pathways, RIG-I signaling, alphavirus structural proteins or the viral entry process. We hypothesize that the increase in eEF2 phosphorylation is triggered by a special/unique feature of the viral genome (e.g. RNA structure) or by the translation or transcription of (alpha)virus RNA as a novel antiviral/stress response that slows down translation.

ACKNOWLEDGMENTS

SUPPORTING INFORMATION

Supporting information tables

Supporting Information Table S1: Complete list of identified phosphorylation sites, the first sheet contains all sites identified in host proteins, the second sheet contains the sites identified in CHIKV proteins. In case this research has not yet been published this table can be requested from the author.

Supporting Information Table S2: Lists with phosphorylation sites that were significantly modulated during CHIKV infection. Each time point is put in a separate sheet. In case this research has not yet been published this table can be requested from the author.

Supporting Information Table S3: GO terms that the list with unique phosphoproteins that contained phosphorylation sites that were significantly modulated during CHIKV infection was enriched for.

GO_id Molecular function	Term	NumberOfGenes	p-value	p-value_fdr	p-value_ bonferroni
GO:0044822	poly(A) RNA binding	42	2.81E-24	1.11E-20	1.11E-20
GO:0003723	RNA binding	44	3.76E-23	7.4E-20	1.48E-19
GO:0003779	actin binding	12	1.97E-07	0.000258	0.000775
GO:0003676	nucleic acid binding	39	2.66E-07	0.000262	0.00105
GO:0005515	protein binding	76	5.75E-07	0.000452	0.00226
GO:0008092	cytoskeletal protein binding	15	2.62E-06	0.00172	0.0103
GO_id Biological processes	**Term**	**NumberOfGenes**	**p-value**	**p-value_fdr**	**p-value_ bonferroni**
GO:0046907	intracellular transport	23	5.70E-08	7.66E-04	7.66E-04
GO:0051641	cellular localization	28	5.11E-07	2.81E-03	6.88E-03
GO:0051649	establishment of localization in cell	25	6.90E-07	2.81E-03	9.29E-03
GO:0007010	cytoskeleton organization	16	8.84E-07	2.81E-03	1.19E-02
GO:0071840	cellular component organization or biogenesis	45	1.04E-06	2.81E-03	1.40E-02
GO:0008380	RNA splicing	11	1.44E-06	3.23E-03	1.94E-02
GO:0007167	enzyme linked receptor protein signaling pathway	17	2.02E-06	3.63E-03	2.71E-02
GO:0033036	macromolecule localization	26	2.16E-06	3.63E-03	2.90E-02
GO:0006996	organelle organization	31	3.49E-06	4.93E-03	4.70E-02
GO:0016043	cellular component organization	43	3.70E-06	4.93E-03	4.98E-02
GO_id Cellular Component	**Term**	**NumberOfGenes**	**p-value**	**p-value_fdr**	**p-value_ bonferroni**
GO:0005925	focal adhesion	27	4.95E-23	3.75E-20	7.84E-20
GO:0005924	cell-substrate adherens junction	27	7.09E-23	3.75E-20	1.12E-19

**Supporting Information Table S3** (continued)

GO_id Cellular Component	Term	NumberOfGenes	p-value	p-value_fdr	p-value_bonferroni
GO:0030055	cell-substrate junction	27	7.09E-23	3.75E-20	1.12E-19
GO:0005912	adherens junction	27	2.40E-21	9.49E-19	3.80E-18
GO:0070161	anchoring junction	27	6.31E-21	2.00E-18	9.99E-18
GO:0030054	cell junction	31	4.69E-15	1.24E-12	7.42E-12
GO:0015629	actin cytoskeleton	17	1.59E-11	3.60E-09	2.52E-08
GO:0044422	organelle part	68	1.33E-09	2.63E-07	2.11E-06
GO:0044446	intracellular organelle part	67	1.51E-09	2.66E-07	2.39E-06
GO:0043228	non-membrane-bounded organelle	42	2.13E-09	3.07E-07	3.38E-06
GO:0043232	intracellular non-membrane-bounded organelle	42	2.13E-09	3.07E-07	3.38E-06
GO:0044428	nuclear part	44	9.81E-09	1.29E-06	1.55E-05
GO:0031988	membrane-bounded vesicle	42	1.72E-08	2.09E-06	2.72E-05
GO:0005829	cytosol	38	4.07E-08	4.60E-06	6.44E-05
GO:0031981	nuclear lumen	40	1.32E-07	1.39E-05	2.09E-04
GO:0043230	extracellular organelle	36	2.26E-07	1.89E-05	3.59E-04
GO:0070062	extracellular exosome	36	2.26E-07	1.89E-05	3.59E-04
GO:0065010	extracellular membrane-bounded organelle	36	2.26E-07	1.89E-05	3.59E-04
GO:0031982	vesicle	40	3.47E-07	2.75E-05	5.50E-04
GO:0005856	cytoskeleton	26	4.92E-07	3.71E-05	7.80E-04
GO:0032991	macromolecular complex	43	5.22E-07	3.76E-05	8.26E-04
GO:0005654	nucleoplasm	35	5.80E-07	3.99E-05	9.18E-04
GO:0005634	nucleus	56	7.48E-07	4.94E-05	1.18E-03
GO:0042641	actomyosin	6	1.31E-06	8.28E-05	2.07E-03
GO:0070013	intracellular organelle lumen	42	1.72E-06	1.05E-04	2.73E-03
GO:0043233	organelle lumen	42	2.54E-06	1.49E-04	4.03E-03
GO:0031974	membrane-enclosed lumen	42	3.55E-06	2.01E-04	5.63E-03
GO:0043227	membrane-bounded organelle	76	3.91E-06	2.14E-04	6.19E-03
GO:0044421	extracellular region part	38	4.85E-06	2.56E-04	7.69E-03
GO:0043234	protein complex	37	7.66E-06	3.92E-04	1.21E-02
GO:0005911	cell-cell junction	10	1.34E-05	6.63E-04	2.12E-02
GO:0016020	membrane	59	2.24E-05	1.08E-03	3.55E-02
GO:0005737	cytoplasm	68	2.55E-05	1.19E-03	4.05E-02
GO:0030529	ribonucleoprotein complex	12	3.03E-05	1.37E-03	4.80E-02

Supporting information figures
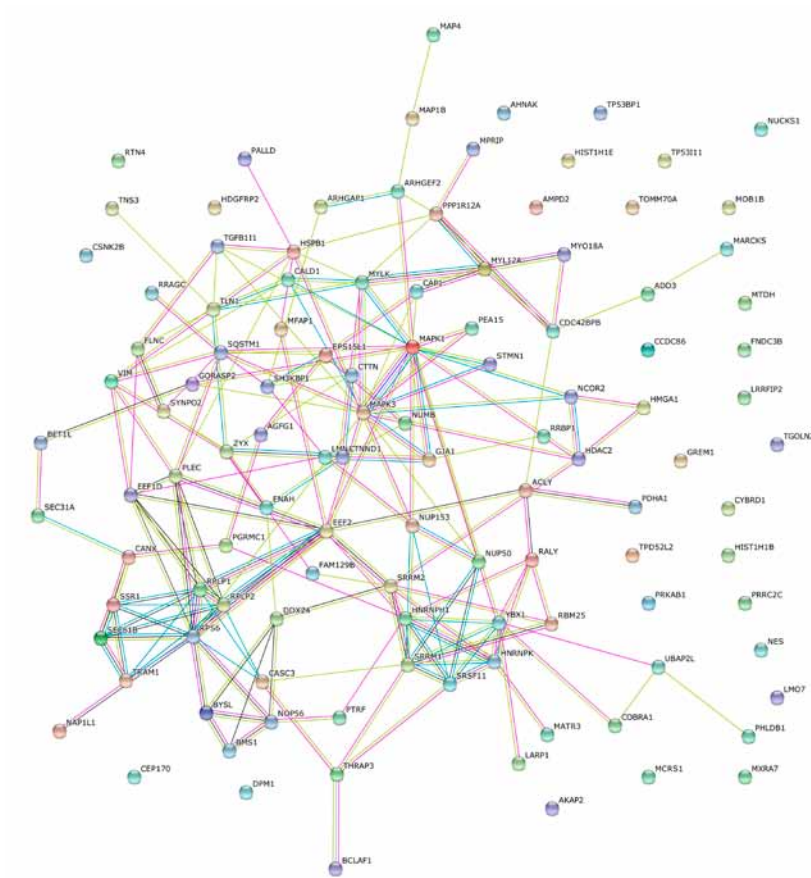
Supporting Information Figure S1: STRING network with medium confidence of the 113 phospho-proteins that contained phosphorylation sites that were significantly modified during CHIKV infection. The dataset was enriched for protein-protein interactions with a p-value of 2.22e-16
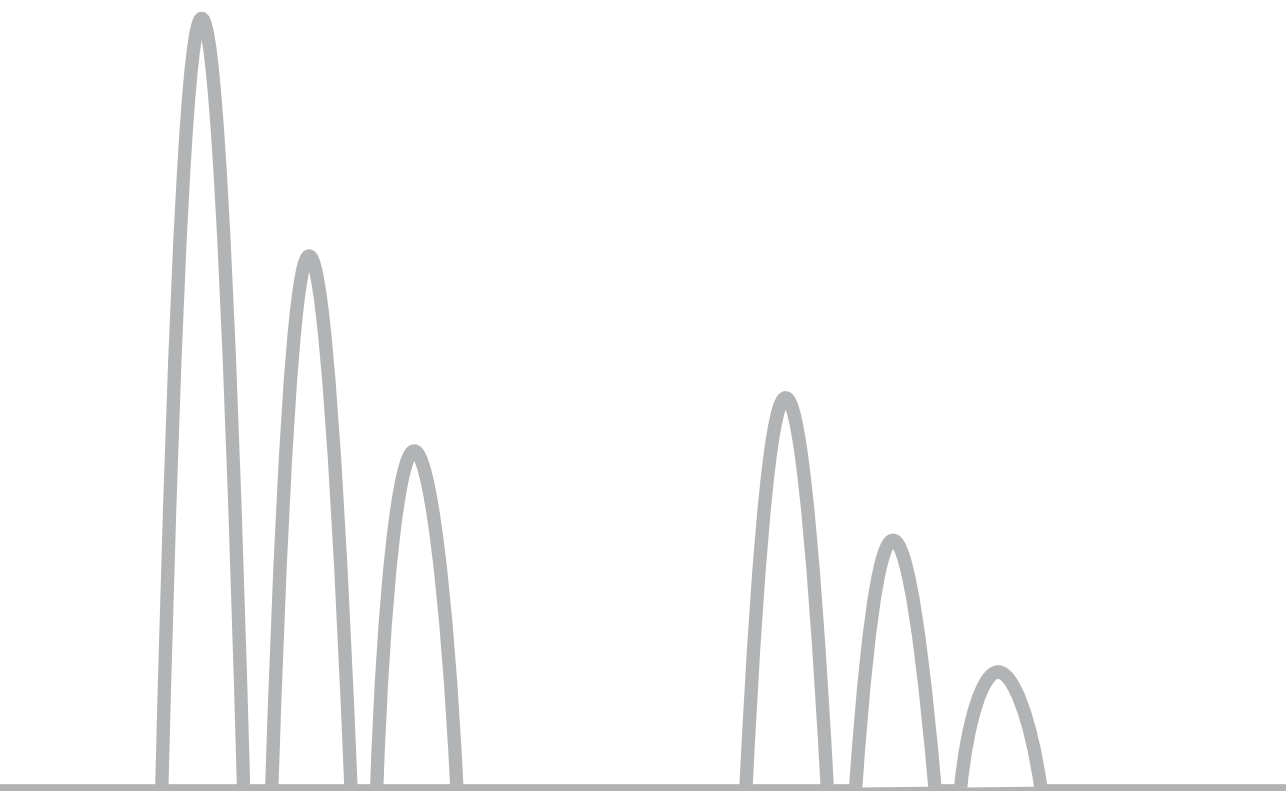
Chapter 4

General discussion part 1

Chapter 2 and 3 of this thesis described two quantitative proteomics studies of CHIKV-infected cells. In chapter 2 the changes in host protein abundance following CHIKV infection were determined, while chapter 3 identified proteins of which the phosphorylation status changed during infection. This chapter summarizes these findings, places them in a broader context and discusses the reproducibility of proteomics studies.

Changes in protein abundance and phosphorylation status during CHIKV infection

As described in chapter 2, the changes in protein abundance in CHIKV-infected cells during the course of infection were rather limited, especially prior to the onset of host translational shut-off. Not only the number of proteins showing a significantly changed abundance was limited, but also the extent of change was quite small (<2 fold) for most proteins. Despite the lack of stronger effects, this study yielded interesting results. The majority of the proteins for which a significant change was observed showed a decreasing abundance after CHIKV infection (**Figure 1, chapter 2**). For many proteins this decrease was probably the result of their normal turn-over, i.e. natural degradation in the absence of translation due to the virus-induced host shut-off that starts around 8 h p.i. However, certain proteins are most likely specifically targeted by the virus for degradation. The best example is the RNA polymerase II (POLR2) complex of which most subunits were progressively degraded during CHIKV infection (**Table 2, chapter 2**). This is in line with a previous study that found that the Rpb1 subunit, which is the catalytic subunit of the POLR2 complex, is targeted for degradation by nsP2 [63]. The same could be true for Rnd3, DDX56 and Plk1, three of the proteins chosen for follow-up research. CHIKV replication was reduced in cells overexpressing these proteins, suggesting that their presence is not beneficial for the virus (**Figure 3, chapter 2**). The decreasing abundance of these proteins might be explained by the viral manipulation of the intracellular environment to create optimal conditions for replication, or could even be part of a specific strategy to evade antiviral responses.

Based on the results in chapter 2, one could argue that a proteomics study studying only changes in protein abundance may not be the best approach to identify host factors that respond to viral infection in the case of viruses that kill their host cell relatively quickly. During the relatively short course of the infection, the time frame for induction of changes in protein abundance is simply not large enough, especially when the virus induces a transcriptional and/or translational host shut-off. For viruses that cause a persistent infection, meaning that cells can be monitored over the course of several days, this strategy seems a better option [202].

The low number of changes observed during CHIKV infection suggested that the cellular response to infection during the first cycle of replication does not affect protein abundance. The much more abundant and larger changes that were observed at the

level of phosphorylation status, described in chapter 3, proved this assumption to be correct. Several phosphorylation sites showed large (>4 fold) increases or decreases in phosphorylation during CHIKV infection, especially at 8 and 12 h p.i. Phosphorylation on one site, T56 on eEF2, even increased >50 fold at 8 and 12 h p.i., and already showed a significant increase in phosphorylation at 2 h p.i. Therefore, the role of eEF2 in the infection with CHIKV and other viruses was studied in more detail. At this point it remains unknown what the trigger is for the induction of eEF2 phosphorylation during alphavirus infection. However, it does not appear to rely on pathogen-associated molecular pattern (PAMP) recognition (**Figure 5, chapter 3**) or energy sensing pathways (**Figure 4, chapter 3**), although additional studies are required to completely exclude their involvement. The presence of the viral structural proteins is not required and merely the presence of a small amount of viral (uncapped) RNA is not enough to induce phosphorylation (**Figure 6, chapter 3**). Perhaps the RNA structures in the viral UTRs are sensed. The reduction of available eEF2 appears to restrict alphavirus replication since siRNA-mediated knockdown of eEF2 reduced translation of viral proteins, although this was a minor effect.

The induction of eEF2 phosphorylation slows down translation, which is unfavorable for a replicating virus, and this might be a general antiviral response. However, while this strong induction of eEF2 phosphorylation in response to infection was also observed with two other alphaviruses and a picornavirus (**Figure 2B and C, chapter 3**), and was previously reported to occur during infection with rift valley fever virus (RVFV), a –RNA virus from the bunyavirus family [235], two other viruses, equine arteritis virus (EAV), a +RNA virus from the arterivirus family and human adenovirus, a dsDNA virus, did not induce a measurable increase in eEF2 phosphorylation (**Figure 2D, chapter 3**). It will be interesting to determine whether viruses from other families also induce eEF2 phosphorylation as this may help to determine what the trigger is. The large increase in phosphorylation of eEF2 residue T56 upon alphavirus infection is induced early, as a 1.6-fold increase could already be observed as early as 2 h p.i. in CHIKV-infected MRC-5 cells in the proteomics analysis. This relatively small increase proved to be relevant since phosphorylation levels strongly increased as infection progressed. The western blot analysis was not equally sensitive as phosphorylated eEF2 could only be detected at 6 h p.i. in these cells (**Figure 2A, chapter 3**). This is an important caveat, since western blot analysis is often used to confirm mass spectrometry findings. However, in case of a phosphorylation site with a low occupancy at the start of the experiment, a relatively small increase in phosphorylation may be hard to detect with phospho-specific antibodies. It proved helpful to analyze different time points as the 2 h p.i. mass spectrometry analysis indicated that the induction of eEF2 phosphorylation started several hours prior to what we would have been able to determine from the WB analysis.

The use of quantitative proteomics approaches to study virus-host interactions

Quantitative proteomics is used for different purposes in the study of virus-host interactions [124]. It is often applied to study global changes during virus infection, either at a single time point or in a time course similar to what is described in this thesis. Some studies involve the overexpression of viral proteins, out of the context of a viral infection, to determine their impact on cellular protein expression [268, 269]. Proteomics studies have also been applied to study the recruitment or depletion of host proteins to/in certain organelles during viral infection [93, 270-272]. The use of e.g. SILAC can help to filter out non-specific interactors when proteomics is used to identify novel interaction partners of viral proteins [273-275].

Of the proteomics studies on CHIKV-infected cells that are described in this thesis, the phosphoproteomics analysis (**chapter 3**) offers a different view into molecular events at the protein level during infection than the analysis of protein abundance (**chapter 2**). However, the findings from the study described in chapter 2 exclude that the observed changes in phosphorylation status in chapter 3 merely resulted from changes in protein abundance. The major advances in post-translational modification (PTM) enrichment methods, mass spectrometer sensitivity and more powerful software tools [276] will probably lead to an increase in studies on proteome-wide changes in PTMs during viral infection. So far, the number of proteome-wide studies that analyzed changes in PTM abundance on host proteins during viral infection is still limited. Only four other studies in which phosphorylation sites were quantified during viral infection have been published so far. The signaling events induced during human immunodeficiency virus (HIV)-1 entry of CD 4+ cells were studied by quantifying phosphorylation sites 1 min p.i. This revealed that abundance of 239 phosphorylation sites out of 1757 quantified sites already changed significantly upon binding of HIV-1 to the cell surface and possibly already modulates the cell to facilitate steps later in the life cycle [239]. For influenza A virus, 366 phosphorylation sites on cellular proteins were identified by 24 h p.i., with phosphorylation on 43 and 35 phosphoproteins being up- or downregulated, respectively, but the majority of these sites changed less than 2-fold [237]. Porcine reproductive and respiratory syndrome virus (PRRSV) WUH3-infected porcine alveolar macrophages (PAMs) were analyzed and of 2125 identified phosphorylation sites, 292 and 373 sites had a significantly changed abundance at 12 and 36 h p.i., respectively. Most of these sites showed increased levels of phosphorylation [236]. Analysis of sendai virus (SeV) infected A549 cells resulted in the identification of 3947 phosphorylation sites, which revealed that mTOR signaling is needed for the host IFN response and viral protein synthesis during SeV infection [238].

Other modifications, such as ubiquitination and lysine acetylation, would also be interesting to study in the context of viral infection. Several virus families encode a deubiquitinating enzyme (DUB) that may serve to suppress the host innate immune

response [167, 277-279] and quantitative proteomics could be used to find out whether specific host cell proteins are targeted by these DUBs. For lysine acetylation it has already been shown that during borna disease virus (BDV) infection the cellular histone lysine acetylation levels are affected and such changes might also occur during infections with other viruses [280].

It will most likely also become more common to study multiple modifications at the same time, as it is becoming increasingly clear that there is cross-talk between different modifications that can modulate protein function. A certain PTM can, for example, signal the addition or removal of a modification on another residue or it could mask the recognition site of a second PTM [146, 148, 281, 282]. One has to keep in mind, however, that it is not only important to make an inventory of PTM changes on proteins, but that for understanding their biological relevance it is also crucial to determine the exact function that a certain PTM or a combination of PTMs has on its activity. Only for a very small subset of the experimentally observed modifications [283] it is currently known what their function is, which hinders interpretation of study results. There are also limitations at the level of the currently available bioinformatics tools. For example, for pathway analyses it is currently only possible to use protein or gene names as input, while for the study of PTMs more specific output might be obtained if the modified site could be included, since modifications often determines the protein's activity and function.

Reproducibility of quantitative proteomics studies

During the past decade, hundreds of studies that used proteomics to study viruses were published. Remarkably, when the different quantitative proteomics studies of CHIKV-infected cells or organisms are compared, there is little overlap between the datasets obtained in different labs (**chapter 2**) [111-114, 116, 119, 120, 208, 284]. Also for other viruses, such as influenza virus A H1N1 [285-287] or HIV-1 [288, 289], there is limited overlap in the lists of host proteins with significantly changed abundance when different studies are compared. Sometimes proteins are even identified as upregulated in one study and downregulated in another. This lack of reproducibility is not unique to proteomics studies as it is also seen with e.g. RNA interference screens [290-292]. The reasons for the limited reproducibility of this type of experiments are explored below.

One of the most important reasons for the limited overlap between various studies probably are the large differences in experimental set-up. Different cell lines and host organisms are used, which may obviously differ in their response to infection (or other stresses and stimuli), including, but not limited to, their ability to mount innate immune responses upon infection. Different virus isolates are used at different MOIs, which influences the dynamics of infection, especially when not every cell becomes infected at the same time. This might for example lead to secondary effects like the responses of non-infected cells to cytokines produced by infected neighboring cells. The timing

of sampling also greatly affects whether a certain protein is identified as displaying significantly changed abundance, e.g. it might be upregulated 10 h p.i. but not yet after 6 h p.i. and no longer after 20 h p.i. Especially for PTMs, timing will most likely be crucial for many proteins since e.g. phosphorylation usually is a transient event and dephosphorylation can also occur within minutes [158]. The differences in the experimental set-up between the various CHIKV studies almost certainly contributed to the limited overlap that was found. For example, the three infected cell culture-based studies each used a different cell line and CHIKV strain. The MOI was either 0.1, 2.5 or 5 and cells were harvested within a day (8, 10 and 12 h p.i.), 24 h p.i. or 48 h p.i. [111, 119, 208]. For viruses such as CHIKV that induce a host transcriptional and translational shut-off, sampling late in infection almost certainly results in the identification of false positive hits, especially when the normal half-life of host proteins is not considered (quick non-specific decrease of proteins with a rapid turnover) (**Figure S2, chapter 2**).

Another aspect that has a serious impact on reproducibility is how many peptides and proteins were identified by mass spectrometry in a given experiment. This number is influenced by many factors. The instrument used for data acquisition already influences how many peptides are identified in a sample [293]. The proteomics method used also influences which proteins are identified [294, 295]. For example, with 2D-DIGE approaches usually a lower number of proteins is identified than when SILAC, iTRAQ or label-free methods are used. In 2D-DIGE studies often only the spots that clearly show a change in abundance are chosen for mass spectrometry analysis and as a result these studies do not provide any information about proteins that were present in the infected cell but did not change abundance or PTM statuswhich can make it difficult to confirm a finding from another study. For CHIKV, the best example of this problem is the study by Fraisier et al. which analyzed the same sample using 2D-DIGE and iTRAQ. In this study only a few proteins of the same sample were identified by both methods, which is likely due to the fact that only 32 protein spots were analyzed in the 2D-DIGE experiment [114]. Several of the CHIKV studies quantified less than 600 proteins [112, 113, 116, 119], which makes it less likely that common proteins with significantly changed abundance are identified, especially since most proteins do not show very large changes in abundance during CHIKV infection. The reproducibility of PTM analyses will be hampered even more by the identification of a limited number of peptides in a study, since PTM identification often depends on the detection of a single peptide while a protein is usually identified by multiple peptides.

The apparent discrepancies between datasets can also be due to differences in analysis methods, software and databases used. Analysis of the same raw data files with a different software package or using a different protein database can already result in differences in protein identifications, which also affects quantification [293, 294]. The method of quantification also determines how accurate peptide and protein ratios are.

In general, metabolic labeling methods such as SILAC generate more precise quantifications than chemical labeling or label-free quantification strategies, since samples are combined immediately after harvesting [140, 294-296].

Another aspect that strongly affects reproducibility is the method used for determining which peptides or proteins display significantly changed abundance. A major factor in the limited reproducibility most likely is the very large number of false-positive hits in many of the published data sets. While it is certainly tempting to have a long list of proteins or peptides that show significantly changed abundance, it is important to be strict when determining which proteins should be on this list. The cut-off is often arbitrary (e.g. >2-fold change), while it depends on how accurate the quantifications are whether this cut-off is strict enough [294]. When statistics methods are employed it is not always clear whether there was any correction applied for multiple testing, while this is certainly necessary when large datasets are analyzed [297]. Sometimes, inappropriate strategies are employed, such as the lack of identification of a peptide or protein in one of the conditions [120, 238]. In shotgun proteomics, peptide lists from technical replicates generally overlap only 35-60% and peptide lists from biological replicates even less [293]. When the lack of its identification is used as a measure to determine whether the abundance of a protein or peptide has changed significantly, a large degree of chance is added to which proteins or peptides end up in the list and will certainly yield less reproducible results. The lists with proteins with significantly changed abundance from the CHIKV studies probably contain many false positives. For example, in a study in which both 2D-DIGE and iTRAQ were used to analyze the same sample, none of the proteins in the list from the 2D-DIGE analysis changed abundance more than 2-fold. Such changes were most likely not large enough to be also significant with the iTRAQ method in case these proteins were also identified in that analysis [114], but this could not be confirmed because the authors only included lists with proteins that significantly changed abundance in the paper.

Some issues are specifically related to certain techniques. Since the research described in chapter 2 and 3 is based on SILAC, some issues related to this technique are discussed in more detail. It was first described in 2003 that the use of arginine for SILAC labeling can result in an underestimation of the intensity of proline-containing peptides in the heavy-labeled sample because certain cell lines can convert arginine into proline [298]. It is simple, and important, to test whether a cell line exhibits this conversion before a large scale SILAC experiment is performed. Moreover, the conversion can easily be prevented by the addition of unlabeled proline and/or the reduction of the arginine concentration in the culture medium [299-301]. Another option is to correct computationally for this conversion after the mass spectrometry data has been acquired [302], but most software platforms used for analyzing proteomics data do not offer this option and it is best to avoid the conversion from happening in the first place. Even though

this conversion issue is well known, in many SILAC papers it is unclear whether the authors tested their cell line for this problem prior to using it for experiments or whether they took any precautions to prevent it. For example, HeLa cells are known to convert arginine into proline [300, 303]. Despite this, there are still SILAC studies published in which HeLa cells were used in which, based on the composition of the culture medium, it is likely that conversion took place [270, 304, 305]. This may have affected protein ratios in these studies and could have made certain findings less reproducible. Another issue with SILAC studies is that a label-swap is often omitted, while it has been shown that this procedure can solve some experimental errors such as incomplete labeling or arginine-to proline conversion that can lead to incorrect identification of proteins with significantly changed abundance [306]. In the studies described in chapter 2 and 3 some proteins or peptides were, for example, upregulated in the infected sample when this sample was labeled with the light amino acids and downregulated with an equally large change when the infected sample was labeled with the heavy amino acids. If a label swap would not have been performed, these proteins would have ended up in the list of proteins with significantly changed abundance, making the results less reproducible by increasing the number of false-positive hits. Using a label swap increases accuracy of the average ratios of biological replicates and should be standard practice when performing SILAC studies.

It would be very helpful to be able to compare the dataset from one proteomics study with those from other relevant studies, for example to determine whether a protein that is phosphorylated upon CHIKV infection is also modified in response to infection with other pathogens. It is currently rather hard to do so and sometimes it is even difficult to determine whether the protein was even identified in other studies. This is, among other issues, due to the fact that datasets are often represented in different ways with inconsistent annotation. It would be helpful if a standard approach for storing and annotating data could be implemented. For some studies only the lists with significantly up- and downregulated proteins are provided in the paper, while for a comparison between studies it is can also be relevant to know whether a certain protein was identified but perhaps did not change in abundance or PTM status. None of the other CHIKV studies supplied a complete list of identified proteins. For a quick comparison between different studies it would be advisable, and most convenient, when such lists of proteins or sites would be supplied as supplemental Excel tables, especially when they contain several thousands of entries, as these can be easily searched for proteins of interest and compared with other datasets. However, it is not uncommon to find them in PDF format. Although tables in PDF files can be converted into Excel tables, conversion is not always flawless which hinders proper searching and comparison.

Nowadays, proteomics journals usually require authors to deposit the raw and supporting data from proteomics studies in public data repositories such as PRIDE to
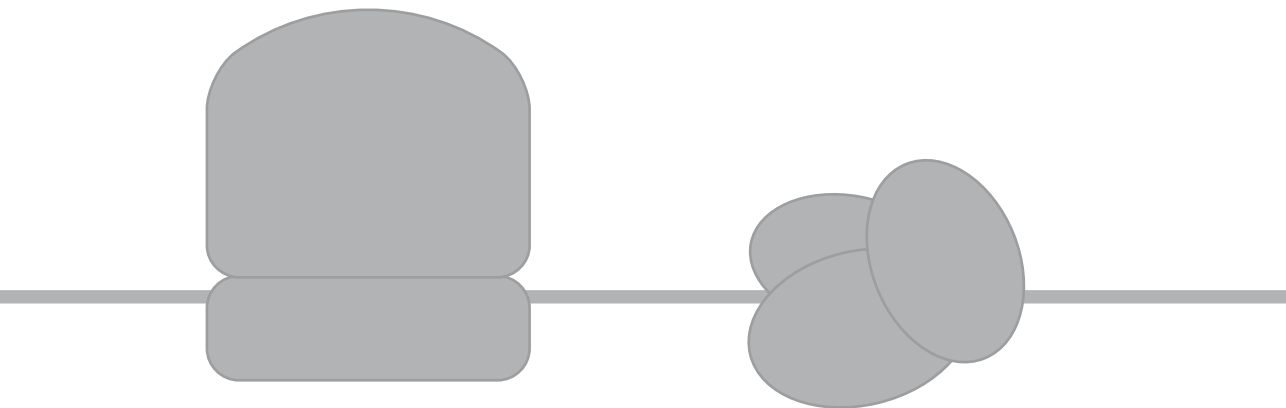
grant other scientists easy access [203]. For older studies or studies published in non-proteomics journals this is usually not the case and as a result it is not possible to find datasets for most of these studies. For the CHIKV proteomics studies, only one other paper mentions depositing the data with PRIDE [114]. Re-analyzing data from another lab using the same, database, search engine and statistics methods can remove some of the aspects that cause variability between different studies.

Conclusion

Quantitative proteomics can aid in the identification of proteins and pathways that were previously not associated with viral infection due to the hypothesis-generating nature of the technique. Nevertheless, after its identification, it can be a challenge to determine the exact function a certain host protein exerts during the viral replication cycle. The increasing number of studies from different labs studying the same virus shows that it remains important to scrutinize the results as many factors can influence reproducibility. Overall, quantitative proteomics, and quantitative phosphoproteomics in particular, remain valuable tools for the analysis of virus-host interactions and will most likely be used increasingly often now that PTM enrichment methods are steadily improving.

PART 2

–2/–1 programmed ribosomal frameshifting in arteriviruses

Chapter 5

General introduction part 2

CANONICAL TRANSLATION IN EUKARYOTES

Ribosomes convert messenger RNA (mRNA) nucleotide sequences into protein amino acid sequences; this process is called translation. Eukaryotic translation is a complex process that is highly regulated and involves many different factors. It can be subdivided into three phases; initiation, elongation and termination/ribosome recycling [307-311].

Translation initiation

The first step in translation initiation is the formation of a ternary complex (TC) that consists of initiator methionyl-transfer RNA (met-tRNA$_i$) and the GTP-bound form of eukaryotic initiation factor 2 (eIF2). The TC assembles with the small (40S) ribosomal subunit and eIF1, 1A, 3 and 5 to form the 43S preinitiation complex (PIC). Initiation factors eIF4B, 4H and 4F (consisting of 4A, 4E and 4G) bind the 5′ end of the mRNA that in eukaryotes contains a 7-methylguanosine (m7G) cap [309, 312]. Eukaryotic translation is dependent on this 5′ cap [313]. The 3′ end of the mRNA is bound by the poly(A)-binding protein (PAPB). The 5′ and 3′ ends of the mRNA are brought together through the interaction of eIF4G with both eIF4E and PAPB, resulting in a circular activated messenger ribonucleoprotein (mRNP). The PIC is then recruited to the activated mRNP at the 5′ end of the mRNA [309, 312]. After the PIC has been loaded onto the 5′ end, it starts to scan the mRNA in a linear base-by-base fashion until it encounters a start codon in a favorable sequence context (Kozak consensus sequence). In eukaryotes the start codon generally is AUG, although there are some exceptions to this rule [313]. Recognition of the start codon halts mRNA scanning by the PIC. The start codon is recognized through base-pairing between the anticodon of Met-RNA$_i$ and the AUG codon in the peptidyl-tRNA (P)-site of the 40S subunit. This results in the dissociation of eIF1 from the ribosome which triggers full accommodation of the tRNA in the P-site [314] and conversion of eIF2 to its GDP-bound state [315]. Release of GDP-bound eIF2 and most other initiation factors from the PIC and binding of GTP-bound eIF5B to the complex facilitate large (60S) subunit joining to form the 80S initiation complex (IC), after which eIF5B hydrolyzes and dissociates. The last subunit to leave the IC is eIF1A, although eIF3 possibly remains associated with the ribosome during elongation [309, 312]. A schematic representation of the translation initiation phase is depicted in Figure 1.
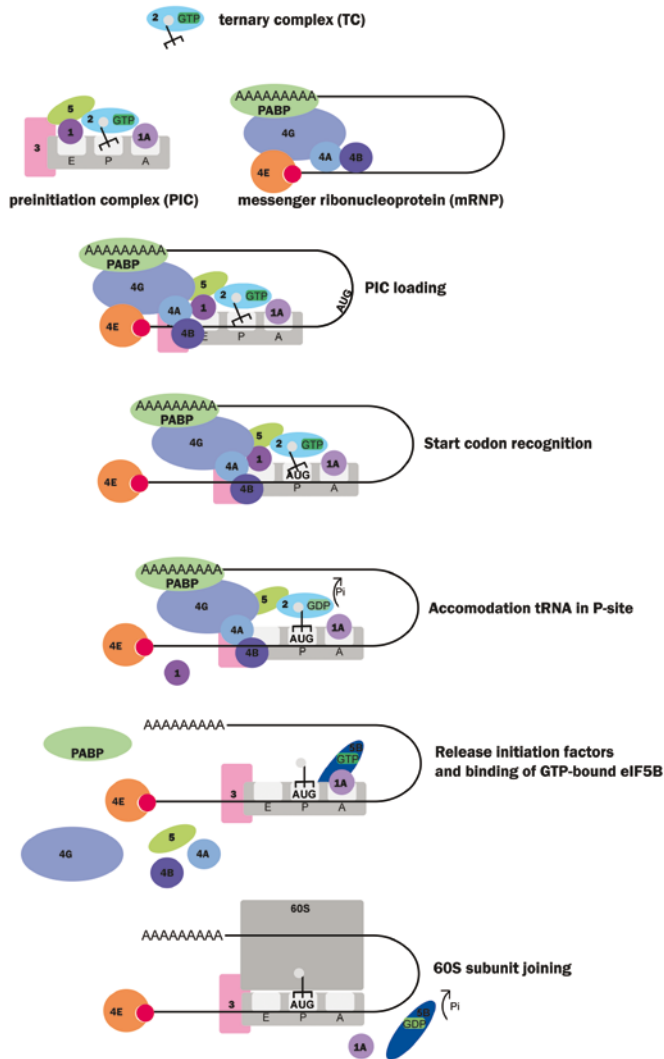
Figure 1: Translation initiation phase. A ternary complex (TC) consisting of initiator methionyl-transfer RNA (met-tRNA) and the GTP-bound form of eIF2 is formed. The TC assembles with the small (40S) ribosomal subunit and eIF1, 1A, 3 and 5 to form the 43S preinitiation complex (PIC). Initiation factors eIF4B, 4H and 4F (consisting of 4A, 4E and 4G) bind the 5' end of a 7-methylguanosine capped mRNA. The 3' end of the mRNA is bound by the poly(A)-binding protein (PABP) and the 3' and 5' ends of the mRNA are brought together, resulting in a circular activated messenger ribonucleoprotein (mRNP). The PIC is recruited to the activated mRNP at the 5' end of the RNA and starts to scan until it encounters an AUG start codon. Start codon recognition halts mRNA scanning by the PIC. The start codon is recognized through base-pairing between the anticodon of the Met-tRNA$_i$ and the AUG codon in the ribosomal P-site. This results in dissociation of eIF1 and conversion of eIF2 to its GDP-bound state. Most of the initiation factors dissociate from the ribosome and GTP-bound eIF5B binds. The 60S ribosomal subunit joins to form the 80S initiation complex (IC) which results in eIF5B hydrolysis and dissociation. eIF1A is the last subunit to leave the IC, although it is possible that eIF3 remains associated during the elongation phase.
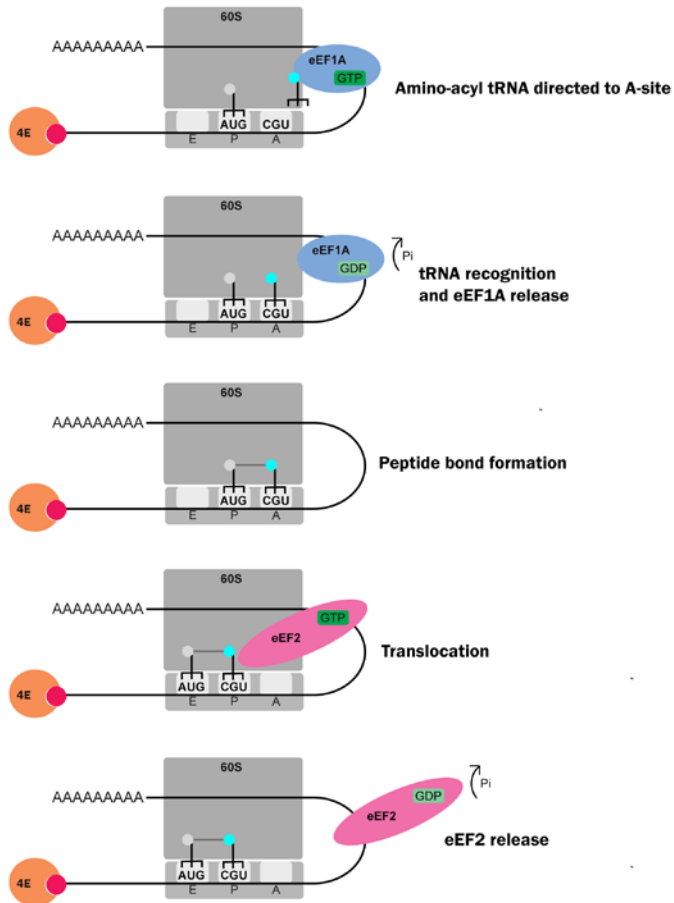
Figure 2: Translation elongation phase. At the start of the elongation phase the 80S ribosome is positioned on the mRNA with the Met-tRNA$_i$ paired to the start condon in the P-site. The A-site contains the second codon of the ORF. Amino-acyl tRNAs are bound by eEF1A in a GTP-dependent manner and directed to the A-site. tRNA recognition results in GTP hydrolysis which releases the elongation factor. Peptide bond formation with the peptidyl-tRNA in the P-site occurs fast. Translocation to the E- and P-site requires binding of eEF2 and GTP. Conformation changes in eEF2 and P$_i$ release following GTP hydrolysis allow movement of tRNA and mRNA. eEF2 is released from the ribosome and the A-site becomes available for the next aminoacyl-tRNA.
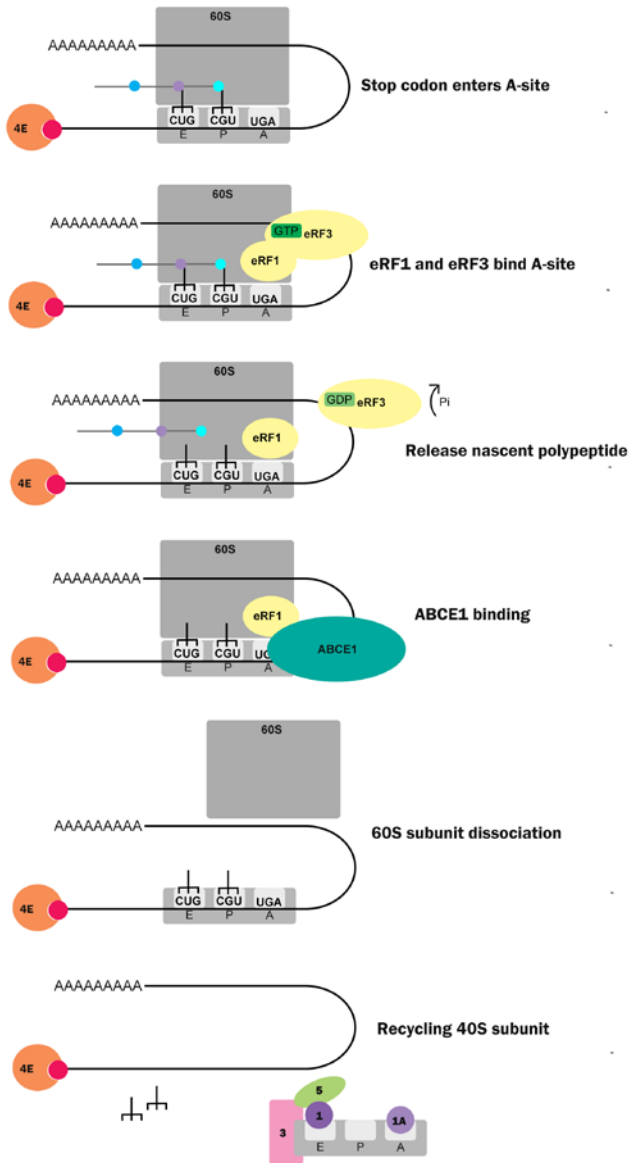
Figure 3: Translation termination/ribosome recycling phase. When one of the stop codons enters the ribosomal A-site, a ternary complex of eRF1 and eRF3 with GTP binds. GTP hydrolysis by eRF3 causes eRF1 to trigger hydrolysis of the polypeptidyl-tRNA in the peptidyl transferase site (PTC) which releases the nascent polypeptide. The 80S ribosome remains associated with the mRNA, the deacetylated tRNAs and eRF1. ABCE1 binds the complex with promotes 60S subunit dissociation. Release of tRNA and 40S subunit is mediated through binding of eIF1A, 1 and 3 to the 40S subunit. These subunits are reused to initiate translation.

Translation elongation

The ribosome has three tRNA binding sites, the acceptor (A)-site for aminoacyl-tRNA, the P-site for peptidyl-tRNA and the exit (E)-site for deacylated tRNA [316]. At the start of elongation, the 80S ribosome is positioned on the mRNA with the Met-tRNA$_i$ anticodon base-paired to the start codon in the P-site. The A-site contains the second codon of the ORF. Amino-acyl tRNAs are bound by eukaryotic elongation factor 1A (eEF1A) in a GTP-dependent fashion and are directed to the A-site of the ribosome. tRNA codon recognition results in GTP hydrolysis by eEF1A, which releases the elongation factor and enables accommodation of the amino-acyl tRNA into the A-site [307, 308]. Peptide bond formation with the peptidyl-tRNA in the P-site occurs fast as the peptidyl transferase center (PTC) in the large ribosomal subunit positions the substrates for catalysis [317]. After peptide bond formation the two ribosomal subunits rotate so that the acceptor ends of the tRNAs are placed in the E- and P-sites while the anticodon loops remain in the P- and A-sites [318]. Translocation to the E and P-sites requires binding of eEF2 (a GTPase) and GTP. Conformational changes in eEF2 and P$_i$ release following GTP hydrolysis allow movement of tRNA and mRNA and locks the subunits in the posttranslocation state [240]. eEF2 is released from the ribosome and the A-site becomes available for binding of the next aminoacyl-tRNA [308]. A schematic representation of the translation elongation phase is depicted in Figure 2.

Translation termination/ribosome recycling

Termination of translation requires two release factors (RFs), eRF1 and eRF3, that form a ternary complex with GTP [319]. When one of the stop codons (UGA, UAG, UAA) enters the ribosomal A-site this complex binds, with eRF1 being responsible for codon recognition [320]. GTP hydrolysis by eRF3 causes eRF1 to trigger hydrolysis of the polypeptidyl-tRNA in the PTC which results in release of the nascent polypeptide [311]. After release of the protein product, the 80S ribosome remains associated with the mRNA, the deacylated tRNA and eRF1. The ribosome dissociates into a free 60S subunit and a 40S subunit that is still associated with the mRNA and tRNA, a process promoted by the ATP-binding cassette protein ABCE1 [321]. Release of tRNA and mRNA from the 40S subunit can be mediated by eIF1A, eIF1 and eIF3 with its loosely associated eIF3j subunit. 40S subunits bound to eIF1A, eIF1 and eIF3 are reused to initiate translation [310, 311]. A schematic representation of the translation termination/ribosome recycling phase is depicted in Figure 3.

Non-canonical translation of RNA virus genomes

Viruses do not encode their own ribosomes and, therefore, for their protein synthesis all viruses are completely dependent on the translational machinery of the host cell. Especially for RNA viruses, this poses a challenge since their genomes are usually short

and often consist of a single strand of RNA from which all replicative, accessory and structural proteins need to be expressed. Additionally, many RNA viruses replicate in the cytoplasm and as a result do not have access to cellular systems for mRNA capping and polyadenylation that reside in the nucleus. Some viruses solve this problem by encoding their own capping enzymes or by snatching caps from cellular mRNAs [322]. Other viruses have evolved non-canonical translation strategies to circumvent the limitations of canonical 5′-end dependent eukaryotic translation, which allows synthesis of just a single protein from a single mRNA. Many +RNA viruses even employ more than one of these mechanisms to expand the possibilities for gene expression from their compact genomes [323-325]. The non-canonical translation strategies that are discussed below are depicted schematically in Figure 4.
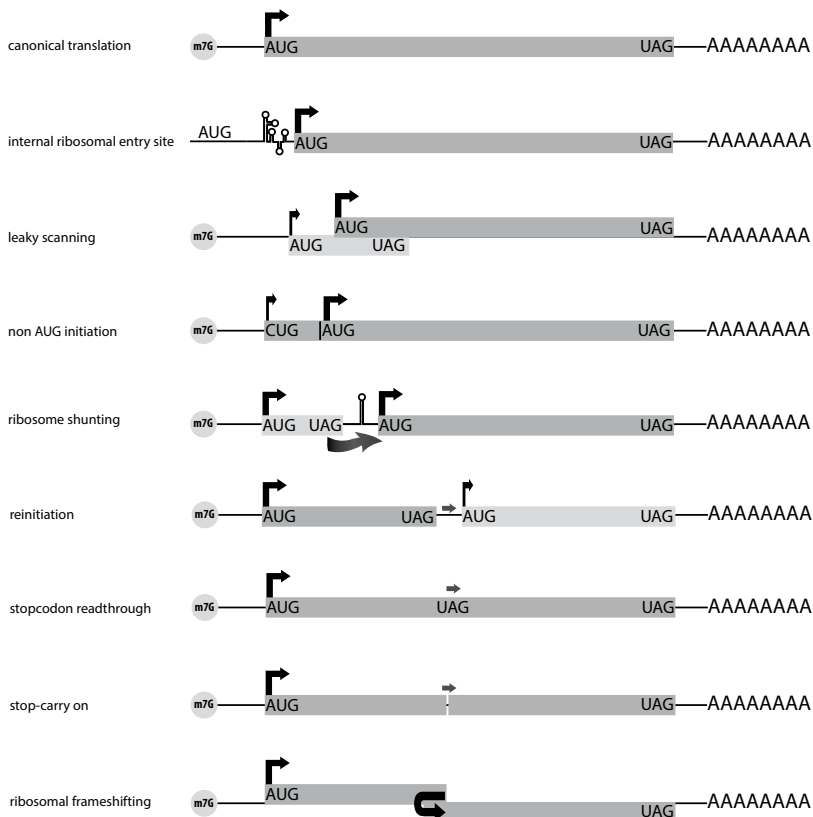
Figure 4: Non-canonical translation mechanisms. Eukaryotic canonical translation is shown at the top. ORFs are indicated as grey bars on the mRNA. The darker grey indicates the main route taken by translating ribosomes. Arrows at the start of the ORF indicate ribosome translation initiation. Horizontal arrows indicate where the ribosome moves non-canonically.

IRES

Several +RNA virus groups employ cap-independent translation via an internal ribosomal entry site (IRES). An RNA structure, usually located in the 5'-untranslated region (UTR) of the genome, replaces the cap and, depending on the type of IRES, some or all of the initiation factors [326], and serves as a binding site for ribosomes. Some viruses can initiate translation at multiple AUGs in the same or different reading frames from one IRES [327, 328]. Other viruses express two polyproteins from separate IRESes [329].

Leaky Scanning

In principle, ribosomes scan mRNAs in a linear fashion, starting at the 5' end and initiating translation at the first AUG start codon they encounter. However, when the context for recognition of that AUG start codon is not optimal (GCCRCCaugG, R is A or G, -3 R and +4 G are most important for efficient initiation) ribosomes can continue scanning and initiate translation at a downstream AUG instead. If the upstream AUG occurs in a very weak context nearly all ribosomes will continue scanning, if it is only somewhat suboptimal most ribosomes will initiate translation [330]. Leaky scanning is widely employed by viruses and allows the translation of two proteins from a single mRNA in the same or a different reading frame [323].

Initiation at non-AUG codons

In a strong context, near-cognate codons, such as CUG and ACG, can be recognized by the Met-tRNA$_i$ to initiate translation. Non-AUG initiation is often inefficient and, as a result, frequently occurs in combination with leaky scanning [330]. In viruses these combined mechanisms can enable the expression of 3 or 4 different proteins from a single strand of RNA [323].

Ribosome shunting

Ribosome shunting allows ribosomes to translate downstream ORFs in a 5' cap-dependent manner. The RNA usually contains a first short ORF followed by a stem-loop structure and a downstream main ORF. Translation of the first ORF is initiated at the start codon and terminates just before the stem-loop. It is thought that the small subunit of the ribosome retains certain initiation factors during translation of the first ORF. Ribosomes are then capable of bypassing the stem-loop and resume scanning at the landing site 3' of the stem-loop [311, 331].

Reinitiation

Certain initiation factors (most likely eIF3 and eIF4G) remain briefly associated with the 40S ribosomal subunit after joining of the 40S and 60S subunits. If translation of a short ORF is terminated before these factors are released, the 40S subunit can resume scan-

ning after dissociation of the 60S subunit. The rescanning ribosome initially does not have an associated TC and successful reinitiation of translation depends on the distance between stop codon and reinitiation site and TC availability [311]. In mammalian systems it is very rare for translation reinitiation to occur after translation of a long ORF. Mammalian caliciviruses have bicistronic subgenomic mRNAs with a short overlap region of the two ORFs. Translation reinitiation at the second ORF is dependent on translation of the upstream ORF. It requires the presence of a sequence element upstream of the second initiation site that is called the "termination codon upstream ribosome-binding site" (TURBS) [332]. Motifs within the TURBS sequence hybridize with 18S rRNA [333]. The TURBS also interacts with eIF3 and 40S ribosomal subunits. It is thought that the interaction with the TURBS secures 40S subunits that have terminated translation of the upstream ORF to the mRNA. Subsequent recruitment of the TC results in translation initiation to express the downstream ORF [334].

Stop codon read-through
The occurrence of a stop codon signals translation termination and induces the release of the polypeptide from the ribosome. However, the efficiency of translation termination is influenced by the type of stop codon (UGA and UAG are more leaky than UAA) and its context. When stop codon read-through occurs the stop codon is decoded by a near-cognate or "suppressor" tRNA, and translation continues until the next stop codon enters the ribosomal A-site. The UGA and UAA codons can induce the incorporation of tryptophan, arginine or cysteine. The UAG codon can induce the incorporation of tyrosine, glutamine or leucine [335]. Read-through results in the synthesis of a C-terminally extended polypeptide. Read-through occurs at a defined frequency and many viruses use it to express their polymerase at a lower level than other replicative proteins or to produce an extended version of their coat protein [323, 335].

Stop-carry on
Many viruses translate proteins from a single ORF as a polyprotein that is subsequently cleaved into the individual protein products by host and/or viral proteases. Stop-carry on is an alternative to proteolytic cleavage that also allows the expression of multiple proteins from a single ORF with near to 100% efficiency. The amino acid motif D(V/I)ExNPGP and the upstream amino acids that are present in the ribosome exit tunnel prevent formation of the peptide bond between glycine and the final proline. Instead, the nascent peptide is released from the ribosome after which the proline tRNA can bind the A-site and translation continues with proline as the N-terminal amino acid of the downstream-encoded product [210, 323].

Ribosomal frameshifting

In certain contexts, ribosomes can shift 1 or 2 nucleotides and continue translation in a different reading frame, which is called programmed ribosomal frameshifting (PRF). −1 PRF is widely used by +RNA viruses and allows them to produce proteins at a defined ratio or translate transframe proteins that share the N-terminal sequence [323]. A bipartite signal in the mRNA determines the frameshifting efficiency. A slippery sequence X_XXY_YYZ (_ separates 0 frame codons, X is any 3 identical nucleotides, YYY is AAA or UUU, Z is A, C or U) on which the ribosome backs up 1 nucleotide and a downstream stimulatory element, in most cases either a pseudoknot or a very stable RNA stem-loop structure [336]. The distance between slippery sequence and downstream element (6-8 nt) is also important [336]. It has been proposed that the resistance of the stimulatory RNA structure to unwinding induces tension in the mRNA. This causes ribosomes to pause at the slippery sequence with XXY in the P-site and YYZ in the A-site. Unpairing of mRNA and tRNA could release this tension, which then causes the ribosome to move to the −1 frame while still retaining base-pairing in the non-wobble positions [337, 338]. +1 and −2 frameshifting mechanisms are less common and their efficiency is usually very low [323].

PRRSV

The work in this second part of my thesis focuses on porcine reproductive and respiratory syndrome virus (PRRSV), a single-stranded enveloped +RNA virus belonging to the family *Arteriviridae* of the order *Nidovirales [339]*. For the expression of its polycistronic genome PRRSV and other arteriviruses employ a combination of strategies, including multiple non-canonical translation mechanisms.

PRRSV currently is one of the economically most important viruses in the swine industry [340]. In the late 1980's, distantly related PRRSVs, which are likely to be reclassified as separate viral species in the near future, simultaneously emerged in Europe [10] and the United States [11, 341]. The two current genotypes, type 1 (European) and type 2 (North-American), probably share a common ancestor but diverged extensively, retaining a nucleotide identity of only 60-70% [342, 343]. Within each genotype the nucleotide similarity is >80% [344].

Order Nidovirales

The viruses belonging to the *Nidovirales* have a similar genome organization and expression strategy and their key replicative enzymes, including the RNA-dependent RNA polymerase (RdRp) and helicase, presumably share a common ancestor [345, 346]. The order *Nidovirales* currently contains four families, the *Coronaviridae*, *Roniviridae*,

Mesoniviridae and *Arteriviridae* [347, 348]. Its coronavirus branch includes several human pathogens, in particular four "established" human coronaviruses (HCoVs), which generally cause common colds [349], and highly pathogenic zoonotic agents like severe acute respiratory syndrome coronavirus (SARS-CoV) [2, 4] and Middle East respiratory syndrome coronavirus (MERS-CoV) [350].

The arterivirus family currently consists of four recognized species that infect mammals: equine arteritis virus (EAV), lactate dehydrogenase-elevating virus (LDV) of mice, simian hemorrhagic fever virus (SHFV) and PRRSV [339]. The recently discovered wobbly possum disease virus (WPDV) will most likely be classified as an arterivirus [351] and several newly discovered monkey viruses appear to be distantly related to SHFV [352].

The PRRSV replicative cycle

The 13-16 kb arterivirus genome is encapsidated by the N protein and wrapped in a lipid envelope that is cell-derived and contains seven viral membrane proteins. The two major envelope proteins are the ORF5-encoded glycoprotein (GP5) and the membrane (M) protein. The five minor envelope proteins are GP2, GP3, GP4, envelope (E) protein and the ORF5a product [353-355]. For PRRSV all envelope proteins are required for the production of infectious virus particles [356, 357]. Cell entry occurs through clathrin-mediated endocytosis. Following endosome acidification and membrane fusion, the viral nucleocapsid is released into the cytosol [358, 359]. It is still uncertain which of the viral envelope proteins mediates fusion but host receptors CD163 and CD169 are likely required for PRRSV internalization and uncoating [360].

Arterivirus genomes are poly-adenylated and presumed to have a 5' cap. Upon cell entry the PRRSV replicase proteins are expressed from two 5' ORFs (ORF1a and ORF1b) as two precursor polyproteins (pp1a and pp1ab) that are proteolytically processed into 14 individual nonstructural proteins (nsps) by four viral proteases residing in nsp1α, nsp1β, nsp2 and nsp4 [361-364]. There is a small overlap between ORF1a and ORF1b and translation of pp1ab depends on a –1 PRF event that is directed by a "slippery sequence" and a downstream pseudoknot structure [342, 365]. The structural and accessory proteins are translated from the 3'ORFs 2-7 as a nested (Latin nidus = nest) set of 5'- and 3'-co-terminal subgenomic mRNAs (sgmRNAs) [218, 366]. The termini of most of the ORFs 2a to 7 overlap with neighboring genes and several sets of proteins are translated from the same sgmRNA, presumably through leaky scanning [354, 366, 367].

The arteriviral replication and transcription complex (RTC) is associated with membrane structures. These are complex networks of modified ER-derived structures in the perinuclear region of the cell that consist of double-membrane vesicles [368-370]. The RTC first synthesizes full-length and subgenome-length minus strand RNAs, which are used as templates for the synthesis of new genomes and the sgmRNAs, respectively [371]. Newly synthesized genomes become encapsidated by N protein [372]. Subse-

quently, nucleocapsids are enwrapped by membranes from smooth ER and/or Golgi complex [373, 374]. Virions are then transported to the plasma membrane by the exocytic pathway and released from the cell [375]. A schematic overview of the arterivirus replicative cycle is shown in figure 5.
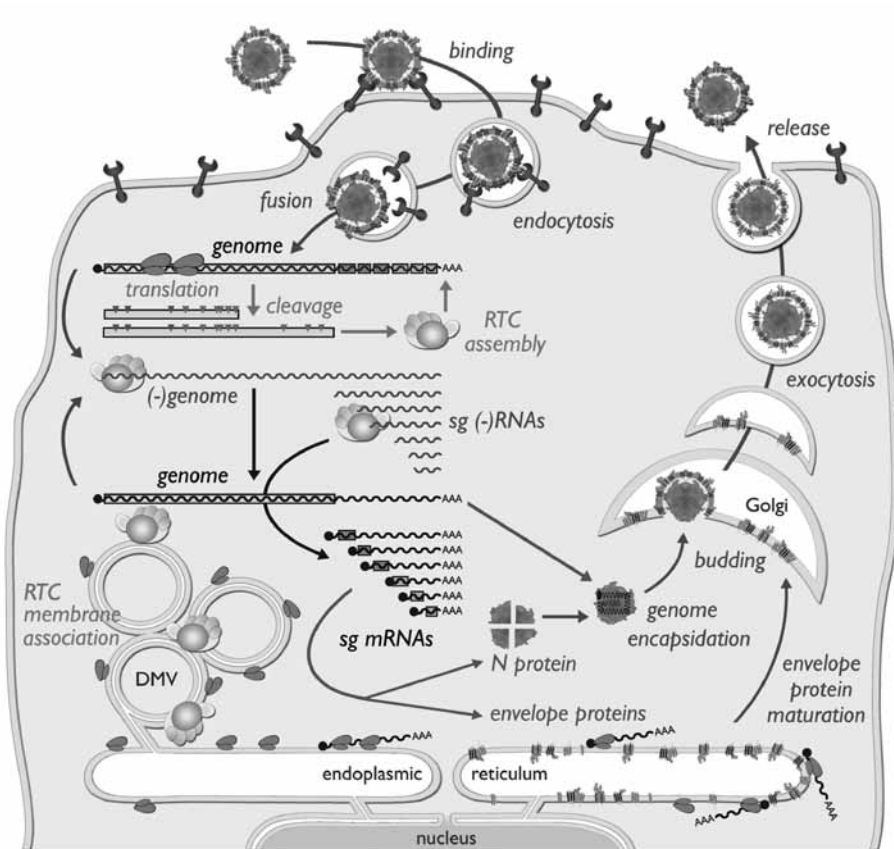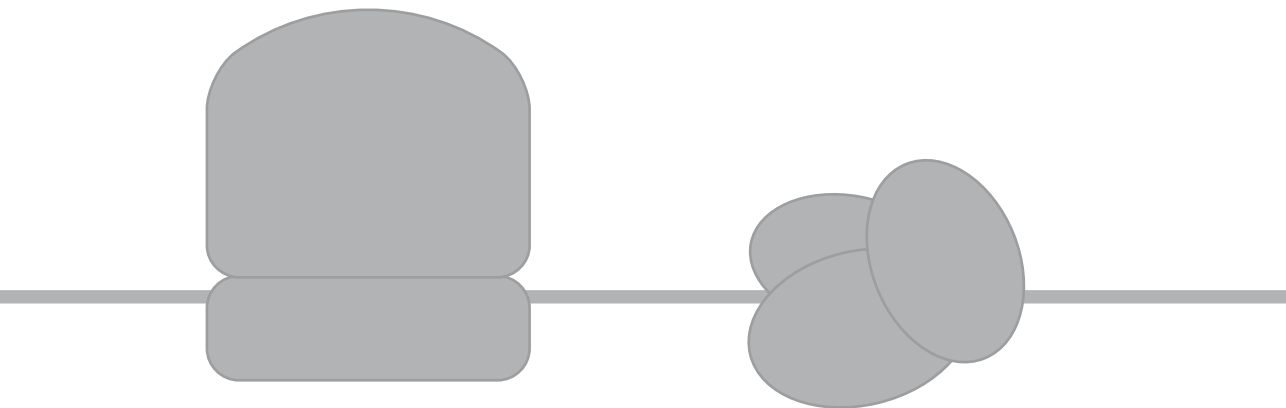

Figure 5: Overview of the arterivirus replicative cycle. Following entry by receptor-mediated endocytosis and endosomal membrane fusion the nucleocapsid is released into the cytosol. Genome translation yields replicase polyproteins pp1a and pp1ab that are cleaved by internal proteinases. The viral nonstructural proteins assemble into a replication and transcription complex (RTC) that first engages in minus-strand RNA synthesis. Both full-length and subgenome-length minus strands are produced, the latter serving as templates for the synthesis of sg mRNAs required to express the structural protein genes, which reside in the 3'-proximal quarter of the genome. Novel genomes are packaged into nucleocapsids that become enveloped by budding from smooth intracellular membranes, after which the new virions leave the cell using the exocytic pathway. For further details, see text. Reprinted with permission from [392].

Pathogenesis of PRRSV infection in swine

PRRSV can be transmitted between swine through semen, milk, direct contact and aerosol transmission [376-378]. Type 1 PRRSV mainly causes reproductive failure in sows late during gestation, while type 2 also causes respiratory disease in growing pigs [379]. The virulence of different PRRSV isolates can vary from very mild to highly pathogenic [380, 381]. Primary target cells for viral replication are porcine alveolar macrophages (PAMs) [10, 382] and the virus replicates and persists mostly in lungs and lymphoid organs [383, 384]. The acute phase of PRRSV infection is usually characterized by a high viral load and the presence of clinical symptoms that can last up to one month post infection. This is followed by a phase with low replication levels of persisting virus, until the virus is eventually cleared, which can take up to 150 days [385]. Both innate and adaptive immune responses against PRRSV are generally weak. PRRSV modulates the innate immune response by suppressing the production of type I interferons [386, 387] and regulating the expression of other cytokines [388]. The production of neutralizing antibodies occurs only late and their levels are generally low [385]. Development of cell-mediated immunity is weak and slow [388, 389]. The immunity acquired after clearance is often not effective against reinfection with a heterologous strain [389]. Vaccines against PRRSV are available, but are not completely effective [390, 391].

OUTLINE PART 2

In part 2 of my thesis, the identification of a new PRF mechanism in arteriviruses is described, which was explored in detail using PRRSV. **Chapter 6** presents the discovery of this mechanism, a highly efficient form of –2 PRF that ensures the expression of a previously unknown transframe protein, nsp2TF, that shares its N-terminal sequence with nsp2 but has a different C-terminal domain. In infected cells, nsp2 and nsp2TF localize to different cellular compartments and virus mutants that are incapable of expressing nsp2TF are seriously crippled. In **chapter 7** –2 PRF in PRRSV is shown to depend on the expression of the PRRSV nsp1β protein. This chapter further shows that –1 PRF takes place at the same frameshift site, resulting in the translation of a truncated form of nsp2, nsp2N. **Chapter 8** shows that –2/–1 PRF in PRRSV is also dependent on the presence of specific host proteins, namely poly (C) binding proteins (PCBP) 1 and 2, which appear to form a PRF-inducing protein complex together with nsp1β. In **chapter 9** the findings of part 2 of this thesis are summarized.

Chapter 6

Efficient –2 frameshifting by mammalian ribosomes to synthesize an additional arterivirus protein

Ying Fang[a,b#], Emmely E. Treffers[c,d,‡], Yanhua Li[a,3‡], Ali Tas[c], Zhi Sun[a], Yvonne van der Meer[c], Arnoud H. de Ru[d], Peter A. van Veelen[d], John F. Atkins[e], Eric J. Snijder[c#], and Andrew E. Firth[f#]

Departments of [a]Veterinary and Biomedical Science and [b]Biology/Microbiology, South Dakota State University, Brookings, SD 57007; [c]Molecular Virology Laboratory, Department of Medical Microbiology, Center of Infectious Diseases and [d]Department of Immunohematology and Blood Transfusion, Leiden University Medical Center, 2333 ZA, Leiden, The Netherlands; [e]BioSciences Institute, University College Cork, Cork, Ireland; and [f]Department of Pathology, University of Cambridge, Cambridge CB2 1QP, United Kingdom

[#]These authors contributed equally
[‡]These authors contributed equally

ABSTRACT

Programmed −1 ribosomal frameshifting (−1 PRF) is a gene-expression mechanism used to express many viral and some cellular genes. In contrast, efficient natural utilization of −2 PRF has not been demonstrated previously in eukaryotic systems. Like all nidoviruses, members of the *Arteriviridae* (a family of positive-stranded RNA viruses) express their replicase polyproteins pp1a and pp1ab from two long ORFs (1a and 1b), where synthesis of pp1ab depends on −1 PRF. These polyproteins are posttranslationally cleaved into at least 13 functional nonstructural proteins. Here we report that porcine reproductive and respiratory syndrome virus (PRRSV), and apparently most other arteriviruses, use an additional PRF mechanism to access a conserved alternative ORF that overlaps the nsp2-encoding region of ORF1a in the +1 frame. We show here that this ORF is translated via −2 PRF at a conserved G_GUU_UUU sequence (underscores separate ORF1a codons) at an estimated efficiency of around 20%, yielding a transframe fusion (nsp2TF) with the N-terminal two thirds of nsp2. Expression of nsp2TF in PRRSV-infected cells was verified using specific Abs, and the site and direction of frameshifting were determined via mass spectrometric analysis of nsp2TF. Further, mutagenesis showed that the frameshift site and an unusual frameshift-stimulatory element (a conserved CCCANCUCC motif 11 nucleotides downstream) are required to direct efficient −2 PRF. Mutations preventing nsp2TF expression impair PRRSV replication and produce a small-plaque phenotype. Our findings demonstrate that −2 PRF is a functional gene-expression mechanism in eukaryotes and add another layer to the complexity of arterivirus genome expression.

INTRODUCTION

In eukaryotes, translation initiation largely involves 5′ end-dependent scanning of mRNAs during which the small ribosomal subunit, in a complex with initiation factors, first binds to the 5′ cap structure and then scans in a 5′-to-3′ direction until it encounters the first suitable initiation codon, at which point translation commences [310]. Consequently, the vast majority of cellular mRNAs are essentially monocistronic (although efficient reinitiation can occur after translation of very short ORFs). The fact that the cellular translational machinery essentially only decodes the 5′-most long ORF of an mRNA imposes a considerable constraint on nonsegmented RNA viruses, which must express a number of enzymatic and structural proteins to complete their replicative cycle. Strategies to overcome this limitation include the synthesis of functionally monocistronic subgenomic mRNAs, the production of precursor polyproteins that are subsequently cleaved by virus- and/or host-encoded proteases, and the use of non-canonical translational mechanisms (such as internal ribosomal entry, leaky scanning, ribosomal frameshifting, and stop codon readthrough) [323] by which additional ORFs may be translated from polycistronic mRNAs.

Members of the order Nidovirales (*Arteriviridae*, *Coronaviridae*, and *Roniviridae*), which includes the RNA viruses with the largest genomes currently known, use many of the above strategies (including polyprotein expression, subgenomic mRNA synthesis, ribosomal frameshifting, and leaky scanning) (Figure 1A) to organize one of the most complex RNA virus replication cycles described to date [347, 393, 394]. The replicative enzymes of nidoviruses are encoded in ORF1a and ORF1b, which occupy the 5′-proximal three quarters of their positive-stranded RNA genome. ORF1a and ORF1b encode two large replicase precursor polyproteins, pp1a and pp1ab, with expression of the latter depending on a −1 ribosomal frameshift in the short ORF1a/ORF1b overlap region [345, 395]. Following their synthesis from the genomic mRNA template, pp1a and pp1ab are processed into at least 13–16 functional nonstructural proteins (nsps) by a complex proteolytic cascade that is directed by two to four ORF1a-encoded proteinase domains (Figure 1A) [361, 362]. The 3′-proximal region of the nidovirus genome contains the genes encoding the viral structural proteins, which are translated from a nested set of 5′- and 3′-coterminal subgenomic mRNAs [371].

As indicated above, the genomes of a variety of viruses, including all currently known nidoviruses, harbor sequences that induce a proportion of translating ribosomes to frameshift −1 nt and continue translating in an alternative reading frame [345, 365, 395-397]. Where functionally important, this process may be termed "−1 programmed ribosomal frameshifting" (−1 PRF). The eukaryotic −1 frameshift site typically consists of a "slippery" heptanucleotide fitting the consensus motif X_XXY_YYZ, where XXX normally represents any three identical nucleotides (although certain exceptions have been
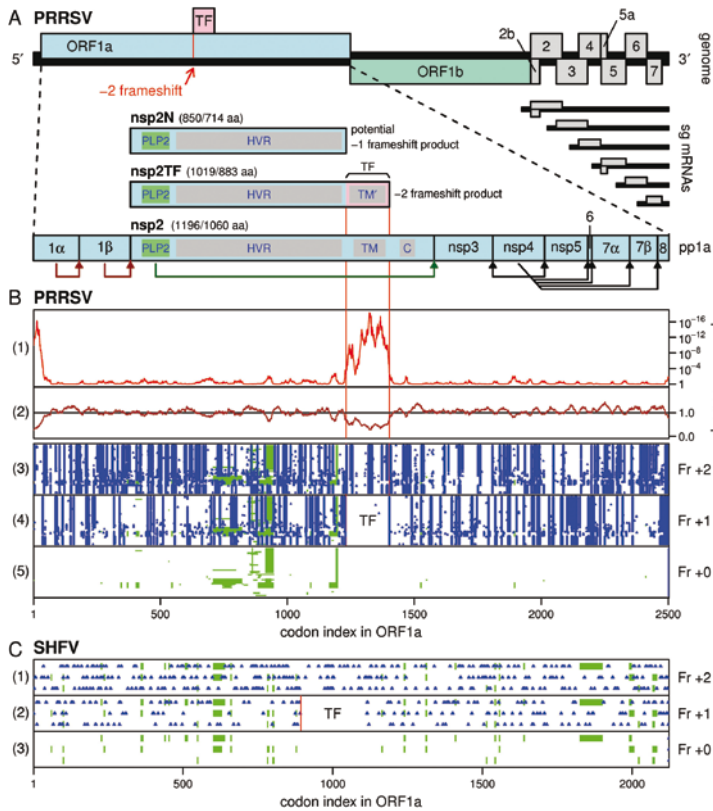
Figure 1: Arterivirus genome organization and expression mechanisms. (*A*) Map of the ~15-kb PRRSV genome. Two long 5' ORFs encode nonstructural polyproteins, and at least eight shorter 3' ORFs encode structural proteins. The 3' ORFs are translated from a nested set of 3'-coterminal subgenomic mRNAs, two of which are bicistronic. ORF1a and ORF1b are translated from the genomic RNA, where translation of ORF1b depends on −1 PRF at the end of ORF1a. The newly described TF ORF overlaps a central region of ORF1a in the +1 reading frame and is accessed via −2 PRF. Domains in nsp2/nsp2TF are annotated as PLP2 (papain-like protease), HVR (hypervariable region), TM/TM' (putative TM domains), and C (Cys-rich domain). Predicted sizes (in aa) for nsp2-related products are shown for GenBank sequences NC_001961 (blue) and DQ489311 (red; the isolate SD01-08 used in this study). (*B*) Bioinformatic analysis of PRRSV ORF1a. Panels 1 and 2 depict the conservation at ORF1a-frame synonymous sites in an alignment of 212 PRRSV sequences using a 25-codon sliding window. Panel 2 shows the ratio of the observed number of substitutions to the number expected under a null model of neutral evolution at synonymous sites. Panel 1 shows the corresponding P value. Summed over the whole TF ORF, the corresponding P value is 5.7×10^{-65}. To map the conservation statistic onto the coordinates of a specific sequence, all alignment columns with gaps in a chosen reference sequence, NC_001961, were removed (note that the original alignment is gap-free within the TF ORF itself). Panels 3–5 show the positions of stop codons (blue) in the three possible reading frames, and alignment gaps (green) in all 212 aligned sequences. Note the conserved absence of stop codons in the +1 reading frame in the TF region. (*C*) Positions of stop codons (blue) in an alignment of the three available SHFV sequences. The vertical red line indicates the location of the G_GUU_UUU/G_GUC_UCU motif. Note the conserved absence of stop codons in the +1 reading frame for 220 codons immediately following this site.

found); YYY represents strictly AAA or UUU; and Z represents A, C, or U, and underscores separate zero-frame codons. This consensus motif generally is followed by a stimulatory element that comprises a stable RNA secondary structure, such as a pseudoknot or stemloop, beginning 5–9 nt downstream of the shift site [396, 397]. In contrast, very little is known about the utilization of −2 programmed ribosomal frameshifting (−2 PRF) in eukaryotic systems, including the potential shift sites and stimulatory elements. Here we describe the identification of a short arterivirus ORF (TF) that is translated via efficient −2 PRF. In porcine reproductive and respiratory syndrome virus (PRRSV), frameshifting (with an efficiency of about 20%) occurs at a conserved G_GUU_UUU sequence (underscores separate ORF1a codons) in a central region of ORF1a and results in the expression of a transframe protein, nsp2TF, that comprises the N-terminal two thirds of nsp2 fused to a 169-aa C-terminal region encoded by the TF ORF. Mutations that prevent expression of nsp2TF seriously impair PRRSV replication in cell culture. Because ribosomes translating the TF ORF will not decode the remainder of the replicase gene, the combination of this −2 PRF mechanism and the downstream ORF1a/ORF1b −1 PRF results in differential expression of three replicase gene segments, suggesting that the regulation of genome translation plays a key role in fine-tuning the replicative cycle of arteriviruses, including PRRSV, one of the economically most important swine pathogens.

RESULTS

Computational analysis reveals a conserved ORF overlapping the Arterivirus nsp2-coding sequence

Arteriviruses comprise a family of small, enveloped positive-stranded RNA viruses that currently includes PRRSV, equine arteritis virus (EAV), lactate dehydrogenase-elevating virus (LDV), and simian hemorrhagic fever virus (SHFV) [394]. The ORF1a sequences of 212 PRRSV isolates available in GenBank as of February 26, 2012 were extracted, translated, aligned, and back-translated to a nucleotide sequence alignment. Both European (EU; type I) and North American (NA; type II) genotype isolates, which typically share 50–55% amino acid identity within pp1a, were included. Next, the alignment was analyzed for conservation at ORF1a-synonymous sites, as described previously [398]. The analysis revealed a striking and highly statistically significant ($P < 10^{-64}$) increase in synonymous-site conservation in a region covering around 170 codons toward the 3′ end of the nsp2-encoding sequence (Figure 1B). Within this region the mean synonymous substitution rate was reduced to 47% of the ORF1a average. Such peaks in synonymous-site conservation generally are indicative of functionally important overlapping elements, either coding or noncoding [398-400]. In this case, inspection of the +1 and +2 reading frames relative to ORF1a, in all 212 sequences, revealed an almost

complete absence of stop codons in the +1 reading frame in a region corresponding precisely to the region of enhanced conservation (Figure 1B). This finding suggested an overlapping ORF in the +1 reading frame as a possible explanation for the enhanced conservation at ORF1a synonymous sites.

An inspection of other arterivirus genomes revealed further evidence for a +1 frame ORF overlapping the equivalent region of ORF1a. Currently, three SHFV sequences are available [401]. With pairwise amino acid identities within pp1a of just 35–37%, these sequences are too divergent for the analysis of synonymous-site conservation. However, the conserved presence of a 219–225 codon ORF in the +1 frame in such divergent sequences is, in itself, statistically significant ($P < 10^{-10}$) (Figure 1C; also see SI Materials and Methods). A 169-codon ORF also is present in one (LDV-P) of two published LDV sequences, but in the second sequence (LDV-C) the ORF is disrupted by a single stop codon. To assess the likeliness of a sequencing error in the LDV-C sequence, we sequenced the relevant region of an additional LDV isolate (795 nt, including the last 158 codons of the ORF; GenBank accession no. JX258842). This sequence is divergent from both LDV-P and LDV-C (locally 83–90% nucleotide identity) and lacks the interrupting stop codon that is present in LDV-C. Remarkably, no evidence for a corresponding ORF was found in EAV. In fact, in this part of the genome, EAV is highly divergent from other arteriviruses, and the nsp2 region is greatly reduced in size.

At the 5' end of the conserved ORF, there is a G_GUU_UUU sequence that is present in 206/212 PRRSV sequences (4/212 have G_GUU_UUC), both LDV sequences, and one SHFV sequence. The other two SHFV sequences have G_GUC_UCU at the corresponding position. Significantly, the G at position 1 is conserved although the corresponding ORF1a codon is CAG (Gln), UGG (Trp), or CGG (Arg) in different PRRSV isolates (Figure 2A). We hypothesized that this motif could facilitate −2 PRF from ORF1a into the overlapping ORF. We were not able to predict convincing downstream RNA secondary structures at a distance of 5–9 nt that might (by analogy to −1 PRF sites) be expected to be present, nor could we definitively rule out the existence of such structures. On the other hand, we did observe a highly conserved CCCANCUCC motif beginning 11 nt downstream of the G_GUU_UUU sequence. This motif is present in both LDV, all three SHFV, and 211/212 PRRSV sequences. Such high conservation could reflect protein sequence constraints (in two overlapping reading frames) but also might be part of a frameshift-stimulatory RNA sequence.

In PRRSV, −2 frameshifting at the G_GUU_UUU sequence would produce a transframe fusion protein comprising the N-terminal 65–72% (typically 714–850 amino acids, depending on isolate) of nsp2 fused to the 169 amino acids encoded by the overlapping ORF (Fig. 1A). We refer to the overlapping ORF as "TF ORF" and to the predicted transframe fusion protein as "nsp2TF." In PRRSV, LDV, and SHFV, nsp2TF is 14–19% shorter than full-length nsp2, and the TF ORF overlaps the part of ORF1a that encodes the predicted nsp2
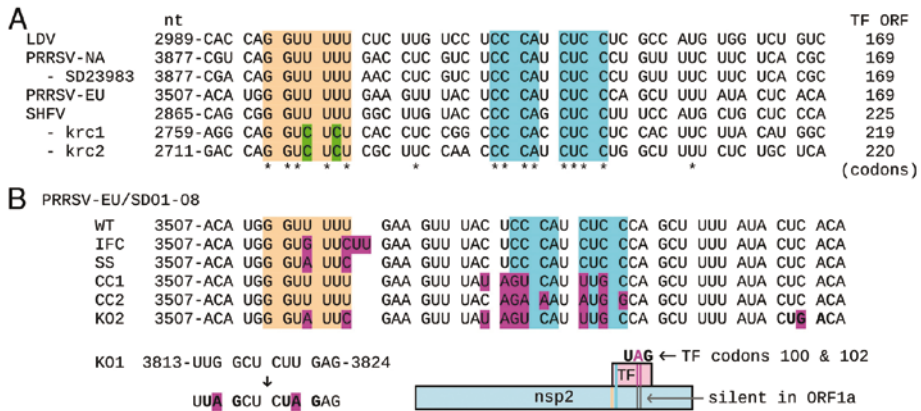
Figure 2: Nucleotide sequences in the vicinity of the TF frameshift site. (*A*) Sequences from representative arteriviruses [GenBank accession numbers: NC_001639, LDV; NC_001961, PRRSV NA (type II); JX258843, PRRSV NA (type II) isolate SD23983; DQ489311, PRRSV EU (type I) isolate SD01-08; NC_003092, SHFV; HQ845737-8. SHFV strains krc1-2]. The proposed frameshift site (confirmed in SD01-08 and SD23983) is highlighted in orange, and nucleotide variations in SHFV isolates krc1-2 are indicated in green. The highly conserved downstream CCCANCUCC motif is highlighted in cyan. Spaces separate ORF1a codons. The length of the +1 frame TF ORF is indicated at the right. (*B*) Overview of mutants used to investigate TF expression and function by ORF1a expression and reverse genetics. Non-WT nucleotides are shown in pink. Coordinates of terminal nucleotides refer to sequence DQ489311. IFC, in-frame control; SS, shift-site mutant; CC1 and CC2, disrupted CCCANCUCC motif; KO1, knockout mutant 1 (premature termination codons in TF); KO2, knockout mutant 2 (premature termination codon and disrupted frameshift cassette). Only the IFC and CC2 mutations are nonsynonymous with respect to the nsp2 amino acid sequence.

transmembrane (TM) domain. The TF ORF appears to encode an alternative TM domain containing four or more potential TM regions, depending on species and isolate, but no other conserved amino acid motifs that might give further indications as to protein function were identified.

Immunodetection of nsp2T in PRRSV-infected cells

To confirm expression of the predicted nsp2TF frameshift product in PRRSV-infected cells, a polyclonal Ab (pAb-TF) was raised against the C-terminal peptide (CPKGVVTSVGESV) of nsp2TF of PRRSV type I isolate SD01-08 [402]. We also used mAbs 36-19 and 58-46, raised against the N-terminal 436 amino acids of SD01-08 nsp2 [402, 403] and therefore expected also to recognize nsp2TF, which (in SD01-08) would share its N-terminal 714 amino acids with nsp2. Expression of nsp2TF was first analyzed by immunoprecipitation (IP) and Western blot analysis. Lysates of SD01-08–infected MARC-145 cells were harvested at 48 h postinfection (p.i.). PRRSV proteins were immunoprecipitated using mAb36-19 and were separated by SDS/PAGE. Four high-molecular-mass bands with apparent masses between 100–150 kDa were detected by Coomassie Brilliant Blue
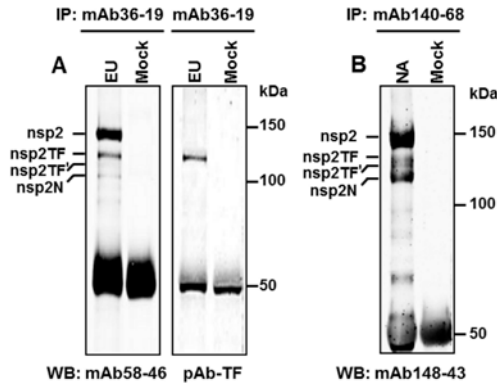
Figure 3: Analysis of nsp2 and nsp2TF expression in PRRSV-infected cells. MARC- 145 cells were infected with type I (EU) PRRSV isolate SD01-08 (*A*) or type II (NA) PRRSV isolate SD23983 (*B*) or were mock infected. Proteins were immunoprecipitated with mAbs specific for the common N-terminal domain of nsp2 and nsp2TF, separated by SDS/PAGE, and probed by Western blotting using an alternative mAb recognizing the common N-terminal domain or a TF-specific pAb (SD01-08 only), as indicated at the bottom. Size markers and putative PRRSV proteins are indicated.

staining, with the two smallest products clearly being less abundant. In addition, various lower-molecular-mass products were observed (Figure S1A). To confirm that the high-molecular-mass bands represented nsp2-related products, Western blot analysis was performed using anti-nsp2 mAb58-46 and pAb-TF. All four high-molecular-mass products were specifically recognized by mAb58-46 (Figure 3A), indicating that they must share N-terminal sequences. However, only the second- largest protein labeled nsp2TF was recognized by pAb-TF. The third protein labeled nsp2TF′ was not detected in a Western blot using pAb-TF, possibly because of its lower abundance and/or lower affinity for the Ab; this product was, however, detected in IP using pAb-TF (see below). We speculate that nsp2TF′ could be either a precursor to or a modified form of nsp2TF. The smallest product of the four proteins labeled nsp2N might derive from a −1 frameshift at the same conserved G_GUU_UUU sequence. Such a frameshift would lead to an immediate termination, because there is a −1/+2 frame stop codon adjacent to the shift site (Figure 2A; see also Figure S1D and Discussion).

A similar analysis for the PRRSV type II isolate SD23983 also revealed multiple products in the 100–150 kDa range (Figure 3B). The available SD01-08 pAb-TF did not cross-react with any of these bands, but this result was not surprising, because the sequence of the 13-aa peptide used to produce this Ab is not conserved in type II viruses.

Mass spectrometric analysis of the site and direction of frameshifting

Although bioinformatic analysis suggested the site (G_GUU_UUU) and direction (−2) of frameshifting, we sought to confirm both predictions with direct protein sequence analysis. To this end, proteins from SD01-08–infected and mock-infected MARC-145 cell lysates were immunoprecipitated, resolved by SDS/PAGE, and stained with Coomassie Brilliant Blue (Figure S1A). The gel slice containing the putative nsp2TF band was analyzed by LC/MS/MS. Four peptides specific for the 169-aa TF region were identified in addition to many peptides from the common N-terminal domain of nsp2 and nsp2TF (Figure 4B). One of the peptides, LMTW**VF**LK, spanned the frameshift site itself (bold), and its sequence is fully compatible with −2 PRF (after decoding GUU_UUU as VF) (Figure 4A and Figure S1C) but not with +1 PRF, which would produce a shift-site peptide that is shorter by one amino acid (e.g., LMTW**VF**K). To verify the correct identification of the frameshift peptide, a synthetic version of the peptide was subjected to the same LC/MS/MS analysis. The tandem mass spectrum of the synthetic peptide was identical to that of the peptide derived from the gel slice, confirming that nsp2TF is indeed translated via −2 PRF at the G_GUU_UUU motif (Figure S1B). Again, the analysis was repeated for the type II isolate SD23983, for which a −2 frameshift tryptic peptide with a different sequence (Q**VF**LTSSPISLFSSHAFSTR; shift site-encoded amino acids in bold) was predicted. Likewise, this sequence was identified in mass spectrometric analysis of the presumed nsp2TF band excised from an SDS/PAGE gel (Figure S2).

Estimation of the frameshifting efficiency

The frameshifting efficiency and turnover of nsp2 and nsp2TF were investigated in a pulse-chase labeling experiment in SD01-08–infected MARC-145 cells. After a 1-h pulse labeling with ^{35}S-labeled amino acids, the incorporated label was chased for various periods (up to 24 h), and proteins were immunoprecipitated with nsp2- and nsp2TF-specific Abs. As shown in Figure 5A, this analysis revealed the existence of two smaller products (labeled "nsp2'" and "nsp2TF'") that apparently are about 10 kDa smaller than nsp2 and nsp2TF. Its disappearance during the chase suggested that nsp2' is a direct precursor of nsp2, although we cannot rule out the possibility that it may be a degradation product of nsp2. Nsp2 itself also appeared to be subject to further modification during the chase period, as indicated by its slight size increase and more heterogeneous migration in the gel (Figure 5A; compare C0h with later time points). The possible precursor status of nsp2TF' was less obvious, because the amount of this product was more or less stable throughout the chase period, during which both the nsp2TF and nsp2TF' bands also appeared to convert into doublets. In terms of protein turnover, the amount of nsp2TF declined much more rapidly than that of nsp2 during the chase period. This difference in turnover was even more pronounced for nsp2N, the putative −1 frameshift product (see above), which seems to be the least stable of the various nsp2 forms described here.
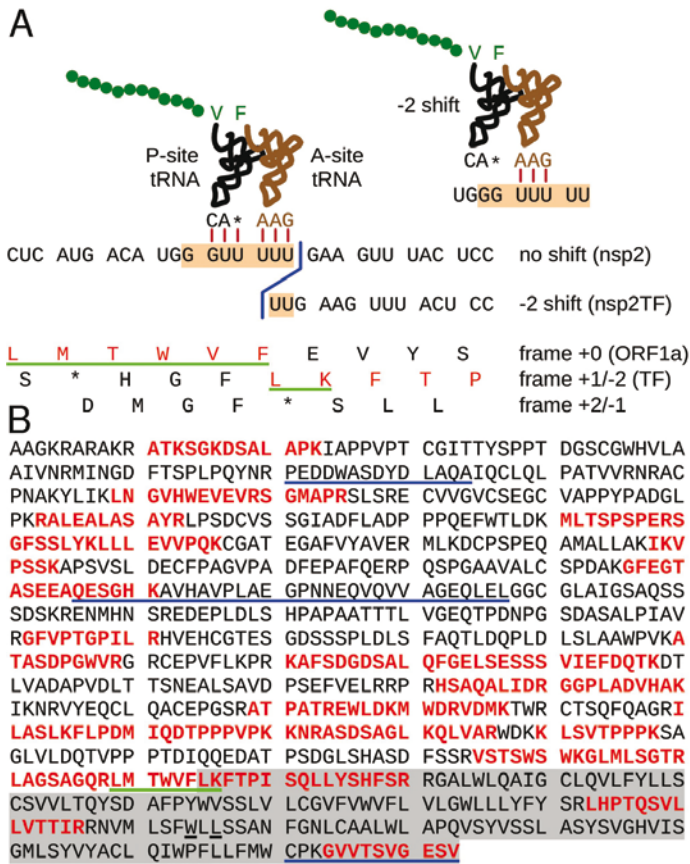
Figure 4: Mass spectrometric analysis of nsp2TF purified from MARC-145 cells infected with type I PRRSV isolate SD01-08. (*A*) Nucleotide sequence in the vicinity of the TF shift site G_GUU_UUU, with conceptual amino acid translations in all three reading frames shown. The product of −2 frameshifting is indicated in red. Consecutive tryptic peptides covering these amino acids were detected by mass spectrometric analysis. The peptide underlined in green, which spans the shift site, is compatible with −2 but not +1 frameshifting. Stylized P- and A-site tRNAs illustrate expected codon:anticodon duplexes before and after frameshifting (see also main text). In eukaryotes, GUU is expected to be decoded by the valine tRNA with anticodon 5'-IAC-3' (I = inosine), but it is possible that it also is decoded by the valine tRNA with anticodon 5'-ncm⁵UAC-3' (ncm⁵U = 5-carbamoylmethyluridine). Because currently it is not known whether one or both of these, or perhaps a hypomodified form, are compatible with frameshifting, the valine anticodon is indicated by 3'-CA*-5' in the schematic. (*B*) Complete amino acid sequence of nsp2TF, with peptides identified by mass spectrometry indicated in red. The C-terminal 169 amino acids encoded by the +1 reading frame are highlighted in gray. The N-terminal 714 amino acids are shared with nsp2. The epitopes recognized by mAbs 36-19 and 58-64 and pAb-TF, in order, are underlined in blue. Pink underlining indicates the locations of the premature termination codons of mutant KO1 (Figure 2B).
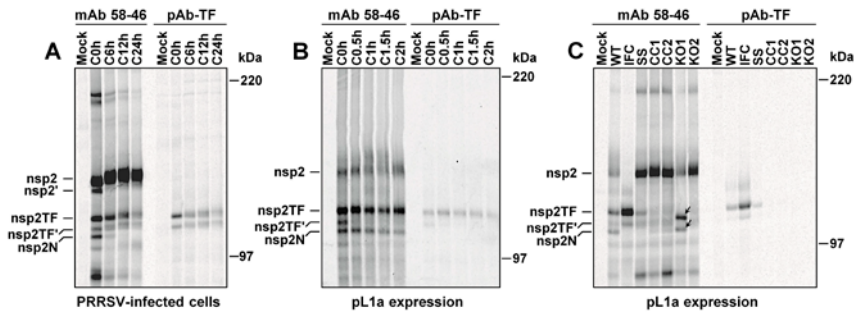
Figure 5: Analysis of nsp2-related products in PRRSV SD01-08–infected cells and transient ORF1a expression. Following metabolic labeling with ^{35}S, proteins were immunoprecipitated with mAb58-46 or pAb-TF and were analyzed by SDS/PAGE and autoradiography. Size markers and putative nsp2-related products are indicated at the side of each panel. (*A*) Pulse-chase experiment with PRRSV-infected MARC-145 cells, using a 1-h pulse labeling and various chase periods as indicated at the top. (*B*) Pulse-chase analysis of transient ORF1a expression in the recombinant vaccinia virus/T7 RNA polymerase system. RK-13 cells were infected with recombinant vaccinia virus and either were mock transfected or were transfected with pL1a plasmid DNA. At 5 h post vaccinia virus infection, protein synthesis was labeled for 30 min and chased for up to 2 h. (*C*) Analysis of −2 PRF and nsp2TF mutants expressed using pL1a and the recombinant vaccinia virus/T7 RNA polymerase expression system as described for B. Arrows point to the C-terminally truncated nsp2TF products from mutant KO1.

Several smaller nsp2-specific products were observed in the C0h sample, including a prominent band migrating at around 90 kDa, a fainter band migrating slightly behind the 90-kDa band, and a faint band migrating at around 98 kDa (Fig. 5A). The prominent 90-kDa product was not detected by pAb-TF and most likely derives from internal cleavage of nsp2. Two additional nsp2-specific products were observed to migrate at around 200 kDa. Based on their size, these may represent nsp2–8 and a modified or precursor form of nsp2–8. The definitive identity of these products cannot be explained at this point in time, and they add to the puzzling complexity of nsp2 expression and post-translational modification.

To assess the frameshifting efficiencies in the context of PRRSV infection, we measured the radioactive incorporation into the nsp2+nsp2', nsp2TF+nsp2TF', and nsp2N bands directly after the pulse labeling (Figure 5A; lane C0h) and corrected these numbers for the Met and Cys content of the different proteins (Table S1). Notwithstanding potential differences in Ab affinity for the different nsp2-related products and other potential confounding factors, these measurements suggested −2 and (putative) −1 frameshifting efficiencies of ~20% and 7%, respectively. Because the presence and identity of the minor products, including those migrating at around 90, 98, and 200 kDa, could affect the calculated frameshifting efficiencies to a limited extent, we repeated the calculations under the conservative assumption that these minor products are all derived from non-

frameshift products. These results, combined with those above, indicate frameshifting efficiencies in the range of 16–20% and 6–7% for −2 and −1 frameshifting, respectively.

The G_GUU_UUU and CCCANCUCC motifs are required for efficient frameshifting

The transient expression of ORF1a in the recombinant vaccinia virus/T7 polymerase expression system [403] was used to develop an assay for PRRSV −2 PRF in uninfected cells. A T7 promoter-driven, full-length ORF1a expression vector (pL1a) was constructed, and the synthesis of nsp2, nsp2TF, and nsp2N was monitored by radiolabeling (30-min pulse, 0- to 120-min chase) and IP (Figure 5B). With the exception of nsp2′, all products immunoprecipitated from PRRSV-infected cell lysates (Figure 5A) could be identified. Although differences in abundance, time of appearance, and stability were observed, these data clearly demonstrated that translation of the PRRSV ORF1a sequence is sufficient to allow efficient −2 frameshifting. In fact, the efficiency of frameshifting was estimated to be even higher in this system than in virus-infected cells (with an estimated −2 shift efficiency of around 50%).

Subsequently, using this ORF1a expression system, we investigated the requirement for an intact shift site and downstream elements for efficient frameshifting. First we engineered a shift-site mutant (SS) (Figure 2B) that contains two mutations in the shift site (G_GUU_UUU to G_GUA_UUC) and therefore is expected to express only nsp2. To mark the position at which WT nsp2TF migrates in gels, an in-frame control (IFC) was constructed in which the shift site was mutated synonymously and an extra 2 nt were inserted to force expression of the TF reading frame (G_GUU_UUU to G_GUG_UUC_UU) (Figure 2B). As shown in Figure 5C, WT pL1a produced both nsp2 and nsp2TF, and expression of the latter protein was confirmed by IP with pAb-TF. As expected, the IFC mutant produced only nsp2TF, as detected by both pAb-TF and mAb58-46. For the SS mutant, production of nsp2TF was greatly reduced, although, unexpectedly, frameshifting was not totally inhibited (Figure 5C). Next we investigated whether the conserved CCCANCUCC motif (see above and Figure 2A) is involved in the stimulation of frameshifting at the G_GUU_UUU shift site by constructing two additional mutants in which the CCCANCUCC motif was disrupted (CC1 and CC2) (Figure 2B). The CC1 mutations are synonymous in the nsp2 frame; the CC2 mutations disrupt the motif more thoroughly but include substitutions that are not synonymous with respect to nsp2. The CC1 and CC2 mutants produced nsp2, but no nsp2TF was detected (Figure 5C). These results indicate that both the G_GUU_UUU shift site and the downstream CCCANCUCC motif are required for ribosomal frameshifting at the WT efficiency. When frameshifting is prevented (in the mutants SS, CC1, and CC2), expression of nsp2 is expected to increase substantially, because ribosomes no longer are being diverted into the alternative reading frame (Figure 5C).

Inactivation of nsp2TF expression affects PRRSV replication

To investigate whether nsp2TF expression and/or the diversion of a proportion of the ribosomes out of ORF1a are relevant for virus replication, two mutants were generated (KO1 and KO2) (Figure 2B) in which nsp2TF expression was partially or completely knocked out with mutations that are translationally silent with respect to ORF1a. KO1 makes a truncated nsp2TF protein because of the mutagenesis of codons 100 and 102 of the 169-codon TF ORF into stop codons (nsp2-frame UUG_GCU_CUU_GAG to U**UA_G**CU_C**UA_G**AG; stop codons indicated in bold) (Figure 2B and 4B). Consequently, the truncated nsp2TF lacks the C-terminal pAb-TF epitope. KO2 contains nine mutations that disrupt the frameshift site and the downstream CCCANCUCC motif, besides introducing a stop codon into the TF ORF (Figure 2B). This mutant was intended to knock out the frameshift signal completely and to express only nsp2. KO1 and KO2 differ from WT virus by two and nine nucleotide substitutions, respectively, none of which affect the encoded nsp2 amino acid sequence (Figure 2B).

The mutants were tested in the ORF1a expression system to verify synthesis of the expected proteins. As expected, the KO1 frameshift product could not be detected using pAb-TF because of the truncation of the C-terminal epitope region. However, IP with mAb58-64 revealed the synthesis of truncated forms of nsp2TF and nsp2TF' (Figure 5C, arrows) and a ratio between full-length nsp2 and the frameshift product similar to that in WT. On the other hand, the mutations introduced in KO2 indeed were found to eliminate frameshifting, and, as anticipated, a much larger amount of full-length nsp2 was produced (Figure 5C).

The KO1 and KO2 mutations subsequently were transferred to a PRRSV full-length cDNA infectious clone, and the resulting recombinant viruses, vSD-KO1 and vSD-KO2, were found to be viable. However, both mutants produced plaques that were clearly smaller than those of the WT virus (vSD-WT), with vSD-KO2 producing somewhat smaller plaques than vSD-KO1 (Figure 6A). Analysis of growth kinetics consistently showed that replication of both vSD-KO1 and vSD-KO2 is seriously impaired in MARC-145 cells, with peak titers of both mutants being 50–100-fold lower than those of the WT control, and vSD-KO2 again displaying the larger reduction (Figure 6B).
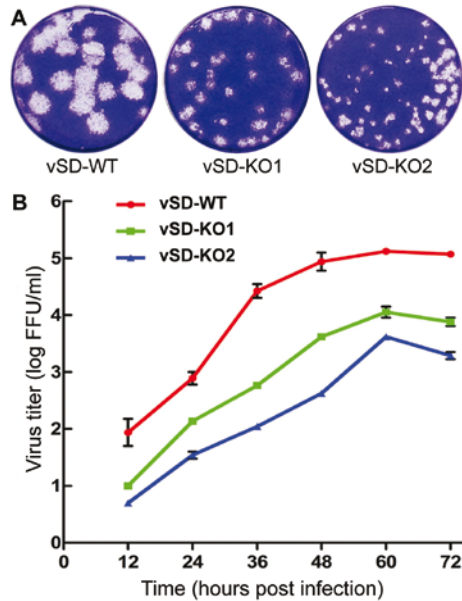
6

Figure 6: Comparison of growth characteristics of WT PRRSV and nsp2TF knockout mutants. (*A*) Plaque morphology of WT and nsp2TF knockout PRRSVs on MARC-145 cells. Confluent cell monolayers were infected with 10-fold serial dilutions of the virus suspension, and after 2 h an agar overlay was applied. Plaques were detected after 4 d of incubation at 37 °C and stained with 0.1% crystal violet. (*B*) Growth kinetics of WT and nsp2TF-knockout PRRSV on MARC-145 cells. The results shown are mean values from three replicates. Virus titers are expressed as numbers of fluorescent-focus units (FFU) per milliliter.

PRRSV nsp2 and nsp2TF localize to different intracellular compartments

The localization of nsp2TF in PRRSV SD01-08–infected cells was investigated using immunofluorescence (IF) microscopy. As shown in Figure 7A, pAb-TF labeled specific foci mainly localizing to the perinuclear region of infected cells. This region also is known to contain the arterivirus replication structures, modified membranes that label abundantly for most of the viral nsps [369, 403, 404]. We used an mAb that recognizes nsp4 [403] to visualize these structures, which revealed a labeling pattern clearly different from that observed with pAb-TF. Surprisingly, a control labeling with nsp2-specific mAbs (36-19 and 58-46, recognizing different epitopes in the common N-terminal domain of nsp2 and nsp2TF) (Figure 4B) revealed that the pAb-TF–specific foci were not, or were barely, recognized by nsp2-specific mAbs (Figure 7B), even though these mAbs did detect nsp2TF convincingly in other immuno-assays (Figure 3 and 5). Apparently, this part of nsp2TF is not accessible in these formaldehyde-fixed, Triton X-100-permeabilized cells. Confocal microscopy further corroborated that the labeling patterns obtained with nsp2-specific mAbs and pAb-TF in PRRSV-infected cells did not overlap (Figure 7C).
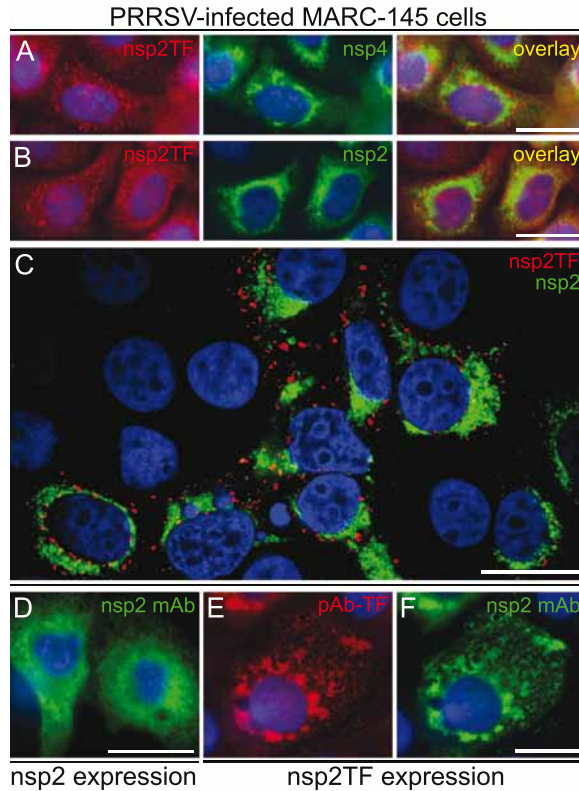
PRRSV-infected MARC-145 cells

Figure 7: IF microscopy analysis of the subcellular localization of nsp2TF in PRRSV-infected MARC-145 cells at 18–24 h p.i. (*A*–*C*) and a CMV promoter-driven expression system for nsp2 and nsp2TF (*D*–*F*). (*A*) Infected cells were double labeled with pAb-TF and a PRRSV nsp4-specfic mAb (mAb54-19) (19), revealing that nsp2TF does not colocalize with the viral replication structures. (*B*) Double labeling with pAb-TF and PRRSV nsp2-specific mAb36-19 results in nonoverlapping labeling patterns suggesting that the nsp2-specific mAb cannot access the N-terminal domain of nsp2TF in fixed PRRSV-infected cells. (*C*) Confocal microscopy of a 0.8-μm slice through the nucleus confirming the nearly complete spatial separation of the structures labeled with pAb-TF and mAb36-19. (*D*) Nsp2-expressing MARC-145 cells labeled with mAb58-46. (E and F) Nsp2TF-expressing MARC-145 cells double labeled with pAb-TF (*E*) and mAb58-46 (*F*) show that, in contrast to PRRSV-infected cells, the latter antibody is able to recognize nsp2TF in this expression system. Cell nuclei were stained with Hoechst 33342. (Scale bars, 20 nm.) See also Figure S3.

To investigate this phenomenon in more detail, nsp2 and nsp2TF were expressed individually in MARC-145 cells (nsp2TF was expressed using the in-frame control sequence to mimic −2 frameshifting) (Figure 2B). In the absence of replication structures and other nsps, arterivirus nsp2 localizes to endoplasmic reticulum (ER) (Figure 7D), as was confirmed using an Ab recognizing an ER-specific marker protein (Figure S3F). Individual expression of nsp2TF (Figure 7 E and F and Figure S3E) again yielded a pattern clearly different from that observed for nsp2 expression. It resembled the foci found in infected

cells, but in this case these foci could be labeled with both pAb-TF and mAbs recognizing the N-terminal domain of nsp2, suggesting a conformational difference between nsp2TF (or nsp2TF-containing structures) in infected cells and the expression system. In both systems, however, the foci labeled by pAb-TF partially overlapped with the staining for exocytic pathway markers (specifically intermediate compartment and Golgi complex) (Figure S3 A–D), an observation that currently is being investigated in more detail. Together, our microscopy studies strongly suggest that, unlike nsp2, nsp2TF is targeted not to the arterivirus replication structures but instead to an alternative intracellular destination.

DISCUSSION

RNA viruses of eukaryotes use a variety of non-canonical translational mechanisms to express multiple proteins from a limited number of transcripts, regulate gene expression, and otherwise manipulate the host cell translational machinery for their own specific needs. However, in contrast to −1 PRF, little was known about the potential for efficient functional utilization of −2 PRF. We now have demonstrated that PRRSV uses this mechanism to produce an nsp2-related transframe protein, nsp2TF. Bioinformatic analysis strongly suggests that such an nsp2TF product is encoded by all arteriviruses, with the striking exception of EAV, in which the relevant region of ORF1a has diverged dramatically. Efficient expression of nsp2TF in virus-infected cells was verified experimentally in multiple independent immunoassays using nsp2- and nsp2TF-specific Abs. Mass spectrometric analysis of proteins purified from infected cells further confirmed the expression of nsp2TF and the exact site and direction of frameshifting. Both nsp2TF knockout mutants exhibited a crippled phenotype with a smaller plaque size, indicating an important role for nsp2TF (KO1) and possibly also for frameshifting per se (compare KO1 and KO2) in virus replication.

Currently, there are very few, if any, other examples of natural utilization of −2 PRF in eukaryotic systems. In diverse animals and fungi, expression of the cellular gene antizyme involves +1 PRF [405]. However, when mammalian antizyme is expressed artificially in the yeast *Saccharomyces cerevisiae*, the full-length antizyme product is expressed via −2 frameshifting [406]. Here the stimulatory elements comprise a stop codon (bold, lower case) 3' adjacent to the shift site (UGC_UCC_**uga**) and a 3'-proximal RNA pseudoknot structure. The −2 frameshift translation reads CSPD, and frameshifting is thought to involve mainly P-site slippage on GC_UCC with an empty A-site. Recently it was found that by artificially reducing the spacer length to the downstream stimulator (either an RNA secondary structure or an antisense oligonucleotide), the slippery site for −1 PRF in HIV, i.e., U_UUU_UUA, also could serve as a slippery site for efficient −2 frame-

shifting [407]. In prokaryotic systems, relatively inefficient −2 PRF (about 2.2%) is used in the expression of the gpGT tail-assembly protein in phage Mu, although the majority of dsDNA phages that express gpGT via frameshifting appear to use −1 PRF instead [408]. Finally, although not yet formally demonstrated, sequence analysis suggests that −2 PRF on CC_CUU_UUU is used in the expression of Gag-Pol in *Trichomonas vaginalis* virus 1 (TVV1) [409, 410]. However, PRF in TVV1 is likely to be relatively inefficient, because the Gag-Pol:Gag ratio in virions is extremely low (e.g., 1–2%) [409], in sharp contrast to the high efficiencies observed here for PRRSV −2 PRF (Figure 5A).

Although typical −1 PRF sites (X_XXY_YYZ) allow codon: anticodon re-pairing in both the P- and A-sites with mismatches only at the wobble positions, for −2 PRF on G_GUU_UUU perfect re-pairing is maintained only in the A-site (Figure 4A). The nucleotide preceding the heptanucleotide is typically a G or an A in different arterivirus sequences; thus the post-shift P-site anticodon:codon duplex, if a duplex forms at all, may have multiple mismatches. However, although the integrity of the A-site duplex is strictly monitored by the translating ribosome [411], the P-site duplex is not monitored so strictly, and, even for −1 PRF, a number of variations on the canonical XXX are al-lowed, including UCC, GGA, GUU, and GGU [323, 345, 399]. The potential −2 PRF site in TVV1 (CC_CUU_UUU) and the site of presumed −2 PRF in SHFV isolates HQ845737 and HQ845738 (G_GUC_UCU) also concur with the theme of perfect re-pairing in the A-site but reduced potential for re-pairing in the P-site.

With few exceptions, eukaryotic −1 PRF is stimulated by a 3′-proximal RNA secondary structure [396, 397]. Whether such structures are sufficient and/or necessary for the stimulation of −2 frameshifting remains to be determined. However, our results indicate that −2 PRF in PRRSV, and by implication other arteriviruses, may be stimulated instead by unstructured 3′ sequence elements including a highly conserved CCCANCUCC motif. Unstructured 3′ sequences also have been implicated in the stimulation of +1 PRF in yeast [412] and −1 PRF in Semliki Forest alphavirus [413]. Depending on their distance from the shift site, such sequences may exert their action via mRNA:rRNA base-pairing or interaction with other translational components, although precise mechanisms have not yet been elucidated.

A variety of nsp2-related proteins were observed in our analysis, as previously re-ported for type II PRRSV isolate VR2332, in which such products were assumed to derive from the use of alternative N- and/or C-terminal cleavage sites [414]. At least one of these products now seems to correspond to nsp2TF; some others may correspond to the nsp2- and nsp2TF-related products observed in the pulse-chase experiments in Figure 5. Moreover, the G_GUU_UUU sequence is also a suitable shift site for −1 frameshifting [399], and a potential −1 frameshift product (nsp2N) was observed (Figure 3 and 5 and Figure S1D). In the vast majority of PRRSV sequences (205/212, including isolate SD01-08), such a frameshift would result in immediate termination at a −1 frame stop codon

(with G_GUU_UU**U**_ga, G_GUU_UU**U**_ag, and G_GUU_UU**U**_aa all found in different PRRSV isolates; stop codons in −1 frame indicated in bold). Mass spectrometry of the nsp2N band identified a peptide corresponding to the predicted C terminus of a −1 frameshift product (Figure S1D), although formally such a peptide also could derive from internal cleavage of nsp2 or nsp2TF at this position. The identification of the nsp2N band as the −1 frameshift product was supported further by the fact that this protein was not observed upon expression of the SS mutant, in which the frameshift site is mutated (Figure 5C). The identities and posttranslational modifications of the various nsp2-related products are currently under further investigation, but the data accumulated thus far leave no doubt that efficient −2 PRF occurs, likely accompanied by a lower level of −1 PRF at the same nucleotide sequence.

As previously described, the balance between the synthesis of the arterivirus pp1a and pp1ab replicase polyproteins is regulated by another ribosomal frameshift event, a −1 PRF (Figure 1A), leading to an estimated pp1a:pp1ab ratio of about 4:1 [345]. It now is apparent that the expression level of the different replicase proteins also is affected by efficient −2 PRF (and probably also −1 PRF) at the TF shift site, leading to a more complex series of ratios. Of the ribosomes that translate nsp1α/nsp1β, ~20% synthesize nsp2TF, ~7% synthesize nsp2N, the other ~73% synthesize nsp2–8, and only ~15% subsequently translate the ORF1b-encoded proteins nsp9–12. Interestingly, betaretroviruses and deltaretroviruses also use two ribosomal frameshifts (both −1 PRF) in the expression of their Gag-Pro-Pol polyprotein [415]. Here Gag, Pro, and Pol are encoded by consecutive terminally overlapping ORFs, and, in contrast to the arteriviruses, the polymerase is translated by ribosomes that have frameshifted twice.

PRRSV nsp2, the largest replicase cleavage product, is released by the autoproteolytic activities of the upstream papain-like protease (PLPβ) in nsp1β and the PLP2 protease residing in the N-terminal domain of nsp2 (Figure 1A) [416, 417]. Nsp2 is a multi-domain and multifunctional protein. Besides cleaving the nsp2/3 site, nsp2 functions as a cofactor for the nsp4 serine protease during processing of the C-terminal half of pp1a [418]. The C-terminal domain of nsp2, but not nsp2TF, is a highly conserved Cys-rich domain of unknown function [419]. Furthermore, nsp2 is predicted to be a multispanning TM protein that contributes to the formation of the membranous structures that scaffold the assembly of the viral replication complex [369, 404]. Recent studies also have implicated nsp2 in viral pathogenesis, specifically by virtue of PLP2's deubiquitinating and deISGylating activities [420-423]. The biological significance of these activities was supported by the ability of PLP2 to inhibit type I IFN activation and antagonize the antiviral effect of ISG15. Finally, certain regions of nsp2 that appear to be less or nonessential for PRRSV replication are thought to play a role in the modulation of host immune responses in vivo [424].

The nsp2TF protein adds to the complexity of functions potentially encoded in this region of the genome. Its conservation in three of four distantly related arteriviruses and our reverse-genetics studies (Figure 6) suggest it is an important protein. The frameshift site is located just upstream of the region encoding the predicted nsp2 TM domain. Thus, nsp2 and nsp2TF share the PLP2 domain and the hypervariable region of nsp2 but have distinct C-terminal segments of different sizes that appear to constitute alternative TM domains. Strikingly, IF microscopy revealed that, in both infected cells and expression systems, nsp2TF and nsp2 are targeted to different locations. Although nsp2TF's specific destination and function in replication or pathogenesis requires further study, these data may be a first step in understanding the presence and conservation of this additional ORF in most arteriviruses. If a fraction of the frameshifting ribosomes indeed make a −1 rather than a −2 shift (see above), the production of an nsp2 variant lacking either TM domain would add further to the complexity of nsp2 expression. Such a presumably cytosolic version of PLP2 may have major implications for the interactions of this protease with host cell targets.

Our reverse-genetics analysis suggests that, although nsp2TF is not essential for replication, it is crucial for maximum virus fitness. When nsp2TF expression was prevented (mutant KO2) or a C-terminally truncated nsp2TF was produced (mutant KO1), the virus exhibited a lower growth rate and a clearly reduced plaque size. KO1 was designed to truncate the TF ORF without disrupting frameshifting per se (Figure 5C). That this mutation still resulted in a crippled phenotype highlights the functional importance of full-length nsp2TF. In contrast, KO2 was intended to knock out frameshifting completely. There can be multiple reasons why KO2 was more crippled than KO1. KO2 does not express nsp2TF, and the absence of nsp2TF, as demonstrated by KO1, is sufficient to impair virus replication. However, knocking out the frameshift signal also would be expected to upregulate expression of all the downstream nsps (nsp3–12), perhaps disturbing the balance of viral protein synthesis in a way that is detrimental to virus growth. Potential disruptive effects might involve replicase proteolytic processing by nsp4 (also using nsp2 as a cofactor), the formation of replication complexes, or the overexpression of RdRp and helicase, all of which individually have the potential to affect viral RNA synthesis directly or indirectly.

Because of its highly immunogenic nature, the PRRSV nsp2 region has been explored for the development of a diagnostic assay [425]. Although the immunological properties of nsp2TF have not yet been determined, it might represent a viral antigen and/or a potential target for diagnostic assay development. The nsp2-coding region also was explored for its potential application in the development of PRRSV vaccines (reviewed in [361]). Using reverse-genetics approaches, modified live viruses with engineered deletions and foreign inserts were created in an attempt to generate differentiable marker vaccines. In addition, mutations and deletions were introduced into certain nsp2 regions

during attempts to attenuate the pathogenicity of the virus. The identification of −2 PRF and the TF ORF has important consequences for the rational design of such recombinant viruses. Engineering of the PRRSV nsp2-encoding sequence may unintentionally affect the integrity and/or expression level of nsp2TF and all downstream replicase subunits, perhaps crippling virus replication, as discussed above. On the other hand, such effects also could be useful in the context of modified live virus vaccine design.

MATERIALS AND METHODS

Computational analysis

A total of 255 arterivirus nucleotide sequences in GenBank with full coverage of ORF1a (listed in SI Materials and Methods) were identified by applying tBLASTn [426] to the pp1a peptide sequence derived from GenBank sequences NC_001961 (PRRSV-NA), NC_001639 (LDV), NC_003092 (SHFV), and NC_002532 (EAV). Six ORF1a-defective PRRSV sequences were excluded from subsequent analysis. For each virus, the ORF1a sequences were extracted, translated, aligned, and back-translated to produce nucleotide-sequence alignments using EMBOSS and Clustal [427, 428]. Synonymous-site conservation was calculated as described previously [398].

Viruses and cells

BHK-21, RK-13, and MARC-145 cells were cultured as described previously [402, 419]. The type I PRRSV isolate, SD01-08 (GenBank accession DQ489311) [402], and type II PRRSV isolate, SD23983 (GenBank accession JX258843), were used. Recombinant vaccinia virus vTF7-3 [429] was propagated in RK-13 cells.

Immunoassays

To detect nsp2 and nsp2TF expression in infected cells, proteins were immunoprecipitated with mAb36-19 and analyzed by Western blot as described previously [403, 430]. To determine the nsp2TF subcellular localization, IF microscopy was conducted essentially as described previously [369, 403]. The frameshifting efficiency and turnover of nsp2 and nsp2TF were investigated in a pulse-chase experiment using a method modified from that described by Snijder et al. [419]. Transient PRRSV ORF1a expression in RK-13 cells, using plasmid pL1a and the recombinant vaccinia virus/T7 polymerase expression system, was performed as described previously [419]. See Supporting Information Materials and Methods for detailed procedures.

Mass spectrometry

Nsp2TF was immunoprecipitated from SD01-08–infected cell lysate using mAb36-19. Proteins from IP were separated on a 6% SDS/PAGE gel, which subsequently was fixed and stained with Coomassie Brilliant Blue G-250 (Bio-Rad). The gel was destained, and the band expected to contain nsp2TF was excised. Trypsin digestion and LC-MS/MS analysis were performed as described previously [431]. MS spectra were searched against a custom-made protein database containing possible nsp2 frameshift proteins.

DNA constructs

Figure 2B lists the constructs used in this study. All constructs were made by standard PCR mutagenesis and recombinant DNA techniques and were verified by DNA sequencing. Further details are given in Supporting Information Materials and Methods.

Rescue and characterization of recombinant PRRSVs

Recombinant PRRSVs were recovered from WT (pSD01-08) or mutant (pSD01-08-KO1 or pSD01-08-KO2) full-length cDNA clones as described previously [402]. Growth kinetics was examined by infecting MARC-145 cells with passage 2 WT or mutant virus at a multiplicity of infection of 0.1. Supernatants from infected cells were collected at 12, 24, 36, 48, 60, and 72 h p.i., and virus titers were determined by fluorescent focus or plaque assay as described previously [402].

6

SUPPORTING INFORMATION

Supporting information materials and methods

GenBank accession numbers of sequences used in the bioinformatic analysis.

Porcine reproductive and respiratory syndrome virus (European genotype, type I)

M96262, AY366525, AY375474, AY588319, DQ489311, DQ864705, EU076704, FJ349261, GQ461593, GU047344, GU047345, GU067771, GU737264, JF276430, JF276431, JF276432, JF276433, JF276434, JF276435, JF802085.

Porcine reproductive and respiratory syndrome virus (North American genotype, type II)

NC_001961 = AF046869, AB288356, AF066183, AF159149, AF176348, AF184212, AF303354, AF303355, AF303356, AF303357, AF325691, AF331831, AF494042, AY032626, AY150312, AY150564, AY262352, AY424271, AY457635, AY545985, AY585241, AY612613, DQ056373, DQ176019, DQ176020, DQ176021, DQ217415, DQ459471, DQ473474, DQ779791, DQ988080, EF075945, EF112445, EF112446, EF112447, EF153486, EF484031, EF484033, EF488048, EF488739, EF517962, EF532801, EF532802, EF532803, EF532804, EF532805, EF532806, EF532807, EF532808, EF532809, EF532810, EF532811, EF532812, EF532813, EF532814, EF532815, EF532816, EF532817, EF532818, EF532819, EF535999, EF536000, EF536001, EF536002, EF536003, EF635006, EF641008, EU097706, EU097707, EU106888, EU109502, EU109503, EU144079, EU187484, EU200961, EU200962, EU236259, EU262603, EU360128, EU360129, EU360130, EU624117, EU678352, EU708726, EU807840, EU825723, EU825724, EU860248, EU860249, EU864231, EU864232, EU864233, EU880431, EU880432, EU880433, EU880434, EU880435, EU880436, EU880437, EU880438, EU880439, EU880440, EU880441, EU880442, EU880443, EU939312, FJ175687, FJ175688, FJ175689, FJ393456, FJ393457, FJ393458, FJ393459, FJ394029, FJ536165, FJ548851, FJ548852, FJ548853, FJ548854, FJ548855, FJ797690, FJ889129, FJ895329, FJ899592, GQ330474, GQ351601, GQ359108, GQ374441, GQ374442, GQ475526, GQ499193, GQ499194, GQ499195, GQ499196, GQ857656, GU143913, GU168567, GU168568, GU168569, GU169411, GU232735, GU232736, GU232737, GU232738, GU269541, GU454850, GU461292, HM011104, HM016158, HM016159, HM189676, HM214913, HM214914, HM214915, HM853673, HQ233604, HQ233605, HQ315835, HQ315836, HQ315837, HQ401282, HQ416720, HQ699067, HQ843178, HQ843179, HQ843180, HQ843181, JF268672, JF268673, JF268674, JF268675, JF268676, JF268677, JF268678, JF268679, JF268680, JF268681, JF268682, JF268683, JF268684, JF748717, JF748718, JF796180, JF800911, JN256115, JN387271, JN387272, JN387273, JN387274, JN626287, JN662424, U87392.

Lactate dehydrogenase-elevating virus

NC_001639 = U15146, L13298.

Simian hemorrhagic fever virus

NC_003092 = AF180391, HQ845737, HQ845738.

Equine arteritis virus

NC_002532 = X53459, AY349167, AY349168, DQ846750, EU252113, EU252114, EU586273, EU586274, EU586275, GQ903794, GQ903795, GQ903796, GQ903797, GQ903798, GQ903799, GQ903800, GQ903801, GQ903802, GQ903803, GQ903804, GQ903805, GQ903806, GQ903807, GQ903808, GQ903809, GQ903810, GQ903811, JN211316, JN211317, JN211318, JN211319, JN211320.

Statistical significance of the long ORF in simian hemorrhagic fever virus
The statistical significance of the conserved presence of the long TF ORF (Figure 1C) in the three highly divergent simian hemorrhagic fever virus (SHFV) sequences was evaluated by randomly shuffling ORF1a-frame codon columns within the TF ORF region and calculating what fraction of shuffled alignments preserve an ORF in the +1 frame. This procedure controls for any bias for or against random long +1 frame ORFs resulting from ORF1a-frame amino acid use, codon use, or nucleotide biases and also controls for phylogenetic nonindependence. In fact, the proportion of randomizations that preserve a +1 frame ORF was too small to estimate directly (0 occurrences in 4,000 randomizations) and so was estimated instead from the mean number (per randomized alignment) of +1 frame alignment codon columns containing stop codons in one or more sequences (viz., 24.70), assuming Poisson statistics. Using this method, the P value for such a long +1 frame ORF occurring by chance in this region of the SHFV alignment is 1.9×10^{-11}. Neither this statistic for SHFV (1.9×10^{-11}) nor the conservation statistic quoted in the main text for porcine reproductive and respiratory syndrome virus (PRRSV) (1×10^{-64}) has been corrected for multiple tests. In principle one might consider testing the whole genome (~15,000 nt) for conserved regions and/or conserved ORFs of ~200 codons, and, in principle, one might apply such an analysis to the ~1,000 RNA virus species represented in GenBank, making a total of ~25,000 independent tests. Thus, the P values should be scaled by ~25,000 (giving 2.5×10^{-60} and 4.8×10^{-7}, respectively), although a correction for multiple testing is not, in fact, required for the SHFV statistic, because the location of the 5′ end of the TF ORF in SHFV is known a priori (it aligns to the 5′ end of the TF ORF in PRRSV).

DNA constructs

Plasmid pL1a was a derivative of a similar equine arteritis virus (EAV) ORF1a expression vector, in which the foreign gene is under the control of a T7 RNA polymerase promoter and an encephalomyocarditis virus internal ribosomal entry site and is followed by a downstream T7 terminator sequence [419]. The EAV ORF1a sequence was replaced by PRRSV SD01-08 ORF1a (nucleotides 1–7137 of the genome). To construct the nsp2TF knockout mutants (Fig. 2B), mutations were introduced into the nsp2 region of the expression vector pL1a or a PRRSV full-length cDNA infectious clone plasmid pSD01-08 [402]. Except for the KO2 mutant, for which a synthesized oligonucleotide was used, all mutations were introduced by site-directed mutagenesis using the Quick-Change site-directed PCR mutagenesis kit (Stratagene). The pCAGGS-nsp2 and pCAGGS-nsp2TF plasmids were constructed by PCR amplification of the nsp2- and nsp2TF-encoding regions from plasmid DNA of pL1a and pL1a-IFC, respectively, and cloning of PCR products into a eukaryotic expression vector, pCAGGS [432].

Antibodies

For type I PRRSV SD01-08, anti-nsp2 mAbs 36-19 and 58-46 were produced by immunizing mice with amino acids 386–821 of pp1a. Anti-nsp4 mAb 54-19 was produced by immunizing mice with amino acids 1677–1879 of pp1a [403]. For type II PRRSV SD23983, anti-nsp2 mAbs 140-68 and 148-43 were produced by immunizing mice with amino acids 435–514 of SD23983 pp1a. A polyclonal affinity-purified rabbit Ab, pAb- TF, was produced by GenScript using the synthetic peptide CPKGVVTSVGESV (C-terminal 13 amino acids of nsp2TF). Abs for detection of cellular marker proteins comprised a pAb against protein disulfide isomerase (PDI) (Enzo), anti-giantin mAb G1/93 (Alexis), and anti-ERGIC-p53 mAb G1/133 (Alexis).

Immunofluorescence microscopy

MARC-145 cells were infected with PRRSV [multiplicity of infection (MOI) 0.1] or were transfected with pCAGGS-nsp2 or pCAGGS-nsp2TF. At 18–24 h postinfection (p.i.) or 24–48 h posttransfection (p.t.) cells were fixed with paraformaldehyde, and single or double immunolabeling and immunofluorescence microscopy were performed essentially as described previously [369, 403].

Immunoprecipitation and SDS/PAGE

Whole-cell lysates of PRRSV- infected MARC-145 cells were suspended in RIPA buffer [0.5% (wt/vol) sodium deoxycholate, 1% (wt/vol) SDS, 1% (vol/vol) Nonidet P-40, 100 mM NaCl, 1 mM EDTA (pH 8.0), 10 mM Tris·HCl (pH 7.5)]. To reduce nonspecific background, cell lysates were precleared with preimmune rabbit sera or nonspecific mouse ascites. Protein A-Sepharose CL-4B beads (Pharmacia Biotech) and an nsp2-specific mAb were

added to precleared cell lysates. After incubating overnight at 4°C, immune complexes were washed three times with RIP buffer (10 mM Tris·HCl, 150 mM NaCl, 1% Nonidet P-40, 0.5% sodium deoxycholate) and three times with deionized H_2O. After boiling in 2× Laemmli sample buffer for 5 min, proteins were separated on a 6% SDS/PAGE gel.

Western blot

Western blot was performed as described previously [403]. The membrane was probed with primary nsp2- and/or nsp2TF-specific Abs. IRDye 680-conjugated goat anti-rabbit Ab and/or IRDye 800CW-conjugated goat anti-mouse Ab (LI- COR Biosciences) were used as secondary Abs. Imaging of the blot was performed using an Odyssey infrared imaging system (LI-COR-Biosciences).

Radioactive labeling and radioimmunoprecipitation analysis

To analyze PRRSV SD01-08 nsp2TF expression in infected MARC- 145 cells (MOI 0.1), cells were starved for 30 min in Met- and Cys-free medium, and at 24 h p.i. proteins were pulse-labeled for 1 h with 500 μCi/mL of a [^{35}S]Met/Cys mixture (EXPRE^{35}S^{35}S Protein Labeling Mix; Perkin-Elmer). Following two PBS washes, cells were incubated for various chase periods up to 24 h in DMEM containing 2 mM each of unlabeled Met and Cys and 2% FCS. Transient PRRSV ORF1a expression in RK-13 cells, using plasmid pL1a and the recombinant vaccinia virus/T7 polymerase expression system, was performed as described previously [419]. Pulse-chase experiments were performed as described above at 4 h p.t. using a 30-min pulse of 500 μCi of [35S]Met/Cys per mL. Protocols for cell lysis, immunoprecipitation, SDS/PAGE, and imaging (Typhoon Variable Mode Imager; GE Healthcare) have been described previously [403].

Calculation of frameshifting efficiencies

Band intensities (nsp2, nsp2′, nsp2TF, nsp2TF′, and nsp2N) were quantified with ImageQuant TL (GE Healthcare) and normalized by the Met+Cys content of the respective products assuming that ^{35}S Met and ^{35}S Cys are incorporated with an efficiency ratio of 73:22 (the Met:Cys ratio in the mixture according to the manufacturer's documentation). (Note that calculated frameshifting efficiencies were only 1.06–1.07 times higher if equal incorporation efficiencies were assumed instead.) Radioactivity in the two bands at around 200 kDa was quantified and normalized assuming that these products represent nsp2–8; the product migrating at around 98 kDa was assumed to have Met and Cys content similar to that of nsp2N; and the two products migrating at around 90 kDa were assumed, based on migration position, to have half as many Met and Cys residues as nsp2. Using these values (Table S1), frame- shifting efficiencies were calculated as (nsp2TF + nsp2TF′)/(nsp2 + nsp2′ + nsp2TF + nsp2TF′ + nsp2N) (upper bound) and (nsp2TF + nsp2TF′)/(nsp2 + nsp2′ + nsp2TF + nsp2TF′ + nsp2N + 200K products + 98K

product + 90K products) (lower bound) for −2 frameshifting, and similarly, using nsp2N as the dividend, for −1 frameshifting.
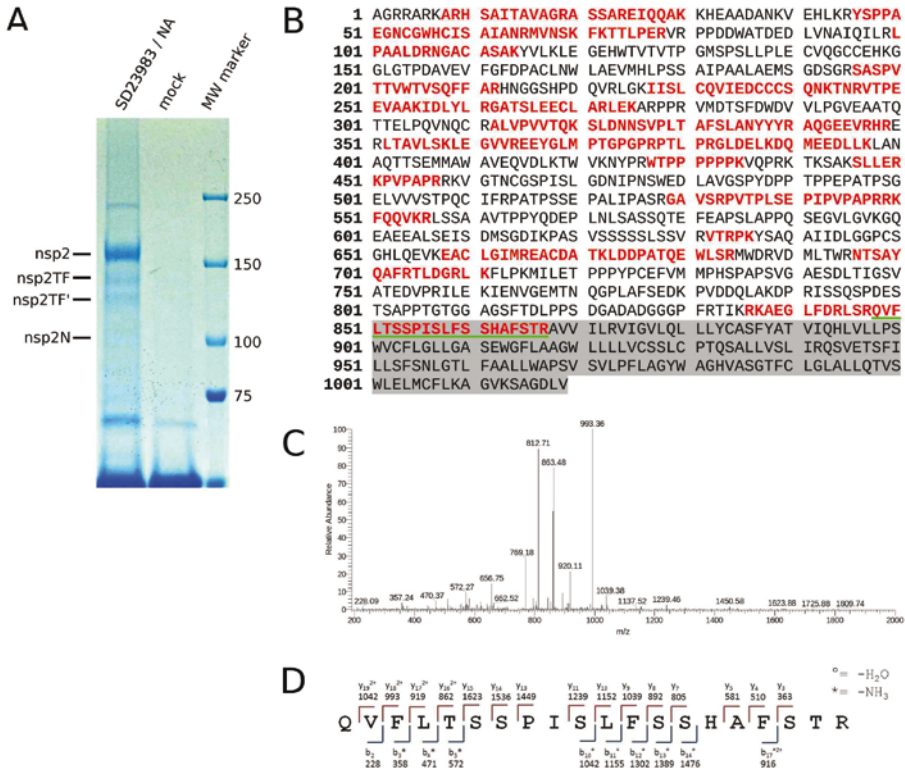
Supporting information tables

Band	% radioactive label	Adjusted for Met/Cys* (%s)	Adjusted for Met/Cys† (%s)
200K (upper)	2.4	1.1	1.0
200K (lower)	0.5	0.3	0.2
nsp2	59.1	50.7	51.8
nsp2′	5.2	4.5	4.6
nsp2TF	13.3	13.8	13.0
nsp2TF′	2.8	2.9	2.8
nsp2N	4.4	6.0	5.6
98K	1.2	1.6	1.5
90K (upper)	1.9	3.3	3.3
90K (lower)	9.2	15.8	16.1

Supporting Information Table S1. Radioactive incorporation into nsp2-related products in virus-infected cells. *Assuming equal incorporation efficiencies for ^{35}S Met and ^{35}S Cys. Met and Cys content for the unidentified nsp2-related products migrating at around 200K, 98K, and 90K are estimated as described in SI Materials and Methods. †Assuming ^{35}S Met and ^{35}S Cys are incorporated with an efficiency ratio of 73:22.
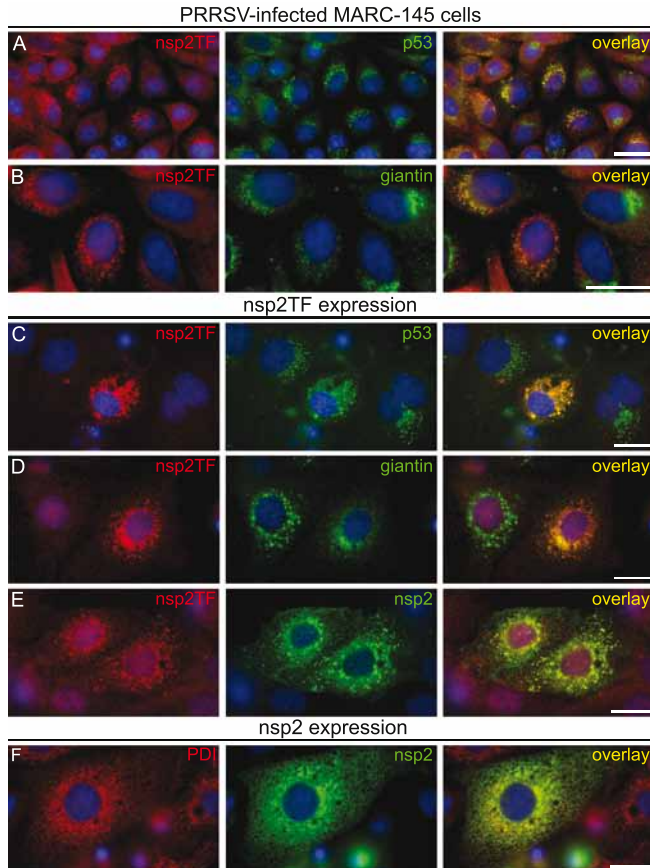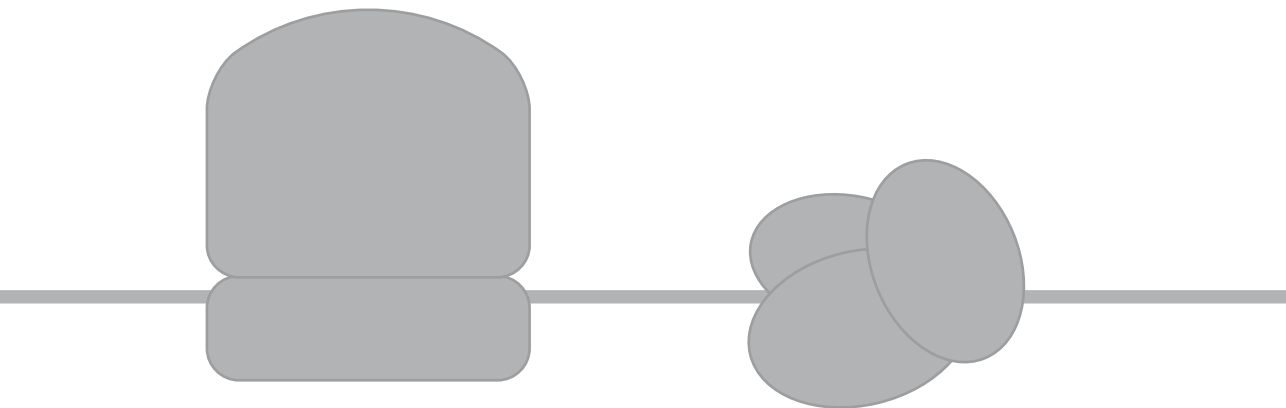
Supporting information figures
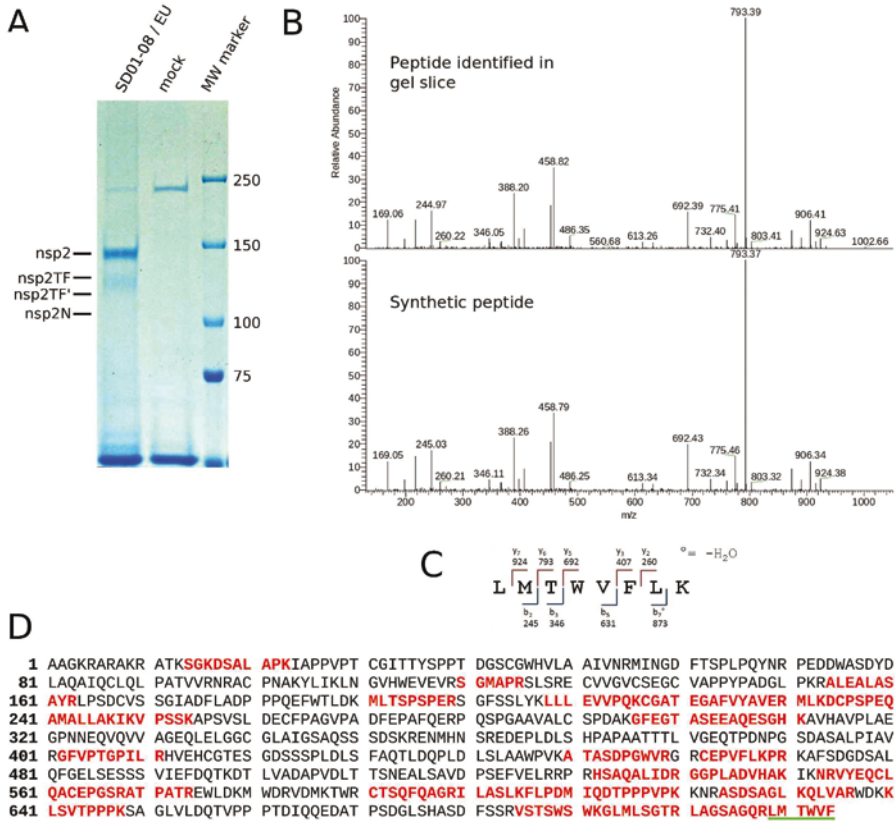
Supporting Information Figure S1: Mass spectrometric analysis of nsp2TF and nsp2N purified from cells infected with type I PRRSV isolate SD01-08. (*A*) PRRSV-infected or mock-infected MARC-145 cell lysates were immunoprecipitated with nsp2-specific mAb36-19. Immunoprecipitated proteins were separated by SDS/PAGE and stained with Coomassie Blue. Positions of molecular mass markers and putative PRRSV proteins are indicated. (*B*) Fragmentation spectra of the shift-site peptide LMTW**VF**LK (shift site-encoded amino acids in bold) identified from the gel slice containing the band labeled nsp2TF (*Upper*) and the synthetic peptide (*Lower*). (*C*) Peptide sequence of the nsp2TF shift-site peptide. The fragment ions that were identified in the LC-MS/MS analysis of the gel slice are indicated. (*D*) Complete amino acid sequence of the predicted −1 frameshift product, nsp2N (see main text for the corresponding figure, Figure 4B, for nsp2TF). Peptides identified by mass spectrometry of the gel slice containing the band labeled nsp2N are in red. The peptide compatible with −1 frameshifting and termination at the −1 frame stop codon is underlined in green.

Supporting Information Figure S2: Mass spectrometric analysis of nsp2TF purified from cells infected with type II PRRSV isolate SD23983. (*A*) PRRSV-infected or mock-infected MARC-145 cell lysates were immunoprecipitated with nsp2-specific mAb140-68. Immunoprecipitated proteins were separated by SDS/PAGE and stained with Coomassie Blue. Positions of molecular mass markers and putative PRRSV proteins are indicated. (*B*) Complete amino acid sequence of nsp2TF. Peptides identified by mass spectrometry are in red. The C-terminal 169 amino acids encoded by the +1 reading frame are highlighted in gray. The N-terminal 850 amino acids are shared with nsp2. The peptide spanning the frameshift site is underlined in green. (*C*) Fragmentation spectrum of the shift-site peptide Q**VF**LTSSPISLFSSHAFSTR (shift site-encoded amino acids in bold). (*D*) Peptide sequence of the nsp2TF shift-site peptide. The fragment ions that were identified in the LC-MS/MS analysis of the gel slice are indicated.

PRRSV-infected MARC-145 cells

nsp2TF expression

nsp2 expression

Supporting Information Figure S3: IF microscopy analysis of the subcellular localization of nsp2TF in PRRSV-infected MARC-145 cells at 24 h p.i. (*A and B*) and a CMV promoter-driven expression system for nsp2TF (*C–E*) and nsp2 (*F*). (*A*) Infected cells were double labeled with pAb-TF and an mAb recognizing the ERGIC marker p53, revealing a substantial overlap between the two labeling patterns. (*B*) Double labeling with pAb-TF and an mAb directed against the Golgi marker giantin. Partial colocalization was observed, and the two labelings seemed to influence each other, because cells with a strong nsp2TF signal showed a weaker giantin signal (e.g., the cell in the middle). (*C*) Nsp2TF-expressing MARC-145 cells double labeled with pAb-TF and the anti-p53 mAb. As in infected cells, a considerable overlap was observed. (*D*) Double labeling with pAb-TF and the anti-giantin mAb, again showing partial colocalization. (*E*) Double labeling with pAb-TF and anti-nsp2 mAb 58-46 showing identical labeling patterns, indicating that mAb 58-46 does recognize nsp2TF in nsp2TF-expressing MARC-145 cells, in contrast to observations in PRRSV-infected cells (Fig. 7). (*F*) Nsp2-expressing MARC-145 cells double labeled for the endoplasmic reticulum marker PDI (pAb) and nsp2 (mAb58-46), showing partial colocalization. Cell nuclei were stained with Hoechst 33258. (Scale bars, 20 nm.)

Chapter 7

Transactivation of programmed ribosomal frameshifting by a viral protein

Yanhua Li[a,b,1], Emmely E. Treffers[c,d], Sawsan Napthine[e], Ali Tas[c], Longchao Zhu[a,b,1], Zhi Sun[a], Susanne Bell[e], Brian L. Mark[f], Peter A. van Veelen[d], Martijn J. van Hemert[c], Andrew E. Firth[e], Ian Brierley[e], Eric J. Snijder[c], and Ying Fang[a,b,1]

[a]Department of Veterinary and Biomedical Science, Department of Biology/Microbiology, South Dakota State University, Brookings, SD, USA, [b]Department of Diagnostic Medicine/Pathobiology, College of Veterinary Medicine, Kansas State University, Manhattan, KS, USA, [c]Molecular Virology Laboratory, Department of Medical Microbiology, Leiden University Medical Center, Leiden, The Netherlands, [d]Department of Immunohematology and Blood transfusion, Leiden University Medical Center, Leiden, The Netherlands, [e]Division of Virology, Department of Pathology, University of Cambridge, Cambridge, CB2 1QP, UK, [f]Department of Microbiology, University of Manitoba, Winnipeg, Canada, R3T 2N2

[1]Present address: Department of Diagnostic Medicine/Pathobiology, College of Veterinary Medicine, Kansas State University, Manhattan, KS 66506-5705, USA

ABSTRACT

Programmed -1 ribosomal frameshifting (−1 PRF) is a widely used translational mechanism facilitating the expression of two polypeptides from a single mRNA. Commonly, the ribosome interacts with an mRNA secondary structure that promotes −1 frameshifting on a homopolymeric slippery sequence. Recently, we described an unusual −2 frameshifting (−2 PRF) signal directing efficient expression of a transframe protein [nonstructural protein 2TF (nsp2TF)] of porcine reproductive and respiratory syndrome virus (PRRSV) from an alternative reading frame overlapping the viral replicase gene. Unusually, this arterivirus PRF signal lacks an obvious stimulatory RNA secondary structure, but as confirmed here, can also direct the occurrence of −1 PRF, yielding a third, truncated nsp2 variant named "nsp2N." Remarkably, we now show that both −2 and −1 PRF are transactivated by a protein factor, specifically a PRRSV replicase subunit (nsp1β). Embedded in nsp1β's papain-like autoproteinase domain, we identified a highly conserved, putative RNA-binding motif that is critical for PRF transactivation. The minimal RNA sequence required for PRF was mapped within a 34-nt region that includes the slippery sequence and a downstream conserved CCCANCUCC motif. Interaction of nsp1beta with the PRF signal was demonstrated in pull-down assays. These studies demonstrate for the first time, to our knowledge, that a protein can function as a transactivator of ribosomal frameshifting. The newly identified frameshifting determinants provide potential antiviral targets for arterivirus disease control and prevention. Moreover, protein-induced transactivation of frameshifting may be a widely used mechanism, potentially including previously undiscovered viral strategies to regulate viral gene expression and/or modulate host cell translation upon infection.

INTRODUCTION

Among the repertoire of mechanisms that viruses use to control or regulate their gene expression, non-canonical translation plays an important role, in particular for positive-strand RNA viruses whose genomic RNA serves a dual function as mRNA and genome [reviewed in [323]]. A commonly used strategy is –1 programmed ribosomal frame-shifting (–1 PRF), in which mRNA signals induce a significant proportion of translating ribosomes to change reading frame, with ribosomes slipping back (in the 5' direction) by one nucleotide (nt) into an overlapping open reading frame (ORF) before continuing translation, generating a fusion protein composed of the products of both upstream and downstream ORFs [reviewed in [323, 336, 396, 433]]. PRF was first described as the mechanism by which the Gag-Pol polyprotein of the retrovirus Rous sarcoma virus (RSV) is expressed from overlapping *gag* and *pol* ORFs [434, 435] and related signals have since been documented in many other viruses of medical, veterinary, and agricultural importance [436-440]. PRF has also been increasingly recognized in cellular genes of both prokaryotes and eukaryotes as well as in other replicating elements, such as insertion sequences and transposons [441].

Recently, we identified an unusual –2 programmed ribosomal frameshifting (–2 PRF) event that operates during the translation of the genome of porcine reproductive and respiratory syndrome virus (PRRSV), a member of the arterivirus family in the order *Nidovirales* [442]. PRRSV can be divided into distinct European (EU, type 1) and North American (NA, type 2) genotypes. The viral genome comprises a positive-sense RNA molecule, ~15 kb in length [339]. As in other nidoviruses, its 5' proximal region contains two large replicase open reading frames [ORF1a and ORF1b; [361]], with the ORF1b product being expressed as a fusion with the ORF1a product following –1 PRF in the short ORF1a/ORF1b overlap region (see Figure 1). Four ORF1a-encoded proteinases (residing in nsp1α, nsp1β, nsp2, and nsp4) subsequently cleave the pp1a and pp1ab polyproteins into (at least) 14 different nonstructural proteins (nsps; Figure 1A). The recently identified –2 PRF signal is located several kilobases upstream of the ORF1a/ORF1b –1 PRF signal, and maps to the part of ORF1a that encodes nsp2. This large, multifunctional replicase subunit is involved in diverse steps of the arterivirus replicative cycle, including replicase polyprotein processing [418], the formation of replication structures [369, 443], and innate immune evasion [167, 420-422]. At the PRRSV –2 PRF signal, a proportion of ribosomes back up two nucleotides, to generate a transframe fusion protein (nsp2TF) comprising the N-terminal two-thirds of nsp2 and the product encoded by a conserved alternative ORF (transframe; TF) in the –2 reading frame. Compared to full-length nsp2, the nsp2TF product is truncated, equipped with an alternative C-terminal transmembrane domain (Figure 1A), and targeted to a different subcellular compartment [442]. Mutations preventing nsp2TF expression reduce PRRSV replication
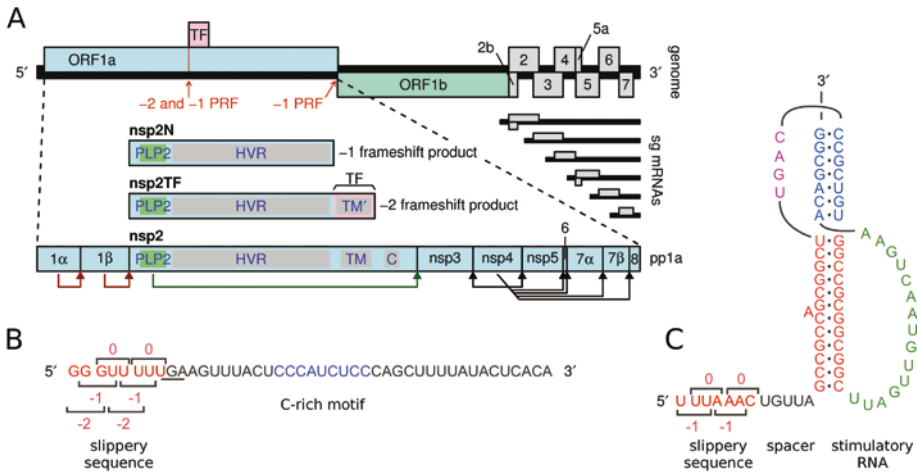
Figure 1: PRRSV genome organization and location of ribosomal frameshifting signals. (*A*) Overview of the ~15-kb PRRSV genome. The long 5′ ORFs 1a and 1b encode nonstructural polyproteins, and at least eight shorter 3′ ORFs (2a-7) encode structural proteins. The 3′ ORFs are translated from a nested set of subgenomic mRNAs, two of which are bicistronic. ORF1a and ORF1b are translated from the genomic RNA, with translation of ORF1b depending on −1 PRF at the end of ORF1a. The TF ORF overlaps the central ORF1a region in the −2 reading frame and is accessed via −2 PRF [442]. A −1 frameshift at the same site generates the nsp2N product (see text). The vertical red line indicates the location of the RG_GUU_UUU shift site (R = A or G, in different arteriviruses). Domains in nsp2/nsp2TF: PLP2, papain-like proteinase; HVR, hypervariable region; TM/TM′, (putative) trans-membrane domains; C, Cys-rich domain. (*B*) Sequence of the SD01-08 RNA in the region of the −2/−1 PRF signal, with the slippery sequence (red) and C-rich motif (blue) highlighted. The −1 reading frame stop codon is underlined and codons for each of the reading frames are indicated. (*C*) Features of the canonical −1 PRF signal present in the PRRSV ORF1a/ORF1b overlap region. The stimulatory RNA pseudoknot is composed of two stems connected by single-stranded loops.

efficiency in cell culture 50- to 100-fold, highlighting the biological importance of the frameshifting event and nsp2TF expression. The −2 PRF takes place at a highly conserved RG_GUU_UUU slippery sequence (R = G or A), and frameshifting is remarkably efficient [around 20% in virus-infected cells and up to 50% in expression systems; [442]].

As depicted in Figure 1B-C, the elements that promote PRF in PRRSV are quite distinct. The −1 PRF signal at the ORF1a/1b junction comprises a "slippery" sequence (generally U_UUA_AAC) where the ribosome changes frame, and a stimulatory RNA pseudoknot structure immediately downstream, an organization that is conserved throughout the *Nidovirales* order [345, 395] and widely used in other viral −1 PRF mechanisms. It is thought that interaction of the translating ribosome with the pseudoknot confounds its RNA-unwinding activity [444, 445] and may induce tension in the mRNA that assists in the uncoupling of codon:anticodon interactions at the shift site [337, 407, 446]. In contrast, only a few cases of −2 PRF in mammalian cells have been documented thus far [407, 442] and the elements involved are poorly understood. Our previous computer-

based RNA folding analysis suggested that the RNA downstream of the slippery sequence (RG_GUU_UUU) used for –2 PRF in PRRSV is rather unstructured and does not fold into a structure compatible with canonical RNA-structure-stimulated PRF. However, mutations within a conserved CCCANCUCC motif located 11 nt downstream of the shift site can reduce or inhibit frameshifting, consistent with the presence of a 3' stimulatory element of some form [442]. Remarkably, our previous study also provided indications for the occurrence of efficient –1 frameshifting at (or near) the same slippery sequence. Due to the presence of a translation termination codon in the –1 reading frame immediately following the slippery sequence, this would yield a truncated form of nsp2, termed nsp2N (Figure 1A).

   In this report, we identify PRRSV replicase subunit nsp1β as a transactivator of efficient –2 and –1 PRF at the same slippery sequence and provide evidence that its frameshift-stimulatory activity requires interaction with the viral mRNA. In support of this, a highly conserved putative RNA-binding motif (GKYLQRRLQ), integrated into the structure of nsp1β's papain-like autoproteinase domain, was found to be critical for the stimulation of frameshifting and for interacting with the RNA sequence of the PRRSV PRF signal. The minimal RNA sequence required to direct efficient PRF was mapped within a 34-nt region of the PRRSV nsp2-coding sequence that includes the shift site and the conserved CCCANCUCC motif. Our findings reveal an unusual non-canonical translation mechanism in which a viral protein functions as a transactivator of efficient –2 and –1 PRF. This study advances our understanding of non-canonical translation, suggests that viruses may employ additional strategies to modulate viral and potentially host cell translation during infection, and has practical implications in biotechnology and the design of antiviral strategies.

RESULTS

Alternative –2 and –1 PRF at the same PRRSV slippery sequence

Previously [442], we demonstrated expression of the PRRSV TF ORF (Figure 1A) using a rabbit antiserum raised against the epitope on the C-terminus of the polypeptide it encodes. Subsequently, the frameshift product was immunopurified from infected cells and mass spectrometry (MS) was employed to identify both the site (RG_GUU_UUU) and direction (–2, rather than +1) of ribosomal frameshifting. In both PRRSV-infected cells and an ORF1a expression system, and using distantly related type 1 and type 2 PRRSV isolates, the same studies revealed an additional nsp2-related product (nsp2N) with a size consistent with –1 PRF occurring at the same site [estimated efficiency ~7%; [442]]. However, a stop codon is present in the –1 frame immediately downstream of the RG_GUU_UUU slippery sequence (Figure 1A) and consequently, if nsp2N were derived

from −1 frameshifting, it would lack a unique C-terminal sequence that could be used to discriminate it from a product derived through the internal proteolytic cleavage of full-length nsp2. In an attempt to confirm the occurrence of −1 PRF by IP-MS, we sought to extend the potential −1 frameshift product with a unique C-terminal signature. In a full-length cDNA clone of the previously used PRRSV isolate SD01-08 [a type 1 virus; [402]], the −1 frame stop codon (UGA) was replaced by a tryptophan codon (UGG), extending the −1 frame by an additional 87 codons (Figure S1, SD01-08-M1). However, this point mutation unavoidably also introduced amino acid substitutions in the overlapping 0 and −2 frames encoding nsp2 and nsp2TF (Glu→Gly and Lys→Glu, respectively), and perhaps as a consequence, the resulting recombinant virus was severely crippled (titer reduced to 10^3 FFU/ml), preventing us from immunopurifying sufficient nsp2N for reliable MS analysis. We, therefore, reverted to a type 2 PRRSV isolate [SD95-21; [447]] and introduced the same A to G mutation, which in this case extended the −1 ORF by 23 additional codons to generate mutant SD95-21-M1 (Figure S1). Fortunately, despite carrying Asp→Gly and Thr→Ala mutations in the nsp2 and nsp2TF products respectively, this recombinant virus replicated to much higher titers ($10^{6.2}$ FFU/ml) and the C-terminally extended nsp2N product (nsp2N*) could be immunopurified from infected MARC-145 cells. A gel slice containing the nsp2N* band was analyzed by LC/MS/MS and a QVFWPR tryptic peptide that spanned the frameshift site and is compatible with −1 PRF at the RG_GUU_UUU sequence was identified (Figure S2). To verify correct identification of this peptide, a synthetic version was subjected to the same LC/MS/MS analysis. The tandem mass spectrum of this synthetic peptide was found to be identical to that of the peptide derived from the nsp2N*-containing gel slice (Figure S2D), confirming that nsp2N is indeed translated via −1 PRF at the RG_GUU_UUU slippery sequence, which is, therefore, able to direct both −1 and −2 PRF.

PRRSV nsp1β is required for efficient −1 and −2 frameshifting in the nsp2-coding region

Previously we demonstrated that translation of the complete PRRSV ORF1a sequence is sufficient to allow efficient −2 PRF [442]. To define the minimal sequence requirements for −2/−1 PRF in PRRSV isolate SD01-08, we focused our attention on the N-terminal half of ORF1a (the nsp1α-nsp3 region) and generated a panel of truncated ORF1a constructs (Figure 2A) for expression in the recombinant vaccinia virus/T7 RNA polymerase system [429]. Following radiolabeling of proteins synthesized in transfected RK-13 cells, expression of nsp2, nsp2TF, and nsp2N was analyzed by immunoprecipitation using monoclonal antibody (mAb) α-EU-nsp2 and rabbit antiserum α-EU-TF, recognizing all three nsp2-related products and the unique C-terminal epitope of nsp2TF, respectively (see Figure S1B for a summary of antibody nomenclature and epitopes recognized). As shown in Figure 2B, constructs lacking the nsp1α- and/or nsp3-coding region still ef-
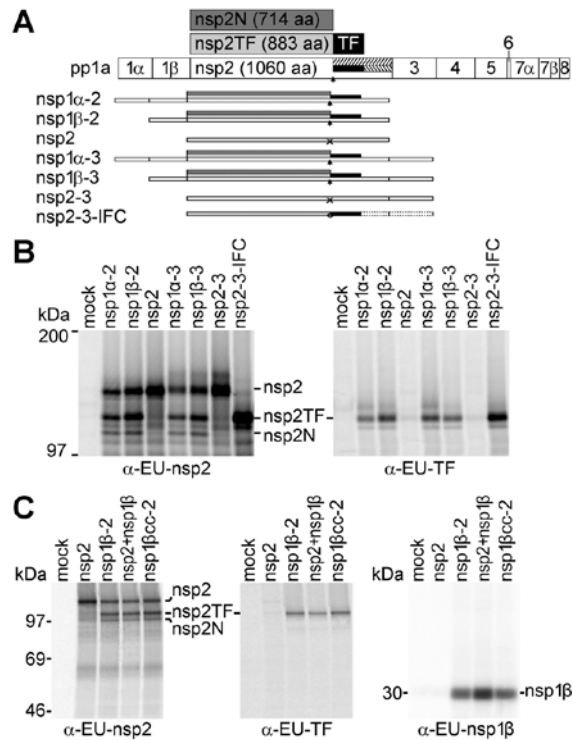
Figure 2: PRRSV nsp1β transactivates −2/−1 PRF. (*A*) Schematic representation of expression products from vectors encoding different combinations of replicase subunits from the nsp1α-nsp3 region of type 1 PRRSV (isolate SD01-08), expressed as single nsps or self-cleaving multi-nsp polyproteins. PRRSV pp1a and its processing scheme are shown at the top and the −2 and −1 frameshift products, nsp2TF and 2N, are shown in light and dark grey with the polypeptide encoded by the TF ORF indicated in black. The arrow indicates the PRF site and untranslated parts of the −2 and −1 reading frames are hatched. The scheme below, for each expression vector used in panel B, shows which nsps were expressed, with arrows indicating the occurrence of −2/−1 PRF and "X" indicating lack of efficient frameshifting. Nsp2-3-IFC (13) is an engineered in-frame control construct that expresses nsp2TF only, due to the insertion of two nucleotides at the PRF site (circle). (*B*) Expression of different protein combinations (see panel A) using the recombinant vaccinia virus/T7 RNA polymerase expression system and RK-13 cells, revealing that nsp1β expression is required for efficient −2/−1 PRF. After metabolic labeling, expression products were immunoprecipitated with the antibodies indicated below each panel; mAb α-EU-nsp2 recognizes the common N-terminal domain of nsp2, nsp2TF, and nsp2N, while α-EU-TF recognizes the C-terminal domain of nsp2TF. Immunoprecipitated proteins were separated by SDS-PAGE and visualized by autoradiography. (*C*) Analysis of −2/−1 PRF transactivation by nsp1β using the recombinant vaccinia virus/T7 RNA polymerase expression system as described for *B*. RK-13 cells were transfected with plasmid DNAs expressing nsp2, nsp1β-2, nsp2+nsp1β (from separate plasmids), or nsp1βcc-2, with the latter containing an nsp1β-coding sequence in which the large majority of codons had been synonymously mutated (Figure S3). Expression products were immunoprecipitated using specific antibodies indicated at the bottom of each panel and visualized by SDS-PAGE and autoradiography.

ficiently expressed nsp2TF and nsp2N. In contrast, constructs lacking the nsp1β-coding region expressed nsp2 but only trace amounts of nsp2TF or nsp2N were detected. This indicates that nsp1β, or the RNA sequence encoding nsp1β, is required for efficient −2/−1 PRF at the RG_GUU_UUU slippery sequence in the nsp2-coding region, located some 2.5 kb downstream of the nsp1β-coding region.

Extending this further, using the same expression system, nsp2 and nsp1β were expressed from separate, co-transfected plasmids (pLnsp2 and pLnsp1β) rather than as a self-cleaving nsp1β-2 polyprotein (pLnsp1β-2). Again, both nsp2TF and nsp2N were produced (Figure 2C), indicating that nsp1β can stimulate −2/−1 PRF in the nsp2-coding region *in trans*. To establish whether this effect was mediated by the nsp1β protein or the nsp1β-coding RNA sequence, a drastically altered version of the nsp1β-2 expression vector was produced in which almost every codon of the nsp1β-coding sequence was mutated synonymously, while avoiding rare codons (mutant pLnsp1βcc-2; Figure S3). This pLnsp1βcc-2 construct expresses an unaltered nsp1β protein, but the nucleotide sequence encoding it is changed to such an extent that we would expect to have disrupted any primary sequence or RNA secondary structure elements that might be involved in −2 PRF (for example, an element having a long-range interaction with the PRF region in the nsp2-coding sequence). Immunoprecipitation analysis revealed that nsp2TF and nsp2N were expressed with equal efficiency in cells transfected with pLnsp1βcc-2 and wild-type pLnsp1β-2 (Figure 2C), indicating that PRF stimulation involves the nsp1β protein rather than an RNA signal in the nsp1β-coding sequence.

Minimal RNA sequence requirements for −2/−1 PRF

We next set out to define the minimal RNA sequences in the nsp2-coding region that are required for efficient −2/−1 PRF. To this end, we prepared a reporter gene construct in which PRRSV RNA sequences from the PRF-inducing region were placed between two luciferase genes [pDluc [448, 449]; Figure 3A]. Whereas the ORF1a frame of the PRRSV insert was placed in-frame with the upstream (*Renilla*) luciferase gene, the downstream (firefly) luciferase was in the −2 frame and thus its expression depended on the occurrence of −2 frameshifting. Also −1 PRF could be monitored, since the native stop codon in the −1 frame was retained and −1 PRF would, therefore, yield a polypeptide slightly shorter than the product resulting from translation termination in the zero reading frame. As controls, an in-frame control (IFC) construct was also prepared in which the two luciferase genes were aligned in the same frame by inserting two nucleotides (CU) immediately downstream of the slippery sequence. A previously described PRF knockout construct [KO2; Figure S1; [442]] containing point mutations within the slippery sequence and downstream C-rich region was also included in the analysis.

Initially, a 79-nt region spanning 5-nt upstream of the slippery sequence to 66-nt downstream (including the conserved CCCANCUCC motif) was cloned between the
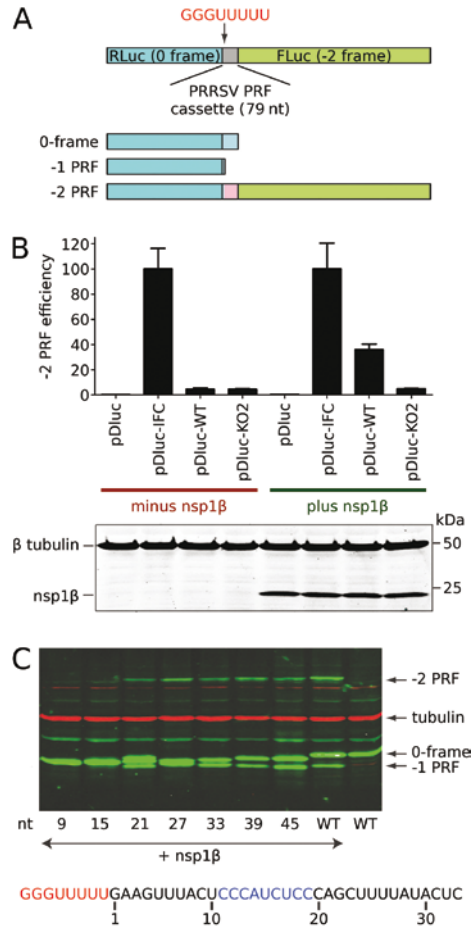
Figure 3: Delineation of RNA elements required for PRRSV −2/−1 PRF. (*A*) Schematic representation of the pDluc dual luciferase construct. The GG_GUU_UUU shift site, 5 upstream nucleotides and 66 downstream nucleotides (79 nt in total) were inserted between the *Renilla* and firefly luciferases genes such that −2 PRF is required for firefly luciferase expression. (*B*) Dual luciferase reporter assay showing that efficient −2 PRF depends on co-expression of nsp1β. For type1 PRRSV (isolate SD01-08), nsp1β was co-expressed with dual luciferase constructs containing a WT or −2 PRF knock out (KO2) frameshift signal. Mutant KO2 [Figure S1; [442]] contains point mutations within both slippery sequence and downstream C-rich motif. The −2 PRF efficiencies were calculated by comparing the ratio of firefly and *Renilla* luciferase activities, using the in-frame control mutant (IFC, Figure S1A) as a reference. Error bars represent the standard deviation of three independent experiments, in which each construct was transfected in duplicate. The bottom panel shows a Western blot analysis confirming equal expression of nsp1β and equal loading (β tubulin). (*C*) Delineation of the minimal RNA sequence requirements for efficient −2/−1 PRF. Starting from a construct containing the 66 nt downstream of the slippery sequence, a series of 3′ truncations was engineered in pDluc. Upon co-expression with nsp1β, cell lysates were analyzed by Western blot, using an antibody recognizing the common *Renilla* luciferase part of all pDluc translation products (see panel A). The number below each lane represents the remaining PRRSV-specific RNA sequence downstream of the slippery sequence, of which 21 nt were sufficient for efficient −2/−1 PRF in this assay.

two luciferase genes (construct pDluc-WT). Frameshifting efficiencies were determined by comparing the ratio of enzymatic activities of firefly and *Renilla* luciferase in parallel HEK-293T cell cultures transfected with individual pDluc constructs with or without co-transfection of the plasmid expressing nsp1β. As shown in Figure 3B, in comparison to the IFC control, the wild-type PRRSV −2 PRF efficiency was ~38%, and this high-level of −2 PRF was only observed in cells co-transfected with the nsp1β-expressing plasmid; in the absence of the transactivator, only low levels of −2 PRF (<5%) were observed. As expected, frameshifting was not observed in cells transfected with pDluc-KO2. Western blot analysis of transfected cell lysates revealed that both efficient −2 PRF and efficient −1 PRF could be observed with pDluc-WT provided that an nsp1β expression plasmid was co-transfected (Figure 3C). These data indicated that the 79-nt PRRSV sequence included in pDluc-WT contains all *cis*-acting sequences required for efficient −2/−1 PRF, and that, as documented above, both types of frameshift depend on the presence of nsp1β. In the absence of this transactivator, only low levels of PRF were observed.

To further investigate the key RNA sequences required for PRF, in-frame deletions were introduced into pDluc-WT, starting from the 3′ end of the PRRSV insert. As shown in Figure 3C, an initial deletion that reduced the PRRSV sequence downstream of the shift site to 45 nt (pDluc-45) led to a small reduction in −2 PRF (about 2-fold), albeit with a concurrent increase in −1 PRF. Subsequent deletions had no further effect until part of the conserved CCCANCUCC motif was removed (Figure 3C; compare pDluc-21 and pDluc-15). In pDluc-15, which lacked the second half (CUCC) of the conserved motif, the capacity for transactivation of PRF by nsp1β was lost. These data provided further support for a role of the C-rich motif in PRF, and allowed us to define the functional PRRSV −2/−1 PRF cassette as a 34-nt region containing the slippery sequence and the 3′ C-rich motif.

Identification of a conserved nsp1β motif that is critical for PRF transactivation

The nsp1α-nsp1β region has previously been implicated in a variety of processes in the arterivirus replicative cycle, including replicase polyprotein processing [416], transcriptional control [450, 451], and innate immune evasion [447, 452]. An analysis of nsp1β sequence conservation (Figure 4A), together with the published crystal structure of nsp1β from a type 2 PRRSV isolate [[453; Figure 4B-C], pointed towards a previously identified conserved sequence motif as a potential RNA interaction domain. This sequence, GKYLQRRLQ in both type 1 and type 2 PRRSV, forms one of three alpha-helices [labeled α4 in [453]] in the region between the active site Cys and His residues of the papain-like proteinase domain (PLP1β) that constitutes the C-terminal two-thirds of nsp1β. Interestingly, compared to the active site of the PLP1β proteinase, helix α4 maps to the other side of the molecule and, in the available crystal structure, the three conserved basic residues of the GKYLQRRLQ motif are exposed on the nsp1β surface. Moreover, in
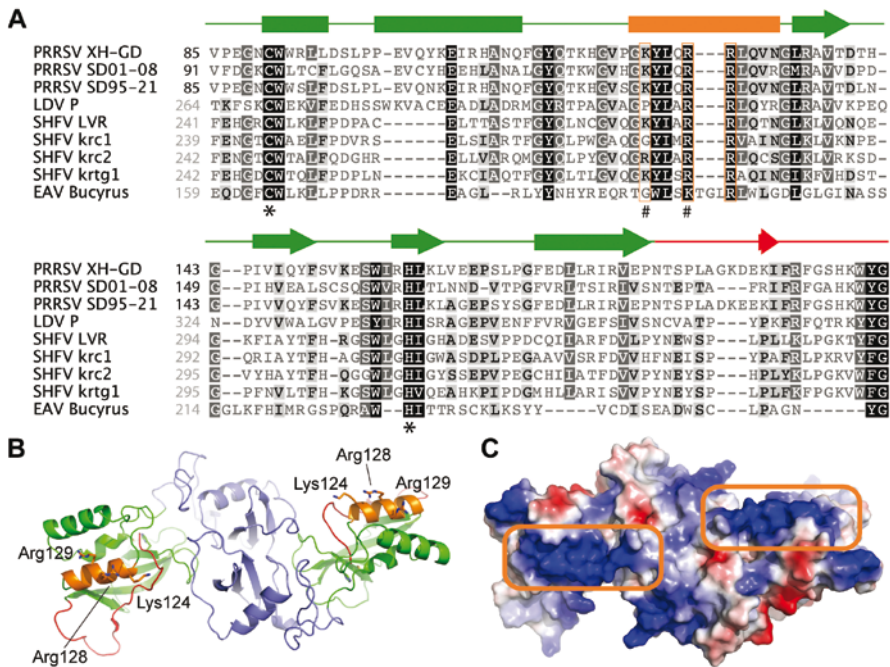
Figure 4: PRRSV nsp1β sequence and structure. *(A)* Amino acid sequence alignment of the PLP1β domains from selected arterivirus nsp1β proteins. Secondary structure elements [based on the published crystal structure from type 2 PRRSV isolate XH-GD; [453]] are shown above the alignment and are color-matched to the nsp1β structure in panel B. Conserved basic residues in PLP1β helix α4 are boxed in orange; #, residues mutated in mutant 1βKO (see text for details); *, PLP1β active site residues. PRRSV sequences are numbered (black) from the nsp1α/nsp1β cleavage site, whereas all other sequences are numbered (grey) starting from the N-terminus of the pp1a polyprotein. EAV, equine arteritis virus; LDV, lactate dehydrogenase elevating virus; SHFV, simian hemorrhagic fever virus; names of specific isolates used are indicated. Genbank accession numbers of sequences used: EU624117 (PRRSV XH-GD), DQ489311 (PRRSV SD01-08), KC469618 (PRRSV SD95-21), NC_001639 (LDV P), NC_003092 (SHFV LVR), HQ845737 (SHFV krc1), HQ845738 (SHFV krc2), JX473847 (SHFV krtg1), NC_002532 (EAV Bucyrus). *(B)* Cartoon representation of the crystal structure of the nsp1β dimer from a type 2 PRRSV isolate [PRRSV XH-GD; PDB entry 3MTV; [453]]. For both monomers, the N-terminal domain is colored purple, whereas the PLP1β domain and the C-terminal extension (leading up to the nsp1β/nsp2 site cleaved by PLP1β) are colored green and red, respectively. Helix α4 of PLP1β, containing the conserved GKYLQRRLQ motif, is colored orange with basic residues represented as sticks. *(C)* Electrostatic surface representation of the nsp1β dimer showing the positively charged (blue) patches on helix α4 of PLP1β (boxed in orange) created by the basic residues of the GKYLQRRLQ motif. Both patches reside on the same side of the structure, potentially allowing for RNA to bind across the entire dimer surface.

the nsp1β homodimer that was the basis for structural studies, the α4 helices of both monomers map to the same side of the dimer and may form a continuous surface across the protein that binds nucleic acid (Figure 4C; see Discussion).

In a recent study [447], the GKYLQRRLQ motif was targeted by site-directed mutagenesis and the Lys and the first Arg of the motif were replaced with Ala (mutant 1βKO, Figure S1). For both PRRSV genotypes, the replication of the 1βKO mutant in MARC-145 cells was found to be seriously crippled. The fact that we had observed similar defects in mutants in which the −2/−1 PRF signal had been inactivated, or in which the expression of a functional nsp2TF was prevented [442], prompted us to investigate whether this KR→AA double mutation affected nsp2TF/nsp2N expression. Strikingly, upon expression of nsp1β-nsp2 from either PRRSV genotype carrying these nsp1β mutations, neither nsp2TF nor nsp2N could be detected (Figure 5). These data indicate that the GKYLQRRLQ motif plays a key role in PRF activation.

To investigate nsp1β transactivation of PRF in the context of PRRSV infection, we analyzed nsp2 expression using the 1βKO mutant of both PRRSV genotypes. As controls, we included the corresponding KO2 mutants, which carry mutations within the slippery sequence and C-rich region that eliminate frameshifting [Figure S1 [442]]. Using reverse genetics, KO2 and 1βKO mutant viruses were recovered from full-length infectious clones of the two PRRSV genotypes. Both mutants replicated poorly in MARC-145 cells, but for the type 2 PRRSV isolate (SD95-21), they produced titers ($10^{5.1}$ and $10^{5.3}$ FFU/ml for KO2 and 1βKO, respectively) that sufficed for the subsequent experiments of infection, metabolic labeling and radioimmunoprecipitation analysis. As expected (Figure 6A), the expression of nsp2, nsp2TF, and nsp2N was detected in SD95-21-WT-infected cells,
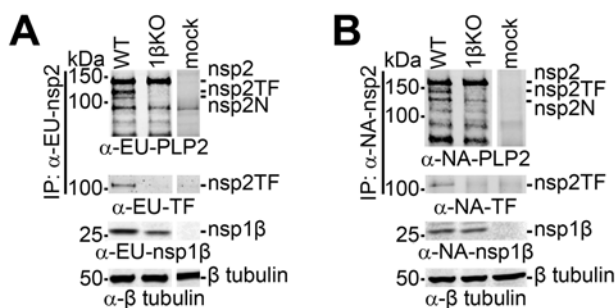

Figure 5: A conserved motif in PRRSV PLP1β is critical for trans-activation of −2/−1 PRF in an expression system. The recombinant vaccinia virus/T7 RNA polymerase expression system and HEK-293T cells were used to express WT and 1βKO mutant nsp1β-nsp2 polyproteins from *(A)* type 1 and *(B)* type 2 PRRSV. The 1βKO mutant carried a double Ala substitution of basic residues in the highly conserved GKYLQRRLQ motif of nsp1β (see also Figure 4 and S1). Expression products were immunoprecipitated with mAbs recognizing the common N-terminal domain of the nsp2-related products. Following SDS-PAGE, they were identified in Western blot analysis using antibodies recognizing the common nsp2 domain, the C-terminus of nsp2TF, or nsp1β. A tubulin antiserum was used for a loading control.
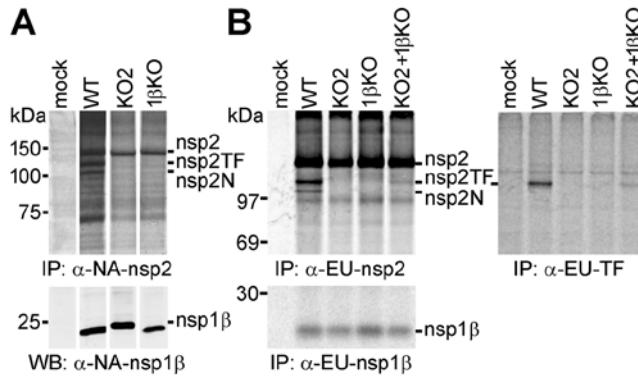
Figure 6: A conserved motif in PRRSV PLP1β is critical for transactivation of −2/−1 PRF in infected cells. (*A*) Analysis of nsp2-related products in MARC-145 cells infected with WT type 2 PRRSV (isolate SD95-21) or mutants KO2 and 1βKO. Mutant KO2 (Figure S1) contained PRF-inactivating point mutations in both slippery sequence and downstream C-rich motif, whereas 1βKO carried a double Ala substitution of basic residues in the highly conserved GKYLQRRLQ motif of nsp1β. Following metabolic labeling, proteins were immunoprecipitated using mAb α-NA-nsp2 and visualized by SDS-PAGE and autoradiography. The expression of nsp1β was monitored by Western blot analysis. (*B*) BHK-21 cells were transfected with *in vitro* transcribed full-length RNA of WT, KO2 or 1βKO PRRSV SD01-08 (type 1) or were double-transfected with equal amounts of KO2 and 1βKO RNA to demonstrate complementation between these two virus mutants. Following metabolic labeling, viral proteins were immunoprecipitated using specific mAbs that recognize a common nsp2 domain (α-EU-nsp2 panel), the polypeptide encoded by the TF ORF (α-EU-TF panel), or nsp1β (α-EU-nsp1β panel). Protein products were visualized using SDS-PAGE and autoradiography.

whereas only nsp2 was recovered from cells infected with either SD95-21-KO2 or SD95-21-1βKO, while their nsp1β was expressed at a level similar to that observed with the wild-type (WT) virus.

Unfortunately, the 1βKO mutant of the PRRSV type 1 isolate (SD01-08) yielded very low titers in MARC-145 cells (10^2 FFU/ml). Considering the number of viral functions and properties potentially affected by nsp1β mutations (see Discussion), we therefore performed a so-called 'first-cycle analysis' of the phenotypes of SD01-08 WT, KO2, and 1βKO. The three viruses were launched by transfecting *in vitro* transcribed full-length RNA into BHK-21 cells, which support replication of transfected PRRSV RNA but cannot be infected by the progeny virus released from the transfected cells, due to the lack of the appropriate receptor(s) on their surface [454]. Moreover, BHK-21 cells have a defect in interferon production [455], thus minimizing the (potential) impact of host innate responses on the comparison of viral replication phenotypes. Following metabolic labeling of protein synthesis in transfected cells, a radioimmunoprecipitation analysis revealed that SD01-08-1βKO produced large amounts of nsp2, whereas the production of nsp2TF was greatly reduced and nsp2N was not detected (Figure 6B). As previously established, SD01-08-KO2 produced only nsp2, while SD01-08-WT produced all three nsp2 variants. Equal expression of nsp1β in WT-, KO2-, and 1βKO-transfected cells was

confirmed by immunoprecipitation with an nsp1β-specific mAb. We also investigated whether the mutations in 1βKO affected the activity of the PLP1β protease or the (potential) involvement of nsp1β in the control of viral subgenomic mRNA synthesis. While the total amount of nsp1β and viral RNA was somewhat reduced in 1βKO-transfected cells, cleavage of the site between nsp1β and nsp2 and subgenomic mRNA production (Figure S4B-C), were not affected by the mutations in the GKYLQRRLQ motif, nor did they affect nsp1β stability (Figure S4D). Finally, we included a double-transfection of BHK-21 cells with KO2 and 1βKO full-length RNA (Figure 6B) and demonstrated complementation between the two PRF-negative mutants leading to reactivation of nsp2TF/nsp2N expression. As expected, the wild-type nsp1β expressed by mutant KO2 was able to transactivate -2/-1 PRF on the wild-type ORF signal in the 1βKO genome, again confirming that the GKYLQRRLQ motif plays a critical role in the PRF stimulatory activity of nsp1β in PRRSV-infected cells.

PRRSV nsp1β interacts with the RNA signals that direct –2/–1 PRF

To test the hypothesis that nsp1β, and specifically its GKYLQRRLQ motif, interacts with the PRRSV RNA sequences that direct –2/–1 PRF, we developed an RNA-binding protein immunoprecipitation assay. To produce an RNA target, we engineered plasmid pR79WT-EGFP yielding an RNA in which a 79-nt PRRSV SD01-08 RNA sequence (Figure 3A) containing the shift site and conserved CCCAUCUCC motif was fused to the enhanced green fluorescent protein (EGFP) open reading frame (Figure 7A). The latter served as a target for quantitative RT-PCR (qRT-PCR) amplification of target RNA bound to nsp1β. As controls, we included plasmids pR79KO2-EGFP and pR79CC2-EGFP, containing combinations of point mutations in the shift site and/or CCCAUCUCC motif that were previously demonstrated to completely inactivate PRF [Figure S1; [442]]. To express the nsp1β bait, we used constructs pFLAG-nsp1β-WT and pFLAG-nsp1β-KO, producing wild-type and mutant (K130A/R134A) nsp1β, respectively, each fused to an N-terminal triple FLAG tag. The empty vectors pFLAG and pEGFP were included as negative controls.

Following co-transfection of vectors expressing RNA target and nsp1β into 293T cells, cell lysates were prepared. Western blot and qRT-PCR analysis (Figure 7B) were first used to determine the expression levels of nsp1β bait and target RNA, respectively, and confirmed the presence of similar amounts of both molecules in all co-transfection samples. Subsequently, we immunoprecipitated FLAG-nsp1β using an anti-FLAG mAb and analyzed these samples for co-immunoprecipitation of target RNA using the same qRT-PCR method, while verifying successful immunoprecipitation of nsp1β with a specific mAb (Figure 7C). A strong and specific RNA co-immunoprecipitation signal was detected only in samples from cells co-transfected with pFLAG-nsp1β-WT and pR79WT-EGFP. In contrast, when mutant 1βKO carrying the K130A/R134A double mutation in the GKYLQRRLQ motif was used, only very low levels of target RNA were pulled down, suggesting
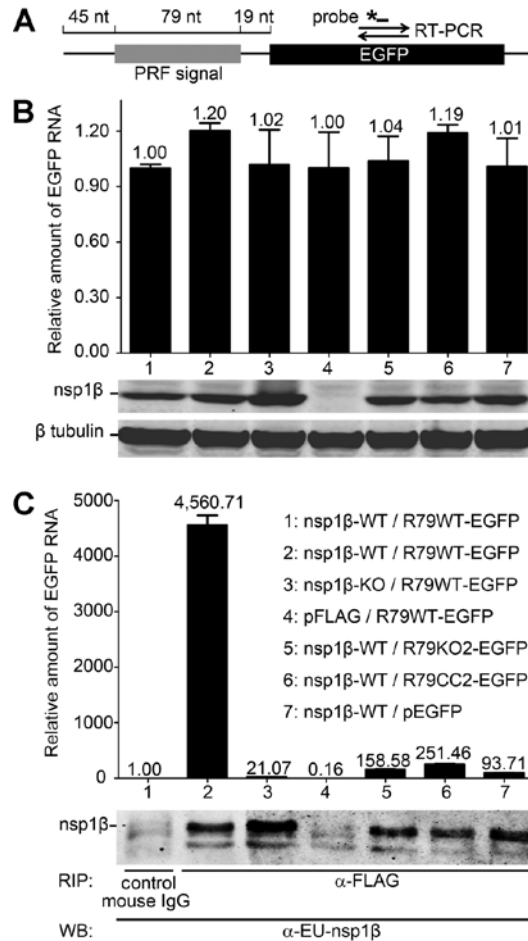
Figure 7: An RNA carrying the PRRSV −2/−1 PRF signal co-immunoprecipitates with nsp1β. A protein-RNA interaction assay was designed based on co-immunoprecipitation of nsp1β and RNA transcripts containing 79-nt of type1 PRRSV sequence, including the −2/−1 PRF signal. (A) Schematic representation of the target RNA in which a WT 79-nt PRF signal (R79WT-EGFP), or its mutant KO2 or CC2 derivatives (see Figure S1A), was fused to the EGFP sequence. The latter served as a target for qRT-PCR amplification (primer set and TaqMan probe indicated), which was used to quantify the amount of RNA target bound to nsp1β. (B-C) HEK-293T cells were co-transfected with a plasmid expressing WT or 1βKO FLAG-tagged nsp1β and a plasmid expressing a WT or mutant target RNA. Empty vectors (pFLAG and pEGFP) were included as negative controls. (*B*) Detection of input levels of nsp1β bait and RNA target in cell lysates prior to the co-immunoprecipitation assay. Quantitative qRT-PCR was used to determine the levels of WT or mutant R79-EGFP mRNA in transfected cells (top panel). Western blot analysis was used to monitor the input of 1βKO or WT nsp1β bait (middle panel) and to verify the use of equal amounts of cell lysate (β-tubulin control; bottom panel). Lane numbers are explained in panel B. (*C*) Following FLAG-nsp1β immunoprecipitation, the amount of co-precipitating target RNA was determined by qRT-PCR (see panel A). Western blot analysis using a mAb α-EU-nsp1β was used to monitor the amount of immunoprecipitated nsp1β. A legend explaining the co-transfected plasmids for each lane number is given on the right side of the panel.

that the K130A/R134A mutations impaired the interaction of nsp1β with PRRSV RNA. Only background signal was detected when using a negative control mouse IgG for immunoprecipitation, or when expressing the pFLAG empty vector control, demonstrating specificity for nsp1β. When the PRRSV PRF-specific sequences in the RNA target were mutated (R79KO2-EGFP or R79CC2-EGFP) or the pEGFP empty vector control was used, only trace amounts of RNA (6% or less of the R79WT-EGFP signal) could be captured, thus demonstrating that co-immunoprecipitation of the target RNA strongly depends on the presence of the CCCAUCUCC motif.

To further corroborate the interaction between nsp1β and the 79-nt RNA sequence from the PRF region, we employed a complementary assay [RiboTrap system; [456]] in which RNA transcripts were labeled with 5-bromo-uridine, facilitating their immunopurification using a 5-bromo-U-specific mAb, and subsequent analysis of immunoprecipitates for the presence of proteins binding to the RNA bait. Using PCR amplicons containing T7 promoter and 79-nt PRRSV RNA sequence (R79WT, R79KO2, or R79CC2) as template, 5-bromo-U-labeled RNA transcripts were produced *in vitro* and incubated with lysates of 293T cells transfected with the plasmid expressing 1βKO or wild-type nsp1β. Following immunoprecipitation with the 5-bromo-U-specific mAb, samples were analyzed for the presence of nsp1β using SDS-PAGE and Western blot analysis (Figure 8). A strong and specific nsp1β signal was detected only when the R79WT bait was incubated with cell lysates containing nsp1β-WT. When using lysates containing 1βKO of nsp1β, only a very small fraction of the available protein was bound to the RNA. Likewise, only trace amounts of wild-type nsp1β were pulled down when 5-bromo-U-labeled R79KO2 or R79CC2 RNA was used as bait. These data are consistent with those obtained in the RNA-binding protein immunoprecipitation assay presented in Figure 7
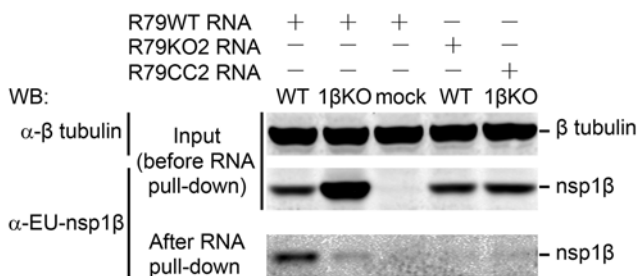

Figure 8: PRRSV nsp1β can be pulled-down using an RNA carrying the −2/−1 PRF signal. Cell lysates from 1βKO or WT nsp1β-expressing HEK-293T cells were incubated with *in vitro* produced BrU-labeled RNA transcripts containing WT or mutant (KO2 or CC2, Figure S1) versions of a 79-nt sequence from the −2/−1 PRF region of type 1 PRRSV (isolate SD01-08). RNA-protein complexes were immunoprecipitated with an anti-BrU antibody and subjected to Western blot analysis using an nsp1β-specific mAb. The amount of β-tubulin in the initial samples was monitored to verify equal loading.

and further support a key role for the GKYLQRRLQ motif in the specific *trans*-activation of the PRRSV −2/−1 PRF by nsp1β.

DISCUSSION

In this paper, we report the remarkable discovery that efficient ribosomal frameshifting in the expression of the PRRSV nsp2TF and nsp2N proteins requires the viral nsp1β protein as a transactivator. Protein-stimulated PRF is unprecedented. It has been reported that cellular annexin A2 may interact with the −1 PRF signal of the coronavirus infectious bronchitis virus, but its role appears to be to down-regulate frameshifting through destabilization of the stimulatory pseudoknot [457], and no specific frameshift-stimulatory protein factors have been identified to date. While downregulation of eukaryotic translation release factor levels can lead to a low-level stimulation of −1 PRF [458, 459], this is a poorly-characterized phenomenon, likely to be a rather non-specific effect brought about by changes in translation rates [460]. It is known that −1 PRF at the human immunodeficiency virus type 1 slippery sequence can be promoted by replacing the natural stimulatory RNA with a combination of the iron-responsive element (IRE) RNA and its cognate binding partner (the IRE-binding protein), but this is a highly artificial experimental system and the stimulation of −1 PRF is very weak [461].

Exactly how nsp1β stimulates frameshifting remains to be determined. Based on the RNA binding experiments, we propose that the region 3' of the PRRSV PRF slippery sequence acts to recruit nsp1β, or an nsp1β-containing protein complex, which modulates ribosome function to promote frameshifting. Our evidence to date supports the view that nsp1β binds directly to the C-rich region, since point mutations within this region strongly reduce RNA binding (Figure 7). However, we cannot rule out the involvement of other factors, for example, poly (C) binding proteins [PCBPs; [462]], which are known to interact with C-rich regions. Moreover, certain PCBPs have been reported to bind to PRRSV nsp1β in pull-down assays [463]. As the C-rich region is located only 11 nt downstream of the slippery sequence, this would likely place bound nsp1β, or an nsp1β-containing complex, in close proximity to a ribosome decoding the slippery sequence, permitting interactions that may lead to frameshifting. While this model is speculative, there is growing evidence that proteins can modulate the elongation step of protein synthesis. The fragile X mental retardation protein reversibly stalls ribosomes on its target mRNAs [464] and the HIF-1α mRNA-associated cytoplasmic polyadenylation element binding protein 2 binds eEF2 and slows elongation [465]. Conceivable routes through which bound proteins could modulate ribosomal function include induction of ribosomal pausing by acting as a roadblock, recruitment of, or localized depletion of translation factors, and direct interaction with a ribosomal component(s). It may be

significant that nsp1β was reported to interact with rpS14 [463], a protein immediately adjacent to rpS3 of the ribosomal helicase [466], and PCBP1, which is known to interact with RACK1, a ribosome-associated protein located close to the mRNA entry channel [467]. These features are consistent with a role for nsp1β in modulating the ribosomal helicase, the suspected target for the stimulatory RNA sequences of canonical −1 PRF signals.

An interesting aspect of this PRRSV PRF signal is that both −2 and −1 frameshifting events are promoted. Tandem slippage of ribosome-bound tRNAs on RG_GUU_UUU would allow complete A-site re-pairing in both −1 and −2 frames (tRNA anticodon:mRNA codon pairing in 0-frame is 3′-AAG-5′ : 5′-UUU-3′; single tRNA^{Phe} isoacceptor AAG), but especially for −2 PRF, P-site re-pairing appears to be compromised [at least at the 2nd and 3rd positions (lower case), 3′-Cai-5′ : 5′-Ggg-3′; i is inosine]. Some tolerance for P-site mispairing has been noted at certain viral −1 PRF signals, but usually, these are associated with single mismatches in the anticodon:codon interaction [323, 468]. It may be that in the particular context of a −2 PRF, stable P-site repairing is not required, reminiscent of the unusual "single-tRNA" slippage events seen in prokaryotic systems with slippery sequences ending in AAG, where P-site repairing does not appear to be present [469, 470]. In a recent study on RNA secondary structure-stimulated −1 and −2 PRF on a GU_UUU_UUA slippery sequence [407], it was noted that the length of the spacer between slippery sequence and secondary structure affected the relative utilisation of −1 or −2 PRF modes, perhaps a reflection of differences in mRNA tension arising as the ribosomal helicase unwinds the secondary structure, with increased tension forcing the ribosome into the −2 rather than the −1 frame. Interestingly, the relative levels of the PRRSV −2 and −1 PRF products were seen to change as the length of the PRRSV region 3′ of the slippery sequence was shortened from 67 to 45 nt (Figure 3C), although the sum total of PRF was similar. This may hint at the involvement of additional factors, or some subtle effect on the positioning of a bound protein (or complex), that can influence frameshift magnitude, although it remains to be determined whether this is linked to mRNA tension.

Analysis of the published structure of nsp1β from a type 2 PRRSV isolate provides insights into the mechanism of how the protein may interact with viral RNA. The PLP1β domain of nsp1β adopts a papain-like fold consisting of three α-helices that pack against a β-sheet of 4 antiparallel strands [453] (Figure 4B). One of the helices (helix α4 in the overall nsp1β structure) contains a conserved GKYLQRRLQ motif that we now show plays a critical role in the transactivation of frameshifting. The crystal structure of nsp1β suggests that the protein exists as a homodimer [453] and, interestingly, helix α4 of both nsp1β monomers resides on the same side of the dimer, which may generate a continuous, positively charged surface that could bind a long single- or double-stranded RNA molecule (Figure 4C). The involvement of an α-helix in RNA binding is consistent with

the observation that nucleoproteins of many RNA viruses encapsidate the viral genome using domains of α-helical structure [471, 472]. Thus, it is plausible that the GKYLQRRLQ motif of helix α4 directly binds viral RNA, although we cannot exclude the possibility that this helix may be a binding site for a cellular protein that in turn could bind to the PRF signal in the viral RNA.

Except for equine arteritis virus (EAV), the −2 PRF mechanism seems to be conserved in all currently known arteriviruses as judged by the presence of a TF ORF overlapping ORF1a and a conserved slippery sequence and downstream C-rich region [442]. In PRRSV and lactate dehydrogenase elevating virus (LDV) −1 PRF can occur, but in contrast the RG_GUC_UCU shift site in some of the recently identified simian haemorrhagic fever virus (SHFV)-like viruses [352] would preclude −1 PRF while still allowing −2 PRF. It is expected that such PRF events in LDV and SHFV would also be controlled by nsp1β, and indeed the transactivating motif in nsp1β was found to be largely conserved in these viruses (Figure 4A). Possibly, the nsp1β component of the frameshift mechanism, which is encoded several kilobases upstream of the PRF site, evolved secondarily, for example to enhance the efficiency of nsp2TF/nsp2N expression, since in the absence of nsp1β low levels of PRF could still be observed (Figure 2A). Amino acid sequence comparisons reveal that the GKYLQRRLQ motif-containing helix is highly conserved in the PLP1β domains of PRRSV, SHFV, and LDV, but the motif is lacking in EAV. For the latter virus, the three helices of the PLP1β domain are predicted to be present, but with an insertion of three amino acids in the EAV equivalent of the α4 helix compared to the other arteriviruses. The nsp2-encoding region of EAV lacks an equivalent of the (overlapping) TF ORF and produces a substantially smaller nsp2. Assuming the TF ORF was lost at some point during the evolution of the EAV lineage, changes in this helix may have been tolerated when it was no longer required to stimulate PRF *in trans*. Although an alternative evolutionary scenario (*i.e.* a common ancestor of PRRSV, SHFV, and LDV independently acquiring a TF ORF) cannot be excluded, loss of the requirement to transactivate PRF may also explain a second remarkable difference between the nsp1 region of EAV and other arteriviruses: the inactivation of the proteolytic activity of the PLP1α proteinase, resulting in the synthesis of a single nsp1 protein rather than nsp1α and nsp1β [416, 473]. In particular, the N-terminal zinc finger of nsp1 (EAV) or nsp1α (PRRSV) has been implicated in the control of viral subgenomic mRNA synthesis [450, 474-476], a function that may not be compatible with a role in PRF transactivation, thus requiring the internal cleavage of nsp1 by PLP1α in arteriviruses that employ nsp1β-mediated transactivation of TF ORF expression.

The capacity of nsp1β to stimulate both −1 and −2 PRF suggests that protein transactivation could be employed more widely in the induction of programmed frameshifting events in diverse systems. With regards to arteriviruses, it is possible that nsp1β might also modulate translation of host cell mRNAs containing appropriate signals. A cursory

search of porcine mRNAs revealed hundreds of −1 and/or −2 frameshift-compatible shift sites followed by C-rich motifs at an appropriate spacing, though no site that is exactly identical to the PRRSV minimal PRF cassette (8-nt shift site plus the downstream 21 nt). Whether and to what extent frameshifting occurs at such sites remains to be investigated. While the occurrence of nsp1β-responsive frameshift signals in host mRNAs would presumably be spurious, the overall effect may perturb cellular gene expression, thus adding an extra dimension to virus-host interactions.

When screening PRRSV nonstructural proteins for their capacity to suppress type I IFN expression, both nsp1β and nsp2 were found to possess such activities [167, 421, 447, 452, 463]. In reporter gene-based assays, nsp1β had the strongest potential to inhibit IFN-β promoter activity and could also inhibit downstream IFN-induced signaling pathways for expression of IFN-stimulated genes (ISGs), including ISG15 [447, 452, 463, 477, 478]. On the other hand, the PLP2 activity of nsp2 is able to disrupt innate immune signalling by removing ubiquitin (Ub) and Ub-like modifiers from host cell substrates, exhibiting a general deubiquitinating (DUB) activity towards cellular ubiquitin conjugates and also cleaving the ubiquitin homolog ISG15 [167, 420-422]. As documented here, nsp1β transactivates both nsp2TF and nsp2N expression, resulting in the synthesis of three nsp2-related proteins (nsp2, nsp2TF, and nsp2N) that have the N-terminal PLP2-DUB domain in common. Thus, it remains to be established to which extent nsp1β directly modulates the innate immune response or does so by stimulating the expression of nsp2TF and nsp2N. Furthermore, nsp1β may affect the immune response through modulation of host cell mRNA translation. The identification of viral/host elements responsible for innate immune evasion is fundamental for the development of modified live virus vaccines. As illustrated by our reverse genetics studies, mutagenesis of key residues in nsp1β and the PRF site could attenuate virus growth and improve host innate immune responses [442, 447]. Since the GKYLQRRLQ motif and PRF site are highly conserved, technologies developed in this study may have broad application in the field.

MATERIALS AND METHODS

Cells and viruses

HEK-293T, RK-13, BHK-21, and MARC-145 cells were cultured as described previously [402, 419]. The US type 1 PRRSV isolate SD01-08 (GenBank accession #DQ489311) and type 2 PRRSV isolate SD95–21 (GenBank accession #KC469618) were used in all experiments.

Antibodies

Antibodies recognizing PRRSV proteins (see also Figure S1B for the nomenclature used in this paper), including mAb 22–28 (α-EU-nsp1β), mAb 123-128 (α-NA-nsp1β), mAb 36-19 (α-EU-PLP2), mAb 58-46 (α-EU-nsp2), mAb140-68 (α-NA-PLP2), mAb 148-43 (α-NA-nsp2), and a rabbit antiserum recognizing the C-terminal part of nsp2TF (α-EU-TF) were produced as described previously [442]. A rabbit antiserum (α-NA-TF) recognizing the C-terminal epitope (CFLKVGVKSAGDLV) of nsp2TF of type 2 PRRSV was generated by GenScript (Piscataway, NJ). For detection of FLAG-tagged proteins, an anti-FLAG mAb was obtained from Sigma® Life Science. Anti-β-tubulin and anti-dsRNA (J2-0601) mAbs were obtained from Lamda Biotech and English & Scientific Consulting, respectively.

DNA constructs and reverse genetics

Except for the KO2 (Figure S1) and pLnsp1βcc–2 (Figure S3) mutants, for which synthetic DNA was used, all other constructs were made by standard PCR-based mutagenesis and recombinant DNA techniques. Procedures for the construction of plasmids are provided in SI Materials and Methods. Methods for *in vitro* transcription, virus rescue from full-length cDNA clones, and virus titration were described previously [402, 447], [442].

Mass spectrometry

Nsp2N was immunoprecipitated from SD95-21-M1–infected MARC-145 cell lysate using mAb α-NA-PLP2 and samples were separated on a 6% SDS-PAGE gel, which was fixed and stained with Coomassie Brilliant Blue G-250 (Bio-Rad). The band expected to contain nsp2N* (based on predicted protein size) was excised. Trypsin digestion and LC-MS/MS analysis were performed as described previously [431]. MS spectra were searched against a custom-made protein database containing the nsp2N* sequence. As positive control, a synthetic version of the identified frameshift peptide was made and analyzed by LC-MS/MS.

Immunoassays

Different regions of PRRSV ORF1a were transiently expressed in RK-13 or HEK-293T cells using truncated derivatives of expression plasmid pL1a and the recombinant vaccinia virus/T7 polymerase expression system [419]. Expression products were ^{35}S-labeled, immunoprecipitated, and analyzed by SDS-PAGE and autoradiography as described previously [442]. Alternatively, nsp1β and nsp2-related products were detected by consecutive immunoprecipitation of (unlabelled) proteins and Western blot analysis, using a combination of PRRSV nsp-specific monoclonal antibodies as described previously [442, 447]. Wild-type and mutant SD01-08 viruses were launched by transfecting *in vitro* transcribed full-length RNA into BHK-21 cells, and radioimmunoprecipitation was

conducted to detect the expression of nsp1β and nsp2-related products (See SI Materials and Methods for detailed procedures).

Dual luciferase assay

Using FuGENE HD transfection reagent (Roche Molecular Biochemicals), HEK-293T cells were co-transfected with 0.2 µg of dual luciferase plasmid containing the PRRSV PRF sequence and 50 ng of pFLAG-nsp1β. At 24h post-transfection, cells were harvested and luciferase expression was measured using the Dual Luciferase Stop & Glo® Reporter Assay System (Promega) and a luminometer (Bethold). Frameshifting efficiencies were calculated from the ratio of firefly to *Renilla* luciferase activities, using the IFC control construct as the standard.

Analysis of protein sequences and structure

Sequence alignment of the PLP1β domain of PRRSV, LDV, and SHFV nsp1β and EAV nsp1 was performed using the MUSCLE algorithm in Geneious 6 (Biomatters Ltd, Auckland, NZ). Potential RNA-binding residues in nsp1β were identified using the program BindN [479]. Images of the crystal structure of the PRRSV nsp1β dimer (PDB: 3MTV) [453] were created using PyMOL [480].

Assays for detecting interactions between nsp1β and viral RNA

Immunoprecipitation assays to detect RNA-binding proteins were performed using the Magna RIP™ kit (Millipore) and Ribo Trap kit (Medical & Biological Laboratories) following the manufacturer's instructions. The amount of target mRNA bound to nsp1β was determined by quantitative RT-PCR, and the presence of nsp1β in RNA-protein complexes was verified by Western blot. Detailed experimental procedures are presented in SI Materials and Methods.

Acknowledgments

Conflict of interest statement: The authors have filed a patent application that relates to some aspects of this work.

SUPPORTING INFORMATION

Supporting information materials and methods

DNA constructs
Plasmids for expression of full-length or partial PRRSV ORF1a were constructed by RT-PCR amplification of corresponding regions from genomic RNA (nt 191-7702 nt of the SD95–21 genome; nt 222-7361 of the SD01-08 genome). The PCR product was digested with *Nco*I and *Not*I restriction enzymes and ligated into a pL1a backbone digested with the same enzymes. The design of pL1a was described previously [442]. Except for the KO2 (Figure S1) and nsp1βcc–2 (Figure S3) constructs, which were generated by gene synthesis, all mutations were introduced by using the Quick-Change site-directed muta-genesis kit (Stratagene) following the manufacturer's instruction.

Plasmid pLnsp1βcc–2 was constructed by extensively mutating the SD01-08 PRRSV nsp1β-coding region with synonymous replacements (Figure S3), while keeping the encoded amino acid sequence intact and avoiding rare codons. The modified sequence (nsp1βcc) was produced as a synthetic gene and fused back with the nsp2-coding re-gion to generate the nsp1βcc–2 construct. To construct plasmids for the dual-luciferase assay, a 79-nt oligonucleotide (3506 to 3584 nt of SD01-08 genome) containing the WT sequence or mutations (Figure S3, IFC and KO2) at the PRF region was synthesized and cloned into the dual luciferase vector pDluc as described previously [448, 449].

The plasmid expressing FLAG-tagged nsp1β (pFLAG-nsp1β) was generated by PCR amplification of the nsp1β-coding region (nt 762-1376 of the SD01-08 genome) and cloned into the plasmid vector p3xFLAG-CMV–24 (Sigma). Plasmids expressing EGFP-tagged PRF sequence (pR79WT-EGFP, pR79KO2-EGFP, and pR79CC2-EGFP) were gener-ated by cloning the PRF region (nt 3506 to 3584 of SD01-08) into the plasmid vector pEGFP-N1 (Clontech).

In vitro transcription and radioimmunoprecipitation analysis of SD01-08 WT and mutants
Full length SD01-08-WT, KO2, or 1βKO RNA was transcribed from 1 ug of linearized plas-mid DNA using the mMESSAGE mMACHINE T7 kit (Ambion) following the manufacturer's instructions. BHK-21 cells (4×10^6) were electroporated with 8 ug of *in vitro* transcribed RNA using program T-020 of the Amaxa Nucleofector and kit T (Lonza). Newly synthe-sized proteins were labeled from 16.5 to 18.5 hours post transfection in cysteine- and methionine-free medium containing 200 µCi/ml [35S]-methionine and [35S]-cysteine (Perkin-Elmer). Cell lysis and immunoprecipitation analysis were performed as described previously [419]. The mAbs α-EU-nsp1β and α-EU-nsp2 were used to immunoprecipitate nsp1β and nsp2/nsp2TF/nsp2N, and precipitated proteins were separated on a 12% SDS-PAGE gel. Protein bands were visualized using phosphorimaging and a Typhoon Variable

Mode Imager (GE Healthcare). Image analysis was performed with the ImageQuant TL software (GE Healthcare).

Pulse-chase analysis of nsp1β expression by SD01-08 WT and mutants KO2 and 1βKO
Following electroporation with *in vitro* transcribed SD01-08 WT, KO2, 1βKO RNA, or a 1:1 mixture of KO2 and 1βKO RNA, 0.3×10^6 BHK-21 cells were labeled at 16.5h post transfection for 15' in cysteine- and methionine-free medium containing 500 µCi/ml [^{35}]S-methionine and [^{35}]S-cysteine (Perkin-Elmer). After removal of the label medium, cells were either lysed immediately or chased for 1h in the presence of an excess of unlabeled methionine and cysteine. Immunoprecipitation was performed as described previously [419] with mAb α-EU-nsp1β. Precipitated proteins were separated on a 12% SDS-PAGE gel and phosphorimaging was performed as described above.

Immunofluorescence microscopy of transfected BHK-21 cells
Following electroporation with *in vitro* transcribed SD01-08 WT, KO2, or 1βKO RNA, 0.15×10^6 BHK-21 cells were seeded on glass coverslips. At 18 h post transfection, cells were fixed in 3% para-formaldehyde in PBS. Cells were double-labeled with a mAb recognizing dsRNA and Hoechst 33342 to stain nuclear DNA.

*RNA isolation, denaturing formaldehyde gel electrophoreses and in gel hybridization*
BHK cells (0.75×10^6) were electroporated with *in vitro* transcribed RNA of SD01-08 WT, KO2, and 1βKO. At 18 h post transfection, cells were lysed in 20 mM Tris-HCl (pH 7.4), 100 mM LiCl, 2 mM EDTA, 5 mM DTT, 5% (w/v) lithium dodecyl sulfate, and 100 µg/ml proteinase K. Total RNA was extracted and separated on a denaturing formaldehyde gel. Positive-stranded viral RNA was visualized by gel drying and hybridization with a ^{32}P-labeled oligonucleotide probe (PRRSV-hyb1 5'-TCGCCCTAATTGAATAGGTG-3') that is complementary to the 3' end of the viral genome and therefore recognizes all viral mRNAs. 18S ribosomal RNA was used as a loading control and was detected with probe 5'-ATGCCCCCGGCCGTCCCTCT-3'. Hybridized gels were analyzed by phosphorimaging as described above. Correction for loading variations was performed using the amount of 18S RNA in the same lane. The sum of the signal for all viral mRNAs in each lane was used to calculate the relative abundance of each individual mRNA.

RNA-binding protein immunoprecipitation and RNA pull-down assay
RNA-binding protein immunoprecipitation was performed using a Magna RIP™ kit (Millipore) according to the manufacturer's instruction. Briefly, HEK-293T cells seeded in 10-cm petri dishes were co-transfected with plasmids expressing the RNA bait (R79WT-EGFP, R79KO2-EGFP, R79CC2-EGFP, or pEGFP; 8 ug) and the nsp1β bait (nsp1β-WT, nsp1β-KO, or pFLAG; 2 µg). At 24 h post transfection, cell lysates were prepared for co-

immunoprecipitation. FLAG-tagged nsp1β was immunoprecipitated using an α-FLAG mAb, and co-precipitating target RNA was quantified by qRT-PCR using TaqMan® Gene Expression Assay kit (Life technologies) targeting the EGFP RNA sequence. The expression of nsp1β in all co-transfected samples was determined by Western blot analysis using an nsp1β-specific mAb. Ribo Trap kit (MBL International Corporation) was used to further confirm the interaction between nsp1β and the 79-nt RNA sequence from the PRRSV PRF region. The R79WT, R79KO2, or R79CC2 RNA was labeled with 5-bromo-U and *in vitro* synthesized using the MEGAscript® T7 Kit (Life technologies). The 5-bromo-U-labeled RNA transcripts were bound to magnetic beads conjugated with anti-BrU mAb. Subsequently, these magnetic beads were incubated with lysates of HEK-293T cell expressing FLAG-tagged 1βKO or WT of nsp1β. The amount of nsp1β pulled-down with the RNA bait was determined by western blot analysis using an nsp1β-specific mAb.

Supporting information figures
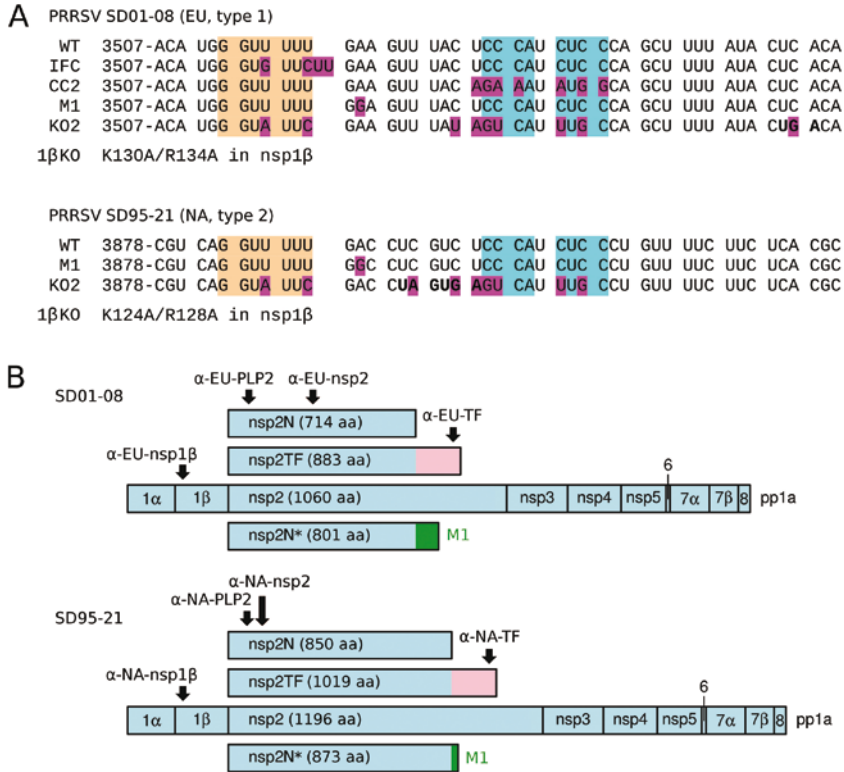
Supporting Information Figure S1: Overview of mutants and antibodies used in this study. (*A*) List of WT and mutant sequences of the PRRSV PRF region (GGUUUUU shift site and conserved CCCANCUCC motif indicated with orange and magenta boxes, respectively). Mutated nucleotides are highlighted in cyan. Coordinates of starting nucleotides refer to PRRSV sequences DQ489311 (type 1 PRRSV) and KC469618 (type 2 PRRSV). IFC, in-frame control; CC2, disrupted CCCANCUCC motif; M1, mutated –1 frame termination codon to C-terminally extend nsp2N; KO2, knockout mutant 2 (premature –2 frame termination codon and disrupted frameshift cassette); 1βKO, nsp1β knockout mutant (double mutation introduced into the nsp1β GKYLQRRLQ motif). (*B*) Sizes of nsp2-related polypeptides described in this study and location of epitopes recognized by the PRRSV-specific antibodies in the nsp1β-2 region. Nsp2N* refers to the C-terminally extended version of nsp2N that is produced by mutant M1, due to removal of the stop codon (see panel A). Origin and original name of each antibody are provided in the Materials and Methods section. Sizes (in aa) for nsp2-related products are shown for GenBank sequences DQ489311 (SD01-08, PRRSV type 1) and KC469618 (SD95-21, PRRSV type 2).
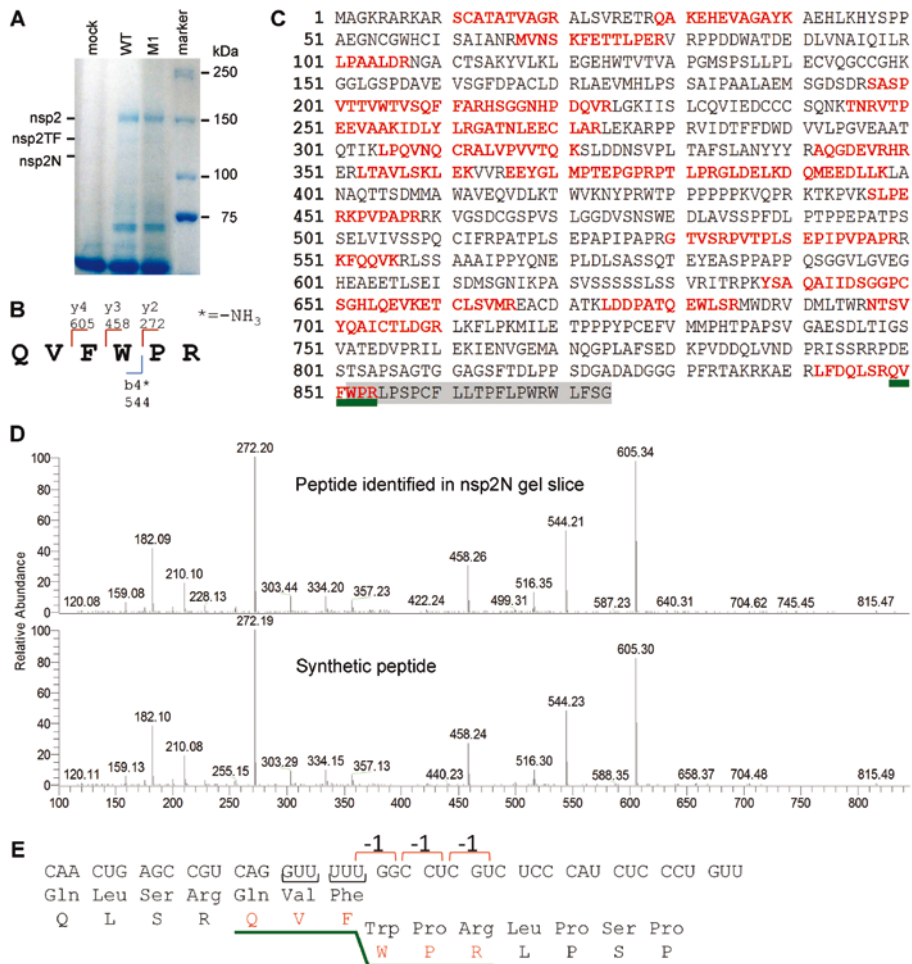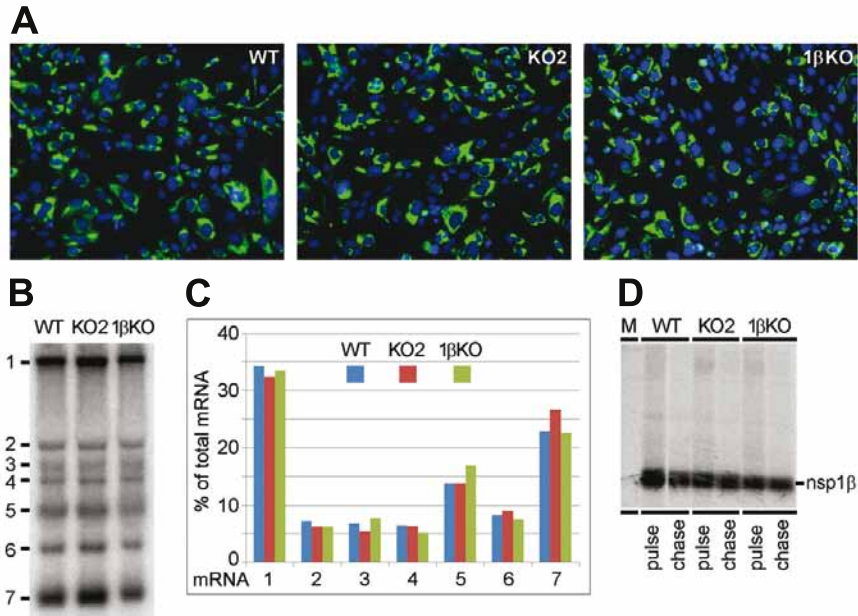
Supporting Information Figure S2: Mass spectrometric analysis of nsp2N* (a C-terminally extended version of nsp2N) purified from cells infected with mutant SD95-21-M1. (*A*) PRRSV-infected or mock-infected MARC-145 cell lysates were immunoprecipitated with nsp2-specific α-NA-PLP2. Immunoprecipitated proteins were separated by SDS-PAGE and stained with Coomassie Blue. The positions of nsp2, nsp2TF, and nsp2N are indicated. (*B*) Fragmentation spectrum of the −1 frameshift-specific peptide QVFWPR. (*C*) Complete amino acid sequence of nsp2N* comprising nsp2N and a 23-amino-acid C-terminal extension (highlighted in gray). Peptides identified by mass spectrometry are depicted in red. The peptide spanning the −1 frameshift site is underlined in green. (*D*) Peptide sequence of the nsp2N* −1 frameshift-specific peptide. The fragment ions that were identified in the LC-MS/MS analysis of the gel slice are indicated. (*E*) Nucleotide sequence and -1 PRF-directed translation of nsp2N* at the frameshift site.

```
  1 TCTGACGTTTACAGGTGGAAGAAATTTGTGATTTTTACGGACTCCTCTCCCAACGGTCGATTTCGC
  1 TCGGATGTGTATAGATGGAAAAAGTTCGTCATCTTCACCGATAGCAGCCCTAATGGCAGATTCAGA
    S  D  V  Y  R  W  K  K  F  V  I  F  T  D  S  S  P  N  G  R  F  R

 67 ATGATGTGGACGCCGGAATCCGATGACTCAGCCGCCCTGGAGGTGCTGCCGCCCGAGTTAGAACGT
 67 ATGATGTGGACCCCCGAGAGCGACGATAGCGCTGCTCTCGAAGTCCTCCCCCCTGAACTGGAGAGA
    M  M  W  T  P  E  S  D  D  S  A  A  L  E  V  L  P  P  E  L  E  R

133 CAGGTCGAGATCCTCACTCGGAGTTTTCCCGCTCATCACCCTATCAACCTAGCTGACTGGGAGCTC
133 CAAGTGGAAATTCTGACCAGAAGCTTCCCTGCCCACCATCCCATTAATCTGGCCGATTGGGAACTG
    Q  V  E  I  L  T  R  S  F  P  A  H  H  P  I  N  L  A  D  W  E  L

199 ACTGAGTCCCCTGAGAACGGTTTTTCTTTCGGCACGTCCCATTCTTGCGGCCACATCGTCCAGAAC
199 ACCGAAAAGCCCCGAAAATGGCTTCAGCTTTGGGACCAGCCACAGCTGTGGGCATATTGTGCAAAT
    T  E  S  P  E  N  G  F  S  F  G  T  S  H  S  C  G  H  I  V  Q  N

265 CCCAACGTGTTTGACGGCAAGTGCTGGCTCACCTGCTTTTTTGGGCCAATCGGCTGAAGTGTGCTAC
265 CCTAATGTCTTCGATGGGAAATGTTGGCTGACATGTTTCCTGGGGCAGAGCGCCGAGGTCTGTTAT
    P  N  V  F  D  G  K  C  W  L  T  C  F  L  G  Q  S  A  E  V  C  Y

331 CACGAGGAACATCTAGCTAACGCCCTCGGTTACCAAACCAAGTGGGGCGTGCATGGTAAGTACCTC
331 CATGAAGAGCACCTGGCCAATGCTCTGGGCTATCAGACAAAATGGGGGGTCCACGGCAAATATCTG
    H  E  E  H  L  A  N  A  L  G  Y  Q  T  K  W  G  V  H  G  K  Y  L

397 CAACGCAGGCTTCAAGTCCGCGGCATGCGTGCTGTGGTCGATCCTGACGGCCCTATTCACGTTGAA
397 CAGAGAAGACTGCAGGTGAGAGGGATGAGAGCCGTCGTGGACCCCGATGGGCCCATCCATGTGGAG
    Q  R  R  L  Q  V  R  G  M  R  A  V  V  D  P  D  G  P  I  H  V  E

463 GCGCTGTCTTGCTCCCAGTCTTGGGTCAGGCACCTGACTCTGAATAATGATGTCACCCCAGGATTC
463 GCCCTCAGCTGTAGCCAAAGCTGGGTGAGACATCTCACCCTCAACAACGACGTGACACCCGGCTTT
    A  L  S  C  S  Q  S  W  V  R  H  L  T  L  N  N  D  V  T  P  G  F

529 GTTCGCCTGACATCCATCCGCATTGTGTCCAACACAGAACCCACCGCTTTCCGGATCTTTCGGTTT
529 GTGAGACTCACCAGCATTAGAATCGTCAGCAATACCGAGCCTACAGCCTTTAGAATTTTCAGATTC
    V  R  L  T  S  I  R  I  V  S  N  T  E  P  T  A  F  R  I  F  R  F

595 GGAGCACATAAGTGGTATGGC
595 GGCGCCCACAAATGGTACGGG
    G  A  H  K  W  Y  G
```

Supporting Information Figure S3: Nucleotide sequences of the WT and synonymously mutated nsp1β-coding region. For each block, the top, middle, and bottom line gives the wild type nsp1β-coding sequence (black), the synonymously mutated sequence (mutations present in pLnsp1βcc-nsp2 given in red), and the (unchanged) translation into amino acids (blue).

Supporting Information Figure S4: First-cycle analysis of RNA synthesis and nsp1β stability of WT and mutant (KO2 or 1βKO) SD01-08 virus in BHK-21 cells. (*A*) IF microscopy analysis of transfection rate in BHK-21 cells electroporated with *in vitro* transcribed full-length PRRSV RNA. Transfected cells were double labeled with a mAb specific for viral dsRNA (green) and Hoechst 33342 (blue) for staining of DNA in cell nuclei. (*B, C*) Gel hybridization analysis and quantification of PRRSV-specific mRNA accumulation in cells transfected with mutants KO2 and 1βKO or a WT control. (*B*) Total intracellular RNA was isolated at 18 h post-transfection and resolved by denaturing formaldehyde agarose gel electrophoresis. PRRSV-specific mRNAs were detected by hybridization of the dried gel with a ^{32}P-labelled probe complementary to the 3'-end of the viral genome and subsequent phosphorimaging. The positions of the PRRSV genome (RNA1) and the six sg mRNAs (RNA2 to RNA7) are indicated. (*C*) The volume of the bands corresponding to each of the viral mRNAs was quantified by phosphorimaging and adjusted for the control 18S ribosomal RNA band in the same lane. The sum of the signals for all viral mRNA bands in each lane was used to calculate the relative abundance of each individual mRNA. (*D*) Pulse-chase analysis of nsp1β expression. BHK-21 cells were transfected with RNA transcribed from WT or mutants of PRRSV full-length cDNA clones. At 16.5 h post transfection, protein synthesis was labeled for 15 min and chased for 1h. Cells were lysed and, following immunoprecipitation with mAb α-EU-nsp1β, the production and turn-over of nsp1β were analyzed using SDS-PAGE and autoradiography.
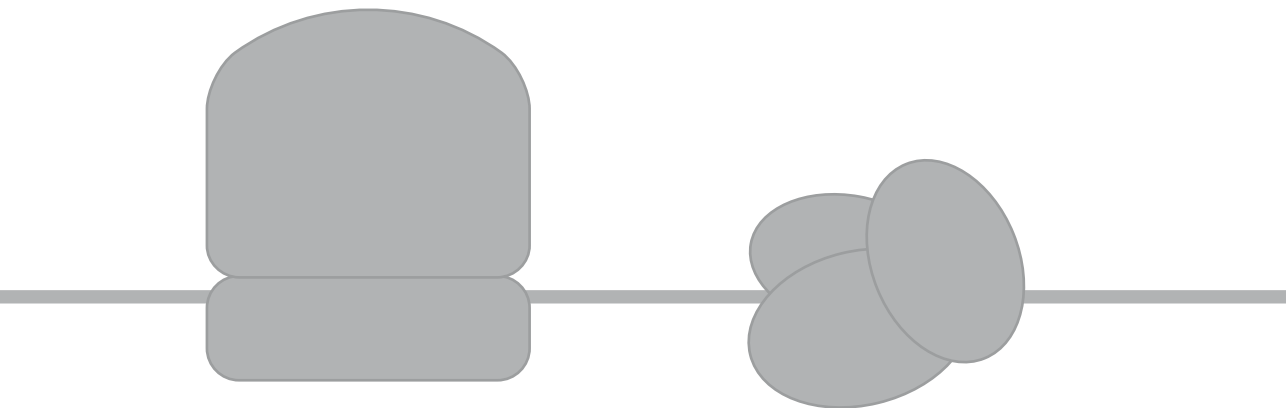
Chapter 8

Arterivirus non-structural protein 1β co-operates with cellular poly (C) binding proteins to transactivate –2/–1 programmed ribosomal frameshifting

Emmely E. Treffers[1#], Sawsan Napthine[2#], Susanne Bell[2], Ian Goodfellow[2], Ying Fang[3], Andrew E. Firth[2], Eric J. Snijder[1‡], Ian Brierley[2‡]

[1]Department of Medical Microbiology, Center for Infectious Diseases, Leiden University Medical Center, Leiden, The Netherlands, [2]Department of Pathology, University of Cambridge, Cambridge, U.K., [3]Department of Diagnostic Medicine and Pathobiology, College of Veterinary Medicine, Kansas State University, Manhattan, U.S.A.

[#]These authors contributed equally
[‡]These authors contributed equally

ABSTRACT

Porcine reproductive and respiratory syndrome virus (PRRSV) utilizes programmed ribosomal frameshifting (PRF) to direct efficient expression of a transframe protein (nsp2TF) from an alternative reading frame overlapping the viral replicase gene. This arterivirus frameshifting signal induces both −2 and −1 PRF and, unusually, lacks an obvious stimulatory RNA structure downstream of the shift site. The minimal RNA sequence required for frameshifting maps to a 34-nucleotide region that includes the slippery sequence (GG_GUU_UUU) and a downstream conserved C-rich motif (CCCAUCUCC). Unusually, efficient frameshifting is also dependent upon expression of a viral protein, non-structural protein (nsp) 1β, one of the 14 subunits produced from the PRRSV replicase polyproteins. Nsp1β is released by the combined action of two papain-like protease (PLP) domains, which reside in nsp1α and nsp1β and each cleave at their own C-terminus. Here we show that, in addition to this viral transactivator, frameshifting also requires the participation of cellular poly (C) binding proteins (PCBPs), which were previously identified as nsp1β interaction partners. *In vitro* translation assays demonstrated that both nsp1β and either PCBP1 or PCBP2 are required for efficient −2/−1 PRF. When cells were depleted for PCBP1 and PCBP2 by siRNA-mediated knockdown and subsequently transfected with a plasmid expressing nsp1β and nsp2, we observed a ~40% reduction of the expression of nsp2TF and nsp2N, the respective −2 and −1 PRF products. PCBP1 predominantly stimulates −2 PRF, while PCBP2 stimulates −1 PRF. We hypothesize that a complex of nsp1β and PCBP binds to the RNA signal downstream of the slippery sequence and here mimics the action of the more typical RNA pseudoknot stimulators of PRF. This unprecedented viral frameshift-stimulatory signal may provide new insights as to how the ribosomal elongation cycle can be modified by *trans*acting protein factors. Furthermore, it broadens the repertoire of activities associated with poly (C) binding proteins and prototypes a new class of arterivirus-host interactions.

INTRODUCTION

Programmed −1 ribosomal frameshifting (−1 PRF) is a commonly used translational control strategy in which mRNA signals induce ribosomes to change reading frame at high frequency, allowing the co-ordinated expression of two or more proteins from a single mRNA (reviewed in: [323, 336, 396, 433]). In −1 PRF, the ribosome slips backwards (in the 5' direction) by one nucleotide (nt) into an overlapping open reading frame (ORF) and continues translation, generating a fusion protein composed of the products of both upstream and downstream ORFs. First described as the mechanism by which the Gag-Pol polyprotein of the retrovirus Rous sarcoma virus (RSV) is expressed from overlapping *gag* and *pol* ORFs [434, 435], related −1 PRF signals have been documented in many other viruses of clinical, veterinary and agricultural importance [436-440]. Programmed ribosomal frameshifting has also been increasingly recognized in the expression of conventional cellular genes of both prokaryotes and eukaryotes as well as in other replicating elements, such as insertion sequences and transposons [441, 481-484].

Central to almost all examples of −1 PRF is the interaction of the ribosome with a stimulatory mRNA structure - a stem-loop (SL) or RNA pseudoknot (PK) - which acts to promote the −1 frameshift on a stretch of homopolymeric bases known as the slippery sequence. How these RNA structures act is incompletely understood, but accumulating evidence supports the view that by presenting an unusual topology [92, 336, 337, 396, 446], they confound an intrinsic unwinding activity of the ribosome with consequent effects on the elongation cycle and frame maintenance [444, 445, 485]. Indeed, kinetic analyses indicate that stimulatory RNAs can impair movements of the ribosomal small subunit (30S) head, delaying dissociation of EF-G, and the release of tRNA from the ribosome [338, 486, 487].

Recently, we identified a novel, highly efficient −2/−1 PRF event that functions without a recognizable stimulatory RNA structure [442, 488]. This signal operates during translation of the genome of porcine reproductive and respiratory syndrome virus (PRRSV), a member of the arterivirus family in the order *Nidovirales* [442, 488]. The PRRSV genome (Figure 1A) is a ~15 kb positive-sense RNA molecule, and its 5'-proximal ORF (the large replicase ORF1a) in fact harbors two consecutive PRF signals. A "canonical" −1 PRF signal, conserved in all nidoviruses, is located in the short overlap region of ORF1a and the downstream replicase ORF1b [361]. It facilitates expression of the ORF1b-encoded sequence as a fusion with the ORF1a product and defines the ratio of the synthesis of the two viral replicase polyproteins, pp1a and pp1ab, produced by nidovirus genome translation. The second signal, which stimulates both −2 PRF and −1 PRF, is only found in (most) arterivirus genomes, where it is located within the region of ORF1a that encodes a large, multifunctional replicase subunit, nonstructural protein 2 (nsp2). Here, about 20% of ribosomes translating nsp2 frameshift into the −2 reading frame to generate a
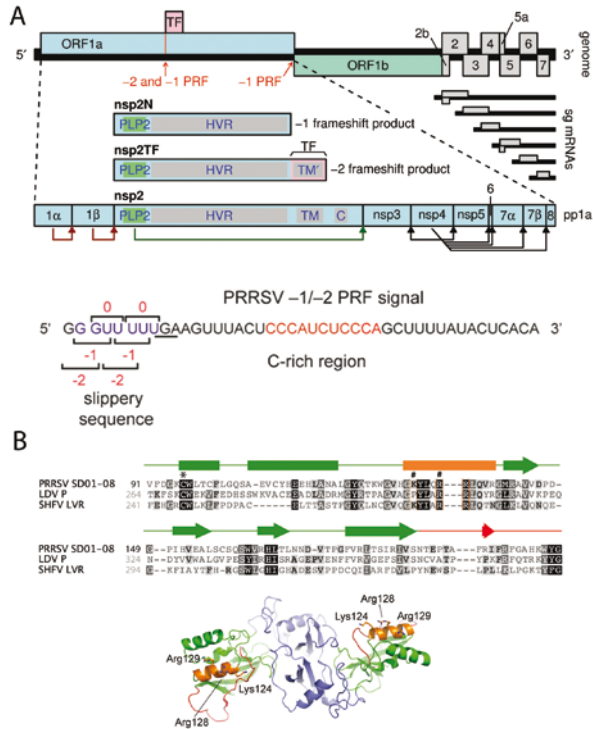
Figure 1: (*A*) Overview of the PRRSV genome organization and localization of ribosomal frameshifting signals. The long 5′ ORFs 1a and 1b encode nonstructural polyproteins, and at least eight shorter 3′ ORFs (2a-7) encode structural proteins. The 3′ ORFs are translated from a nested set of subgenomic mRNAs, two of which are bicistronic. ORF1a and ORF1b are translated from the genomic RNA, with translation of ORF1b depending on −1 PRF at the end of ORF1a. The TF ORF overlaps the central ORF1a region in the −2 reading frame and is accessed via −2 PRF. A −1 frameshift at the same site generates the nsp2N product. The vertical red line indicates the location of the RG_GUU_UUU shift site (R = A or G, in different arteriviruses). Domains in nsp2/nsp2TF: C, Cys-rich domain HVR, hypervariable region; PLP2, papain-like proteinase; TM/ TM′, (putative) transmembrane domains. Below the genome organization, the sequence of the SD01-08 RNA in the region of the −2/−1 PRF signal is shown, with the slippery sequence (red) and C-rich motif (blue) highlighted. The −1 reading frame stop codon is underlined and codons for each of the reading frames are indicated. Figure reproduced from [488] (*B*) PRRSV nsp1β sequence and structure. Amino acid sequence alignment of the PLP1β domains from selected arterivirus nsp1β proteins. Secondary structure elements (based on the published crystal structure from type 2 PRRSV isolate XH-GD) [453] are shown above the alignment and are color matched to the nsp1β structure shown below. Conserved basic residues in PLP1β helix α4 are boxed in orange. #, residues mutated in RBD- mutant. *, residue mutated in PR- mutant. The PRRSV sequence is numbered (black) from the nsp1α/nsp1β cleavage site, whereas the two other sequences are numbered (gray) starting from the N terminus of the pp1a polyprotein. The names of specific isolates used are indicated. GenBank accession nos. of sequences used are as follows: DQ489311 (PRRSV SD01-08), NC_001639 (LDV P), NC_003092 (SHFV LVR). In the cartoon representation of the nsp1β dimer structure, the N-terminal domains are colored purple, whereas the PLP1β domains and the C-terminal extensions (leading up to the nsp1β/nsp2 site cleaved byPLP1β) are colored green and red, respectively. Helix α4 of PLP1β, containing the conserved GKYLQRRLQ motif, is colored orange with basic residues represented as sticks. Figure modified from [488].

transframe fusion protein (nsp2TF) comprising the N-terminal two-thirds of nsp2 and the product encoded by a conserved alternative ORF (transframe; TF) in the −2 reading frame. An estimated further 7% of ribosomes shift into the −1 reading frame where they immediately encounter a stop codon resulting in the synthesis of a truncated version of nsp2 termed nsp2N [442]. As depicted in Fig. 1A, the RNA downstream of the slippery sequence (GG_GUU_UUU) used for −2/−1 PRF in PRRSV does not harbor an obvious higher-order structure compatible with canonical RNA-structure-stimulated PRF. However, mutations within a conserved CCCAUCUCC motif located 11 nt downstream of the shift site reduce or inhibit frameshifting, consistent with the presence of a 3′ stimulatory element of some form [442]. A further novelty of the PRRSV −2/−1 PRF signal is a requirement for the presence of viral protein, nsp1β, which functions as a transactivator of both −2 and −1 PRF [488]. The 205-amino acid nsp1β contains a papain-like protease domain (PLP1β) that cleaves the nsp1β/nsp2 junction in the arterivirus nonstructural polyprotein [363]. How nsp1β acts to stimulate frameshifting is unclear, although basic residues in a highly conserved putative RNA-binding motif (GKYLQRRLQ) (Figure 1B), integrated into the structure of nsp1β's papain-like autoproteinase, were found to be critical for the stimulation of frameshifting [488].

In PRRSV-infected cells, in addition to its role in PRF control and the PLP1β–mediated cleavage of the nsp1β-nsp2 junction, nsp1β has also been implicated in arteriviral innate immune evasion (reviewed by [361, 489]). Part of nsp1β can be detected in the nucleus of infected cells and ectopic expression of nsp1β can suppress the induction of innate immune responses [452, 490, 491], possibly by inhibiting downstream interferon-induced signaling pathways [452, 478]. The highly conserved nsp1β motif that was implicated in the control of PRF was also linked to innate immune evasion [447], as mutagenesis of its conserved basic residues yielded viable recombinant viruses inducing increased expression of IFN-α, IFN-β, and ISG15. Whether this effect should be attributed directly to nsp1β's role in mediating the expression of nsp2TF/nsp2N or may (in part) be linked to the protein's other functions remains to be investigated in more detail.

We previously determined the −2 and −1 PRF frameshifting efficiencies in different cell culture settings. In PRRSV SD01-08-infected MARC-145 cells the −2 PRF efficiency was around 20%, while in RK13 cells transiently expressing PRRSV ORF1a a much higher −2 PRF efficiency of about 50% was observed [442]. These large differences led us to hypothesize that frameshifting may be modulated by host factors expressed at different levels in these cells. Given its involvement in PRF outlined above, cellular proteins interacting with nsp1β were obvious candidates for such a role. During a previous study that used GST-tagged nsp1β as bait and aimed to discover cellular factors involved in PRRSV replication, a number of nsp1β interaction partners was identified, including the poly(C) binding proteins 1 and 2 (PCBP1 and PCBP2) [463]. PCBP1 and PCBP2 belong to a family of proteins that is characterized by three nucleic acid-binding hnRNP K

homology (KH) domains. The KH1 and KH2 domains are grouped near the N-terminus and KH3 is located at the C-terminus, separated from KH1 and KH2 by a sequence of variable length [492]. The other members of this family are PCBP3, PCBP4 and hnRNP K [492]. Yeast two-hybrid binding assays revealed that the interaction between nsp1β and PCBP2 requires minimally the PLP1β and C-terminal extension (CTE) domains of nsp1β and the KH2 domain of PCBP2 [493]. Interestingly, the conserved C-rich motif required for −2/−1 PRF in PRRSV shares some similarities with previously established PCBP-binding consensus sequences, although these sequences commonly have three (or more) C-triplets, each potentially binding one KH domain. Thus, PCBPs could bind to the PRRSV mRNA either directly or indirectly, by virtue of an association with nsp1β. In PRRSV-infected cells, overexpressed recombinant PCBP1 and PCBP2 were reported to partially localize to the region of the cell in which viral RNA synthesis is presumed to occur [463]. Endogenous PCBP2 was shown to be predominantly present in the nucleus of mock-infected cells, but to translocate to the cytoplasm in PRRSV-infected cells, after which it partially co-localized with nsp1β [493]. Furthermore, recombinant PCBPs were found to bind RNA transcripts representing the 5' untranslated region (5'-UTR) of the PRRSV genome [463]. Finally, siRNA-mediated knockdown of the two PCBPs moderately inhibited PRRSV replication. In these PCBP-depleted cells, viral genome translation was claimed to be not affected, but this was concluded on the basis of measuring nsp1α and nsp1β expression levels only (the occurrence of −2 and −1 PRF in the nsp2-coding region was unknown at the time; [463]). In another study in which PCBP2 was depleted, both viral replication and nsp1β expression were decreased during PRRSV infection [493]. Beura *et al.* postulated that PCBP1 and PCBP2 are important host cell factors for PRRSV RNA synthesis [463], but our subsequent discovery of nsp1β-mediated PRF, at a signal containing a conserved poly(C)-containing RNA motif, prompted us to revisit this hypothesis. If indeed PCBP1 and PCBP2 are interaction partners of nsp1β and have affinity for poly(C)-containing RNA sequences, their presence or absence may (also) affect the occurrence of −2/−1 PRF. As previously described, the inactivation of −2/−1 PRF in itself suffices to severely cripple PRRSV replication [442, 488], a defect that should also be reflected in viral RNA accumulation levels, but is in fact determined at the translational level.

In this paper, through reconstitution of the PRRSV −2/−1 PRF mechanism *in vitro* and studying the PRF-specific effects of PCBP knockdown in an expression system, we have established that efficient frameshifting indeed depends on the presence of cellular PCBPs, and that their presence is essentially equally important as that of the viral transactivator nsp1β. Through *in vitro* translation assays, we demonstrate that the combined presence of nsp1β and PCBP1 or PCBP2 mimics the action of the more typical RNA pseudoknot stimulators of programmed frameshifting. PCBP depletion in a cell culture-based expression system resulted in a significant reduction of nsp2TF and nsp2N expression.

This unprecedented arteriviral frameshift-stimulatory signal may provide new insights as to how the ribosomal elongation cycle can be modified by transacting protein factors. Furthermore, our study confirms and extends the important role of PCBPs as pro-viral host factors during PRRSV replication and prototypes a new class of arterivirus-host interactions.

MATERIALS AND METHODS

Viruses and cells

MARC-145 and RK-13 cells were cultured in DMEM (Lonza) containing 8% fetal calf serum (FCS; PAA), 100 IU/ml penicillin and 100 µg/ml streptomycin at 37 °C and 5% CO_2. Recombinant vaccinia virus vTF7-3 [429] was propagated in RK-13 cells.

Protein expression and purification

Full-length, N-terminally hexa-histidine tagged nsp1β, PCBP1 and PCBP2 were expressed in *Escherichia coli* Rosetta2(DE3)pLysS (Novagen). Bacteria were grown in 2× TY medium to an A^{600} of 0.8 at 37 °C and cooled to 22 °C, and protein expression induced by the addition of 0.2 mM isopropyl β-d-thiogalactopyranoside. After 16 h, cells were harvested by centrifugation at 5000 × g for 15 min at 4 °C, and the pellet was stored at −20 °C until required.

  Cells were thawed and resuspended in 20 mM Tris, 500 mM NaCl, 30 mM imidazole, 1.4 mM β-mercaptoethanol, 0.05% Tween 20, pH 7.5, supplemented with 400 units of bovine DNase I (Sigma-Aldrich) and 200 µl of EDTA-free protease inhibitor mixture (Sigma-Aldrich) before lysis at 165.5 MPa using a TS series cell disruptor (Constant Systems) and centrifugation at 40,000 × g for 30 min at 4 °C. Cleared lysate was incubated with Ni^{2+}-NTA-agarose (Qiagen) for 1 h at 4 °C, the beads were washed, and the bound protein eluted in 20 mM Tris, 500 mM NaCl, 250 mM imidazole, pH 7.5, before dialysis against 20 mM Tris, pH 7.6, 200 mM NaCl, 2 mM DTT. Purified proteins were concentrated, snap-frozen in liquid nitrogen, and stored at −80 °C until required.

Frameshift reporter plasmids and *in vitro* translation

Dual luciferase reporter plasmids were prepared by inserting a 79-nt sequence from PRRSV isolate SD01-08 containing the GG_GUU_UUU shift site, 5 upstream nucleotides and 66 downstream nucleotides between the *Renilla* and firefly luciferase genes in pDluc [448, 449] to create pDluc-PRRSV/wt so that −2PRF is required for firefly luciferase expression. To create pDuc-PRRSV/stop, the stop codon (UGA) in the −1 frame was changed to UUA, extending the −1 reading frame by a further 56 codons, whilst terminating the Rluc gene slightly earlier. *In vitro* transcribed RNA from the pDluc constructs was translated in

rabbit reticulocyte lysates (RRL) and wheat germ extracts (WG) as previously described [407] with or without recombinant nsp1β, PCBP1 and PCBP2 present. Translation products were separated by SDS-PAGE and visualized by WB using an antibody recognizing *Renilla* luciferase which is common to all pDluc translation products.

siRNA-mediated knockdown

To determine frameshifting efficiencies after siRNA-mediated knockdown of PCBP expression, 6×10^4 MARC-145 or 4.8×10^4 RK-13 cells were seeded per well in 12-well clusters in DMEM containing 8% FCS. MARC-145 cultures were transfected with siGENOME Human siRNA SMARTpools targeting PCBP1, PCBP2, PCBP3, PCBP4, or hnRNPK (final concentration 10 nM) using 2 ul of Dharmafect1 lipofection reagent (Dharmacon) per well. RK-13 cultures were transfected with siRNA pools targeting PCBP1 or PCBP2 (final concentration 25 nM) using 3 ul Lipofectamine 2000 (LifeTechnologies) per well. A non-targeting pool (NTP) of "scrambled" siRNAs (Dharmacon) was used as a negative control. At 24 h post transfection (p.t.), the transfection medium was replaced with DMEM containing 8% FCS. Possible cytotoxic effects of siRNA transfection were monitored at 48 h p.t., using the CellTiter 96® AQueous Non-Radioactive Cell Proliferation Assay (Promega). After 120 min, the reaction was stopped by the addition of 25 µl of 10% SDS and absorbance at 490 nm (A490) was measured using a 96-well plate reader (Berthold). At 48 h p.t., the cells were infected with a T7 RNA polymerase-expressing recombinant vaccinia virus [429] and 1 h later transfected with plasmid pL-nsp1β-2 [363], which encodes a self-cleaving nsp1β-nsp2 polyprotein under the control of a T7 promoter. Four hours later, the cells were starved for 30 min in methionine- and cysteine-free DMEM (Gibco) containing 2% FCS prior to a 45- (MARC-145) or 30-min (RK-13) metabolic labeling with 500 µCi/mL of a [^{35}S]Met/Cys mixture (EXPRE^{35}S^{35}S Protein Labeling Mix; Perkin-Elmer). Subsequently, cells were washed twice with PBS and harvested in lysis buffer (20 mM Tris, pH 7.6, 150 mM NaCl, 1% v/v NP-40, 0.1% DOC, 0.1% SDS and Complete protease inhibitor (Roche)).

The human sequences targeted by the siRNA pools were compared with the sequences of the African green monkey (Chlorocebus sabaeus) genes encoding PCBP1 (Genbank accession XM_007970341), PCBP2 (Genbank accession XM_008003436), PCBP3 (Genbank accession XM_007970531), PCBP4 (Genbank accession XM_007984435) and hnRNPK (Genbank accession XM_007969613) and the rabbit (Oryctolagus cuniculus) PCBP1 and PCBP2 genes (Genbank accession NM_001082124 and XM_002711018, respectively).

Immunoassays

Following siRNA pool transfection, MARC-145 and RK-13 cells were harvested at 48 h p.t. or 72 h p.t./24 h p.i. by first washing with PBS and then lysing in 4× Laemmli sample buffer (100 mM Tris-HCl, pH 6.8, 40% glycerol, 8% SDS, 40 mM DTT, 0,04 mg/

ml bromophenol blue). Proteins were visualized by Western blot analysis as described previously [208] using primary antibodies mAb-nsp2 Eu58-46, mAb-nsp1beta Eu22-28 [403], rabbit polyclonal anti-hnRNP E1 (C-terminal) (PCBP1) (Sigma-Aldrich), mouse monoclonal anti-PCBP2 (M07) (Abgent), rabbit polyclonal anti-human MGC10 (PCBP4) (PromoKine), mouse monoclonal anti-hnRNPK (Abcam), or mouse monoclonal antibody H68.4 against the transferrin receptor (Invitrogen) diluted in PBST containing 1% casein. Biotin-conjugated swine-a-rabbit (DAKO) or goat-a-mouse (DAKO), and Cy3-conjugated mouse-a-biotin (Jackson, Pennsylvania, USA) diluted in PBST containing 0.5% casein were used for fluorescent detection with a Typhoon-9410 imager (GE Healthcare, UK).

Lysates from [^{35}S]Met/Cys-labelled cells were used to immunoprecipitate nsp2, nsp2TF and nsp2N using mAb-nsp2 Eu58-46 [403], which recognizes the common N-terminal domain of the three products. Proteins were separated on a 6% SDS-PAGE gel and imaged as described previously [403]. Band intensities of nsp2, nsp2′, nsp2TF, nsp2TF′, nsp2N and nsp2N′ (nsp2′, nsp2TF′ and nsp2N′ are faster migrating forms of the three products; [442]) were quantified using ImageQuant TL (GE Healthcare) and normalized using the Met+Cys content of the respective products, while assuming that [^{35}S]Met and [^{35}S]Cys are incorporated with an efficiency ratio of 73:22 (the Met:Cys ratio in the mixture according to the manufacturer's documentation). Using these values, −2 PRF efficiencies were calculated as (nsp2TF + nsp2TF′)/(nsp2 + nsp2′ + nsp2TF + nsp2TF′ + nsp2N + nsp2N′) and −1 PRF efficiencies were calculated as (nsp2N + nsp2N′)/(nsp2 + nsp2′ + nsp2TF + nsp2TF′ + nsp2N + nsp2N′). Quantification was performed in triplicate.

RESULTS

Reconstitution of the PRRSV −2/−1 PRF signal *in vitro*

Transactivation of the PRRSV −2/−1 PRF signal by nsp1β was previously demonstrated by co-expression of nsp1β and nsp2 in cultured cells and by site-directed mutagenesis of the viral genome to inactivate the critical RNA and protein signals described above [488]. To study the phenomenon *in vitro*, a 79-bp cDNA fragment encompassing the slippery sequence and C-rich region was subcloned between the *Renilla* and *firefly* genes of the frameshift-reporter plasmid pDluc [448] such that expression of the downstream cistron (*fluc*) was dependent upon a −2 PRF within the inserted PRRSV sequences (Figure 2A). Upon *in vitro* translation in the rabbit reticulocyte system (RRL), mRNAs transcribed from pDluc/PRRSV/wt specified the synthesis of only the product of the 5′ cistron of the reporter mRNA (Rluc) and no frameshifting was evident (Figure 2B). However, in translation reactions supplemented with recombinant, purified His$_6$-tagged nsp1β, two additional bands were observed, the most abundant corresponding to the product of −2 ribosomal frameshifting. The second band migrated more rapidly than that of the

Rluc product (generated by ribosomes that do not frameshift*)* and we surmised that this protein corresponded to the product of a −1 PRF, as a −1 frame stop codon is present immediately downstream of the slippery sequence (underlined in Fig. 1A; see also [488]. To confirm this, the stop codon (UGA) in the −1 frame was changed to UUA, extending the −1 reading frame by a further 56 codons, whilst terminating the *Rluc* gene slightly earlier (Figure 2A). As anticipated, this point mutation in pDLuc/PRRSV/stop increased the size of the putative −1 PRF product, and slightly reduced the size of the Rluc protein (stop) (Figure 2B). These experiments indicate that the 79-nt PRRSV sequence cloned into pDluc likely contains all of the *cis*-acting elements required for −2/−1 frameshift-ing and that supplementation of the RRL translations with nsp1β alone is sufficient to transactivate frameshifting.

The previously reported crystal structure of nsp1β reveals an overall elliptical structure consisting of six α-helices and seven β-strands [453]. The protein is composed of two
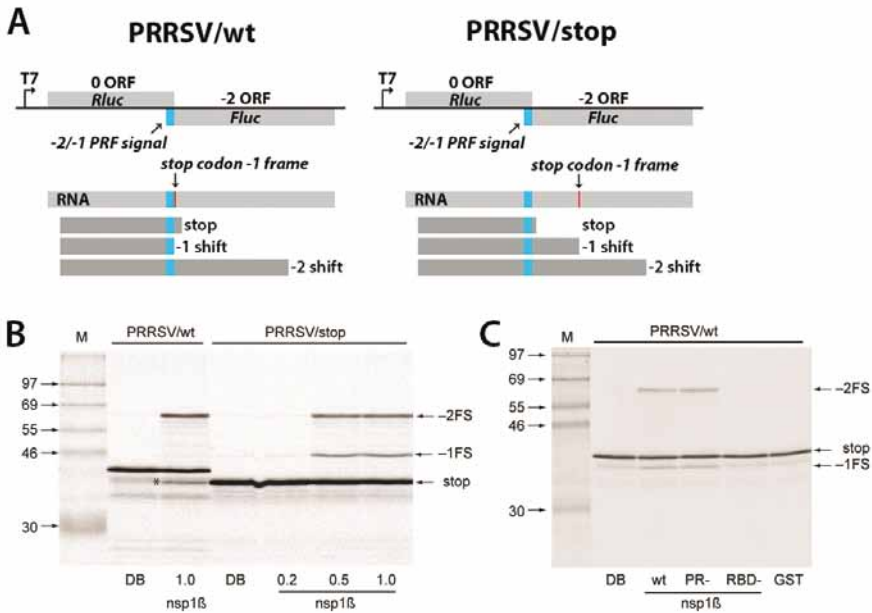

Figure 2: Nsp1β is required for reconstitution of −2/−1 PRF *in vitro* in RRL (*A*) Schematic representa-tion of the pDluc dual luciferase constructs used. **For** PRRSV/wt the GG_GUU_UUU shift site, 5 upstream nucleotides and 66 downstream nucleotides (79 nt in total) from the SD01-08 sequence were inserted be-tween the *Renilla* and firefly luciferase genes such that −2PRF is required for firefly luciferase expression. For PRRSV/stop the stopcodon in the −1 frame immediately downstream from the slippery sequence was removed to extend the −1 PRF reading frame and the *Rluc* gene (zero frame) terminates earlier than in the WT sequence. (*B*) *In vitro* transcribed PRRSV/wt and PRRSV/stop RNA were translated *in vitro* in RRL with and without supplementation of purified His$_6$-tagged nsp1β. DB; dialysis buffer. (*C*) The PRRSV/wt RNA was translated in RRL in the presence of WT nsp1β, a putative RNA binding mutant (RBD-; K124A/R128A) or a protease-defective mutant (PR-; C96S). GST: his-tagged GST negative control.

major domains, a 48-amino acid N-terminal domain (NTD), which adopts an architecture similar to that of several metal ion-dependent nucleases, and a C-terminal papain-like cysteine protease (PLP1β) domain [453]. Within the latter domain, we previously identified a conserved sequence motif $GK_{124}YLQR_{128}RLQ$ as a potential RNA binding domain (RBD) (Figure 1B). This sequence forms one of three α helices in the region between active site C_{96} and H_{165} residues of PLP1β [488]. To investigate the importance of protease or RNA binding activities to PRF transactivation, amino acid substitutions were introduced into the protease active site (C96S; mutant PR-) or within the putative nsp1β RBD (K124A/R128A; mutant RBD-, previously described as mutant 1βKO [488]) of the recombinant His_6-tagged nsp1β used to transactivate frameshifting in vitro (Figure 1B). These variants were expressed and tested in the RRL-based PRF assay (Figure 2C). In this experiment, the protease-defective variant retained full activity, ruling out the involvement of PLP1β's protease activity in the stimulation of PRF by nsp1β. The RBD- mutation, on the other hand, inactivated nsp1β's potential to induce frameshifting, supporting a role for RNA binding by nsp1β.

A requirement for poly (C) binding proteins in PRRSV −2/−1 PRF *in vitro*

To explore the potential role of PCBPs in PRRSV PRF, we translated the pDluc/PRRSV/wt mRNA in RRL reactions supplemented with nsp1β, His_6-tagged PCBP2 or both proteins (Figure 3A). We found that PCBP2 alone did not stimulate PRF, but when added together with nsp1β, a substantial increase in the synthesis of the −1 PRF product was observed. To account for this observation, we reasoned that if the "active" transacting stimulator of PRRSV PRF is indeed a complex of nsp1β and PCBP, then the RRL system must already contain an abundant form of PCBP, but one which would, in complex with nsp1β, preferentially lead to −2 PRF. In this scenario, supplementation with PCBP2 may have generated nsp1β-PCBP2 complexes that could preferentially promote −1 PRF. We therefore translated the PRRSV frameshift reporter mRNA in the wheat germ (WG) *in vitro* translation system in the hope that this lysate would contain fewer endogenous PCBPs, or PCBPs of sufficient evolutionary diversity to preclude any interactions with nsp1β. Consistent with this expectation, ribosomal frameshifting at the PRRSV signal in WG was found to be completely dependent upon the simultaneous addition of both nsp1β and PCBP2, with neither protein alone having any frameshift-stimulatory activity in this system (Figure 3B). In further support of the hypothesis above, supplementation of WG translations with PCBP2 led preferentially to a −1 PRF, whereas with PCPB1, the −2 PRF was most evident (Figure 4), suggesting that the abundant form in RRL is PCBP1.
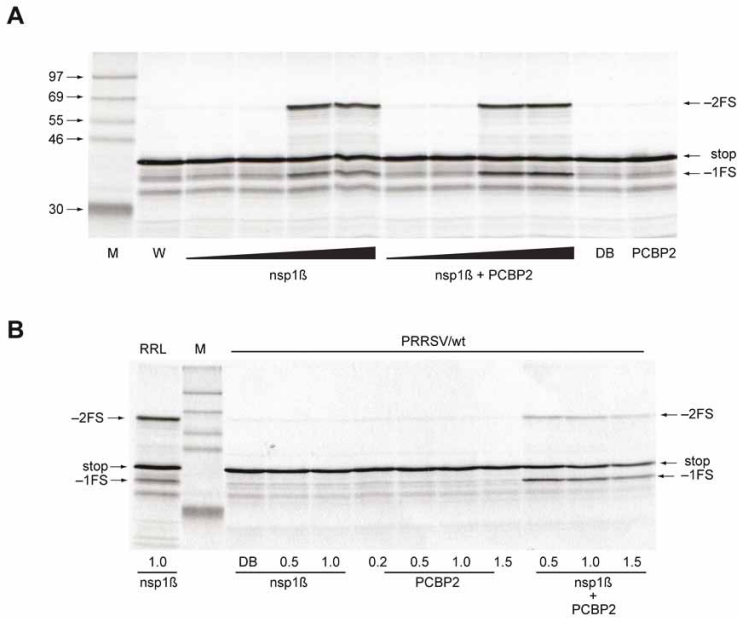
Figure 3: PCBPs are required for reconstitution of −2/−1 PRF *in vitro* in WG. (*A*) *In vitro* translation reactions with RRL and PRRSV/wt RNA were supplemented with nsp1β, his-tagged PCBP or both proteins. (*B*) *In vitro* translation reactions with WG and PRRSV/wt RNA were supplemented with nsp1β, his-tagged PCBP2 or both proteins.
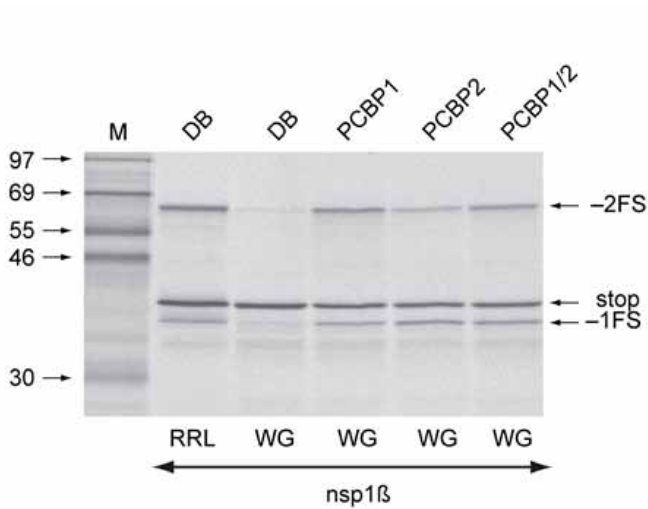

Figure 4: PCBP1 and PCBP2 can reconstitute −2/−1 PRF in WG. *In vitro* translation reactions with RRL or WG and *in vitro* transcribed PRRSV/wt RNA were supplemented with nsp1β and his-tagged PCBP1 or his-tagged PCBP2 or both PCBPs.

Knockdown of poly (C) binding proteins reduces frameshifting efficiency in cultured cells

To explore the potential involvement of PCBPs on −2/−1PRF in living cells, siRNA-mediated knockdown of PCBP1, PCBP2, PCBP3, PCBP4, and hnRNPK in MARC-145 cells was combined with transient expression of nsp1β and nsp2 using the recombinant vaccinia virus/T7 polymerase expression system (Figure 5ABC). For PCBP1 and PCBP2 the findings from MARC-145 cells were further confirmed in RK-13 cells (Figure 5DEF).

The siRNA pools that were used target human sequences, but most of these pools are expected to also mediate knockdown in MARC-145 (African green monkey) and RK-13 (rabbit) cells. For example, the four sequences targeted by the siRNA pool against PCBP1 are fully conserved in African green monkey, and only one of these has a 1-nt mismatch with the rabbit PCBP1 sequence. Likewise, only one of the four sequences targeted by the siRNA pool against PCBP2 has a 1-nt mismatch in African green monkey, while two of these sequences have a 1-nt mismatch in rabbit. The four sequences targeted by the siRNA pools for PCBP3 and hnRNPK are fully conserved in African green monkey, but one of the sequences targeted by the siRNA pool against PCBP4 contains a 1-nt mismatch and one sequence contains three 1-nt mismatches. In MARC-145 cells, knockdown of PCBP1, PCBP2, PCBP4 and hnNRPK was successful, as established by Western blot analysis (Figure 5A). Unfortunately, a suitable antibody to monitor PCBP3 knockdown was not available. Also in RK-13 cells knockdown of PCBP1 and PCBP2 was successful (Figure 5D). When assessed via an MTS-based cell viability assay, knockdown of none of these targets was toxic to uninfected cells. However, knockdown of PCBP3 in MARC-145 cells resulted in accelerated cell death during infection with recombinant vaccinia virus vT7-3, resulting in lower band intensities for the three nsp2 products upon immunoprecipitation compared to the other conditions (Figure 5B).

In line with the observations made during the *in vitro* translation experiments outlined above (Figure 4), PCBP1 knockdown mostly reduced the efficiency of the −2 shift, while PCBP2 knockdown mostly impaired the −1 shift (Figure 5C and F). Whereas knockdown of a single PCBP reduced −1 or −2 PRF by about 40%, simultaneous knockdown of PCBP1 and PCBP2 reduced the total amount of −2/−1 PRF by ~50% in MARC-145 cells (Figure 5C) and ~70% in RK13 cells (Figure 5F). In multiple independent experiments, PCBP1 depletion, while decreasing −2 PRF, at the same time increased the −1 PRF efficiency in RK-13 cells, but not in MARC-145 cells. Interestingly, PCBP2 knockdown also induced knockdown of PCBP1 in RK-13 cells (Fig. 5D) and as a result the −2 shift was also reduced in these cells when only siRNAs targeting PCBP2 were used (Figure 5F). This might be caused by siRNA off-target effects because the PCBP1 and PCBP2 sequences are very similar and one of the siRNAs targeting PCBP2 has only a single mismatch with rabbit PCBP1, while there are two mismatches with African green monkey PCBP1. The mismatch is located at position 14 of the siRNA and siRNAs with a mismatch at this posi-
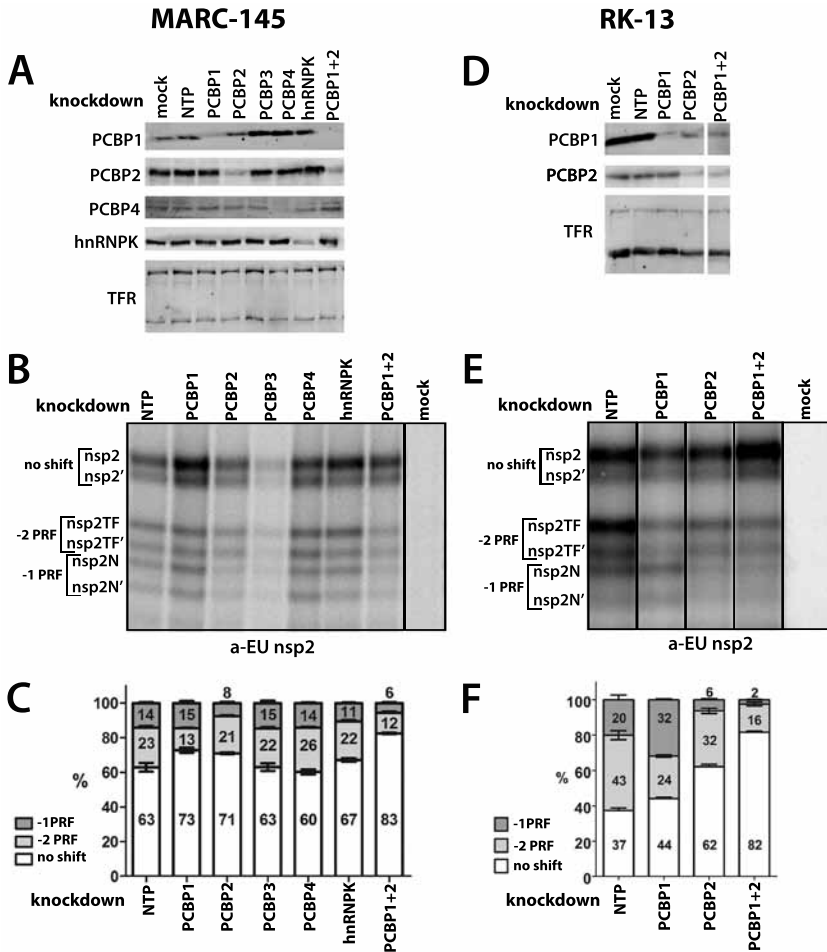
Figure 5: siRNA knockdown of PCBP1 and PCBP2 decreases −2/−1 frameshifting efficiency in cell culture. (*A & D*) MARC-145 (*A*) and RK-13 cells (*D*) were transfected with siRNA pools targeting PCBP1, PCBP2, PCBP3, PCBP4, hnRNPK. Knockdown of proteins targeted by siRNAs was assessed using WB analysis. A suitable antibody for PCBP3 was not available. (*B & E*) MARC-145 (*B*) and RK-13 cells (*E*) were transfected with siRNA pools targeting PCBP1, PCBP2, PCBP3, PCBP4, hnRNPK. The recombinant vaccinia virus–T7 RNA polymerase expression system was used to express pLnsp1β-2. After metabolic labeling, expression products were immunoprecipitated with mAb α-EU-nsp2, this antibody recognizes the common N-terminal domain of nsp2, nsp2TF, and nsp2N. Immunoprecipitated proteins were separated by SDS/PAGE and visualized by autoradiography. (*C & F*) Protein bands were quantified (triplicates) using ImageQuant TL software for MARC-145 cells (*C*) and RK-13 cells (*F*). The total intensity for all protein bands (nsp2, nsp2′, nsp2TF, nsp2TF′, nsp2N and nsp2N′) was set as 100%. Efficiencies were corrected for the methionine and cysteine content of each protein product. nsp2′, nsp2TF′ and nsp2N′ are precursors.

tion are often still able to induce knockdown [494]. Knockdown of PCBP3, PCBP4 and hnRNPK in MARC-145 cells increased PCBP1 expression, most likely due to some kind of a compensatory mechanism at the level of genome expression (Figure 5A). Knockdown of PCBP3, PCBP4 and hnRNPK had little or no effect on frameshifting (Figure 5BC), but such effects may have been masked by the increased PCBP1 levels in these cells (Figure 5A). If this is the case, these proteins are more likely to affect −2 PRF than −1 PRF.

DISCUSSION

In this paper we describe the discovery that −2/−1 PRF in arteriviruses does not only depend on the presence of viral protein nsp1β, but both frameshifts also require PCBPs as additional, host cell-encoded protein transactivators.

RNA-protein complex

In mammalian cells, two PCBP subsets have been described, hnRNPs K/J [495] and the αCP proteins αCP1 and αCP2, commonly referred to as PCBP1 and PCBP2 [496, 497]. The latter group also includes the more recently described isoforms PCBP3 and PCBP4 [498]. Being members of the KH domain superfamily of nucleic acid binding proteins, PCBPs have been implicated in a wide spectrum of biological activities, including the regulation of RNA splicing, the stabilization of cellular and viral mRNAs, transcriptional activation and inhibition, and translational silencing and enhancement (reviewed in: [492, 499]. The PCBPs are ubiquitously expressed across many tissues, but the level of expression and isoform(s) predominantly expressed varies per cell type [496, 500-502]. This probably explains why we observed different frameshifting efficiencies when expressing pLnsp1β-2 in MARC-145 and RK-13 cells. Likewise, the levels of PCBP1 and PCBP2 in a PRRSV-infected cell will most likely determine the relative efficiencies of the two PRF events, which must to a certain extent be connected because they are mutually exclusive translational events. In the siRNA-based knockdown experiments, PCBP1 depletion mainly reduced −2 PRF, while PCBP2 knockdown mainly affected −1 PRF. In the *in vitro* translation assays PCBP1 and PCBP2 were both able to induce −2 as well as −1 PRF, although PCBP1 was concluded to be more efficient at promoting −2 PRF, whereas PCBP2 more efficiently induced −1 PRF. In an artificial experimental system in which alternative shifts at a single slippery sequence could be observed, the spacing between slippery sequence and downstream stimulatory RNA structure was found to determine whether −1 or −2 PRF was more prominent. The optimal distance for −2 PRF was 1-2 nt shorter than the optimal distance for −1 PRF [407]. This suggests that the RNA-protein complexes containing PCBP1 might have a slightly different orientation or conformation than those containing PCBP2, resulting in more frequent induction of −2 PRF. For −1

PRF it has been shown that roadblocks, such as RNA pseudoknots downstream of the slippery site, stall ribosomes in a metastable conformational state. Slippage into the −1 frame accelerates completion of translocation [337, 338, 486, 487, 503]. The complexes containing PCBP1 might induce more tension on a paused ribosome, thus favoring slippage into the −2 rather than the −1 frame to allow completion of translocation. The exact composition of the RNA-protein complex will likely have to be determined by structural biology techniques to understand exactly how the proteins interact with each other and the viral RNA and whether the complexes formed by PCBP1 have a different conformation than the ones formed by PCBP2. Another explanation for the different effects that PCBP1 and PCBP2 have on inducing PRF would be that the affinity of the KH domains of PCBP1 for the C-rich motif is different than the affinity of those present in PCBP2, which could also result in variable degrees of tension on the ribosome. We postulate that PCBPs interact with the C-rich motif through KH1 and KH3 and that KH2 interacts with nsp1β since it has previously been shown that this domain is required for nsp1β binding [504]. For PCBP1 it has been determined that for optimal binding to its KH1 domain, cytosine is preferred in all four positions in the oligonucleotide binding cleft and that a C-tetrad binds KH1 with 10 times higher affinity than a C-triplet [505]. The PRRSV C-rich motif, the evolution of which may have been restricted by its presence within overlapping ORFs, only contains two sets of three consecutive cytosines and this probably does not result in a high enough affinity of PCBPs for the RNA to induce frameshifting on their own, but requires the additional interaction with nsp1β.

Potential interaction of other PCBP family members

The involvement of PCBP1 and PCBP2 as protein transactivators of PRF is clear, but for the other members of the PCBP family their involvement is still uncertain. The study of separate PCBPs in living cells was complicated by an apparently compensatory mechanism that resulted in an increase in PCBP1 levels when PCBP3, PCBP4 or hnRNPK were knocked down in MARC-145 cells. It seems likely that PCBP3, PCBP4 and hnRNPK are also capable of stimulating frameshifting, but additional studies are required to determine their potential role in more detail. Knockdown of these targets in combination with knockdown of PCBP1 could perhaps clarify their involvement. For PCBP3 we will, however, need to use another system since knockdown of this factor appeared to influence vaccinia virus replication resulting in accelerated cell death. Supplementation of the WG *in vitro* translation assay with these proteins could also help to assess their (potential) involvement.

The involvement of PCBPs in frameshifting in infected cells

Experiments to determine the impact of PCBP knockdown on −2/−1 PRF in PRRSV-infected cells are currently ongoing. However, combining efficient siRNA-mediated

knockdown with a single-cycle PRRSV infection is a technical challenge. During the primary phase of PRRSV infection in MARC-145 cells only a subset of cells is permissive to infection, even when an amount of virus is used that should result in a high MOI. The majority of the cells become infected in subsequent rounds of infection, by cell-to-cell transmission to clusters of neighboring cells [506]. We hope to be able to overcome this issue by efficient transfection of *in vitro* transcribed, full-length viral RNA into siRNA-treated cells. An additional complicating factor is the reported possible interaction of PCBP1 and PCBP2 with sequences in the 5'-UTR of the PRRSV genome, which might influence viral replication and transcription, even though this interaction did not appear to be very strong and remains to be corroborated in infected cells [463]. Beura *et al*, hypothesized that PCBP1 and PCBP2 play a role in regulating PRRSV RNA synthesis because knockdown of both proteins resulted in a reduction of genomic and subgenomic RNA synthesis [463]. However, this effect might also result from the reduced expression of nsp2TF and nsp2N when PCBP1 and PCBP2 are depleted. We previously showed that PRRSV mutants lacking a functional −2/−1 PRF mechanism are seriously crippled in their replication [442, 488]. The siRNA-mediated depletion of PCBP1 and PCBP2 was reported to result in a ~ 0.5-1 log reduction in viral progeny titers [463, 493] which, for example, is slightly less than what was observed with a mutant virus (KO2) in which −2/−1 PRF is completely knocked out by mutating the RNA signals involved (1.5 log reduction) [442]. This difference may be explained by the incomplete depletion of PCBPs when performing siRNA-mediated knockdown. In our transient expression system, we were able to reduce the total amount of frameshifting by only ~50-70% when PCBP1 and PCBP2 were depleted simultaneously. The small amount of PCBP1 and PCBP2 that remains present in the cell after knockdown, or the presence of PCBP3, PCBP4 and hnRNPK if these proteins are also capable of transactivating PRF, most likely still suffices to induce a reasonable level of frameshifting and, consequently, expression of nsp2TF and nsp2N during PRRSV infection. We may be able to distinguish between the role PCBPs play in PRF transactivation and additional functions during PRRSV infection by using this KO2 mutant virus. During a KO2 infection, any effects of PCBP knockdown on viral RNA or protein synthesis should be explained as an involvement of PCBPs in other processes than PRF transactivation, since −2/−1 PRF is already completely knocked out in this mutant virus.

There have been some reports that PCBP1 and PCBP2 co-localize with nsps during PRRSV infection. Overexpressed recombinant PCBP1 and PCBP2 were reported to partially co-localize with nsp1β and nsp2/3 during PRRSV infection [463] and endogenous PCBP2 was shown to translocate from the nucleus to the cytosol upon PRRSV infection where it co-localized with nsp1β [493]. A proportion of nsp1β localizes to the nucleus during infection [403] and we hypothesize that PCBP1 and PCBP2 are recruited to the cytosol through their interaction with nsp1β. If this were true, most of the PCBPs that remain after siRNA-mediated knockdown might still be recruited to the site of viral

genome translation and stimulate frameshifting. This would explain the relatively high level of frameshifting observed even after efficient siRNA-mediated knockdown of PCBPs. It will be interesting to determine whether PCBP1 and PCBP2 are still recruited to the cytosol during infection with the virus mutant containing the nsp1β RBD- mutations, since this mutant is no longer able to transactivate −2/−1 PRF. This could be because it can no longer interact with the PRRSV RNA, but might also be because its interaction with PCBPs is abolished.

Binding partners of nsp1β and PCBPs

It might be possible that other binding partners of either nsp1β or PCBPs are also part of the frameshift stimulatory complex that presumably interacts with the translating ribosome to induce frameshifting. Nsp1β has been reported to interact with rpS14 [463], which is located adjacent to rpS3 of the ribosomal helicase [466]. Cytosolic PCBP1 has been shown to interact with RACK1 [467], a protein that interacts with the head region of the 40S ribosome close to the mRNA exit channel [507]. It was recently shown that ribosomes lacking Asc1, a yeast homolog of RACK1, frameshift more often at CGA codon repeats [508], suggesting that an interaction of PCBP1 with RACK1 could potentially modulate frameshifting frequency on the PRRSV mRNA as well.

PCBP interactions with viral genomes

The interaction of PCBPs with the PRRSV (C-rich) mRNA is not unique, as interactions with other viral RNAs have been described previously and such interactions often modulate viral protein translation and RNA replication. PCBP1 and PCBP2 interact with the terminal clover leaf structure of the poliovirus (PV) genome and that of other picornaviruses, forming a ribonucleoprotein complex with viral protein 3CD [509-511]. This interaction promotes mRNA stability by protecting it from degradation by 5′ exonucleases [512, 513]. The interaction of PCBPs with the clover leaf structure also plays a role in circularization of the viral genome through an RNA-protein-protein-RNA bridge that is required for the initiation of negative strand RNA synthesis [514]. PCPB2 also interacts with stem-loop IV of the type 1 picornavirus IRES to mediate translation initiation [515, 516]. During mid-to-late phase of PV infection, PCBP1 and PCBP2 are cleaved by viral proteinases 3C/3CD resulting in a truncated protein lacking the KH3 domain [517]. The same has been reported for PCBP2 and the 3C protease of hepatitis A virus, an atypical picornavirus [518]. Cleaved PCBP2 can no longer function in translation initiation but retains its activity in viral RNA replication. It was suggested that this cleavage mediates the shift from viral translation to RNA replication [517, 518].

Binding of PCBPs to 5′ UTRs has also been reported for hepatitis C virus (HCV) and Norwalk virus [519, 520]. For HCV, PCBP2 is also required for circularization of the viral

genome [521]. Finally, PCBP1 has been reported to interact with N^{pro}, a cysteine-like autoprotease, of classical swine fever (CSFV) and promotes virus growth [522].

Protein-dependent recoding

This is the first time that cellular proteins have been shown to stimulate PRF. Previously, Annexin A2 was shown to inhibit −1 PRF in the coronavirus infectious bronchitis virus by destabilizing the stimulatory pseudoknot [457]. In a very artificial system the −1 HIV-1 PRF signal could be stimulated by replacing the RNA structure with the iron-responsive-element (IRE) from ferritin mRNA. Binding of iron regulatory proteins to the IRE did increase frameshifting efficiency about 2-3 fold, but efficiency was still less than 10% [461]. While this was a very unnatural example, it did show that RNA-protein interactions can stimulate frameshifting.

Only one other example of protein-dependent recoding where a protein interacts directly with the mRNA has been described. The vascular endothelial growth factor A (VEGFA) mRNA undergoes stop codon readthrough which generates an isoform containing a unique 22-amino acid C terminus extension called VEGF-Ax. This readthrough is transactivated by a high affinity interaction of hnRNPA2/B1 with a near-consensus hnRNPA2/B1 recognition element (A2RE) in a 63-nt long Ax element located in between the two stop codons in the VEGFA 3' UTR RNA sequence. The A2RE starts 11 nt downstream from the stopcodon and the authors suggest that bound hnRNPA2/B1 interacts with ribosomal proteins to induce translational pausing allowing the incorporation of serine at the stop codon position *[523]*. Like hnRNPK, hnRNPA2/B1 belongs to the heterogeneous ribonucleoproteins, but RNA recognition of hnRNPA2/B1 goes through two RNA recognition motifs (RRMs) and an arginine/glycine-rich (RGG) box instead of KH domains [524].

Over 1500 RNA binding proteins have been described [525] and it does seem likely that more of these proteins might interact with RNA sequences downstream from slippery sequences forming RNA-protein complexes that could potentially block translating ribosomes.

CONCLUSION

The requirement of a host protein in the RNA-protein complex that induces ribosomal frameshifting in this, so far, unique arterivirus mechanism was a surprise. This mechanism thus constitutes a novel type of arterivirus-host interaction and shows that protein complexes can replace RNA structures to stimulate frameshifting with high efficiency. It raises the question whether protein-mediated frameshifting might be more common, since many different RNA binding proteins that interact with RNA through various RNA

binding domains could potentially interact with RNA motifs downstream of slippery sequences to induce shifting on their own (or as a protein-complex) or interact with (viral) RNA structures to modulate the shifting frequency.

The WG *in vitro* translation assay allowed us to study the effects of the protein factors of the PRRSV mechanism separately. For other eukaryotic frameshifting sequences this might also be a good system to test whether these are also modulated by cellular or viral protein factors. This unique frameshift-stimulatory signal may also provide novel insights into how transacting protein factors can modify the ribosomal elongation cycle.

The siRNA-mediated knockdown experiments clearly showed that the expression level of individual PCBPs determines the direction and efficiency of frameshifting in cells. PCBP1 predominantly stimulated −2 PRF and PCBP2 stimulated −1 PRF, but other PCBP family members may be able to stimulate −2/−1 PRF as well. This study adds yet another function to the already broad repertoire of cellular PCBP activities.

ACKNOWLEDGMENTS
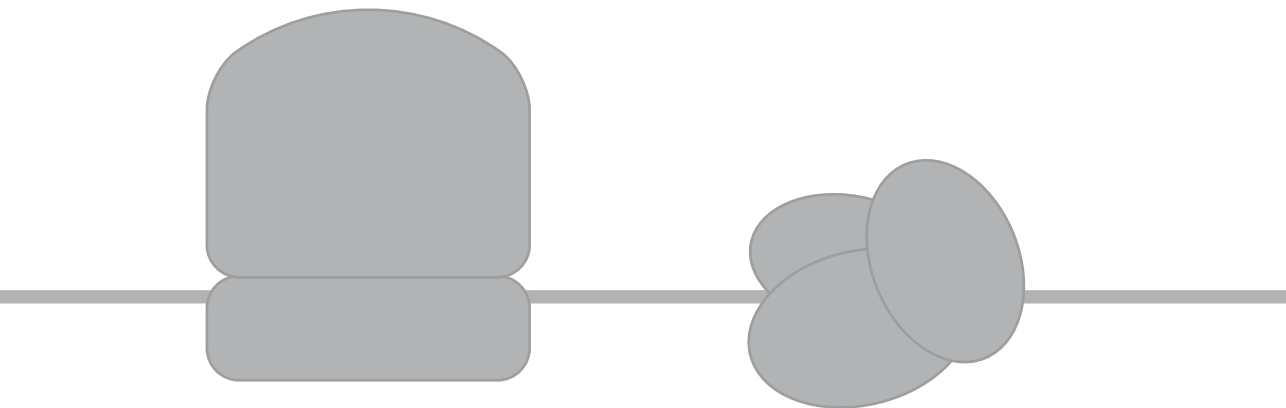
Chapter 9

General discussion part 2

Chapters 6-8 of this thesis describe a novel programmed ribosomal frameshifting (PRF) mechanism in arteriviruses. This mechanism involves a slippery sequence, C-rich motif, and – uniquely – a protein complex consisting of the viral protein nsp1β and the host poly (C) binding proteins PCBP1 and/or PCBP2. These signals and factors interact to stimulate translating ribosomes to slip backwards 1 or 2 nucleotides at a relatively high frequency. This results in the synthesis of two previously unknown nsp2 variants, nsp2TF (the –2 frameshift product) and nsp2N (the –1 frameshift product), which each share their N-terminal domain with the full-length nsp2 encoded by ORF1a.

While most examples of ribosomal frameshifting have been found in viruses, ribosomal frameshifting mechanisms are employed by organisms from all kingdoms of life to regulate gene expression [526]. Especially in +RNA- and retroviruses, –1 PRF is a quite common strategy to regulate the expression of their RNA-dependent RNA polymerase (RdRp) or reverse transcriptase [323].

For –1 PRF, the frameshifting signals are commonly embedded in the mRNA in the form of a bipartite signal consisting of a heptanucleotide slippery sequence and a downstream stimulatory RNA structure, either a pseudoknot or stable RNA stem-loop structure, which are separated by a 6-8 nt spacer sequence [336, 527]. The mechanism described in this thesis contains several atypical aspects compared to these well-characterized –1 PRF mechanisms. These features are the occurrence of efficient –2 PRF, alternative shifts (–2 or –1) occurring at the same slippery site and the requirement for a transactivating protein complex that appears to replace the stimulatory RNA structure downstream of the shift site, presumably by forming a similar kind of roadblock for the translating ribosome. This chapter summarizes the findings of chapters 6-8 and assesses whether certain features of this novel frameshifting mechanism are unique for arteriviruses or may be employed more commonly, without having been discovered so far. Furthermore, the function of PCBPs during PRRSV infection and possible roles of the frameshift products nsp2TF and nsp2N are discussed.

How widespread is –2 ribosomal frameshifting?

The first time –2 PRF was shown to be possible was almost 30 years ago, in an artificial system using a synthetic mRNA sequence on which a 2% PRF efficiency was observed [528]. Since then, only a handful of other examples have been described, with only three of them, including the arterivirus mechanism discussed here, occurring in natural situations.

In an *in vitro* system, homopolymeric stretches of nucleotides, in particular a U_8 stretch, were able to induce both –1 and –2/+1 frameshifting (direction of frameshift in the latter case not confirmed, but not unlikely that it was –2) [468]. The construct containing the U_8 stretch was even able to induce 20% of –2/+1 shifting when the downstream RNA pseudoknot was destabilized [468]. It has been shown that naturally occurring +1 and

−1 PRF signals can also induce −2 PRF in artificial systems. Expression of mammalian ornithine carboxylase antizyme completely depends on +1 PRF and frameshifting efficiencies are regulated by polyamine levels [529, 530]. However, when a cassette containing the antizyme frameshifting signals was expressed in yeast, translation occurred via −2 instead of +1 PRF. When the distance between the shift site and downstream pseudoknot was increased by 3 nucleotides both +1 and −2 PRF were stimulated [406]. The HIV-1 U_6A slippery sequence induced both −1 and −2 frameshifting in an *in vitro* system with a pseudoknot, stem-loop or antisense oligonucleotide (AON) as stimulator. The spacer length influenced frameshift frequencies, with the optimal spacer length for −2 PRF being 1-2 nucleotides shorter than for −1 PRF [407].

The first natural occurrence of −2 PRF was described in 2004 in bacteriophage Mu, regulating translation of two tail assembly genes with an efficiency of 2.2%. For this frameshifting mechanism a slippery site was identified, but no obvious downstream RNA secondary structure could be found. [408]. Translation of the RdRp from the dsRNA Trichomonas vaginalis viruses (TVV) as a fusion with the capsid protein depends on a ribosomal frameshift. In TVV2 and 3 this shift is −1 and in TVV1 it was thought to be +1 [531]. In **chapter 6** we predicted that TVV1 might employ −2 PRF instead of +1 because the slippery site CC_CUU_UUU is very similar to the −2/−1 shift site in PRRSV (GG_GUU_UUU) [442]. Parent *et al.* later used mass spectrometry to show that it is indeed a −2 shift [532]. If the ribosome would slip −1 at this site in TVV1, it would also run into a stop codon [533], as in the case of PRRSV SD01-08. Parent *et al.* were, however, not able to verify the synthesis of a −1 PRF product [532]. The efficiency of −2 PRF in this system is supposedly much lower (1-2%)[409] than that observed for PRRSV, which would probably make −1 PRF an even rarer event and would complicate the detection of the corresponding frameshift product. Interestingly, the RNA downstream of the slippery site in TVV also lacks an obvious secondary RNA structure and the downstream sequences are not very conserved across the three species [531] which suggests the downstream RNA sequence may not be crucial for stimulating PRF in this virus.

So far, the high efficiency of −2 PRF that is induced by the interaction of nsp1β and PCBPs with the C-rich motif in arteriviruses appears to be unique as in the other two known cases of naturally occurring −2 PRF only 2% of the ribosomes make the shift. However, the occurrence of −2 PRF is likely not unique in itself. The examples of naturally occurring −2 PRF described above indicate that −2 PRF can occur in different organisms and most likely also through different mechanisms. It is interesting that in none of these three examples a downstream stimulatory RNA structure could be identified. In arterivirus −2/−1 PRF, the RNA structure is presumably replaced by the formation of an mRNA-protein complex inducing a much higher frameshifting efficiency than what has been observed for the other two viruses. There is, however, some evidence that slippery sequences on their own can also induce −1 PRF, although at low frequencies

[338]. When PRRSV nsp2 was expressed in the absence of nsp1β, frameshifting was almost completely abolished, but a trace amount of the −2 frameshift product (nsp2TF) could still be observed (**Figure 2, chapter 7**). Perhaps −2 PRF in bacteriophage Mu and TVV1 is not stimulated by additional factors, which may explain the low PRF frequencies observed for these viruses. If this is true, it is not unlikely that −2 PRF is much more common, probably also outside of the virus world. However, the low shift efficiencies make it a challenge to find more examples, as it can be difficult to detect such low-abundant frameshift products.

How common are alternative frameshifts at the same site in natural situations?

One of the unusual features of the arterivirus PRF mechanism described in this thesis is the occurrence of alternative frameshifts at the same site, which both occur with relatively high frequency, ~7 and ~20% for −1 and −2 PRF, respectively, in PRRSV-infected MARC-145 cells (**Chapters 6 and 7**). In the case of PRRSV −2/−1 PRF, the −2 shift is the predominant shift while −1 PRF occurs at a lower frequency, although the −2/−1 ratio can be modified by changing the levels of PCBP1 and PCBP2 through siRNA-mediated knockdown (**Figure 5, chapter 8**) and *in vitro* by modulating the relative PCBP concentrations (**Figure 2 and 3, chapter 8**). This is the first example of naturally occurring alternative shifts at a single shift site [488]. In artificial systems it has already been shown that alternative shifts are possible in various contexts [406, 407, 468]. This raises the question whether there could also be natural shift sites where −1 PRF is the predominant shift with occasional occurrence of −2 PRF. At least 22 possible slippery sites have been described for −1 PRF [527] and for some of these, proper base-pairing in the A-site would also be maintained when a −2 shift would take place. However, on most mRNAs the ribosome will in all likelihood quickly run into a termination codon after −2 PRF resulting in a truncated protein that most likely is degraded rapidly, making it a challenge to prove its existence. Whether such truncated products would have any biological relevance remains to be investigated, but there is some evidence that frameshifting can be used as a strategy for the post-transcriptional regulation of gene expression [483]. For +RNA viruses, given their limited genome size, the possibility to express an additional protein, even at low expression levels, may have a selective advantage.

How common is protein-mediated ribosomal frameshifting?

The most surprising feature of the arterivirus −2/−1 frameshifting mechanism is that it is induced by a complex of multiple protein transactivators instead of a stimulatory higher-order RNA structure (**Chapter 7 and 8**). It is now clear that such RNA-protein complexes can have the same function as an RNA structure and are capable of inducing high levels of frameshifting. However, the molecular basis for the interactions between the RNA and protein players involved remains to be elucidated. Do both nps1β and PCBP1 or

2 interact with the C-rich motif in the mRNA, or is it only one of the proteins and does the other only interact with the other protein factor? Furthermore, it is possible that there are multiple copies of each protein in the complex since nsp1β was crystallized as a homodimer [453] and PCBPs are known to be capable of forming homo- or heterodimers through interactions between their KH domains [534]. It is likely that nsp1β and PCBP form a complex prior to binding to the RNA, since these proteins individually could not bind strongly enough to shift an RNA probe in a band shift assay (Napthine et al, unpublished data).

The discovery that proteins are capable of stimulating frameshifting in a natural system raises the question whether this mechanism may be employed more often. The first time it was shown that −1 PRF could be stimulated by an RNA-protein interaction was in an artificial system where the RNA hairpin of the HIV-1 gag-pol frameshift cassette was replaced by the iron-responsive-element (IRE) from ferritin mRNA. The IRE is a stem-loop structure that can bind iron regulatory proteins and in this system the IRE could substitute for the HIV-1 hairpin. When proteins interacted with the IRE frameshifting increased 2- to 3-fold, although the frameshift efficiency was still less than 10% [461]. This was clearly a very artificial set-up, but the possibility of binding of viral or cellular proteins to RNA stem-loops or pseudoknots to stimulate or inhibit frameshifting does not seem implausible. After all, non-canonical translation initiation via viral and cellular internal ribosomal entry sites (IRESes) can also be modulated by the interaction of various proteins, including PCBP1 and PCBP2, with RNA structures [535].

In natural systems, the presence of protein factors has also been shown to modulate −1 PRF efficiency, although in these cases the protein did not stimulate frameshifting but reduced the PRF frequency. Annexin A2 (ANXA2) binds to the infectious bronchitis virus (IBV) RNA pseudoknot. Overexpression of this protein reduced IBV frameshifting, while its RNAi-mediated knockdown had the opposite effect [457]. The HIV-1 Tat protein was shown to decrease the −1 PRF shift that produces the Gag-Pol fusion protein in a dose-dependent manner by increasing the rate of translation initiation. It was proposed that increased translation initiation would decrease the spacing between translating ribosomes, thus reducing the chance that a stimulatory RNA signal will refold in between translating ribosomes. As a result, fewer ribosomes will encounter the signal and make the frameshift [460, 536].

Over 1500 proteins have already been described to bind RNA through various RNA binding domains [525, 537]. If proteins bind RNA sequences close enough to slippery sites, with sufficient affinity, they could form a roadblock for translating ribosomes. If these proteins interact with stimulatory RNA structures this could modulate frameshifting frequencies, either by destabilization of the structure to reduce the efficiency or by doing the opposite and enhance frameshifting by forming more stable roadblocks.

The involvement of a specific protein as a modulator of the frameshifting efficiency is more difficult to prove than the presence of an RNA pseudoknot stimulatory structure. When studying the arterivirus mechanism, substantial differences in frameshifting efficiencies were observed when different cell lines were used for the experiments (**Chapter 6**). This was the first observation that led us to believe that cellular factors may be involved as well, as differences in their expression level may explain the differences in PRF efficiency observed between cell lines. If such an observation would be made for another system, this may also indicate that other factors are involved in stimulating (or inhibiting) frameshifting. The wheat germ-based *in vitro* translation system (**Chapter 8**) made it possible to study the mechanism in the absence of the cognate protein factors. Since the sequence of most eukaryotic, viral and bacterial proteins will not be conserved in plants, this system would be a useful tool to probe whether, besides the RNA sequence, other factors are also required for frameshift induction or stimulation. *Vice versa,* for studying plant frameshifting signals the rabbit reticulocyte lysate system may prove to be a useful alternative to the wheat germ extract.

If additional stimulators are required in a PRF mechanism, this stimulator (or inhibitor) does not necessarily have to be a protein. In artificial systems, metabolite-responsive pseudoknots from RNA riboswitches were able to stimulate –1 PRF when placed in reporter constructs [538, 539]. The pseudoknot derived from the S-adenosylhomocysteine (SAH) riboswitch induced up to 4 % –1 PRF (10-fold increase) in the presence of an optimal concentration of SAH [538]. Pseudoknots derived from 7-aminomethyl-7-deazaguanine (preQ$_1$) riboswitches induced 7-20% –1 PRF in the presence of preQ [539]. These two studies suggest that small molecules are able to interact with PRF-stimulatory RNA structures and could potentially modulate frameshifting frequencies. It has also been shown that the –1 PRF events in cytokine mRNAs, which are directed by a pseudoknot, can be stimulated by miRNAs through mRNA-miRNA interactions. On these mRNAs frameshifting directs ribosomes to a premature termination codon resulting in destabilization of the mRNA through the nonsense-mediated mRNA decay pathway which might be a mechanism to post-transcriptionally regulate gene expression [483, 540].

Many viruses employ PRF as a method to generate a defined ratio of their polymerase relative to other replicase components [323]. For some of these viruses, the artificial alteration of their frameshifting efficiency, by mutating the frameshift signal, resulted in a crippled virus [541-543]. In instances where a fixed frameshifting frequency is essential, it seems unlikely that PRF would be regulated by a protein factor that may not always be present at the same concentration. On the other hand, regulation of PRF efficiency by protein factors would offer a feedback mechanism that could control frameshift product abundance in certain conditions. Such a use of protein-mediated PRF would be more likely to be employed in cellular mRNAs though than in +RNA genomes, since cells may

require much more variation in gene expression levels than most viruses (e.g. during circadian clock and cell cycle regulation [544]).

Protein-mediated frameshifting in PRRSV-infected cells

The cellular expression levels of PCBPs determined the efficiency of PRF in cells transfected with a plasmid expressing nsp1β and nsp2 (**Chapter 8**). It will be interesting to determine to which extent siRNA-mediated knockdown of PCBPs also affects −2/−1 PRF in PRRSV-infected cells and to which extent other processes such as viral replication and protein synthesis are affected. One study reported no effect on viral polyprotein synthesis after siRNA knockdown of PCBP1 and PCBP2 [463], while another study reported a significant reduction after PCBP2 knockdown [493]. Modest reductions (~0.5-1 log) of viral progeny titers upon knockdown of PCBP1 and PCBP2 have been reported [463, 493]. This was slightly less than the ~1.5 log reduction that was observed with the KO2 mutant virus in which frameshifting was knocked out by mutating the PRF signal (**Figure 6, chapter 6**). This difference in titer reduction could be explained by incomplete knockdown of PCBPs, since we observed that total frameshifting could only be reduced by ~50-70% with simultaneous knockdown of PCBP1 and PCBP2. It is also possible that PCBPs have additional functions during the PRRSV replicative cycle that are unrelated to their frameshift stimulatory role. Beura *et al.* reported an interaction of PCBP1 and PCBP2 with the 5′UTR of the PRRSV genome in an *in vitro* mobility shift assay, although this interaction did not appear to be very strong [463]. Additional roles of PCBPs during the viral replicative cycle will complicate our future studies to determine the exact function of PCBPs during ribosomal frameshifting. The potential additional roles of PCBPs during PRRSV infection could be studied by using the KO2 mutant virus in combination with cellular PCBP knockdown, since any observed effects should be mediated through non-frameshift stimulatory functions. Studying the role of PCBPs in PRF, without interference of any other function these proteins may have during PRRSV infection, may prove to be more problematical. If it will be possible to prevent the interaction of PCBPs with the PRRSV 5′UTR while maintaining virus viability, there is no guarantee that other virus-PCBP interactions will not emerge, which may still interfere with our studies.

It is not yet certain whether the other PCBP family members, PCBP3, PCBP4 and hnRNPK can also transactivate −2/−1 PRF. Given their sequence similarity it seems likely that they can. It may not be possible to successfully knock down three or even more proteins from this family simultaneously without causing cellular toxicity. Knockdown of PCBP3 on its own proved to be a problem during vaccinia virus infection as these cells died much faster (**Figure 5, chapter 8**), but this might not be the case when the siRNA-treated cells are infected with PRRSV instead.

The function of the frameshift products, nsp2TF and nsp2N, during the viral replicative cycle is still unknown. Nsp1β has been implicated in innate immune evasion (reviewed by [361, 489]) and it is possible that this is at least partially mediated through the expression of nsp2TF and nsp2N. The mutant viruses KO1 and KO2 which express a truncated form of nsp2TF (KO1), or no frameshift products at all (KO2), were crippled and displayed a small-plaque phenotype, with KO2 being more crippled than KO1 (**Figure 6, chapter 6**). The nsp1βKO mutant was even more crippled and produced tiny plaques that could only be visualized when held up against a light source (**Chapter 7** and unpublished data). These phenotypes highlight that expression of the frameshift products is important during PRRSV infection. It may be possible to distinguish between the innate immune functions of nsp1β itself and any effects mediated through its role as a frameshift stimulator by using the KO2 mutant virus since any potential innate immune modulating effects mediated by nsp2TF and nsp2N have been knocked out in this virus. It will be much more difficult, if not impossible, to determine the functions of nsp2TF and nsp2N in the context of viral infection without nsp1β present. Arteriviruses lacking the nsp1 region are not viable [451, 476], and the expression of the two frameshift products depends on the presence of nsp1β. It may be possible to study some of the nsp2TF and nsp2N functions using transient expression, but when the proteins are expressed from separate plasmids it may be difficult to achieve the correct ratio between them, which normally is determined by the frameshifting mechanism. Another potential problem with nsp2TF overexpression to determine its biological function is the observation that the localization of nsp2TF in this context may not be the same as it is during a PRRSV infection (**Figure 7, chapter 6**). During PRRSV infection nsp2 and nsp2TF appear to be targeted to different cellular compartments, presumably via their different transmembrane domains, and this suggests that they may have different functions (**Figure 7, chapter 6**). Nsp2N does not have a transmembrane domain and probably localizes to the cytosol. It will be interesting to determine whether these different localizations and potential different functions are linked to the deubiquitinating and deISGylating activities of their shared papain-like protease (PLP2) domains [167, 420-422]. Arterivirus nsp2 is also one of the proteins responsible for the formation of the membranous structures that are used to scaffold the viral replication and transcription complex [369, 404]. It may be possible that nsp2TF plays a role in the formation of these complexes as well. Certain domains of nsp2 itself, which are shared with nsp2TF and nsp2N, have also been implicated in modulation of the host immune response [424].

For any virus with mutations in nsp1β or nsp2, it will be important to determine whether its phenotype is not (partially) mediated through disruption of efficient −2/−1 PRF. Mutations in the nsp2 sequence overlapping the nsp2TF sequence may not be silent in nsp2TF (and *vice versa*) and this should be taken into account when mutations are made in this part of the protein.

Conclusion

The novel features of the arterivirus PRF mechanism described in this thesis show that the mechanistic possibilities for inducing efficient PRF are much broader than previously thought. While several features thus far appear to be unique for arteriviruses, it does seem likely that efficient −2 PRF, alternative shifts at the same site and protein-mediated frameshifting are more widely employed in biological systems. Especially the alternative shifts and the use of protein stimulators offer additional options for regulating protein expression.

The stimulation of PRF by PCBPs during PRRSV infection is an example of a novel class of virus-host interactions. Further studies into the functions of PCBPs in the PRRSV replicative cycle may prove PRRSV-PCBP interactions to be even more widespread, as there are multiple potential roles for PCBPs during infection. The biological function of the two frameshift products, nsp2TF and nsp2N remains elusive thus far and their synthesis is inevitably linked to nsp1β expression. While many questions still remain, this frameshifting mechanism is a perfect example of how multifaceted +RNA virus genome translation and virus-host interactions can be.
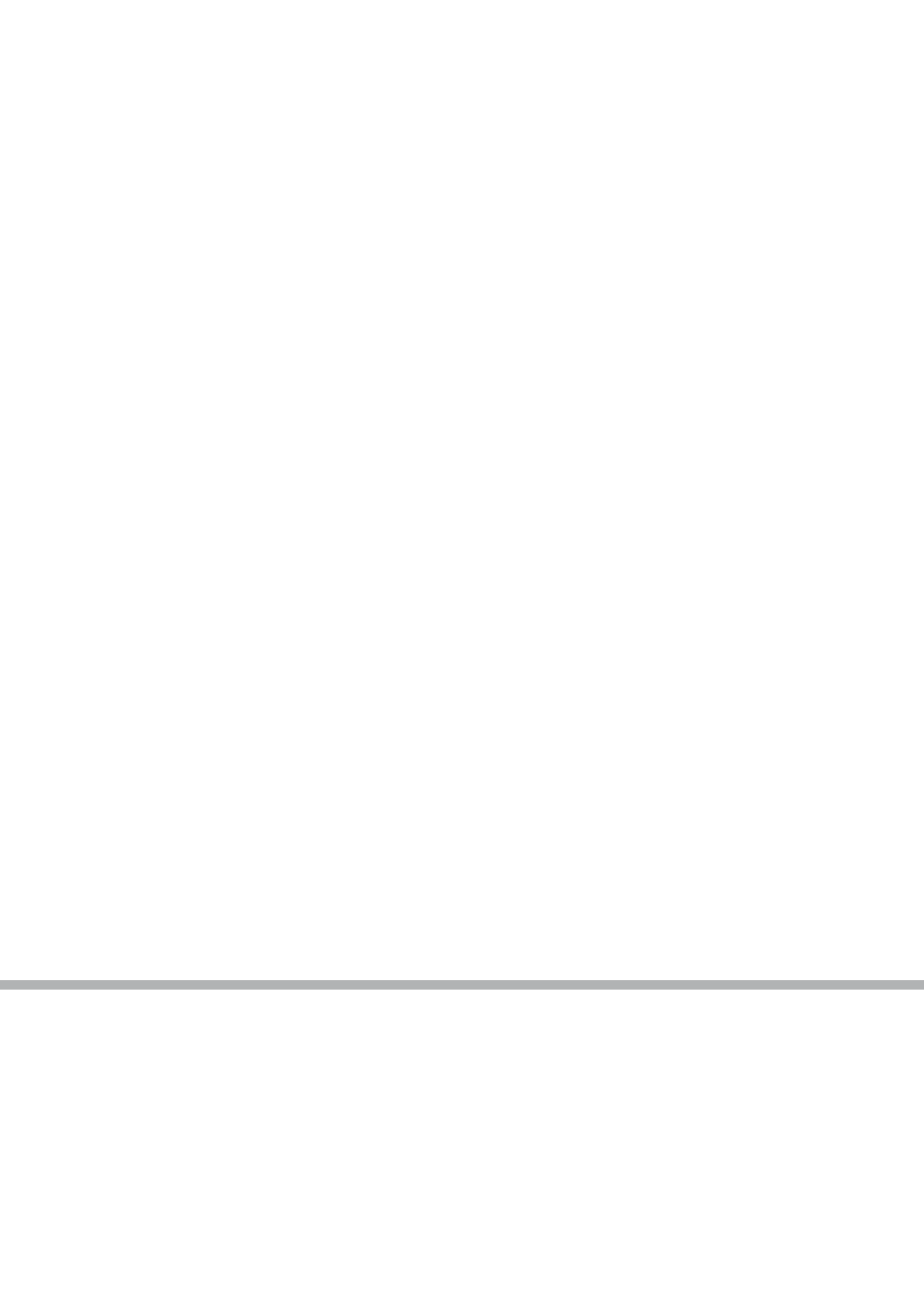
References

1. Baltimore, D., *Expression of animal virus genomes.* Bacteriological Reviews, 1971. **35**(3): p. 235-241.

2. Drosten, C., et al., *Identification of a novel coronavirus in patients with severe acute respiratory syndrome.* N Engl J Med, 2003. **348**(20): p. 1967-76.

3. Kuiken, T., et al., *Newly discovered coronavirus as the primary cause of severe acute respiratory syndrome.* Lancet, 2003. **362**(9380): p. 263-70.

4. Ksiazek, T.G., et al., *A novel coronavirus associated with severe acute respiratory syndrome.* N Engl J Med, 2003. **348**(20): p. 1953-66.

5. Kitamura, N., et al., *Primary structure, gene organization and polypeptide expression of poliovirus RNA.* Nature, 1981. **291**(5816): p. 547-53.

6. Choo, Q.L., et al., *Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome.* Science, 1989. **244**(4902): p. 359-62.

7. Sabin, A.B. and R.W. Schlesinger, *PRODUCTION OF IMMUNITY TO DENGUE WITH VIRUS MODIFIED BY PROPAGATION IN MICE.* Science, 1945. **101**(2634): p. 640-2.

8. Robinson, M.C., *An epidemic of virus disease in Southern Province, Tanganyika Territory, in 1952-53. I. Clinical features.* Trans R Soc Trop Med Hyg, 1955. **49**(1): p. 28-32.

9. Galloway, I.A. and M. Schlesinger, *Purification and concentration of the virus of foot-and-mouth disease by combined centrifugation and ultrafiltration methods.* J Hyg (Lond), 1937. **37**(3): p. 463-70.

10. Wensvoort, G., et al., *Mystery swine disease in The Netherlands: the isolation of Lelystad virus.* Vet Q, 1991. **13**(3): p. 121-30.

11. Benfield, D.A., et al., *Characterization of swine infertility and respiratory syndrome (SIRS) virus (isolate ATCC VR-2332).* J Vet Diagn Invest, 1992. **4**(2): p. 127-33.

12. te Velthuis, A.J., *Common and unique features of viral RNA-dependent polymerases.* Cell Mol Life Sci, 2014. **71**(22): p. 4403-20.

13. Harak, C. and V. Lohmann, *Ultrastructure of the replication sites of positive-strand RNA viruses.* Virology, 2015.

14. Burt, F.J., et al., *Chikungunya: a re-emerging virus.* Lancet, 2012. **379**(9816): p. 662-71.

15. Schwartz, O. and M.L. Albert, *Biology and pathogenesis of chikungunya virus.* Nat Rev Microbiol, 2010. **8**(7): p. 491-500.

16. Tsetsarkin, K.A., et al., *A single mutation in chikungunya virus affects vector specificity and epidemic potential.* PLoS Pathog, 2007. **3**(12): p. e201.

17. Fischer, M. and J.E. Staples, *Notes from the field: chikungunya virus spreads in the Americas - Caribbean and South America, 2013-2014.* MMWR Morb Mortal Wkly Rep, 2014. **63**(22): p. 500-1.

18. Frank, C., I. Schoneberg, and K. Stark, *Trends in imported chikungunya virus infections in Germany, 2006-2009.* Vector Borne Zoonotic Dis, 2011. **11**(6): p. 631-6.

19. Druce, J.D., et al., *Chikungunya virus infection in traveler to Australia.* Emerg Infect Dis, 2007. **13**(3): p. 509-10.

20. Beltrame, A., et al., *Imported Chikungunya Infection, Italy.* Emerg Infect Dis, 2007. **13**(8): p. 1264-6.

21. *Chikungunya fever diagnosed among international travelers--United States, 2005-2006.* MMWR Morb Mortal Wkly Rep, 2006. **55**(38): p. 1040-2.

22. Lanciotti, R.S., et al., *Chikungunya virus in US travelers returning from India, 2006.* Emerg Infect Dis, 2007. **13**(5): p. 764-7.

23. Bottieau, E., et al., *Chikungunya infection confirmed in a Belgian traveller returning from Phuket (Thailand).* Euro Surveill, 2009. **14**(25).

24. Schwartz, K.L., A. Giga, and A.K. Boggild, *Chikungunya fever in Canada: fever and polyarthritis in a returned traveller.* Cmaj, 2014. **186**(10): p. 772-4.

25. Paty, M.C., et al., *Large number of imported chikungunya cases in mainland France, 2014: a challenge for surveillance and response.* Euro Surveill, 2014. **19**(28): p. 20856.

26. Higgs, S. and D. Vanlandingham, *Chikungunya Virus and Its Mosquito Vectors.* Vector Borne Zoonotic Dis, 2015.

27. Vega-Rua, A., et al., *Chikungunya virus transmission potential by local aedes mosquitoes in the americas and europe.* PLoS Negl Trop Dis, 2015. **9**(5): p. e0003780.

28. Rezza, G., et al., *Infection with chikungunya virus in Italy: an outbreak in a temperate region.* Lancet, 2007. **370**(9602): p. 1840-6.

29. Gould, E.A., et al., *First cases of autochthonous dengue fever and chikungunya fever in France: from bad dream to reality!* Clin Microbiol Infect, 2010. **16**(12): p. 1702-4.

30. McCarthy, M., *First case of locally acquired chikungunya is reported in US.* Bmj, 2014. **349**: p. g4706.

31. Queyriaux, B., et al., *Clinical burden of chikungunya virus infection.* Lancet Infect Dis, 2008. **8**(1): p. 2-3.

32. Sissoko, D., et al., *Seroprevalence and risk factors of chikungunya virus infection in Mayotte, Indian Ocean, 2005-2006: a population-based survey.* PLoS One, 2008. **3**(8): p. e3066.

33. Appassakij, H., et al., *Viremic profiles in asymptomatic and symptomatic chikungunya fever: a blood transfusion threat?* Transfusion, 2013. **53**(10 Pt 2): p. 2567-74.

34. Robinson, M., et al., *A model for a chikungunya outbreak in a rural Cambodian setting: implications for disease control in uninfected areas.* PLoS Negl Trop Dis, 2014. **8**(9): p. e3120.

35. Parashar, D. and S. Cherian, *Antiviral Perspectives for Chikungunya Virus.* Biomed Res Int, 2014. **2014**: p. 631642.

36. Kaur, P. and J.J. Chu, *Chikungunya virus: an update on antiviral development and challenges.* Drug Discov Today, 2013. **18**(19-20): p. 969-83.

37. Lauring, A.S. and R. Andino, *Quasispecies theory and the behavior of RNA viruses.* PLoS Pathog, 2010. **6**(7): p. e1001005.

38. Domingo, E., et al., *Quasispecies structure and persistence of RNA viruses.* Emerg Infect Dis, 1998. **4**(4): p. 521-7.

39. Duffy, S., L.A. Shackelton, and E.C. Holmes, *Rates of evolutionary change in viruses: patterns and determinants.* Nat Rev Genet, 2008. **9**(4): p. 267-76.

40. Steinhauer, D.A. and J.J. Holland, *Rapid evolution of RNA viruses.* Annu Rev Microbiol, 1987. **41**: p. 409-33.

41. Gould, E.A., et al., *Understanding the alphaviruses: recent research on important emerging pathogens and progress towards their control.* Antiviral Res, 2010. **87**(2): p. 111-24.

42. Strauss, E.G., C.M. Rice, and J.H. Strauss, *Sequence coding for the alphavirus nonstructural proteins is interrupted by an opal termination codon.* Proc Natl Acad Sci U S A, 1983. **80**(17): p. 5271-5.

43. Hahn, Y.S., E.G. Strauss, and J.H. Strauss, *Mapping of RNA- temperature-sensitive mutants of Sindbis virus: assignment of complementation groups A, B, and G to nonstructural proteins.* J Virol, 1989. **63**(7): p. 3142-50.

44. Hardy, W.R., et al., *Synthesis and processing of the nonstructural polyproteins of several temperature-sensitive mutants of Sindbis virus.* Virology, 1990. **177**(1): p. 199-208.

45. Melancon, P. and H. Garoff, *Processing of the Semliki Forest virus structural polyprotein: role of the capsid protease.* J Virol, 1987. **61**(5): p. 1301-9.

46. Liljestrom, P. and H. Garoff, *Internally located cleavable signal sequences direct the formation of Semliki Forest virus membrane proteins from a polyprotein precursor.* J Virol, 1991. **65**(1): p. 147-54.

47. Firth, A.E., et al., *Discovery of frameshifting in Alphavirus 6K resolves a 20-year enigma.* Virol J, 2008. **5**: p. 108.

48. Voss, J.E., et al., *Glycoprotein organization of Chikungunya virus particles revealed by X-ray crystallography.* Nature, 2010. **468**(7324): p. 709-12.

49. Vaney, M.C., S. Duquerroy, and F.A. Rey, *Alphavirus structure: activation for entry at the target cell surface.* Curr Opin Virol, 2013. **3**(2): p. 151-8.

50. Singh, I. and A. Helenius, *Role of ribosomes in Semliki Forest virus nucleocapsid uncoating.* J Virol, 1992. **66**(12): p. 7049-58.

51. Shirako, Y. and J.H. Strauss, *Regulation of Sindbis virus RNA replication: uncleaved P123 and nsP4 function in minus-strand RNA synthesis, whereas cleaved products from P123 are required for efficient plus-strand RNA synthesis.* J Virol, 1994. **68**(3): p. 1874-85.

52. Lemm, J.A., et al., *Polypeptide requirements for assembly of functional Sindbis virus replication complexes: a model for the temporal regulation of minus- and plus-strand RNA synthesis.* Embo j, 1994. **13**(12): p. 2925-34.

53. Mai, J., S.G. Sawicki, and D.L. Sawicki, *Fate of minus-strand templates and replication complexes produced by a p23-cleavage-defective mutant of Sindbis virus.* J Virol, 2009. **83**(17): p. 8553-64.

54. Froshauer, S., J. Kartenbeck, and A. Helenius, *Alphavirus RNA replicase is located on the cytoplasmic surface of endosomes and lysosomes.* J Cell Biol, 1988. **107**(6 Pt 1): p. 2075-86.

55. Spuul, P., et al., *Phosphatidylinositol 3-kinase-, actin-, and microtubule-dependent transport of Semliki Forest Virus replication complexes from the plasma membrane to modified lysosomes.* J Virol, 2010. **84**(15): p. 7543-57.

56. Ahola, T. and L. Kaariainen, *Reaction in alphavirus mRNA capping: formation of a covalent complex of nonstructural protein nsP1 with 7-methyl-GMP.* Proc Natl Acad Sci U S A, 1995. **92**(2): p. 507-11.

57. Mi, S. and V. Stollar, *Expression of Sindbis virus nsP1 and methyltransferase activity in Escherichia coli.* Virology, 1991. **184**(1): p. 423-7.

58. Cross, R.K., *Identification of a unique guanine-7-methyltransferase in Semliki Forest virus (SFV) infected cell extracts.* Virology, 1983. **130**(2): p. 452-63.

59. Peranen, J., et al., *The alphavirus replicase protein nsP1 is membrane-associated and has affinity to endocytic organelles.* Virology, 1995. **208**(2): p. 610-20.

60. Gomez de Cedron, M., et al., *RNA helicase activity of Semliki Forest virus replicase protein NSP2.* FEBS Lett, 1999. **448**(1): p. 19-22.

61. Rikkonen, M., J. Peranen, and L. Kaariainen, *ATPase and GTPase activities associated with Semliki Forest virus nonstructural protein nsP2.* J Virol, 1994. **68**(9): p. 5804-10.

62. Vasiljeva, L., et al., *Identification of a novel function of the alphavirus capping apparatus. RNA 5'-triphosphatase activity of Nsp2.* J Biol Chem, 2000. **275**(23): p. 17281-7.

63. Akhrymuk, I., S.V. Kulemzin, and E.I. Frolova, *Evasion of the innate immune response: the Old World alphavirus nsP2 protein induces rapid degradation of Rpb1, a catalytic subunit of RNA polymerase II.* J Virol, 2012. **86**(13): p. 7180-91.

64. Utt, A., et al., *Mutations conferring a noncytotoxic phenotype on chikungunya virus replicons compromise enzymatic properties of nonstructural protein 2.* J Virol, 2015. **89**(6): p. 3145-62.

65. Scholte, F.E., et al., *Stress Granule Components G3BP1 and G3BP2 Play a Proviral Role Early in Chikungunya Virus Replication.* J Virol, 2015. **89**(8): p. 4457-69.

66. Lemm, J.A. and C.M. Rice, *Roles of nonstructural polyproteins and cleavage products in regulating Sindbis virus RNA replication and transcription.* J Virol, 1993. **67**(4): p. 1916-26.

67. Malet, H., et al., *The crystal structures of Chikungunya and Venezuelan equine encephalitis virus nsP3 macro domains define a conserved adenosine binding pocket.* J Virol, 2009. **83**(13): p. 6534-45.

68. Foy, N.J., et al., *Hypervariable domains of nsP3 proteins of New World and Old World alphaviruses mediate formation of distinct, virus-specific protein complexes.* J Virol, 2013. **87**(4): p. 1997-2010.

69. Fros, J.J., et al., *Chikungunya virus nsP3 blocks stress granule assembly by recruitment of G3BP into cytoplasmic foci.* J Virol, 2012. **86**(19): p. 10873-9.

70. Kamer, G. and P. Argos, *Primary structural comparison of RNA-dependent polymerases from plant, animal and bacterial viruses.* Nucleic Acids Res, 1984. **12**(18): p. 7269-82.

71. Tomar, S., et al., *Catalytic core of alphavirus nonstructural protein nsP4 possesses terminal adenylyl-transferase activity.* J Virol, 2006. **80**(20): p. 9962-9.

72. de Groot, R.J., et al., *Sindbis virus RNA polymerase is degraded by the N-end rule pathway.* Proc Natl Acad Sci U S A, 1991. **88**(20): p. 8967-71.

73. Hahn, C.S., E.G. Strauss, and J.H. Strauss, *Sequence analysis of three Sindbis virus mutants temperature-sensitive in the capsid protein autoprotease.* Proc Natl Acad Sci U S A, 1985. **82**(14): p. 4648-52.

74. Schlesinger, S. and M.J. Schlesinger, *Formation of Sindbis virus proteins: identification of a precursor for one of the envelope proteins.* J Virol, 1972. **10**(5): p. 925-32.

75. Garoff, H., et al., *The signal sequence of the p62 protein of Semliki Forest virus is involved in initiation but not in completing chain translocation.* J Cell Biol, 1990. **111**(3): p. 867-76.

76. Strauss, J.H. and E.G. Strauss, *The alphaviruses: gene expression, replication, and evolution.* Microbiol Rev, 1994. **58**(3): p. 491-562.

77. Schmidt, M.F., M. Bracha, and M.J. Schlesinger, *Evidence for covalent attachment of fatty acids to Sindbis virus glycoproteins.* Proc Natl Acad Sci U S A, 1979. **76**(4): p. 1687-91.

78. Schlesinger, M.J., S. Schlesinger, and B.W. Burge, *Identification of a second glycoprotein in Sindbis virus.* Virology, 1972. **47**(2): p. 539-41.

79. Strauss, J.H., Jr., B.W. Burge, and J.E. Darnell, *Carbohydrate content of the membrane protein of Sindbis virus.* J Mol Biol, 1970. **47**(3): p. 437-48.

80. de Curtis, I. and K. Simons, *Dissection of Semliki Forest virus glycoprotein delivery from the trans-Golgi network to the cell surface in permeabilized BHK cells.* Proc Natl Acad Sci U S A, 1988. **85**(21): p. 8052-6.

81. Zhang, X., et al., *Furin processing and proteolytic activation of Semliki Forest virus.* J Virol, 2003. **77**(5): p. 2981-9.

82. Simizu, B., et al., *Structural proteins of Chikungunya virus.* J Virol, 1984. **51**(1): p. 254-8.

83. Tellinghuisen, T.L. and R.J. Kuhn, *Nucleic acid-dependent cross-linking of the nucleocapsid protein of Sindbis virus.* J Virol, 2000. **74**(9): p. 4302-9.

84. Linger, B.R., et al., *Sindbis virus nucleocapsid assembly: RNA folding promotes capsid protein dimerization.* Rna, 2004. **10**(1): p. 128-38.

85. Jose, J., et al., *Interactions of the cytoplasmic domain of Sindbis virus E2 with nucleocapsid cores promote alphavirus budding.* J Virol, 2012. **86**(5): p. 2585-99.

86. Smith, A.L. and G.H. Tignor, *Host cell receptors for two strains of Sindbis virus.* Arch Virol, 1980. **66**(1): p. 11-26.

87. Smith, T.J., et al., *Putative receptor binding sites on alphaviruses as visualized by cryoelectron microscopy.* Proc Natl Acad Sci U S A, 1995. **92**(23): p. 10648-52.

88. Kielian, M.C., M. Marsh, and A. Helenius, *Kinetics of endosome acidification detected by mutant and wild-type Semliki Forest virus.* Embo j, 1986. **5**(12): p. 3103-9.

89. Wengler, G. and G. Wengler, *Identification of a transfer of viral core protein to cellular ribosomes during the early stages of alphavirus infection.* Virology, 1984. **134**(2): p. 435-42.

90. Raoult, D. and P. Forterre, *Redefining viruses: lessons from Mimivirus.* Nat Rev Microbiol, 2008. **6**(4): p. 315-9.

91. Bourai, M., et al., *Mapping of Chikungunya virus interactions with host proteins identified nsP2 as a highly connected viral component.* J Virol, 2012. **86**(6): p. 3121-34.

92. Rathore, A.P., et al., *Chikungunya virus nsP3 & nsP4 interacts with HSP-90 to promote virus replication: HSP-90 inhibitors reduce CHIKV infection and inflammation in vivo.* Antiviral Res, 2014. **103**: p. 7-16.

93. Varjak, M., et al., *Magnetic fractionation and proteomic dissection of cellular organelles occupied by the late replication complexes of Semliki Forest virus.* J Virol, 2013. **87**(18): p. 10295-312.

94. Cristea, I.M., et al., *Tracking and elucidating alphavirus-host protein interactions.* J Biol Chem, 2006. **281**(40): p. 30269-78.

95. Li, G.P., et al., *Phosphorylation of Sindbis virus nsP3 in vivo and in vitro.* Virology, 1990. **179**(1): p. 416-27.

96. Peranen, J., et al., *Semliki Forest virus-specific non-structural protein nsP3 is a phosphoprotein.* J Gen Virol, 1988. **69 (Pt 9)**: p. 2165-78.

97. Garoff, H., R. Hewson, and D.J. Opstelten, *Virus maturation by budding.* Microbiol Mol Biol Rev, 1998. **62**(4): p. 1171-90.

98. Ng, C.G., et al., *Effect of host cell lipid metabolism on alphavirus replication, virion morphogenesis, and infectivity.* Proc Natl Acad Sci U S A, 2008. **105**(42): p. 16326-31.

99. Frolova, E.I., et al., *Roles of nonstructural protein nsP2 and Alpha/Beta interferons in determining the outcome of Sindbis virus infection.* J Virol, 2002. **76**(22): p. 11254-64.

100. Kaariainen, L. and M. Ranki, *Inhibition of cell functions by RNA-virus infections.* Annu Rev Microbiol, 1984. **38**: p. 91-109.

101. Ma-Lauer, Y., et al., *Virus-host interactomes--antiviral drug discovery.* Curr Opin Virol, 2012. **2**(5): p. 614-21.

102. Prussia, A., et al., *Systematic Approaches towards the Development of Host-Directed Antiviral Therapeutics.* Int J Mol Sci, 2011. **12**(6): p. 4027-52.

103. Ooi, Y.S., et al., *Genome-wide RNAi screen identifies novel host proteins required for alphavirus entry.* PLoS Pathog, 2013. **9**(12): p. e1003835.

104. Nakaya, H.I., et al., *Gene profiling of Chikungunya virus arthritis in a mouse model reveals significant overlap with rheumatoid arthritis.* Arthritis Rheum, 2012. **64**(11): p. 3553-63.

105. Lee, R.C., et al., *Mosquito cellular factors and functions in mediating the infectious entry of chikungunya virus.* PLoS Negl Trop Dis, 2013. **7**(2): p. e2050.

106. Saxena, T., et al., *Combined miRNA and mRNA signature identifies key molecular players and pathways involved in chikungunya virus infection in human cells.* PLoS One, 2013. **8**(11): p. e79886.

107. Dudha, N., et al., *Host-pathogen interactome analysis of Chikungunya virus envelope proteins E1 and E2.* Virus Genes, 2015.

108. Thomas, S., et al., *Chikungunya virus capsid protein contains nuclear import and export signals.* Virol J, 2013. **10**: p. 269.

109. Rana, J., et al., *Deciphering the host-pathogen protein interface in chikungunya virus-mediated sickness.* Arch Virol, 2013. **158**(6): p. 1159-72.

110. Asnet Mary, J., et al., *Identification of structural motifs in the E2 glycoprotein of Chikungunya involved in virus-host interaction.* J Biomol Struct Dyn, 2013. **31**(10): p. 1077-85.

111. Abere, B., et al., *Proteomic analysis of chikungunya virus infected microgial cells.* PLoS One, 2012. **7**(4): p. e34800.

112. Dhanwani, R., et al., *Differential proteome analysis of Chikungunya virus-infected new-born mice tissues reveal implication of stress, inflammatory and apoptotic pathways in disease pathogenesis.* Proteomics, 2011. **11**(10): p. 1936-51.

113. Dhanwani, R., et al., *Characterization of chikungunya virus induced host response in a mouse model of viral myositis.* PLoS One, 2014. **9**(3): p. e92813.

114. Fraisier, C., et al., *Kinetic analysis of mouse brain proteome alterations following Chikungunya virus infection before and after appearance of clinical symptoms.* PLoS One, 2014. **9**(3): p. e91397.

115. Issac, T.H., T.E. Lee, and J.J. Chu, *Proteomic profiling of chikungunya virus-infected human muscle cells: Reveal the role of cytoskeleton network in CHIKV replication.* J Proteomics, 2014.

116. Puttamallesh, V.N., et al., *Proteomic profiling of serum samples from chikungunya-infected patients provides insights into host response.* Clin Proteomics, 2013. **10**(1): p. 14.

117. Tchankouo-Nguetcheu, S., et al., *Infection by chikungunya virus modulates the expression of several proteins in Aedes aegypti salivary glands.* Parasit Vectors, 2012. **5**: p. 264.

118. Tchankouo-Nguetcheu, S., et al., *Differential protein modulation in midguts of Aedes aegypti infected with chikungunya and dengue 2 viruses.* PLoS One, 2010. **5**(10).

119. Thio, C.L., et al., *Differential proteome analysis of chikungunya virus infection on host cells.* PLoS One, 2013. **8**(4): p. e61444.

120. Wikan, N., et al., *Comprehensive proteomic analysis of white blood cells from chikungunya fever patients of different severities.* J Transl Med, 2014. **12**: p. 96.

121. Patterson, S.D. and R.H. Aebersold, *Proteomics: the first decade and beyond.* Nat Genet, 2003. **33 Suppl**: p. 311-23.

122. Viswanathan, K. and K. Fruh, *Viral proteomics: global evaluation of viruses and their interaction with the host.* Expert Rev Proteomics, 2007. **4**(6): p. 815-29.

123. Zhou, S., et al., *Viral proteomics: the emerging cutting-edge of virus research.* Sci China Life Sci, 2011. **54**(6): p. 502-12.

124. Munday, D.C., et al., *Using SILAC and quantitative proteomics to investigate the interactions between viral and host proteomes.* Proteomics, 2012. **12**(4-5): p. 666-72.

125. Maxwell, K.L. and L. Frappier, *Viral proteomics.* Microbiol Mol Biol Rev, 2007. **71**(2): p. 398-411.

126. Ong, S.E., et al., *Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics.* Mol Cell Proteomics, 2002. **1**(5): p. 376-86.

127. Ibarrola, N., et al., *A proteomic approach for quantitation of phosphorylation using stable isotope labeling in cell culture.* Anal Chem, 2003. **75**(22): p. 6043-9.

128. Zhang, G., D. Fenyo, and T.A. Neubert, *Evaluation of the variation in sample preparation for comparative proteomics using stable isotope labeling by amino acids in cell culture.* J Proteome Res, 2009. **8**(3): p. 1285-92.

129. Mirgorodskaya, O.A., et al., *Quantitation of peptides and proteins by matrix-assisted laser desorption/ionization mass spectrometry using (18)O-labeled internal standards.* Rapid Commun Mass Spectrom, 2000. **14**(14): p. 1226-32.

130. Zhou, H., et al., *Quantitative proteome analysis by solid-phase isotope tagging and mass spectrometry.* Nat Biotechnol, 2002. **20**(5): p. 512-5.

131. Boersema, P.J., et al., *Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics.* Nat Protoc, 2009. **4**(4): p. 484-94.

132. Higgs, R.E., et al., *Comprehensive label-free method for the relative quantification of proteins from biological samples.* J Proteome Res, 2005. **4**(4): p. 1442-50.

133. Cox, J., et al., *Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ.* Mol Cell Proteomics, 2014. **13**(9): p. 2513-26.

134. Ono, M., et al., *Label-free quantitative proteomics using large peptide data sets generated by nanoflow liquid chromatography and mass spectrometry.* Mol Cell Proteomics, 2006. **5**(7): p. 1338-47.

135. Wang, G., et al., *Label-free protein quantification using LC-coupled ion trap or FT mass spectrometry: Reproducibility, linearity, and application with complex proteomes.* J Proteome Res, 2006. **5**(5): p. 1214-23.

136. Wang, W., et al., *Quantification of proteins and metabolites by mass spectrometry without isotopic labeling or spiked standards.* Anal Chem, 2003. **75**(18): p. 4818-26.

137. Wiener, M.C., et al., *Differential mass spectrometry: a label-free LC-MS method for finding significant differences in complex peptide and protein mixtures.* Anal Chem, 2004. **76**(20): p. 6085-96.

138. Huang, Q., et al., *SWATH enables precise label-free quantification on proteome-scale.* Proteomics, 2015.

139. Lavallee-Adam, M., et al., *PSEA-Quant: a protein set enrichment analysis on label-free and label-based protein quantification data.* J Proteome Res, 2014. **13**(12): p. 5496-509.

140. Lau, H.T., et al., *Comparing SILAC- and stable isotope dimethyl-labeling approaches for quantitative proteomics.* J Proteome Res, 2014. **13**(9): p. 4164-74.

141. Kruger, M., et al., *SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function.* Cell, 2008. **134**(2): p. 353-64.

142. Sury, M.D., J.X. Chen, and M. Selbach, *The SILAC fly allows for accurate protein quantification in vivo.* Mol Cell Proteomics, 2010. **9**(10): p. 2173-83.

143. Westman-Brinkmalm, A., et al., *SILAC zebrafish for quantitative analysis of protein turnover and tissue regeneration.* J Proteomics, 2011. **75**(2): p. 425-34.

144. Parekh, R.B. and C. Rohlff, *Post-translational modification of proteins and the discovery of new medicine.* Curr Opin Biotechnol, 1997. **8**(6): p. 718-23.

145. Breker, M. and M. Schuldiner, *The emergence of proteome-wide technologies: systematic analysis of proteins comes of age.* Nat Rev Mol Cell Biol, 2014. **15**(7): p. 453-64.

146. Venne, A.S., L. Kollipara, and R.P. Zahedi, *The next level of complexity: crosstalk of posttranslational modifications.* Proteomics, 2014. **14**(4-5): p. 513-24.

147. Khoury, G.A., R.C. Baliban, and C.A. Floudas, *Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database.* Sci. Rep., 2011. **1**.

148. Woodsmith, J., A. Kamburov, and U. Stelzl, *Dual coordination of post translational modifications in human protein networks.* PLoS Comput Biol, 2013. **9**(3): p. e1002933.

149. Huang, J., et al., *Enrichment and separation techniques for large-scale proteomics analysis of the protein post-translational modifications.* J Chromatogr A, 2014. **1372c**: p. 1-17.

150. Engholm-Keller, K. and M.R. Larsen, *Technologies and challenges in large-scale phosphoproteomics.* Proteomics, 2013. **13**(6): p. 910-31.

151. Witze, E.S., et al., *Mapping protein post-translational modifications with mass spectrometry.* Nat Methods, 2007. **4**(10): p. 798-806.

152. Beausoleil, S.A., et al., *A probability-based approach for high-throughput protein phosphorylation analysis and site localization.* Nat Biotechnol, 2006. **24**(10): p. 1285-92.

153. Cohen, P., *The origins of protein phosphorylation.* Nat Cell Biol, 2002. **4**(5): p. E127-30.

154. Johnson, S.A. and T. Hunter, *Kinomics: methods for deciphering the kinome.* Nat Meth, 2005. **2**(1): p. 17-25.

155. Manning, G., et al., *The Protein Kinase Complement of the Human Genome.* Science, 2002. **298**(5600): p. 1912-1934.

156. Alonso, A., et al., *Protein Tyrosine Phosphatases in the Human Genome.* Cell, 2004. **117**(6): p. 699-711.

157. Shi, Y., *Serine/threonine phosphatases: mechanism through structure.* Cell, 2009. **139**(3): p. 468-84.

158. Olsen, J.V., et al., *Global, in vivo, and site-specific phosphorylation dynamics in signaling networks.* Cell, 2006. **127**(3): p. 635-48.

159. Steen, H., et al., *Analysis of tyrosine phosphorylation sites in signaling molecules by a phosphotyrosine-specific immonium ion scanning method.* Sci STKE, 2002. **2002**(154): p. pl16.

160. Smith, D.R., *Global protein profiling studies of chikungunya virus infection identify different proteins but common biological processes.* Rev Med Virol, 2014.

161. Garmashova, N., et al., *The Old World and New World alphaviruses use different virus-specific proteins for induction of transcriptional shutoff.* J Virol, 2007. **81**(5): p. 2472-84.

162. Kamitani, W., et al., *Severe acute respiratory syndrome coronavirus nsp1 protein suppresses host gene expression by promoting host mRNA degradation.* Proc Natl Acad Sci U S A, 2006. **103**(34): p. 12885-90.

163. McGinnis, S. and T.L. Madden, *BLAST: at the core of a powerful and diverse set of sequence analysis tools.* Nucleic Acids Res, 2004. **32**(Web Server issue): p. W20-5.

164. Chou, H.H., et al., *A mutation in human CMP-sialic acid hydroxylase occurred after the Homo-Pan divergence.* Proc Natl Acad Sci U S A, 1998. **95**(20): p. 11751-6.

165. Yim, H. and R.L. Erikson, *Cell division cycle 6, a mitotic substrate of polo-like kinase 1, regulates chromosomal segregation mediated by cyclin-dependent kinase 1 and separase.* Proceedings of the National Academy of Sciences, 2010. **107**(46): p. 19742-19747.

166. Okamoto, Y., et al., *UbcH10 Is the Cancer-related E2 Ubiquitin-conjugating Enzyme.* Cancer Research, 2003. **63**(14): p. 4167-4173.

167. van Kasteren, P.B., et al., *Deubiquitinase function of arterivirus papain-like protease 2 suppresses the innate immune response in infected host cells.* Proc Natl Acad Sci U S A, 2013. **110**(9): p. E838-47.

168. Scholte, F.E., et al., *Characterization of synthetic Chikungunya viruses based on the consensus sequence of recent E1-226V isolates.* PLoS One, 2013. **8**(8): p. e71047.

169. Wisniewski, J.R., et al., *Universal sample preparation method for proteome analysis.* Nat Methods, 2009. **6**(5): p. 359-62.

170. Horth, P., et al., *Efficient fractionation and improved protein identification by peptide OFFGEL electrophoresis.* Mol Cell Proteomics, 2006. **5**(10): p. 1968-74.

171. Hubner, N.C., S. Ren, and M. Mann, *Peptide separation with immobilized pI strips is an attractive alternative to in-gel protein digestion for proteome analysis.* Proteomics, 2008. **8**(23-24): p. 4862-72.

172. Meiring, H.D., et al., *Nanoscale LC-MS(n): technical design and applications to peptide and protein analysis.* Journal of Separation Science, 2002. **25**(9): p. 557-568.

173. Cox, J. and M. Mann, *MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification.* Nat Biotechnol, 2008. **26**(12): p. 1367-72.

174. Cox, J., et al., *Andromeda: a peptide search engine integrated into the MaxQuant environment.* J Proteome Res, 2011. **10**(4): p. 1794-805.

175. Schwanhausser, B., et al., *Global quantification of mammalian gene expression control.* Nature, 2011. **473**(7347): p. 337-42.

176. Kearse, M., et al., *Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data.* Bioinformatics, 2012. **28**(12): p. 1647-9.

177. Utt, A., et al., *Mutations conferring a non-cytotoxic phenotype on Chikungunya virus replicons compromise enzymatic properties of non-structural protein 2.* J Virol, 2014.

178. Scholte, F.E., et al., *Stress granule components G3BP1 and G3BP2 play a proviral role early in chikungunya virus replication.* J Virol, 2015.

179. Metz, S.W., et al., *Functional processing and secretion of Chikungunya virus E1 and E2 glycoproteins in insect cells.* Virol J, 2011. **8**: p. 353.

180. Vellinga, J., et al., *Efficient incorporation of a functional hyper-stable single-chain antibody fragment protein-IX fusion in the adenovirus capsid.* Gene Ther, 2007. **14**(8): p. 664-70.

181. Pichlmair, A. and C. Reis e Sousa, *Innate recognition of viruses.* Immunity, 2007. **27**(3): p. 370-83.

182. White, L.K., et al., *Chikungunya virus induces IPS-1-dependent innate immune activation and protein kinase R-independent translational shutoff.* J Virol, 2011. **85**(1): p. 606-20.

183. Gibbs, R.A., et al., *Evolutionary and biomedical insights from the rhesus macaque genome.* Science, 2007. **316**(5822): p. 222-34.

184. Chardin, P., *Function and regulation of Rnd proteins.* Nat Rev Mol Cell Biol, 2006. **7**(1): p. 54-62.

185. Frolova, E., et al., *Formation of nsP3-specific protein complexes during Sindbis virus replication.* J Virol, 2006. **80**(8): p. 4122-34.

186. Villalonga, P., S. Fernandez de Mattos, and A.J. Ridley, *RhoE inhibits 4E-BP1 phosphorylation and eIF4E function impairing cap-dependent translation.* J Biol Chem, 2009. **284**(51): p. 35287-96.

187. Jarmoskaite, I. and R. Russell, *RNA Helicase Proteins as Chaperones and Remodelers.* Annu Rev Biochem, 2014.

188. Zirwes, R.F., et al., *A novel helicase-type protein in the nucleolus: protein NOH61.* Mol Biol Cell, 2000. **11**(4): p. 1153-67.

189. Xu, Z., R. Anderson, and T.C. Hobman, *The capsid-binding nucleolar helicase DDX56 is important for infectivity of West Nile virus.* J Virol, 2011. **85**(11): p. 5571-80.

190. Xu, Z. and T.C. Hobman, *The helicase activity of DDX56 is required for its role in assembly of infectious West Nile virus particles.* Virology, 2012. **433**(1): p. 226-35.

191. Yasuda-Inoue, M., M. Kuroki, and Y. Ariumi, *Distinct DDX DEAD-box RNA helicases cooperate to modulate the HIV-1 Rev function.* Biochem Biophys Res Commun, 2013. **434**(4): p. 803-8.

192. Chen, Y.C., et al., *Polo-like kinase 1 is involved in hepatitis C virus replication by hyperphosphorylating NS5A.* J Virol, 2010. **84**(16): p. 7983-93.

193. Gallina, A., et al., *Polo-like kinase 1 as a target for human cytomegalovirus pp65 lower matrix protein.* J Virol, 1999. **73**(2): p. 1468-78.

194. Pan, S.H., et al., *Epstein-Barr virus nuclear antigen 2 disrupts mitotic checkpoint and causes chromosomal instability.* Carcinogenesis, 2009. **30**(2): p. 366-75.

195. Sun, D., et al., *PLK1 down-regulates parainfluenza virus 5 gene expression.* PLoS Pathog, 2009. **5**(7): p. e1000525.

196. Lee, K.S., et al., *Mechanisms underlying Plk1 polo-box domain-mediated biological processes and their physiological significance.* Mol Cells, 2014. **37**(4): p. 286-94.

197. Vitour, D., et al., *Polo-like kinase 1 (PLK1) regulates interferon (IFN) induction by MAVS.* J Biol Chem, 2009. **284**(33): p. 21797-809.

198. Fros, J.J., et al., *Chikungunya virus nonstructural protein 2 inhibits type I/II interferon-stimulated JAK-STAT signaling.* J Virol, 2010. **84**(20): p. 10877-87.

199. Summers, M.K., et al., *The unique N terminus of the UbcH10 E2 enzyme controls the threshold for APC activation and enhances checkpoint regulation of the APC.* Mol Cell, 2008. **31**(4): p. 544-56.

200. Hao, Z., H. Zhang, and J. Cowell, *Ubiquitin-conjugating enzyme UBE2C: molecular biology, role in tumorigenesis, and potential as a biomarker.* Tumour Biol, 2012. **33**(3): p. 723-30.

201. Cramer, P., et al., *Structure of Eukaryotic RNA Polymerases.* Annual Review of Biophysics, 2008. **37**(1): p. 337-352.

202. Weekes, M.P., et al., *Quantitative temporal viromics: an approach to investigate host-pathogen interaction.* Cell, 2014. **157**(6): p. 1460-72.

203. Vizcaino, J.A., et al., *The PRoteomics IDEntifications (PRIDE) database and associated tools: status in 2013.* Nucleic Acids Res, 2013. **41**(Database issue): p. D1063-9.

204. Weaver, S.C. and N.L. Forrester, *Chikungunya: Evolutionary history and recent epidemic spread.* Antiviral Res, 2015. **120**: p. 32-9.

205. Morrison, T.E., *Reemergence of chikungunya virus.* J Virol, 2014. **88**(20): p. 11644-7.

206. PAHO/WHO, *Number of Reported Cases of Chikungunya Fever in the Americas, by Country or Territory 2013-2014*. 2015, Pan American Health Organization/World Health Organization.

207. Panning, M., et al., *Chikungunya fever in travelers returning to Europe from the Indian Ocean region, 2006.* Emerg Infect Dis, 2008. **14**(3): p. 416-22.

208. Treffers, E.E., et al., *Temporal SILAC-based quantitative proteomics identifies host factors involved in chikungunya virus replication.* Proteomics, 2015. **15**(13): p. 2267-80.

209. Newman, R.H., J. Zhang, and H. Zhu, *Toward a systems-level view of dynamic phosphorylation networks.* Frontiers in Genetics, 2014. **5**: p. 263.

210. Donnelly, M.L., et al., *Analysis of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'.* J Gen Virol, 2001. **82**(Pt 5): p. 1013-25.

211. Luke, G.A., et al., *Occurrence, function and evolutionary origins of '2A-like' sequences in virus genomes.* J Gen Virol, 2008. **89**(Pt 4): p. 1036-42.

212. Ulper, L., et al., *Construction, properties, and potential application of infectious plasmids containing Semliki Forest virus full-length cDNA with an inserted intron.* J Virol Methods, 2008. **148**(1-2): p. 265-70.

213. de Vries, A.A., et al., *Structural proteins of equine arteritis virus.* J Virol, 1992. **66**(11): p. 6294-303.

214. Sugiyama, N., et al., *Phosphopeptide enrichment by aliphatic hydroxy acid-modified metal oxide chromatography for nano-LC-MS/MS in proteomics applications.* Mol Cell Proteomics, 2007. **6**(6): p. 1103-9.

215. Szklarczyk, D., et al., *STRING v10: protein-protein interaction networks, integrated over the tree of life.* Nucleic Acids Res, 2015. **43**(Database issue): p. D447-52.

216. Olagnier, D., et al., *Inhibition of dengue and chikungunya virus infections by RIG-I-mediated type I interferon-independent stimulation of the innate antiviral response.* J Virol, 2014. **88**(8): p. 4180-94.

217. van Marle, G., et al., *Regulation of coronavirus mRNA transcription.* J Virol, 1995. **69**(12): p. 7851-6.

218. de Vries, A.A., et al., *All subgenomic mRNAs of equine arteritis virus contain a common leader sequence.* Nucleic Acids Res, 1990. **18**(11): p. 3241-7.

219. Ashburner, M., et al., *Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.* Nat Genet, 2000. **25**(1): p. 25-9.

220. *Gene Ontology Consortium: going forward.* Nucleic Acids Res, 2015. **43**(Database issue): p. D1049-56.

221. Price, N.T., et al., *Identification of the phosphorylation sites in elongation factor-2 from rabbit reticulocytes.* FEBS Letters, 1991. **282**(2): p. 253-258.

222. Kaul, G., G. Pattan, and T. Rafeequi, *Eukaryotic elongation factor-2 (eEF2): its regulation and peptide chain elongation.* Cell Biochem Funct, 2011. **29**(3): p. 227-34.

223. Diggle, T.A., et al., *Phosphorylation of elongation factor-2 kinase on serine 499 by cAMP-dependent protein kinase induces Ca2+/calmodulin-independent activity.* Biochem J, 2001. **353**(Pt 3): p. 621-6.

224. Redpath, N.T. and C.G. Proud, *Cyclic AMP-dependent protein kinase phosphorylates rabbit reticulocyte elongation factor-2 kinase and induces calcium-independent activity.* Biochem J, 1993. **293 (Pt 1)**: p. 31-4.

225. Kenney, J.W., et al., *Eukaryotic elongation factor 2 kinase, an unusual enzyme with multiple roles.* Adv Biol Regul, 2014. **55**: p. 15-27.

226. Chandra, V., et al., *The ORF3 protein of hepatitis E virus delays degradation of activated growth factor receptors by interacting with CIN85 and blocking formation of the Cbl-CIN85 complex.* J Virol, 2010. **84**(8): p. 3857-67.

227. Narita, T., et al., *Overexpression of CIN85 suppresses the growth of herpes simplex virus in HeLa cells.* Exp Cell Res, 2005. **311**(2): p. 265-71.

228. Porter, F.W. and A.C. Palmenberg, *Leader-induced phosphorylation of nucleoporins correlates with nuclear trafficking inhibition by cardioviruses.* J Virol, 2009. **83**(4): p. 1941-51.

229. Bardina, M.V., et al., *Mengovirus-induced rearrangement of the nuclear pore complex: hijacking cellular phosphorylation machinery.* J Virol, 2009. **83**(7): p. 3150-61.

230. Ciomperlik, J.J., H.A. Basta, and A.C. Palmenberg, *Three cardiovirus Leader proteins equivalently inhibit four different nucleocytoplasmic trafficking pathways.* Virology, 2015. **484**: p. 194-202.

231. Porter, F.W., B. Brown, and A.C. Palmenberg, *Nucleoporin phosphorylation triggered by the encephalomyocarditis virus leader protein is mediated by mitogen-activated protein kinases.* J Virol, 2010. **84**(24): p. 12538-48.

232. Vard, C., et al., *A specific role for the phosphorylation of mammalian acidic ribosomal protein P2.* J Biol Chem, 1997. **272**(32): p. 20259-62.

233. Hasler, P., et al., *Ribosomal proteins P0, P1, and P2 are phosphorylated by casein kinase II at their conserved carboxyl termini.* J Biol Chem, 1991. **266**(21): p. 13815-20.

234. Montgomery, S.A., et al., *Ribosomal protein S6 associates with alphavirus nonstructural protein 2 and mediates expression from alphavirus messages.* J Virol, 2006. **80**(15): p. 7729-39.

235. Moser, T.S., D. Schieffer, and S. Cherry, *AMP-activated kinase restricts Rift Valley fever virus infection by inhibiting fatty acid synthesis.* PLoS Pathog, 2012. **8**(4): p. e1002661.

236. Luo, R., et al., *Label-free quantitative phosphoproteomic analysis reveals differentially regulated proteins and pathway in PRRSV-infected pulmonary alveolar macrophages.* J Proteome Res, 2014. **13**(3): p. 1270-80.

237. Dapat, C., et al., *Quantitative phosphoproteomic analysis of host responses in human lung epithelial (A549) cells during influenza virus infection.* Virus Res, 2014. **179**: p. 53-63.

238. Ohman, T., et al., *Phosphoproteome characterization reveals that Sendai virus infection activates mTOR signaling in human epithelial cells.* Proteomics, 2015. **15**(12): p. 2087-97.

239. Wojcechowskyj, J.A., et al., *Quantitative phosphoproteomics reveals extensive cellular reprogramming during HIV-1 entry.* Cell Host Microbe, 2013. **13**(5): p. 613-23.

240. Taylor, D.J., et al., *Structures of modified eEF2 80S ribosome complexes reveal the role of GTP hydrolysis in translocation.* Embo j, 2007. **26**(9): p. 2421-31.

241. Carlberg, U., A. Nilsson, and O. Nygard, *Functional properties of phosphorylated elongation factor 2.* Eur J Biochem, 1990. **191**(3): p. 639-45.

242. Redpath, N.T., et al., *Regulation of elongation factor-2 by multisite phosphorylation.* Eur J Biochem, 1993. **213**(2): p. 689-99.

243. Ryazanov, A.G., et al., *Phosphorylation of the elongation factor 2: the fifth Ca2+/calmodulin-dependent system of protein phosphorylation.* Biochimie, 1988. **70**(5): p. 619-26.

244. Tavares, C.D., et al., *The molecular mechanism of eukaryotic elongation factor 2 kinase activation.* J Biol Chem, 2014. **289**(34): p. 23901-16.

245. Hardie, D.G., F.A. Ross, and S.A. Hawley, *AMPK: a nutrient and energy sensor that maintains energy homeostasis.* Nat Rev Mol Cell Biol, 2012. **13**(4): p. 251-62.

246. Hong-Brown, L.Q., et al., *Alcohol regulates eukaryotic elongation factor 2 phosphorylation via an AMP-activated protein kinase-dependent mechanism in C2C12 skeletal myocytes.* J Biol Chem, 2007. **282**(6): p. 3702-12.

247. Knebel, A., et al., *Stress-induced regulation of eukaryotic elongation factor 2 kinase by SB 203580-sensitive and -insensitive pathways.* Biochem J, 2002. **367**(Pt 2): p. 525-32.

248. Knebel, A., N. Morrice, and P. Cohen, *A novel method to identify protein kinase substrates: eEF2 kinase is phosphorylated and inhibited by SAPK4/p38delta.* Embo j, 2001. **20**(16): p. 4360-9.

249. Wang, X., et al., *Regulation of elongation factor 2 kinase by p90(RSK1) and p70 S6 kinase.* Embo j, 2001. **20**(16): p. 4370-9.

250. Dorovkov, M.V., et al., *Regulation of elongation factor-2 kinase by pH.* Biochemistry, 2002. **41**(45): p. 13444-50.

251. Gschwendt, M., et al., *A type 2A protein phosphatase dephosphorylates the elongation factor 2 and is stimulated by the phorbol ester TPA in mouse epidermis in vivo.* FEBS Lett, 1989. **257**(2): p. 357-60.

252. Janeway, C.A., Jr. and R. Medzhitov, *Innate immune recognition.* Annu Rev Immunol, 2002. **20**: p. 197-216.

253. Kumar, H., T. Kawai, and S. Akira, *Pathogen recognition by the innate immune system.* Int Rev Immunol, 2011. **30**(1): p. 16-34.

254. Geddes, K., J.G. Magalhaes, and S.E. Girardin, *Unleashing the therapeutic potential of NOD-like receptors.* Nat Rev Drug Discov, 2009. **8**(6): p. 465-479.

255. Tsalikis, J., et al., *Nutrient sensing and metabolic stress pathways in innate immunity.* Cell Microbiol, 2013. **15**(10): p. 1632-41.

256. Kemp, B.E., et al., *Dealing with energy demand: the AMP-activated protein kinase.* Trends Biochem Sci, 1999. **24**(1): p. 22-5.

257. Wek, R.C., H.Y. Jiang, and T.G. Anthony, *Coping with stress: eIF2 kinases and translational control.* Biochem Soc Trans, 2006. **34**(Pt 1): p. 7-11.

258. Mankouri, J. and M. Harris, *Viruses and the fuel sensor: the emerging link between AMPK and virus replication.* Rev Med Virol, 2011. **21**(4): p. 205-12.

259. Warden, S.M., et al., *Post-translational modifications of the beta-1 subunit of AMP-activated protein kinase affect enzyme activity and cellular localization.* Biochem J, 2001. **354**(Pt 2): p. 275-83.

260. Dodd, K.M. and A.R. Tee, *Leucine and mTORC1: a complex relationship.* Am J Physiol Endocrinol Metab, 2012. **302**(11): p. E1329-42.

261. Proud, C.G., *mTOR-mediated regulation of translation factors by amino acids.* Biochemical and Biophysical Research Communications, 2004. **313**(2): p. 429-436.

262. Roskoski Jr, R., *ERK1/2 MAP kinases: Structure, function, and regulation.* Pharmacological Research, 2012. **66**(2): p. 105-143.

263. Antoine, A.F., et al., *The alphavirus 6K protein activates endogenous ionic conductances when expressed in Xenopus oocytes.* J Membr Biol, 2007. **215**(1): p. 37-48.

264. Valiente-Echeverría, F., et al., *eEF2 and Ras-GAP SH3 domain-binding protein (G3BP1) modulate stress granule assembly during HIV-1 infection.* Nat Commun, 2014. **5**: p. 4819.

265. Peterhans, E., et al., *Mitochondrial calcium uptake during infection of chicken embryo cells with Semliki Forest virus.* J Virol, 1979. **29**(1): p. 143-52.

266. Ubol, S., et al., *Temporal changes in chromatin, intracellular calcium, and poly(ADP-ribose) polymerase during Sindbis virus-induced apoptosis of neuroblastoma cells.* J Virol, 1996. **70**(4): p. 2215-20.

267. Sanz, M.A., et al., *Dual mechanism for the translation of subgenomic mRNA from Sindbis virus in infected and uninfected cells.* PLoS One, 2009. **4**(3): p. e4772.

268. Barrero, C.A., et al., *HIV-1 Vpr modulates macrophage metabolic pathways: a SILAC-based quantitative analysis.* PLoS One, 2013. **8**(7): p. e68376.

269. Lee, S. and C. Lee, *Functional characterization and proteomic analysis of the nucleocapsid protein of porcine deltacoronavirus.* Virus Res, 2015. **208**: p. 136-145.

270. Vogels, M.W., et al., *Identification of host factors involved in coronavirus replication by quantitative proteomics analysis.* Proteomics, 2011. **11**(1): p. 64-80.

271. Horner, S.M., et al., *Proteomic analysis of mitochondrial-associated ER membranes (MAM) during RNA virus infection reveals dynamic changes in protein and organelle trafficking.* PLoS One, 2015. **10**(3): p. e0117963.

272. Li, M., et al., *Quantitative proteomic analysis of exosomes from HIV-1-infected lymphocytic cells.* Proteomics, 2012. **12**(13): p. 2203-11.

273. Morikawa, K., et al., *Quantitative proteomics identifies the membrane-associated peroxidase GPx8 as a cellular substrate of the hepatitis C virus NS3-4A protease.* Hepatology, 2014. **59**(2): p. 423-33.

274. Jourdan, S.S., F. Osorio, and J.A. Hiscox, *An interactome map of the nucleocapsid protein from a highly pathogenic North American porcine reproductive and respiratory syndrome virus strain generated using SILAC-based quantitative proteomics.* Proteomics, 2012. **12**(7): p. 1015-23.

275. Bard-Chapeau, E.A., et al., *EVI1 oncoprotein interacts with a large and complex network of proteins and integrates signals through protein phosphorylation.* Proc Natl Acad Sci U S A, 2013. **110**(31): p. E2885-94.

276. Olsen, J.V. and M. Mann, *Status of large-scale analysis of post-translational modifications by mass spectrometry.* Mol Cell Proteomics, 2013. **12**(12): p. 3444-52.

277. Bailey-Elkin, B.A., et al., *Crystal structure of the Middle East respiratory syndrome coronavirus (MERS-CoV) papain-like protease bound to ubiquitin facilitates targeted disruption of deubiquitinating activity to demonstrate its role in innate immune suppression.* J Biol Chem, 2014. **289**(50): p. 34667-82.

278. Whitehurst, C.B., et al., *Epstein-Barr virus BPLF1 deubiquitinates PCNA and attenuates polymerase eta recruitment to DNA damage sites.* J Virol, 2012. **86**(15): p. 8097-106.

279. Chenon, M., et al., *A viral deubiquitylating enzyme targets viral RNA-dependent RNA polymerase and affects viral infectivity.* EMBO J, 2012. **31**(3): p. 741-53.

280. Liu, X., et al., *Human borna disease virus infection impacts host proteome and histone lysine acetylation in human oligodendroglia cells.* Virology, 2014. **464-465**: p. 196-205.

281. Minguez, P., et al., *Deciphering a global network of functionally associated post-translational modifications.* Mol Syst Biol, 2012. **8**: p. 599.

282. Hunter, T., *The age of crosstalk: phosphorylation, ubiquitination, and beyond.* Mol Cell, 2007. **28**(5): p. 730-8.

283. Hornbeck, P.V., et al., *PhosphoSite: A bioinformatics resource dedicated to physiological protein phosphorylation.* Proteomics, 2004. **4**(6): p. 1551-61.

284. Smith, D.R., *Global protein profiling studies of chikungunya virus infection identify different proteins but common biological processes.* Rev Med Virol, 2015. **25**(1): p. 3-18.

285. Coombs, K.M., et al., *Quantitative proteomic analyses of influenza virus-infected cultured human lung cells.* J Virol, 2010. **84**(20): p. 10888-906.

286. Dove, B.K., et al., *A quantitative proteomic analysis of lung epithelial (A549) cells infected with 2009 pandemic influenza A virus using stable isotope labelling with amino acids in cell culture.* Proteomics, 2012. **12**(9): p. 1431-6.

287. Liu, L., et al., *Proteome alterations in primary human alveolar macrophages in response to influenza A virus infection.* J Proteome Res, 2012. **11**(8): p. 4091-101.

288. Kramer, G., et al., *Proteomic analysis of HIV–T cell interaction: an update.* Frontiers in Microbiology, 2012. **3**: p. 240.

289. Fahey, M.E., et al., *GPS-Prot: a web-based visualization platform for integrating host-pathogen interaction data.* BMC Bioinformatics, 2011. **12**: p. 298.

290. Bhinder, B. and H. Djaballah, *Systematic analysis of RNAi reports identifies dismal commonality at gene-level and reveals an unprecedented enrichment in pooled shRNA screens.* Comb Chem High Throughput Screen, 2013. **16**(9): p. 665-81.

291. Cherry, S., *What have RNAi screens taught us about viral-host interactions?* Current opinion in microbiology, 2009. **12**(4): p. 446-452.

292. Kaelin, W.G., Jr., *Molecular biology. Use and abuse of RNAi to study mammalian gene function.* Science, 2012. **337**(6093): p. 421-2.

293. Tabb, D.L., et al., *Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry.* J Proteome Res, 2010. **9**(2): p. 761-76.

294. Merl, J., et al., *Direct comparison of MS-based label-free and SILAC quantitative proteome profiling strategies in primary retinal Muller cells.* Proteomics, 2012. **12**(12): p. 1902-11.

295. Liu, N.Q., et al., *Quantitative proteomic analysis of microdissected breast cancer tissues: comparison of label-free and SILAC-based quantification with shotgun, directed, and targeted MS approaches.* J Proteome Res, 2013. **12**(10): p. 4627-41.

296. Oppermann, F.S., et al., *Comparison of SILAC and mTRAQ quantification for phosphoproteomics on a quadrupole orbitrap mass spectrometer.* J Proteome Res, 2013. **12**(9): p. 4089-100.

297. Noble, W.S., *How does multiple testing correction work?* Nature biotechnology, 2009. **27**(12): p. 1135-1137.

298. Ong, S.E., I. Kratchmarova, and M. Mann, *Properties of 13C-substituted arginine in stable isotope labeling by amino acids in cell culture (SILAC).* J Proteome Res, 2003. **2**(2): p. 173-81.

299. Lößner, C., et al., *Preventing arginine-to-proline conversion in a cell-line-independent manner during cell cultivation under stable isotope labeling by amino acids in cell culture (SILAC) conditions.* Analytical Biochemistry, 2011. **412**(1): p. 123-125.

300. Bendall, S.C., et al., *Prevention of amino acid conversion in SILAC experiments with embryonic stem cells.* Mol Cell Proteomics, 2008. **7**(9): p. 1587-97.

301. Blagoev, B. and M. Mann, *Quantitative proteomics to study mitogen-activated protein kinases.* Methods, 2006. **40**(3): p. 243-50.

302. Park, S.K., et al., *A computational approach to correct arginine-to-proline conversion in quantitative proteomics.* Nat Methods, 2009. **6**(3): p. 184-5.

303. Marcilla, M., et al., *A systematic approach to assess amino acid conversions in SILAC experiments.* Talanta, 2011. **84**(2): p. 430-6.

304. Zhang, L.-K., et al., *Identification of Host Proteins Involved in Japanese Encephalitis Virus Infection by Quantitative Proteomics Analysis.* Journal of Proteome Research, 2013. **12**(6): p. 2666-2678.

305. Lam, Y.W., et al., *Proteomics analysis of the nucleolus in adenovirus-infected cells.* Mol Cell Proteomics, 2010. **9**(1): p. 117-30.

306. Park, S.S., et al., *Effective correction of experimental errors in quantitative proteomics using stable isotope labeling by amino acids in cell culture (SILAC).* J Proteomics, 2012. **75**(12): p. 3720-32.

307. Rodnina, M.V. and W. Wintermeyer, *Recent mechanistic insights into eukaryotic ribosomes.* Curr Opin Cell Biol, 2009. **21**(3): p. 435-43.

308. Dever, T.E. and R. Green, *The elongation, termination, and recycling phases of translation in eukaryotes.* Cold Spring Harb Perspect Biol, 2012. **4**(7): p. a013706.

309. Hinnebusch, A.G. and J.R. Lorsch, *The mechanism of eukaryotic translation initiation: new insights and challenges.* Cold Spring Harb Perspect Biol, 2012. **4**(10).

310. Jackson, R.J., C.U. Hellen, and T.V. Pestova, *The mechanism of eukaryotic translation initiation and principles of its regulation.* Nat Rev Mol Cell Biol, 2010. **11**(2): p. 113-27.

311. Jackson, R.J., C.U. Hellen, and T.V. Pestova, *Termination and post-termination events in eukaryotic translation.* Adv Protein Chem Struct Biol, 2012. **86**: p. 45-93.

312. Aitken, C.E. and J.R. Lorsch, *A mechanistic overview of translation initiation in eukaryotes.* Nat Struct Mol Biol, 2012. **19**(6): p. 568-76.

313. Kozak, M., *Initiation of translation in prokaryotes and eukaryotes.* Gene, 1999. **234**(2): p. 187-208.

314. Maag, D., et al., *A conformational change in the eukaryotic translation preinitiation complex and release of eIF1 signal recognition of the start codon.* Mol Cell, 2005. **17**(2): p. 265-75.

315. Algire, M.A., D. Maag, and J.R. Lorsch, *Pi release from eIF2, not GTP hydrolysis, is the step controlled by start-site selection during eukaryotic translation initiation.* Mol Cell, 2005. **20**(2): p. 251-62.

316. Rheinberger, H.J., H. Sternbach, and K.H. Nierhaus, *Three tRNA binding sites on Escherichia coli ribosomes.* Proc Natl Acad Sci U S A, 1981. **78**(9): p. 5310-4.

317. Beringer, M. and M.V. Rodnina, *The ribosomal peptidyl transferase.* Mol Cell, 2007. **26**(3): p. 311-21.

318. Cornish, P.V., et al., *Spontaneous intersubunit rotation in single ribosomes.* Mol Cell, 2008. **30**(5): p. 578-88.

319. Hauryliuk, V., et al., *Class-1 release factor eRF1 promotes GTP binding by class-2 release factor eRF3.* Biochimie, 2006. **88**(7): p. 747-57.

320. Kisselev, L., M. Ehrenberg, and L. Frolova, *Termination of translation: interplay of mRNA, rRNAs and release factors?* Embo j, 2003. **22**(2): p. 175-82.

321. Pisarev, A.V., et al., *The role of ABCE1 in eukaryotic posttermination ribosomal recycling.* Mol Cell, 2010. **37**(2): p. 196-210.

322. Ferron, F., et al., *The viral RNA capping machinery as a target for antiviral drugs.* Antiviral Res, 2012. **96**(1): p. 21-31.

323. Firth, A.E. and I. Brierley, *Non-canonical translation in RNA viruses.* J Gen Virol, 2012. **93**(Pt 7): p. 1385-409.

324. Au, H.H. and E. Jan, *Novel viral translation strategies.* Wiley Interdisc Rev RNA, 2014. **5**(6): p. 779-801.

325. Walsh, D., M.B. Mathews, and I. Mohr, *Tinkering with translation: protein synthesis in virus-infected cells.* Cold Spring Harb Perspect Biol, 2013. **5**(1): p. a012351.

326. Kieft, J.S., *Viral IRES RNA structures and ribosome interactions.* Trends Biochem Sci, 2008. **33**(6): p. 274-83.

327. Kong, W.P. and R.P. Roos, *Alternative translation initiation site in the DA strain of Theiler's murine encephalomyelitis virus.* J Virol, 1991. **65**(6): p. 3395-9.

328. Clarke, B.E., et al., *Two initiation sites for foot-and-mouth disease virus polyprotein in vivo.* J Gen Virol, 1985. **66 (Pt 12)**: p. 2615-26.

329. Woo, P.C., et al., *Natural occurrence and characterization of two internal ribosome entry site elements in a novel virus, canine picodicistrovirus, in the picornavirus-like superfamily.* J Virol, 2012. **86**(5): p. 2797-808.

330. Kozak, M., *Pushing the limits of the scanning mechanism for initiation of translation.* Gene, 2002. **299**(1-2): p. 1-34.

331. Ryabova, L.A., M.M. Pooggin, and T. Hohn, *Viral strategies of translation initiation: ribosomal shunt and reinitiation.* Prog Nucleic Acid Res Mol Biol, 2002. **72**: p. 1-39.

332. Meyers, G., *Translation of the minor capsid protein of a calicivirus is initiated by a novel termination-dependent reinitiation mechanism.* J Biol Chem, 2003. **278**(36): p. 34051-60.

333. Luttermann, C. and G. Meyers, *The importance of inter- and intramolecular base pairing for translation reinitiation on a eukaryotic bicistronic mRNA.* Genes Dev, 2009. **23**(3): p. 331-44.

334. Poyry, T.A., et al., *The mechanism of an exceptional case of reinitiation after translation of a long ORF reveals why such events do not generally occur in mammalian mRNA translation.* Genes Dev, 2007. **21**(23): p. 3149-62.

335. Beier, H. and M. Grimm, *Misreading of termination codons in eukaryotes by natural nonsense suppressor tRNAs.* Nucleic Acids Res, 2001. **29**(23): p. 4767-82.

336. Giedroc, D.P. and P.V. Cornish, *Frameshifting RNA pseudoknots: structure and mechanism.* Virus Res, 2009. **139**(2): p. 193-208.

337. Plant, E.P. and J.D. Dinman, *Torsional restraint: a new twist on frameshifting pseudoknots.* Nucleic Acids Res, 2005. **33**(6): p. 1825-33.

338. Caliskan, N., et al., *Programmed −1 Frameshifting by Kinetic Partitioning during Impeded Translocation.* Cell, 2014. **157**(7): p. 1619-1631.

339. Snijder, E.J., M. Kikkert, and Y. Fang, *Arterivirus molecular biology and pathogenesis.* J Gen Virol, 2013. **94**(Pt 10): p. 2141-63.

340. Holtkamp, D.J., et al., *Assessment of the economic impact of porcine reproductive and respiratory syndrome virus on United States pork producers.* Journal of Swine Health and Production, 2013. **21**(2): p. 72-84.

341. Collins, J.E., et al., *Isolation of swine infertility and respiratory syndrome virus (isolate ATCC VR-2332) in North America and experimental reproduction of the disease in gnotobiotic pigs.* J Vet Diagn Invest, 1992. **4**(2): p. 117-26.

342. Nelsen, C.J., M.P. Murtaugh, and K.S. Faaberg, *Porcine reproductive and respiratory syndrome virus comparison: divergent evolution on two continents.* J Virol, 1999. **73**(1): p. 270-80.

343. Allende, R., et al., *North American and European porcine reproductive and respiratory syndrome viruses differ in non-structural protein coding regions.* J Gen Virol, 1999. **80 (Pt 2)**: p. 307-15.

344. Shi, M., et al., *Molecular epidemiology of PRRSV: a phylogenetic perspective.* Virus Res, 2010. **154**(1-2): p. 7-17.

345. den Boon, J.A., et al., *Equine arteritis virus is not a togavirus but belongs to the coronaviruslike superfamily.* J Virol, 1991. **65**(6): p. 2910-20.

346. Snijder, E.J., et al., *The carboxyl-terminal part of the putative Berne virus polymerase is expressed by ribosomal frameshifting and contains sequence motifs which indicate that toro- and coronaviruses are evolutionarily related.* Nucleic Acids Res, 1990. **18**(15): p. 4535-42.

347. Gorbalenya, A.E., et al., *Nidovirales: evolving the largest RNA virus genome.* Virus Res, 2006. **117**(1): p. 17-37.

348. Lauber, C., et al., *Mesoniviridae: a proposed new family in the order Nidovirales formed by a single species of mosquito-borne viruses.* Arch Virol, 2012. **157**(8): p. 1623-8.

349. Dijkman, R. and L. van der Hoek, *Human coronaviruses 229E and NL63: close yet still so far.* J Formos Med Assoc, 2009. **108**(4): p. 270-9.

350. Zaki, A.M., et al., *Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia.* N Engl J Med, 2012. **367**(19): p. 1814-20.

351. Dunowska, M., et al., *Identification of a novel nidovirus associated with a neurological disease of the Australian brushtail possum (Trichosurus vulpecula).* Vet Microbiol, 2012. **156**(3-4): p. 418-24.

352. Lauck, M., et al., *Exceptional simian hemorrhagic fever virus diversity in a wild African primate community.* J Virol, 2013. **87**(1): p. 688-91.

353. Dokland, T., *The structural biology of PRRSV.* Virus Res, 2010. **154**(1-2): p. 86-97.

354. Firth, A.E., et al., *Discovery of a small arterivirus gene that overlaps the GP5 coding sequence and is important for virus production.* J Gen Virol, 2011. **92**(Pt 5): p. 1097-106.

355. Johnson, C.R., et al., *Novel structural protein in porcine reproductive and respiratory syndrome virus encoded by an alternative ORF5 present in all arteriviruses.* J Gen Virol, 2011. **92**(Pt 5): p. 1107-16.

356. Wissink, E.H., et al., *Envelope protein requirements for the assembly of infectious virions of porcine reproductive and respiratory syndrome virus.* J Virol, 2005. **79**(19): p. 12495-506.

357. Sun, L., et al., *Porcine reproductive and respiratory syndrome virus ORF5a protein is essential for virus viability.* Virus Res, 2013. **171**(1): p. 178-85.

358. Kreutz, L.C. and M.R. Ackermann, *Porcine reproductive and respiratory syndrome virus enters cells through a low pH-dependent endocytic pathway.* Virus Res, 1996. **42**(1-2): p. 137-47.

359. Nauwynck, H.J., et al., *Entry of porcine reproductive and respiratory syndrome virus into porcine alveolar macrophages via receptor-mediated endocytosis.* J Gen Virol, 1999. **80 (Pt 2)**: p. 297-305.

360. Van Gorp, H., et al., *Sialoadhesin and CD163 join forces during entry of the porcine reproductive and respiratory syndrome virus.* J Gen Virol, 2008. **89**(Pt 12): p. 2943-53.

361. Fang, Y. and E.J. Snijder, *The PRRSV replicase: exploring the multifunctionality of an intriguing set of nonstructural proteins.* Virus Res, 2010. **154**(1-2): p. 61-76.

362. Ziebuhr, J., E.J. Snijder, and A.E. Gorbalenya, *Virus-encoded proteinases and proteolytic processing in the Nidovirales.* J Gen Virol, 2000. **81**(Pt 4): p. 853-79.

363. Li, Y., et al., *Proteolytic Processing of the Porcine Reproductive and Respiratory Syndrome Virus Replicase.* Virus Res, 2014.

364. Snijder, E.J., et al., *The arterivirus nsp4 protease is the prototype of a novel group of chymotrypsin-like enzymes, the 3C-like serine proteases.* J Biol Chem, 1996. **271**(9): p. 4864-71.

365. Meulenberg, J.J., et al., *Lelystad virus, the causative agent of porcine epidemic abortion and respiratory syndrome (PEARS), is related to LDV and EAV.* Virology, 1993. **192**(1): p. 62-72.

366. Conzelmann, K.-K., et al., *Molecular Characterization of Porcine Reproductive and Respiratory Syndrome Virus, a Member of the Arterivirus Group.* Virology, 1993. **193**(1): p. 329-339.

367. Snijder, E.J., et al., *Identification of a novel structural protein of arteriviruses.* J Virol, 1999. **73**(8): p. 6335-45.

368. van der Meer, Y., et al., *ORF1a-encoded replicase subunits are involved in the membrane association of the arterivirus replication complex.* J Virol, 1998. **72**(8): p. 6689-98.

369. Knoops, K., et al., *Ultrastructural characterization of arterivirus replication structures: reshaping the endoplasmic reticulum to accommodate viral RNA synthesis.* J Virol, 2012. **86**(5): p. 2474-87.

370. Pol, J.M., F. Wagenaar, and J.E. Reus, *Comparative morphogenesis of three PRRS virus strains.* Vet Microbiol, 1997. **55**(1-4): p. 203-8.

371. Pasternak, A.O., W.J. Spaan, and E.J. Snijder, *Nidovirus transcription: how to make sense...?* J Gen Virol, 2006. **87**(Pt 6): p. 1403-21.

372. Dea, S., et al., *Current knowledge on the structural proteins of porcine reproductive and respiratory syndrome (PRRS) virus: comparison of the North American and European isolates.* Arch Virol, 2000. **145**(4): p. 659-88.

373. Magnusson, P., B. Hyllseth, and H. Marusyk, *Morphological studies on equine arteritis virus.* Arch Gesamte Virusforsch, 1970. **30**(2): p. 105-12.

374. Wood, O., N. Tauraso, and H. Liebhaber, *Electron microscopic study of tissue cultures infected with simian haemorrhagic fever virus.* J Gen Virol, 1970. **7**(2): p. 129-36.

375. Veit, M., et al., *Membrane proteins of arterivirus particles: structure, topology, processing and function.* Virus Res, 2014. **194**: p. 16-36.

376. Christopher-Hennings, J., et al., *Detection of porcine reproductive and respiratory syndrome virus in boar semen by PCR.* J Clin Microbiol, 1995. **33**(7): p. 1730-4.

377.	Kang, I., et al., *Localization of porcine reproductive and respiratory syndrome virus in mammary glands of experimentally infected sows.* Res Vet Sci, 2010. **88**(2): p. 304-6.

378.	Kristensen, C.S., et al., *Experimental airborne transmission of PRRS virus.* Vet Microbiol, 2004. **99**(3-4): p. 197-202.

379.	Murtaugh, M.P., et al., *The ever-expanding diversity of porcine reproductive and respiratory syndrome virus.* Virus Res, 2010. **154**(1-2): p. 18-30.

380.	Halbur, P.G., et al., *Comparative pathogenicity of nine US porcine reproductive and respiratory syndrome virus (PRRSV) isolates in a five-week-old cesarean-derived, colostrum-deprived pig model.* J Vet Diagn Invest, 1996. **8**(1): p. 11-20.

381.	Weesendorp, E., et al., *Comparative analysis of immune responses following experimental infection of pigs with European porcine reproductive and respiratory syndrome virus strains of differing virulence.* Vet Microbiol, 2013. **163**(1-2): p. 1-12.

382.	Duan, X., H.J. Nauwynck, and M.B. Pensaert, *Effects of origin and state of differentiation and activation of monocytes/macrophages on their susceptibility to porcine reproductive and respiratory syndrome virus (PRRSV).* Arch Virol, 1997. **142**(12): p. 2483-97.

383.	Lamontagne, L., et al., *Porcine reproductive and respiratory syndrome virus persistence in blood, spleen, lymph nodes, and tonsils of experimentally infected pigs depends on the level of CD8high T cells.* Viral Immunol, 2003. **16**(3): p. 395-406.

384.	Duan, X., H.J. Nauwynck, and M.B. Pensaert, *Virus quantification and identification of cellular targets in the lungs and lymphoid tissues of pigs at different time intervals after inoculation with porcine reproductive and respiratory syndrome virus (PRRSV).* Vet Microbiol, 1997. **56**(1-2): p. 9-19.

385.	Lopez, O.J. and F.A. Osorio, *Role of neutralizing antibodies in PRRSV protective immunity.* Vet Immunol Immunopathol, 2004. **102**(3): p. 155-63.

386.	Miller, L.C., et al., *Interferon type I response in porcine reproductive and respiratory syndrome virus-infected MARC-145 cells.* Arch Virol, 2004. **149**(12): p. 2453-63.

387.	Zhu, L., Y. Zhou, and G. Tong, *Mechanisms of suppression of interferon production by porcine reproductive and respiratory syndrome virus.* Acta Virol, 2012. **56**(1): p. 3-9.

388.	Darwich, L., I. Diaz, and E. Mateu, *Certainties, doubts and hypotheses in porcine reproductive and respiratory syndrome virus immunobiology.* Virus Res, 2010. **154**(1-2): p. 123-32.

389.	Murtaugh, M.P., Z. Xiao, and F. Zuckermann, *Immunological responses of swine to porcine reproductive and respiratory syndrome virus infection.* Viral Immunol, 2002. **15**(4): p. 533-47.

390.	Charerntantanakul, W., *Porcine reproductive and respiratory syndrome virus vaccines: Immunogenicity, efficacy and safety aspects.* World J Virol, 2012. **1**(1): p. 23-30.

391.	Kimman, T.G., et al., *Challenges for porcine reproductive and respiratory syndrome virus (PRRSV) vaccinology.* Vaccine, 2009. **27**(28): p. 3704-18.

392.	Snijder, E.J., Kikkert, M., *Arteriviruses*, in *Fields Virology*, D.M. Knipe and P.M. Howley, Editors. 2013, Lippincott Williams & Wilkins: Philadelphia, PA. p. 859-879.

393.	de Groot, R.J., J.A. Cowley, and L. Enjuanes, *Order Nidovirales*, in *Virus Taxonomy: Ninth report of teh International Committee on Taxonomy of Viruses*, A.M.Q. King, et al., Editors. 2012, Academic Press: Waltham MA. p. 785-795.

394.	Snijder, E.J. and J.J. Meulenberg, *The molecular biology of arteriviruses.* J Gen Virol, 1998. **79 (Pt 5)**: p. 961-79.

395.	Brierley, I., P. Digard, and S.C. Inglis, *Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot.* Cell, 1989. **57**(4): p. 537-47.

396. Brierley, I., Gilbert, R. J. C. & Pennell, S, *Pseudoknot-dependent -1 programmed ribosomal frame-shifting: structures, mechanisms and models*, in *Recoding: Expansion of Decoding Rules Enriches Gene Expression*, J.F.A.R.F. Gesteland, Editor. 2010, Springer: Heidelberg. p. 149–174.

397. Miller, W.A. and D.P. Giedroc, *Ribosomal frameshifting in decoding plant viral RNAs*, in *Recoding: Expansion of Decoding Rules Enriches Gene Expression*, J.F. Atkins and R.F. Gesteland, Editors. 2010, Springer: Heidelberg. p. 193-220.

398. Firth, A.E., et al., *Stimulation of stop codon readthrough: frequent presence of an extended 3' RNA structural element.* Nucleic Acids Res, 2011. **39**(15): p. 6679-91.

399. Loughran, G., A.E. Firth, and J.F. Atkins, *Ribosomal frameshifting into an overlapping gene in the 2B-encoding region of the cardiovirus genome.* Proc Natl Acad Sci U S A, 2011. **108**(46): p. E1111-9.

400. Jagger, B.W., et al., *An overlapping protein-coding region in influenza A virus segment 3 modulates the host response.* Science, 2012. **337**(6091): p. 199-204.

401. Lauck, M., et al., *Novel, divergent simian hemorrhagic fever viruses in a wild Ugandan red colobus monkey discovered using direct pyrosequencing.* PLoS One, 2011. **6**(4): p. e19056.

402. Fang, Y., et al., *A full-length cDNA infectious clone of North American type 1 porcine reproductive and respiratory syndrome virus: expression of green fluorescent protein in the Nsp2 region.* J Virol, 2006. **80**(23): p. 11447-55.

403. Li, Y., et al., *Identification of porcine reproductive and respiratory syndrome virus ORF1a-encoded non-structural proteins in virus-infected cells.* J Gen Virol, 2012. **93**(Pt 4): p. 829-39.

404. Pedersen, K.W., et al., *Open reading frame 1a-encoded subunits of the arterivirus replicase induce endoplasmic reticulum-derived double-membrane vesicles which carry the viral replication complex.* J Virol, 1999. **73**(3): p. 2016-26.

405. Ivanov, I.P. and J.F. Atkins, *Ribosomal frameshifting in decoding antizyme mRNAs from yeast and protists to humans: close to 300 cases reveal remarkable diversity despite underlying conservation.* Nucleic Acids Res, 2007. **35**(6): p. 1842-58.

406. Matsufuji, S., et al., *Reading two bases twice: mammalian antizyme frameshifting in yeast.* EMBO J, 1996. **15**(6): p. 1360-70.

407. Lin, Z., R.J. Gilbert, and I. Brierley, *Spacer-length dependence of programmed -1 or -2 ribosomal frameshifting on a U6A heptamer supports a role for messenger RNA (mRNA) tension in frameshifting.* Nucleic Acids Res, 2012. **40**(17): p. 8674-89.

408. Xu, J., R.W. Hendrix, and R.L. Duda, *Conserved translational frameshift in dsDNA bacteriophage tail assembly genes.* Mol Cell, 2004. **16**(1): p. 11-21.

409. Liu, H.W., Y.D. Chu, and J.H. Tai, *Characterization of Trichomonas vaginalis virus proteins in the pathogenic protozoan T. vaginalis.* Arch Virol, 1998. **143**(5): p. 963-70.

410. Goodman, R.P., et al., *Clinical isolates of Trichomonas vaginalis concurrently infected by strains of up to four Trichomonasvirus species (Family Totiviridae).* J Virol, 2011. **85**(9): p. 4258-70.

411. Demeshkina, N., et al., *A new understanding of the decoding principle on the ribosome.* Nature, 2012. **484**(7393): p. 256-9.

412. Guarraia, C., et al., *Saturation mutagenesis of a +1 programmed frameshift-inducing mRNA sequence derived from a yeast retrotransposon.* RNA, 2007. **13**(11): p. 1940-7.

413. Chung, B.Y., A.E. Firth, and J.F. Atkins, *Frameshifting in alphaviruses: a diversity of 3' stimulatory structures.* J Mol Biol, 2010. **397**(2): p. 448-56.

414. Han, J., M.S. Rutherford, and K.S. Faaberg, *Proteolytic products of the porcine reproductive and respiratory syndrome virus nsp2 replicase protein.* J Virol, 2010. **84**(19): p. 10102-12.

415. Jacks, T., et al., *Two efficient ribosomal frameshifting events are required for synthesis of mouse mammary tumor virus gag-related polyproteins.* Proc Natl Acad Sci U S A, 1987. **84**(12): p. 4298-302.

416.   den Boon, J.A., et al., *Processing and evolution of the N-terminal region of the arterivirus replicase ORF1a protein: identification of two papainlike cysteine proteases.* J Virol, 1995. **69**(7): p. 4500-5.

417.   Snijder, E.J., et al., *The arterivirus Nsp2 protease. An unusual cysteine protease with primary structure similarities to both papain-like and chymotrypsin-like proteases.* J Biol Chem, 1995. **270**(28): p. 16671-6.

418.   Wassenaar, A.L., et al., *Alternative proteolytic processing of the arterivirus replicase ORF1a polyprotein: evidence that NSP2 acts as a cofactor for the NSP4 serine protease.* J Virol, 1997. **71**(12): p. 9313-22.

419.   Snijder, E.J., A.L. Wassenaar, and W.J. Spaan, *Proteolytic processing of the replicase ORF1a protein of equine arteritis virus.* J Virol, 1994. **68**(9): p. 5755-64.

420.   Frias-Staheli, N., et al., *Ovarian tumor domain-containing viral proteases evade ubiquitin- and ISG15-dependent innate immune responses.* Cell Host Microbe, 2007. **2**(6): p. 404-16.

421.   Sun, Z., et al., *The cysteine protease domain of porcine reproductive and respiratory syndrome virus nonstructural protein 2 possesses deubiquitinating and interferon antagonism functions.* J Virol, 2010. **84**(15): p. 7832-46.

422.   Sun, Z., et al., *Nonstructural protein 2 of porcine reproductive and respiratory syndrome virus inhibits the antiviral function of interferon-stimulated gene 15.* J Virol, 2012. **86**(7): p. 3839-50.

423.   van Kasteren, P.B., et al., *Arterivirus and nairovirus ovarian tumor domain-containing Deubiquitinases target activated RIG-I to control innate immune signaling.* J Virol, 2012. **86**(2): p. 773-85.

424.   Chen, Z., et al., *Immunodominant epitopes in nsp2 of porcine reproductive and respiratory syndrome virus are dispensable for replication, but play an important role in modulation of the host immune response.* J Gen Virol, 2010. **91**(Pt 4): p. 1047-57.

425.   Brown, E., et al., *Antibody response to porcine reproductive and respiratory syndrome virus (PRRSV) nonstructural proteins and implications for diagnostic detection and differentiation of PRRSV types I and II.* Clin Vaccine Immunol, 2009. **16**(5): p. 628-35.

426.   Altschul, S.F., et al., *Basic local alignment search tool.* J Mol Biol, 1990. **215**(3): p. 403-10.

427.   Larkin, M.A., et al., *Clustal W and Clustal X version 2.0.* Bioinformatics, 2007. **23**(21): p. 2947-8.

428.   Rice, P., I. Longden, and A. Bleasby, *EMBOSS: the European Molecular Biology Open Software Suite.* Trends Genet, 2000. **16**(6): p. 276-7.

429.   Fuerst, T.R., et al., *Eukaryotic transient-expression system based on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase.* Proc Natl Acad Sci U S A, 1986. **83**(21): p. 8122-6.

430.   Wu, W.H., et al., *The 2b protein as a minor structural component of PRRSV.* Virus Res, 2005. **114**(1-2): p. 177-81.

431.   van den Akker, J., et al., *The redox state of transglutaminase 2 controls arterial remodeling.* PLoS One, 2011. **6**(8): p. e23067.

432.   Niwa, H., K. Yamamura, and J. Miyazaki, *Efficient selection for high-expression transfectants with a novel eukaryotic vector.* Gene, 1991. **108**(2): p. 193-9.

433.   Atkins, J.F. and G.R. Bjork, *A gripping tale of ribosomal frameshifting: extragenic suppressors of frameshift mutations spotlight P-site realignment.* Microbiol Mol Biol Rev, 2009. **73**(1): p. 178-210.

434.   Jacks, T., et al., *Signals for ribosomal frameshifting in the Rous sarcoma virus gag-pol region.* Cell, 1988. **55**(3): p. 447-58.

435.   Jacks, T. and H.E. Varmus, *Expression of the Rous sarcoma virus pol gene by ribosomal frameshifting.* Science, 1985. **230**(4731): p. 1237-42.

436.   Firth, A.E. and J.F. Atkins, *A conserved predicted pseudoknot in the NS2A-encoding sequence of West Nile and Japanese encephalitis flaviviruses suggests NS1' may derive from ribosomal frameshifting.* Virol J, 2009. **6**: p. 14.

437. Jacks, T., et al., *Characterization of ribosomal frameshifting in HIV-1 gag-pol expression.* Nature, 1988. **331**(6153): p. 280-3.

438. Mador, N., A. Panet, and A. Honigman, *Translation of gag, pro, and pol gene products of human T-cell leukemia virus type 2.* J Virol, 1989. **63**(5): p. 2400-4.

439. Nam, S.H., et al., *Characterization of ribosomal frameshifting for expression of pol gene products of human T-cell leukemia virus type I.* J Virol, 1993. **67**(1): p. 196-203.

440. Thiel, V., et al., *Mechanisms and enzymes involved in SARS coronavirus genome expression.* J Gen Virol, 2003. **84**(Pt 9): p. 2305-15.

441. Atkins, J.F. and R.F. Gesteland, *Recoding: Expansion of Decoding Rules Enriches Gene Expression.* 2010, Heidelberg: Springer.

442. Fang, Y., et al., *Efficient -2 frameshifting by mammalian ribosomes to synthesize an additional arterivirus protein.* Proc Natl Acad Sci U S A, 2012. **109**(43): p. E2920-8.

443. Snijder, E.J., et al., *Non-structural proteins 2 and 3 interact to modify host cell membranes during the formation of the arterivirus replication complex.* J Gen Virol, 2001. **82**(Pt 5): p. 985-94.

444. Qu, X., et al., *The ribosome uses two active mechanisms to unwind messenger RNA during translation.* Nature, 2011. **475**(7354): p. 118-21.

445. Takyar, S., R.P. Hickerson, and H.F. Noller, *mRNA helicase activity of the ribosome.* Cell, 2005. **120**(1): p. 49-58.

446. Namy, O., et al., *A mechanical explanation of RNA pseudoknot function in programmed ribosomal frameshifting.* Nature, 2006. **441**(7090): p. 244-7.

447. Li, Y., et al., *Targeted mutations in a highly conserved motif of the nsp1beta protein impair the interferon antagonizing activity of porcine reproductive and respiratory syndrome virus.* J Gen Virol, 2013. **94**(Pt 9): p. 1972-83.

448. Fixsen, S.M. and M.T. Howard, *Processive selenocysteine incorporation during synthesis of eukaryotic selenoproteins.* J Mol Biol, 2010. **399**(3): p. 385-96.

449. Grentzmann, G., et al., *A dual-luciferase reporter system for studying recoding signals.* RNA, 1998. **4**(4): p. 479-86.

450. Nedialkova, D.D., A.E. Gorbalenya, and E.J. Snijder, *Arterivirus Nsp1 modulates the accumulation of minus-strand templates to control the relative abundance of viral mRNAs.* PLoS Pathog, 2010. **6**(2): p. e1000772.

451. Tijms, M.A., et al., *A zinc finger-containing papain-like protease couples subgenomic mRNA synthesis to genome translation in a positive-stranded RNA virus.* Proceedings of the National Academy of Sciences, 2001. **98**(4): p. 1889-1894.

452. Chen, Z., et al., *Identification of two auto-cleavage products of nonstructural protein 1 (nsp1) in porcine reproductive and respiratory syndrome virus infected cells: nsp1 function as interferon antagonist.* Virology, 2010. **398**(1): p. 87-97.

453. Xue, F., et al., *The crystal structure of porcine reproductive and respiratory syndrome virus nonstructural protein Nsp1beta reveals a novel metal-dependent nuclease.* J Virol, 2010. **84**(13): p. 6461-71.

454. Meulenberg, J.J., et al., *Localization and fine mapping of antigenic sites on the nucleocapsid protein N of porcine reproductive and respiratory syndrome virus with monoclonal antibodies.* Virology, 1998. **252**(1): p. 106-14.

455. Clarke, J.B. and R.E. Spier, *An investigation into causes of resistance of a cloned line of BHK cells to a strain of foot-and-mouth disease virus.* Vet Microbiol, 1983. **8**(3): p. 259-70.

456. Beach, D.L. and J.D. Keene, *Ribotrap : targeted purification of RNA-specific RNPs from cell lysates through immunoaffinity precipitation to identify regulatory proteins and RNAs.* Methods Mol Biol, 2008. **419**: p. 69-91.
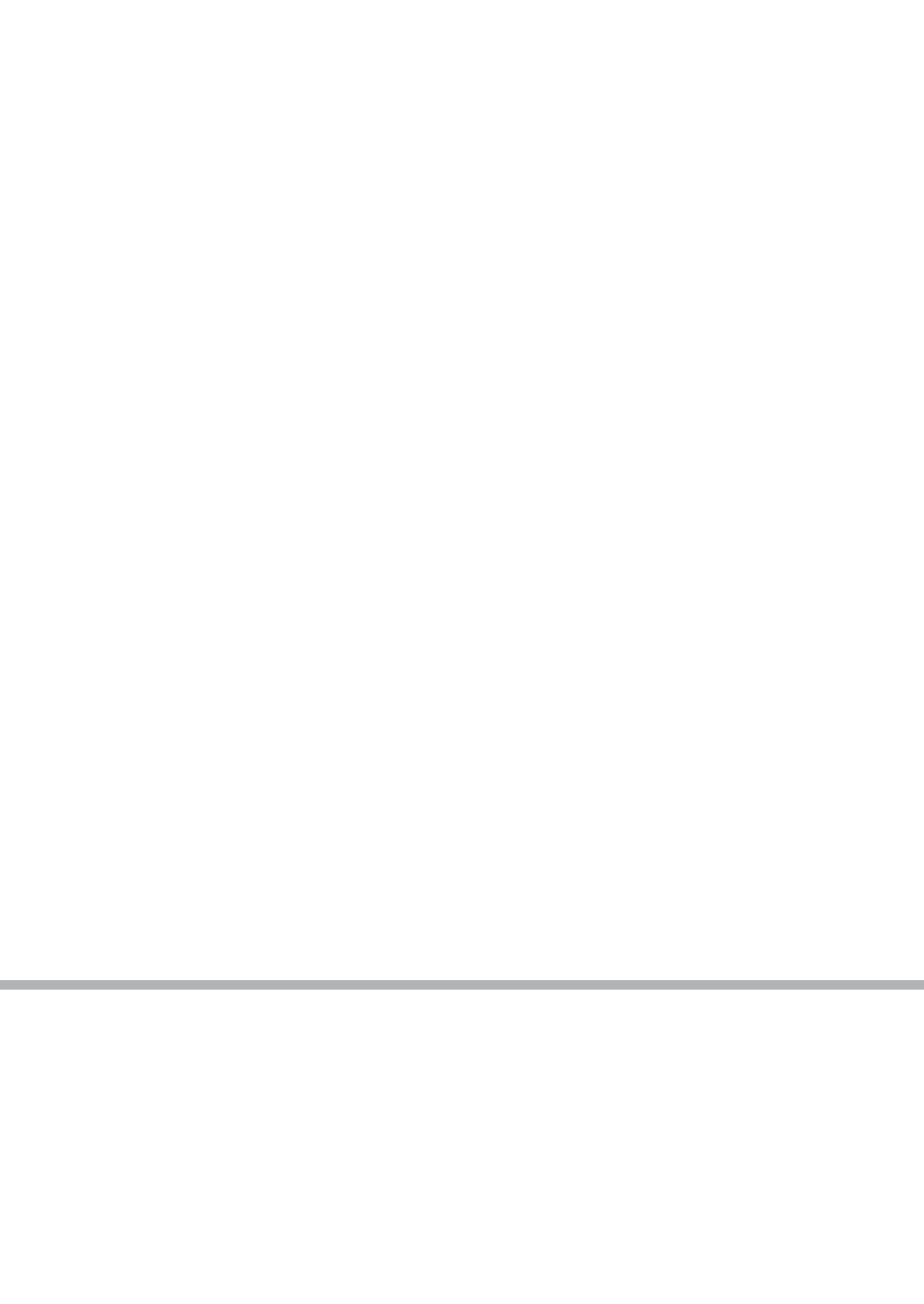
457.   Kwak, H., M.W. Park, and S. Jeong, *Annexin A2 binds RNA and reduces the frameshifting efficiency of infectious bronchitis virus.* PLoS One, 2011. **6**(8): p. e24067.

458.   Kobayashi, Y., et al., *Identification of a cellular factor that modulates HIV-1 programmed ribosomal frameshifting.* J Biol Chem, 2010. **285**(26): p. 19776-84.

459.   Park, H.J., et al., *Increased -1 ribosomal frameshifting efficiency by yeast prion-like phenotype [PSI+].* FEBS Lett, 2009. **583**(4): p. 665-9.

460.   Gendron, K., et al., *The presence of the TAR RNA structure alters the programmed -1 ribosomal frameshift efficiency of the human immunodeficiency virus type 1 (HIV-1) by modifying the rate of translation initiation.* Nucleic Acids Res, 2008. **36**(1): p. 30-40.

461.   Kollmus, H., M.W. Hentze, and H. Hauser, *Regulated ribosomal frameshifting by an RNA-protein interaction.* RNA, 1996. **2**(4): p. 316-323.

462.   Makeyev, A.V., et al., *HnRNP A3 genes and pseudogenes in the vertebrate genomes.* J Exp Zool A Comp Exp Biol, 2005. **303**(4): p. 259-71.

463.   Beura, L.K., et al., *Cellular poly(c) binding proteins 1 and 2 interact with porcine reproductive and respiratory syndrome virus nonstructural protein 1beta and support viral replication.* J Virol, 2011. **85**(24): p. 12939-49.

464.   Darnell, J.C., et al., *FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism.* Cell, 2011. **146**(2): p. 247-61.

465.   Chen, P.J. and Y.S. Huang, *CPEB2-eEF2 interaction impedes HIF-1alpha RNA translation.* EMBO J, 2012. **31**(4): p. 959-71.

466.   Ben-Shem, A., et al., *The structure of the eukaryotic ribosome at 3.0 A resolution.* Science, 2011. **334**(6062): p. 1524-9.

467.   Nahar-Gohad, P., et al., *RACK1 identified as the PCBP1-interacting protein with a novel functional role on the regulation of human MOR gene expression.* J Neurochem, 2013. **124**(4): p. 466-77.

468.   Brierley, I., A.J. Jenner, and S.C. Inglis, *Mutational analysis of the "slippery-sequence" component of a coronavirus ribosomal frameshifting signal.* J Mol Biol, 1992. **227**(2): p. 463-79.

469.   Mejlhede, N., J.F. Atkins, and J. Neuhard, *Ribosomal -1 frameshifting during decoding of Bacillus subtilis cdd occurs at the sequence CGA AAG.* J Bacteriol, 1999. **181**(9): p. 2930-7.

470.   Napthine, S., et al., *Prokaryotic-style frameshifting in a plant translation system: conservation of an unusual single-tRNA slippage event.* EMBO J, 2003. **22**(15): p. 3941-50.

471.   Albertini, A.A., et al., *Crystal structure of the rabies virus nucleoprotein-RNA complex.* Science, 2006. **313**(5785): p. 360-3.

472.   Ruigrok, R.W., T. Crepin, and D. Kolakofsky, *Nucleoproteins and nucleocapsids of negative-strand RNA viruses.* Curr Opin Microbiol, 2011. **14**(4): p. 504-10.

473.   Snijder, E.J., A.L. Wassenaar, and W.J. Spaan, *The 5' end of the equine arteritis virus replicase gene encodes a papainlike cysteine protease.* J Virol, 1992. **66**(12): p. 7040-8.

474.   Sun, Y., et al., *Crystal structure of porcine reproductive and respiratory syndrome virus leader protease Nsp1alpha.* J Virol, 2009. **83**(21): p. 10931-40.

475.   Tijms, M.A., et al., *Arterivirus subgenomic mRNA synthesis and virion biogenesis depend on the multifunctional nsp1 autoprotease.* J Virol, 2007. **81**(19): p. 10496-505.

476.   Kroese, M.V., et al., *The nsp1alpha and nsp1 papain-like autoproteinases are essential for porcine reproductive and respiratory syndrome virus RNA synthesis.* J Gen Virol, 2008. **89**(Pt 2): p. 494-9.

477.   Kim, O., et al., *Modulation of type I interferon induction by porcine reproductive and respiratory syndrome virus and degradation of CREB-binding protein by non-structural protein 1 in MARC-145 and HeLa cells.* Virology, 2010. **402**(2): p. 315-26.

478. Patel, D., et al., *Porcine reproductive and respiratory syndrome virus inhibits type I interferon signaling by blocking STAT1/STAT2 nuclear translocation.* J Virol, 2010. **84**(21): p. 11045-55.

479. Wang, L. and S.J. Brown, *BindN: a web-based tool for efficient prediction of DNA and RNA binding sites in amino acid sequences.* Nucleic Acids Res, 2006. **34**(Web Server issue): p. W243-8.

480. DeLano, W.L., *The PyMOL molecular graphics system.* 2002.

481. Antonov, I., et al., *Identification of the nature of reading frame transitions observed in prokaryotic genomes.* Nucleic Acids Research, 2013. **41**(13): p. 6514-6530.

482. Wilson, W., et al., *Expression strategies of the yeast retrotransposon Ty: a short sequence directs ribosomal frameshifting.* Nucleic Acids Res, 1986. **14**(17): p. 7001-16.

483. Belew, A.T. and J.D. Dinman, *Cell cycle control (and more) by programmed -1 ribosomal frameshifting: implications for disease and therapeutics.* Cell Cycle, 2015. **14**(2): p. 172-8.

484. Kohl, S. and R. Bock, *Transposition of a bacterial insertion sequence in chloroplasts.* Plant J, 2009. **58**(3): p. 423-36.

485. Yusupova, G.Z., et al., *The path of messenger RNA through the ribosome.* Cell, 2001. **106**(2): p. 233-41.

486. Chen, J., et al., *Dynamic pathways of -1 translational frameshifting.* Nature, 2014. **512**(7514): p. 328-32.

487. Kim, H.K., et al., *A frameshifting stimulatory stem loop destabilizes the hybrid state and impedes ribosomal translocation.* Proc Natl Acad Sci U S A, 2014. **111**(15): p. 5538-43.

488. Li, Y., et al., *Transactivation of programmed ribosomal frameshifting by a viral protein.* Proc Natl Acad Sci U S A, 2014. **111**(21): p. E2172-81.

489. Sun, Y., et al., *Interplay between interferon-mediated innate immunity and porcine reproductive and respiratory syndrome virus.* Viruses, 2012. **4**(4): p. 424-46.

490. Beura, L.K., et al., *Porcine reproductive and respiratory syndrome virus nonstructural protein 1beta modulates host innate immune response by antagonizing IRF3 activation.* J Virol, 2010. **84**(3): p. 1574-84.

491. Song, C., P. Krell, and D. Yoo, *Nonstructural protein 1alpha subunit-based inhibition of NF-kappaB activation and suppression of interferon-beta production by porcine reproductive and respiratory syndrome virus.* Virology, 2010. **407**(2): p. 268-80.

492. Makeyev, A.V. and S.A. Liebhaber, *The poly(C)-binding proteins: a multiplicity of functions and a search for mechanisms.* RNA, 2002. **8**(3): p. 265-78.

493. Wang, L., et al., *Interaction of cellular poly(C)-binding protein 2 with nonstructural protein 1beta is beneficial to Chinese highly pathogenic porcine reproductive and respiratory syndrome virus replication.* Virus Res, 2012. **169**(1): p. 222-30.

494. Dahlgren, C., et al., *Analysis of siRNA specificity on targets with double-nucleotide mismatches.* Nucleic Acids Research, 2008. **36**(9): p. e53.

495. Matunis, M.J., W.M. Michael, and G. Dreyfuss, *Characterization and primary structure of the poly(C)-binding heterogeneous nuclear ribonucleoprotein complex K protein.* Mol Cell Biol, 1992. **12**(1): p. 164-71.

496. Leffers, H., K. Dejgaard, and J.E. Celis, *Characterisation of two major cellular poly(rC)-binding human proteins, each containing three K-homologous (KH) domains.* Eur J Biochem, 1995. **230**(2): p. 447-53.

497. Kiledjian, M., X. Wang, and S.A. Liebhaber, *Identification of two KH domain proteins in the alpha-globin mRNP stability complex.* EMBO J, 1995. **14**(17): p. 4357-64.

498. Makeyev, A.V. and S.A. Liebhaber, *Identification of two novel mammalian genes establishes a subfamily of KH-domain RNA-binding proteins.* Genomics, 2000. **67**(3): p. 301-16.

499. Choi, H.S., et al., *Poly(C)-binding proteins as transcriptional regulators of gene expression.* Biochem Biophys Res Commun, 2009. **380**(3): p. 431-6.

500. Boschi, N.M., et al., *Differential expression of polycytosine-binding protein isoforms in adrenal gland, locus coeruleus and midbrain.* Neuroscience, 2015. **286**: p. 1-12.

501. Ghanem, L.R., P. Chatterji, and S.A. Liebhaber, *Specific enrichment of the RNA-binding proteins PCBP1 and PCBP2 in chief cells of the murine gastric mucosa.* Gene Expr Patterns, 2014. **14**(2): p. 78-87.

502. Aasheim, H.C., et al., *Tissue specific expression and cDNA structure of a human transcript encoding a nucleic acid binding [oligo(dC)] protein related to the pre-mRNA binding protein K.* Nucleic Acids Res, 1994. **22**(6): p. 959-64.

503. Caliskan, N., F. Peske, and M.V. Rodnina, *Changed in translation: mRNA recoding by -1 programmed ribosomal frameshifting.* Trends Biochem Sci, 2015. **40**(5): p. 265-274.

504. Wang, L., et al., *Interactome Profile of the Host Cellular Proteins and the Nonstructural Protein 2 of Porcine Reproductive and Respiratory Syndrome Virus.* Plos One, 2014. **9**(6): p. 12.

505. Yoga, Y.M., et al., *Contribution of the first K-homology domain of poly(C)-binding protein 1 to its affinity and specificity for C-rich oligonucleotides.* Nucleic Acids Res, 2012. **40**(11): p. 5101-14.

506. Cafruny, W.A., et al., *Porcine reproductive and respiratory syndrome virus (PRRSV) infection spreads by cell-to-cell transfer in cultured MARC-145 cells, is dependent on an intact cytoskeleton, and is suppressed by drug-targeting of cell permissiveness to virus infection.* Virol J, 2006. **3**: p. 90.

507. Nilsson, J., et al., *Regulation of eukaryotic translation by the RACK1 protein: a platform for signalling molecules on the ribosome.* EMBO Rep, 2004. **5**(12): p. 1137-41.

508. Wolf, A.S. and E.J. Grayhack, *Asc1, homolog of human RACK1, prevents frameshifting in yeast by ribosomes stalled at CGA codon repeats.* RNA, 2015. **21**(5): p. 935-45.

509. Gamarnik, A.V. and R. Andino, *Two functional complexes formed by KH domain containing proteins with the 5' noncoding region of poliovirus RNA.* RNA, 1997. **3**(8): p. 882-92.

510. Pacheco, A., S. Reigadas, and E. Martinez-Salas, *Riboproteomic analysis of polypeptides interacting with the internal ribosome-entry site element of foot-and-mouth disease viral RNA.* Proteomics, 2008. **8**(22): p. 4782-90.

511. Sharma, N., et al., *Functional role of the 5' terminal cloverleaf in Coxsackievirus RNA replication.* Virology, 2009. **393**(2): p. 238-49.

512. Kempf, B.J. and D.J. Barton, *Poly(rC) binding proteins and the 5' cloverleaf of uncapped poliovirus mRNA function during de novo assembly of polysomes.* J Virol, 2008. **82**(12): p. 5835-46.

513. Barton, D.J., B.J. O'Donnell, and J.B. Flanegan, *5' cloverleaf in poliovirus RNA is a cis-acting replication element required for negative-strand synthesis.* EMBO J, 2001. **20**(6): p. 1439-48.

514. Herold, J. and R. Andino, *Poliovirus RNA replication requires genome circularization through a protein-protein bridge.* Mol Cell, 2001. **7**(3): p. 581-91.

515. Blyn, L.B., et al., *Requirement of poly(rC) binding protein 2 for translation of poliovirus RNA.* J Virol, 1997. **71**(8): p. 6243-6.

516. Sweeney, T.R., et al., *The mechanism of translation initiation on Type 1 picornavirus IRESs.* EMBO J, 2014. **33**(1): p. 76-92.

517. Perera, R., et al., *Cellular protein modification by poliovirus: the two faces of poly(rC)-binding protein.* J Virol, 2007. **81**(17): p. 8919-32.

518. Zhang, B., et al., *RNA interaction and cleavage of poly(C)-binding protein 2 by hepatitis A virus protease.* Biochem Biophys Res Commun, 2007. **364**(4): p. 725-30.

519. Spangberg, K. and S. Schwartz, *Poly(C)-binding protein interacts with the hepatitis C virus 5' untranslated region.* J Gen Virol, 1999. **80 (Pt 6)**: p. 1371-6.

520. Gutierrez-Escolano, A.L., et al., *Interaction of cellular proteins with the 5' end of Norwalk virus genomic RNA*. J Virol, 2000. **74**(18): p. 8558-62.

521. Wang, L., K.S. Jeng, and M.M. Lai, *Poly(C)-binding protein 2 interacts with sequences required for viral replication in the hepatitis C virus (HCV) 5' untranslated region and directs HCV RNA replication through circularizing the viral genome*. J Virol, 2011. **85**(16): p. 7954-64.

522. Li, D., et al., *Poly(C)-binding protein 1, a novel N(pro)-interacting protein involved in classical swine fever virus growth*. J Virol, 2013. **87**(4): p. 2072-80.

523. Eswarappa, S.M., et al., *Programmed translational readthrough generates antiangiogenic VEGF-Ax*. Cell, 2014. **157**(7): p. 1605-18.

524. He, Y. and R. Smith, *Nuclear functions of heterogeneous nuclear ribonucleoproteins A/B*. Cell Mol Life Sci, 2009. **66**(7): p. 1239-56.

525. Hennig, J. and M. Sattler, *Deciphering the protein-RNA recognition code: Combining large-scale quantitative methods with structural biology*. Bioessays, 2015.

526. Dinman, J.D., *Programmed Ribosomal Frameshifting Goes Beyond Viruses: Organisms from all three kingdoms use frameshifting to regulate gene expression, perhaps signaling a paradigm shift*. Microbe Wash DC, 2006. **1**(11): p. 521-527.

527. Dinman, J.D., *Mechanisms and implications of programmed translational frameshifting*. Wiley Interdiscip Rev RNA, 2012. **3**(5): p. 661-73.

528. Weiss, R.B., et al., *Slippery runs, shifty stops, backward steps, and forward hops: -2, -1, +1, +2, +5, and +6 ribosomal frameshifting*. Cold Spring Harb Symp Quant Biol, 1987. **52**: p. 687-93.

529. Rom, E. and C. Kahana, *Polyamines regulate the expression of ornithine decarboxylase antizyme in vitro by inducing ribosomal frame-shifting*. Proc Natl Acad Sci U S A, 1994. **91**(9): p. 3959-63.

530. Kurian, L., et al., *Polyamine sensing by nascent ornithine decarboxylase antizyme stimulates decoding of its mRNA*. Nature, 2011. **477**(7365): p. 490-494.

531. Goodman, R.P., et al., *Trichomonasvirus: a new genus of protozoan viruses in the family Totiviridae*. Arch Virol, 2011. **156**(1): p. 171-9.

532. Parent, K.N., et al., *Structure of a protozoan virus from the human genitourinary parasite Trichomonas vaginalis*. MBio, 2013. **4**(2).

533. Su, H.-M. and J.-H. Tai, *Genomic Organization and Sequence Conservation in Type I Trichomonas vaginalis Viruses*. Virology, 1996. **222**(2): p. 470-473.

534. Du, Z., et al., *X-ray crystallographic and NMR studies of protein-protein and protein-nucleic acid interactions involving the KH domains from human poly(C)-binding protein-2*. RNA, 2007. **13**(7): p. 1043-51.

535. Martinez-Salas, E., et al., *RNA-binding proteins impacting on internal initiation of translation*. Int J Mol Sci, 2013. **14**(11): p. 21705-26.

536. Charbonneau, J., et al., *The 5' UTR of HIV-1 full-length mRNA and the Tat viral protein modulate the programmed −1 ribosomal frameshift that generates HIV-1 enzymes*. RNA, 2012. **18**(3): p. 519-529.

537. Cook, K.B., et al., *RBPDB: a database of RNA-binding specificities*. Nucleic Acids Research, 2010.

538. Chou, M.Y., S.C. Lin, and K.Y. Chang, *Stimulation of -1 programmed ribosomal frameshifting by a metabolite-responsive RNA pseudoknot*. RNA, 2010. **16**(6): p. 1236-44.

539. Yu, C.H., et al., *Exploiting preQ(1) Riboswitches To Regulate Ribosomal Frameshifting*. Acs Chemical Biology, 2013. **8**(4): p. 733-740.

540. Belew, A.T., et al., *Ribosomal frameshifting in the CCR5 mRNA is regulated by miRNAs and the NMD pathway*. Nature, 2014. **512**(7514): p. 265-9.

541. Dulude, D., et al., *Decreasing the frameshift efficiency translates into an equivalent reduction of the replication of the human immunodeficiency virus type 1*. Virology, 2006. **345**(1): p. 127-36.

542   Plant, E.P., et al., *Altering SARS coronavirus frameshift efficiency affects genomic and subgenomic RNA production.* Viruses, 2013. **5**(1): p. 279-94.

543   Plant, E.P., et al., *Achieving a golden mean: mechanisms by which coronaviruses ensure synthesis of the correct stoichiometric ratios of viral proteins.* J Virol, 2010. **84**(9): p. 4330-40.

544.   Li, Y., et al., *Integrative analysis of circadian transcriptome and metabolic network reveals the role of de novo purine synthesis in circadian control of cell cycle.* PLoS Comput Biol, 2015. **11**(2): p. e1004086.

Nederlandse samenvatting

Dit proefschrift bestaat uit twee delen en beschrijft onderzoek aan twee virussen die beide een positiefstrengig RNA genoom hebben. Het gaat om het humane pathogeen chikungunya virus (CHIKV) en het economisch belangrijke porcine reproductive and respiratory syndrome virus (PRRSV) dat wereldwijd varkens infecteert.

DEEL 1: KWANTITATIEVE PROTEOMICS ANALYSE VAN CHIKUNGUNYA VIRUS INFECTIE

Dit deel van het proefschrift beschrijft de veranderingen die optreden in het expressieniveau en de fosforyleringsstatus van gastheereiwitten tijdens een CHIKV infectie.

CHIKV is een alphavirus en wordt op mensen overgedragen door muskieten. Een CHIKV infectie veroorzaakt hoge koorts, uitslag en gewrichtspijnen die maanden kunnen aanhouden. CHIKV werd in 1952 voor het eerst beschreven in Tanzania maar tot 2006 veroorzaakte het virus alleen relatief kleine lokale uitbraken in Afrika en Azië. In 2005/2006 was er in het gebied rondom de Indische Oceaan een ongekend grote uitbraak waarbij naar schatting meer dan 4 miljoen mensen werden geïnfecteerd. Eind 2013 werd er voor het eerst een CHIKV uitbraak gemeld in het Caribische gebied en in de daaropvolgende maanden verspreidde het virus zich verder over Zuid- en Centraal Amerika. Tijdens deze laatste uitbraak zijn inmiddels meer dan een miljoen mensen besmet geraakt met CHIKV. In Europa is er een toename van het aantal reizigers dat geinfecteerd met CHIKV terugkeert uit landen waar het virus voorkomt. In regio's waar de muskieten voorkomen die CHIKV kunnen overdragen brengt dit het risico met zich mee dat het virus lokaal verder wordt verspreid. In Italië, Frankrijk en Spanje zijn al kleinschalige uitbraken geweest waarbij patiënten door lokale muskieten werden geïnfecteerd.

CHIKV is, net als andere virussen, sterk afhankelijk van cellulaire factoren tijdens de infectiecyclus. Om meer inzicht te krijgen in deze virus-gastheerinteracties is voor de studies in **hoofdstuk 2 en 3** een kwantitatieve proteomics methode gebruikt. Proteomics is de studie van het proteoom, alle eiwitten die in een cel, weefsel of organisme tot expressie komen. In tegenstelling tot het genoom, de genetische informatie, die voor alle cellen van een organisme hetzelfde is, is het proteoom zeer dynamisch. Het proteoom verandert onder invloed van interne en externe stimuli, zoals een virusinfectie. Voor de analyse van het proteoom worden de eiwitten eerst met enzymen in stukjes geknipt, zodat er korte fragmenten (peptiden) overblijven. Van deze peptiden kan met een massaspectrometer de aminozuursequentie bepaald worden en deze sequentie is meestal uniek voor een specifiek eiwit. Met kwantitatieve proteomics kunnen verschillen in expressieniveau tussen twee of meer monsters bepaald worden. Voor de studies in dit proefschrift is de stabiele isotoop labeling met aminozuren in celkweek (SILAC) methode gebruikt om deze verschillen te kwantificeren. SILAC is een labelingsmethode

waarmee alle eiwitten in de cel voorafgaand aan een experiment gelabeld worden door de cellen te kweken in medium met aminozuren (de bouwstenen van eiwitten) die iets zwaarder zijn dan de standaard aminozuren. Deze aminozuren zijn zwaarder doordat een aantal koolstof-, stikstof- of waterstofatomen zijn vervangen door niet-radioactieve isotopen. De zwaardere aminozuren hebben dezelfde chemische eigenschappen als de lichtere variant en maken het mogelijk om met een massaspectrometer onderscheid te maken tussen peptiden die afkomstig zijn van monsters gelabeld met lichtere dan wel zwaardere aminozuren. Op deze manier kan de relatieve expressie van een eiwit in het geïnfecteerde en niet geïnfecteerde monster bepaald worden.

Hoofdstuk 2 beschrijft een SILAC studie waarin veranderingen in expressieniveau van gastheereiwitten op 8, 10 en 12 uur na een CHIKV infectie zijn bepaald. Hieruit bleek dat het expressieniveau van de meeste gastheereiwitten in de cel tijdens een infectie niet of nauwelijks verandert. In totaal werden meer dan 4700 eiwitten geïdentificeerd en 2800-3600 eiwitten konden op elk van de verschillende tijdstippen worden gekwantificeerd. Op 8, 10 en 12 uur na CHIKV infectie werden respectievelijk 13, 38 en 106 eiwitten ge-identificeerd waarvan de hoeveelheid significant veranderd was tijdens de infectie. Van bijna alle eiwitten waarvan de expressie tijdens de infectie significant veranderde nam de hoeveelheid af. Dit is deels te verklaren doordat de aanmaak van nieuwe (cellulaire) eiwitten geremd wordt door CHIKV vanaf ongeveer 8 uur na infectie, via de blokkade van zowel de cellulaire transcriptie (RNA synthese) als translatie (eiwitsynthese). Het expressieniveau van de meeste onderdelen van het RNA polymerase II complex, verant-woordelijk voor RNA synthese, nam significant af tijdens de CHIKV infectie en dit draagt waarschijnlijk bij aan de door CHIKV geïnduceerde blokkade van cellulaire transcriptie.

Vier eiwitten, Rnd3, DDX56, Plk1 en UbcH10, waarvan de hoeveelheid significant afnam tijdens het SILAC experiment werden geselecteerd voor nader onderzoek. Deze eiwitten werden in cellen tot overexpressie gebracht voorafgaand aan een CHIKV infectie, wat resulteerde in een vermindering van het aantal cellen dat vervolgens geïnfecteerd kon worden. Dit suggereert dat de geobserveerde downregulatie van deze eiwitten tijdens een CHIKV infectie gunstig is voor de virale replicatie.

Hoofdstuk 3 beschrijft een SILAC studie waarin veranderingen in de fosforylerings-status van gastheereiwitten op 2, 8 en 12 uur na een chikungunya infectie zijn bestu-deerd. Fosforylering is een chemische modificatie van eiwitten die onder andere een belangrijke rol speelt in de signaaltransductie, de reactie van cellen op signalen uit de omgeving (b.v. stress, nutriënten, infectie). Een fosfaatgroep kan aan de aminozuren se-rine, threonine en tyrosine van een eiwit worden gekoppeld om de activiteit, functie en/of lokalisatie te reguleren. Tijdens de CHIKV infectie konden bijna 3000 gefosforyleerde aminozuren worden geïdentificeerd en per tijdstip kon van 800-1200 aminozuren de fosforyleringsstatus gekwantificeerd worden. Op 2, 8 en 12 uur na CHIKV infectie werd voor respectievelijk 10, 71 en 136 aminozuren een significante verandering van de fos-

foryleringsstatus vastgesteld, waarbij zowel toenames als afnames werden gemeten. De toename van fosforylering op threonine 56 van eukaryote elongatie factor 2 (eEF2) was zo groot dat dit eiwit werd geselecteerd voor verdere analyse. eEF2 is belangrijk tijdens de translocatiestap van de synthese van eiwitten door ribosomen. Gefosforyleerd eEF2 is inactief en fosforylering zal dus leiden tot remming van (de elongatie fase van) de eiwitsynthese. De toename in fosforylering was met massaspectrometrie al waarneembaar op 2 uur na infectie en nam daarna nog sterk toe, tot een toename van meer dan 50x op 8 en 12 uur na infectie. Deze toename in eEF2 fosforylering bleek ook plaats te vinden tijdens infecties met Sindbis virus en Semliki Forest virus, twee andere alphavirussen. De fosforylering van eEF2 treedt ook op tijdens infectie met coxsackievirus (+RNA virus, picornavirus familie), maar niet tijdens infectie met humaan adenovirus (DNA virus) of equine arteritis virus (+RNA virus, arterivirus familie). Er zijn een aantal experimenten gedaan om het mechanisme achter de sterke toename van eEF2 fosforylering tijdens alphavirusinfecties te identificeren. Het remmen van PKA en AMPK en het activeren van mTORC1, bekende regulatoren van eEF2 fosforylering resulteerde niet in een vermindering van fosforylering tijdens CHIKV infectie. Activatie van RIG-I, een belangrijk sensor eiwit van het aangeboren immuunsysteem, was ook niet verantwoordelijk voor de inductie van eEF2 fosforylering. De expressie van de virale structurele eiwitten en het proces waarbij het virusdeeltje door de cel wordt opgenomen bleken de fosforylering ook niet te veroorzaken. Aangezien de fosforylering van eEF2 wordt geïnduceerd door infecties met diverse virussen zou het een niet eerder beschreven algemeen antiviraal mechanisme kunnen zijn dat de translatie afremt wanneer de cel door een virus wordt geïnfecteerd.

DEEL 2: GEPROGRAMMEERDE −2/−1 RIBOSOMALE LEESRAAMVERSCHUIVING IN ARTERIVIRUSSEN

Dit deel van het proefschrift beschrijft een nieuw en ongebruikelijk transatiemechanisme waarbij het leesraam verschuift tijdens de synthese van het niet-structurele polyproteïne van het arterivirus PRRSV.

Ribosomen zijn de cellulaire eiwitfabrieken. Ze lezen de nucleotidensequentie van RNA af en gebruiken deze informatie om aminozuren aan elkaar te koppelen tot eiwitten. Dit proces heet translatie. Drie nucleotiden, een zogenaamd codon, coderen voor één aminozuur. Normaalgesproken schuift het ribosoom steeds precies drie nucleotiden op om een volgend aminozuur in te bouwen, totdat dit zogenaamde leesraam eindigt met een stopcodon.

+RNA virussen zijn goed in het manipuleren van het gedrag van ribosomen om te zorgen dat hun kleine genoom zo efficiënt mogelijk wordt afgelezen. Hierbij coderen delen
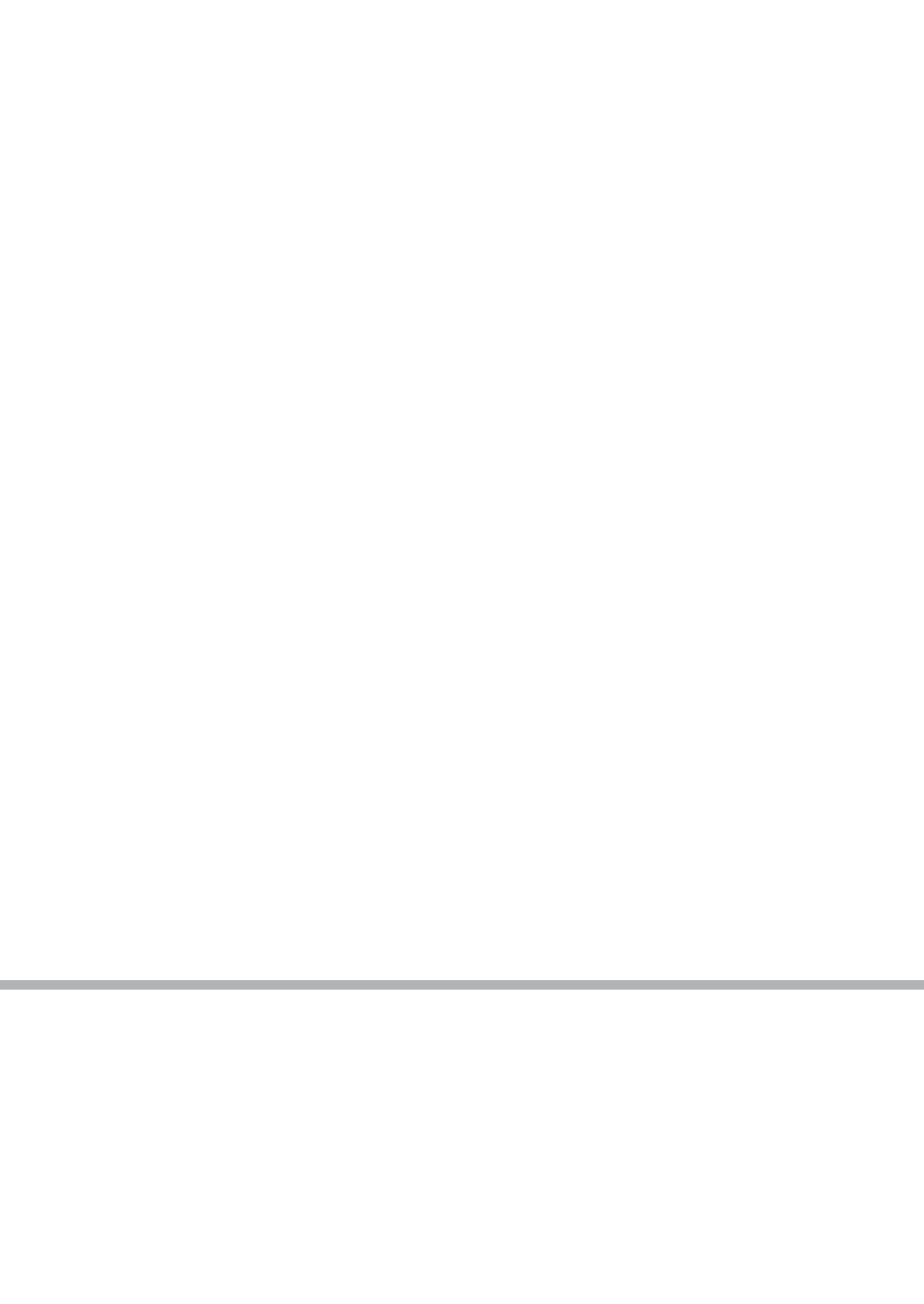
van het genoom soms voor meerdere eiwitten door gebruik te maken van overlappende leesramen. Een van de mechanismen die hierbij gebruikt wordt is ribosomale leesraamverschuiving. Een leesraamverschuiving treedt meestal op doordat een structuur in het RNA een blokkade vormt voor een ribosoom, waardoor dit een nucleotide naar voren (+1) of achteren (–1) schuift. Als het ribosoom daarna verder gaat met het synthetiseren van het eiwit is het leesraam verschoven en wordt een eiwit met een andere sequentie gemaakt. Voor virussen is dit een methode om extra eiwitten te produceren vanaf een enkele RNA sequentie en/of om de expressieratio tussen verschillende eiwitten te reguleren.

De hoofdstukken in deel 2 van dit proefschrift beschrijven een nieuw leesraamverschuivingsmechanisme dat een belangrijke rol speelt tijdens de translatie van het PRRSV genoom. In **hoofdstuk 6** wordt de ontdekking van het mechanisme beschreven. Een bioinformatica analyse van PRRSV genomen liet zien dat in het gebied dat codeert voor het C-terminale deel van nsp2 veel minder synonieme mutaties (mutaties die het gecodeerde aminozuur niet veranderen) voorkomen dan in de rest van het genoom. Dit wijst erop dat er in dit gebied een belangrijke RNA structuur of een tweede eiwit in een ander leesraam zou kunnen worden gecodeerd. In dit geval bleek dat een leesraamverschuiving resulteert in de synthese van een eiwit dat het N-terminale gedeelte gemeen heeft met nsp2, maar dat een andere C-terminale sequentie heeft. Bijzonder aan dit mechanisme is dat het zogenaamde transframe eiwit, nsp2TF, gesynthetiseerd wordt na een –2 leesraamverschuiving. In geïnfecteerde cellen maakt ongeveer 20% van de ribosomen de –2 leesraamverschuiving. Een efficiënte –2 leesraamverschuiving die optreedt in eukaryote cellen was nog niet eerder beschreven. Er zijn aanwijzingen dat op dezelfde locatie in het genoom ook een –1 leesraamverschuiving optreedt, waarna het ribosoom meteen een stopcodon tegenkomt, resulterend in de translatie van een kortere variant van nsp2, nsp2N. De sequentie waarop de leesraamverschuiving plaatsvindt (G_GUU_UUU) en een kort daarachter gelegen C-rijke sequentie (CCCANCUCC) zijn geconserveerd in drie van de vier arterivirussen, maar er zijn geen aanwijzingen dat de RNA sequentie achter de leesraamverschuivingslokatie een structuur vormt die de leesraamverschuiving zou kunnen stimuleren. Het bestaan van nsp2TF werd aangetoond met specifieke antilichamen en de positie en richting van de leesraamverschuiving werden vastgesteld met massaspectrometrie. De expressie van nsp2TF is belangrijk voor PRRSV want virussen waarin de expressie door mutaties in leesraamverschuivingssignalen werd verhinderd repliceerden aanzienlijk minder goed. Tijdens een infectie lokaliseren nsp2 en nsp2TF op verschillende plekken in de cel en dit suggereert dat de twee eiwitten verschillende functies hebben tijdens een infectie.

In **hoofdstuk 7** wordt beschreven dat ook de aanwezigheid van nsp1β, een van de virale eiwitten, noodzakelijk is voor het optreden van de bovengenoemde leesraamverschuiving. Dit was de eerste keer dat is aangetoond dat een eiwit in staat is om een lees-

raamverschuiving te stimuleren. Een geconserveerde sequentie in het papaine-achtige autoprotease domein van nsp1β, die mogelijk RNA zou kunnen binden, werd gemuteerd en deze mutant was niet meer in staat om de leesraamverschuiving te stimuleren. Deze nsp1β mutant was ook niet meer in staat om te binden aan een stuk RNA dat de PRRSV leesraamverschuivingscassette bevatte. De minimale RNA sequentie die nodig is om de leesraamverschuiving te induceren is 34 nucleotiden lang en bevat zowel de eerder beschreven leesraamverschuivingslokatie en de C-rijke sequentie. Het optreden van de −1 leesraamverschuiving en expressie van nsp2N werden bevestigd met massaspectrometrie. Dit was de eerste keer dat werd beschreven dat in een natuurlijke situatie twee verschillende leesraamverschuivingen op dezelfde positie in het RNA kunnen optreden.

In **hoofdstuk 8** wordt beschreven dat niet alleen het virale eiwit nsp1β nodig is om de beide leesraamverschuivingen te induceren, maar dat ook de gastheereiwitten poly(C) binding proteins (PCBP) 1 en 2 betrokken zijn. Van deze eiwitten was eerder beschreven dat ze interactiepartners zijn van nsp1β. In een *in vitro* translatie systeem traden de −2 en −1 leesraamverschuivingen alleen op wanneer gezuiverd nsp1β en PCBP1 en/of PCBP2 tegelijkertijd werden toegevoegd. Vermindering van de hoeveelheid PCBP1 en PCBP2 in een cel door middel van siRNA transfectie en daaropvolgend transfectie van een plasmide dat nsp1β en nsp2 tot expressie brengt resulteerde in een 40% reductie van de expressie van de twee leesraamverschuivingsproducten, nsp2TF en nsp2N. PCBP1 stimuleert voornamelijk de −2 leesraamverschuiving terwijl PCBP2 vooral de −1 leesraamverschuiving stimuleert. We denken dat een complex van nsp1β en nsp2 aan het C-rijke gebied van het PRRSV RNA bindt en daar eenzelfde vertragend effect heeft op de eiwit-synthetiserende ribosomen als de eerder beschreven RNA structuren. Door deze vertraging wordt de efficiëntie van de leesraamverschuiving aanmerkelijk verhoogd. Het stimuleren van een ribosomale leesraamverschuiving is een nieuwe functie voor de PCBPs en dit is ook een nieuw soort virus-gastheerinteractie voor arterivirussen.

Curriculum Vitae

Emmely Treffers werd geboren op 2 juli 1985 in Leiden. Ze behaalde haar VWO diploma in 2003 aan de R.K. Scholengemeenschap Augustinianum in Eindhoven. In 2003 begon zij aan de studie geneeskunde aan de Universiteit Leiden en behaalde in 2009 haar doctoraaldiploma. Van 2007-2009 volgde zij het premaster traject Biomedische Wetenschappen aan de Universiteit Leiden. In 2008 startte ze met de research master Biomedische Wetenshappen aan de Universiteit Leiden en behaalde in 2010 *cum laude* haar master diploma. Tijdens de masteropleiding liep ze twee onderzoeksstages. De eerste stage liep ze bij het hemoglobinopathiënlaboratorium van de afdeling klinische genetica van het LUMC onder begeleiding van Marion Phyllipsen en dr. Kees Harteveld. Ze startte hier met het opzetten van een methode om non-invasieve prenatale diagnostiek voor beta-thalassemie te kunnen uitvoeren. Haar afstudeerstage voerde ze uit bij het massaspectrometrielaboratorium van de afdeling Immunohematologie en Bloedtransfusie van het LUMC onder begeleiding van dr. Peter van Veelen. Tijdens deze stage gebruikte ze kwantitatieve proteomics om 3 aberrante intra-epitheliale T-lymfocytcellijnen van refractaire coeliakie type II patiënten te analyseren. In 2009 kreeg ze een BW-plus beurs toegekend van het LUMC. Deze beurs gebruikte ze om in 2010 de MaxQuant summerschool en de EuPA Scientific Meeting te bezoeken. In 2010 kreeg ze een BW-plus beurs van het LUMC om het in dit proefschrift beschreven promotieonderzoek uit te voeren bij de afdeling Medische Microbiologie van het LUMC en het massaspectrometrielaboratorium van de afdeling Immunohematologie en Bloedtransfusie onder begeleiding van prof.dr. Eric Snijder, dr. Martijn van Hemert en dr. Peter van Veelen. De komende jaren zal Emmely als postdoc bij de afdeling Medische Microbiologie het onderzoek naar het in dit proefschrift beschreven frameshiftmechanisme van arterivirussen voortzetten, mede mogelijk gemaakt door een ECHO subsidie van het gebied Chemische Wetenschappen van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek aan Prof. Dr. Eric Snijder.
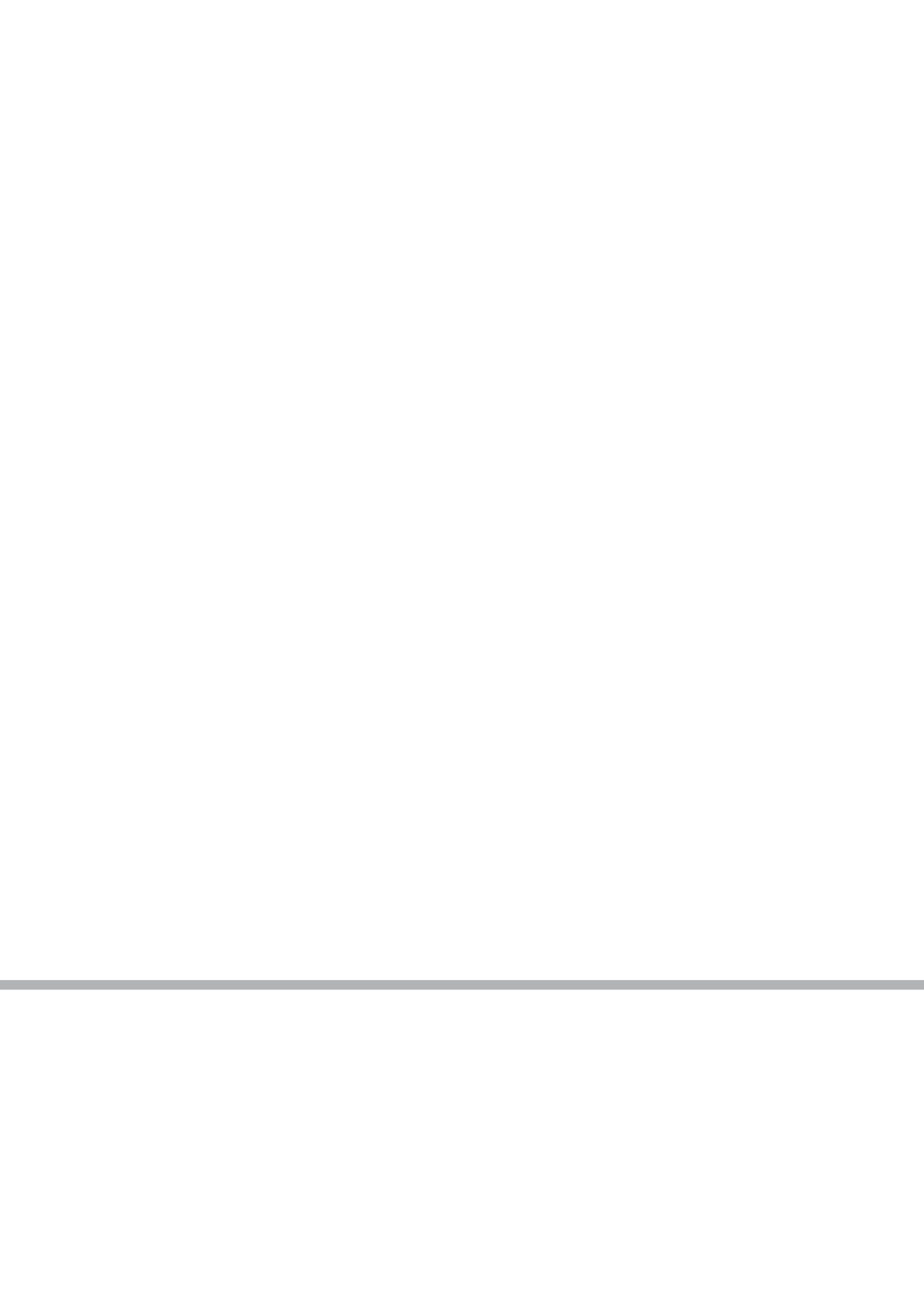
List of publications

1. **Emmely E. Treffers**, Ali Tas, Florine E.M. Scholte, Arnoud H. de Ru, Eric J. Snijder, Peter A. van Veelen2[#], Martijn J. van Hemert[#]. Rapid and strong induction of eEF2 phosphorylation during alphavirus infection; Manuscript in preparation.

2. **Emmely E. Treffers**[*], Sawsan Napthine[*], Susanne Bell, Ian Goodfellow, Ying Fang, Andrew E. Firth, Eric J. Snijder[#], Ian Brierley[#]. Arterivirus non-structural protein 1β co-operates with cellular poly (C) binding proteins to transactivate −2/−1 programmed ribosomal frameshifting; Manuscript in preparation.

3. **Emmely E. Treffers**, Ali Tas, Florine E.M., Scholte, Myrthe Van, Matthias T. Heemskerk, Arnoud H. de Ru, Eric J. Snijder, Martijn J. van Hemert[#], Peter A. van Veelen[#]. Temporal SILAC-based quantitative proteomics identifies host factors involved in chikungunya virus replication; Proteomics, 2015. 15(13): p. 2267-80.

4. Yanhua Li, **Emmely E. Treffers**, Sawsan Napthine, Ali Tas, Longchao Zhu, Zhi Sun, Susanne Bell, Brian L. Mark, Peter A. van Veelen, Martijn J. van Hemert, Andrew E. Firth, Ian Brierley, Eric J. Snijder, Ying Fang. Transactivation of programmed ribosomal frameshifting by a viral protein; Proc Natl Acad Sci U S A, 2014. 111(21): p. E2172-81.

5. Ying Fang[#], **Emmely E. Treffers**[*], Yanhua Li[*], Ali Tas, Zhi Sun, Yvonne van der Meer, Arnoud H. de Ru, Peter A. van Veelen, John F. Atkins, Eric J. Snijder[#], Andrew E. Firth[#]. Efficient -2 frameshifting by mammalian ribosomes to synthesize an additional arterivirus protein; Proc Natl Acad Sci U S A, 2012. 109(43): p. E2920-8.

6. Marion Phylipsen, Supawadee Yamsri, **Emmely E. Treffers**, Diahann T.S.L. Jansen, Warsha A. Kanhai, Elles M.J. Boon, Piero C. Giordano, Supan Fucharoen, Egbert Bakker, Cornelis L. Harteveld. Non-invasive prenatal diagnosis of beta-thalassemia and sickle-cell disease using pyrophosphorolysis-activated polymerization and melting curve analysis; Prenat Diagn, 2012. 32(6): p. 578-87.

[*]These authors contributed equally to this work.
[#]These authors contributed equally to this work.